Data Mining, i kamp mod besvigelser

- Data Minings berettigelse inden for intern kontrol, fokuseret mod opdagelsen af besvigelser

Cand.merc.aud. Kandidatafhandling
Copenhagen Business School
Magnus Stagsted – studienummer: S74107
Kim Røssel Petersen – studienummer: S91182
Vejleder: Karsten Andersen

Anslag: 253.206
Antal sider: 116 ekskl. forside, litteraturliste og bilag
Afleveringsdato: 15. januar 2020

Abstract

In recent years, the world has seen increased investments in the use of technology and data. One aspiring technology is Data Mining. This technology utilises machine learning in order to analyse vast amounts of data and discover patterns that are invisible to the human eye. This technology holds enormous potential for many areas. An area of outmost importance is fraud. Fraud is one of the greatest risks to society in terms of financial losses. The world's largest organisation in the area of fraud investigation, ACFE, estimates that the true cost of fraud could be as high as DKK 27.2 trillion per year.

In this thesis, we investigate if Data Mining can be the means that can limit fraud in corporate enterprises. In our research, we assume that the internal controls as performed today will not be sufficient in the fight against fraud. Therefore, we analyse if Data Mining can be part of the internal control environment and what it takes to implement the technology. The thesis is carried out as a literature review on both fraud and Data Mining. The strategy of the study design is chosen due to the sensitivity of the area.

Firstly, we narrowed the extensive fraud theory down to two areas which are fictitious revenue and procurement process fraud. In order to analyse the two identified fraud risks and to test if Data Mining can identify fraud, we did a large study of red flags in real fraud cases. By applying the concepts of Data Mining to these red flags, we found that the technology can uncover the trends and patterns that fraud holds. However, we found that the combination of variables was very important and that cross-validation between different techniques was needed. If not, false positives and false negatives would make the technology fallible.

By confirming that the technology would potentially be able to detect and prevent fraud, we discussed whether it was possible to integrate Data Mining within the enterprises' internal control environment. Firstly, we discovered that this can be implemented as a preventive control due to its placement in the IT structure as a backend tool. We found that specialists in Data Science were needed and that this would require extensive capabilities. Furthermore, we found that the technology requires vast investments in technology, processes and people. Therefore, we do not believe it is possible to implement in one single organisation. However, it is possible to implement the technology as part of the internal controls by outsourcing fraud detection as a managed service to a service provider. If enterprises outsource fraud detection, it will allow the provider to obtain scale benefits, educate people and build the needed processes around the technology and thereby limit fraud related losses in enterprises.

Indholdsfortegnelse

Kapitel 1 – Indledning

1.1 Indledning

I en verden, med stigende datamængder, hvor tilgængeligheden hertil aldrig har været større, begynder nye teknologier at opstå. Teknologierne er modsvaret af en stigende kompleksitet og en forøgelse af den mængde data, som er til rådighed for virksomhederne. Universiteter, virksomheder og datavidenskabsmænd investerer massivt og til stadighed i, hvordan anvendelse af data kan være med til at løse mange af de udfordringer, som verden står overfor i dag. Som et resultat heraf har et utal af nye og spirende teknologier gjort deres indtog hos de større virksomheder og rådgivningshuse.
En af de nye teknologier er Data Mining. Data Mining er en teknologi, som er i stand til at bearbejde og finde mønstre i store mængder data, som ellers kan være svære at identificere. Data Mining arbejder ud fra matematiske og statistiske algoritmer og kan derved gennemgå store mængder data med langt mindre menneskelig indblanding og arbejdskraft, end tidligere teknologier har været i stand til. Teknologien giver dermed mulighed for at arbejde med områder, som tidligere har været uoverskuelige. Et af de områder, hvor Data Mining har et stort fremtidigt potentiale, er inden for opdagelse af besvigelser.

Besvigelser er et af de problemer, som, set i historisk perspektiv, har haft store konsekvenser for virksomheder, samfund, investorer, leverandører, medarbejdere mfl. og dermed også for den globale økonomi. Virksomheder forsøger i dag at bekæmpe besvigelser ved anvendelse af interne kontroller. Interne kontroller omfatter i dag manuelle og automatiserede kontroller i form af opdagende eller forebyggende kontroller. Kontrollerne er ofte præget af godkendelsesprocedurer eller manuelle stikprøver og afdækker derfor sjældent alle transaktioner. Siden årtusindeskiftet har større virksomheder forsøgt at arbejde intensivt med anvendelse af dataanalyse, hvilket også er begyndt at slå igennem i virksomhedernes interne kontroller. Desværre bevidner flere nye sager inden for det offentlige, at der er foretaget så store og uhensigtsmæssige besparelser på de interne kontroller, at resultatet som konsekvens heraf har været flere store besvigelsessager. Sagerne omfatter svig i flere ministerier, herunder Britta Nielsen-sagen, Forsvarets Ejendomsstyrelse mfl. Disse sager omhandler ikke private virksomheder, da disse sager ofte mørklægges for at begrænse tabene og forsvare virksomhedernes renommé, men det betyder ikke, at problemet ikke eksisterer.

Besvigelser kan til dels skyldes ineffektive interne kontroller, som i Britta Nielsen-sagen, og at kontrollerne ikke er i stand til at opfange den stigende kompleksitet i besvigelser. Det står derfor klart, at besvigelser er et stort og meget aktuelt problem for mange virksomheder og organisationer.

Alene sidste år blev der globalt konstateret 2.690 besvigelsessager, som kostede de berørte virksomheder 48 milliarder kr., hvilket svarer til et gennemsnitstab på 18 millioner kr. for de enkelte virksomheder (ACFE, 2018). Da besvigelsessager i sagens natur i videst muligt omfang forsøges løst internt i virksomhederne for at undgå dårlig omtale, så er der potentielt tale om store, mørklagte tal. I rapporten estimeres de samlede tab at koste samfundet 27,2 billioner kr. årligt. Når besvigelsessager offentliggøres, og særligt den type, der vedrører regnskabsmanipulation, får disse typisk massiv omtale i medierne. Sådanne sager har store konsekvenser for virksomhederne, idet der kan opstå tvivl om ledelsens evne til at drive virksomheden. Besvigelser medfører derved store økonomiske tab for virksomhederne, både direkte såvel som indirekte.

Problemet rangerer højt på mange administrerende direktører og økonomidirektørers agenda, hvilket kommer til udtryk i en undersøgelse foretaget på vegne af Deloitte, hvor direktørerne svarer, at den perfekte revision burde fange langt flere besvigelser, end den gør i dag, og at ny teknologi bør anvendes i denne kamp (Reaseach, 2019). Når virksomhedslederne har et så stort ønske om, at en revision opdager flere besvigelser, antages det reelle ønske at være en reduktion af antallet af besvigelser. Ligeledes ønsker virksomhedslederne, at ny teknologi anvendes til opdagelse af besvigelser, hvilket vidner om, at de interne kontroller ikke har haft en tilstrækkelig effekt i deres nuværende form.

1.2 Problemformulering

Da besvigelser fortsat udgør et enormt problem, og der er et stort ledelsesmæssigt fokus på at få løst eller begrænset problemet, vil vi undersøge, om Data Mining-teknologien kan være løsningen på problemet. Ved at indsnævre besvigelser til at omfatte et begrænset undersøgelsesområde bestående af udvalgte besvigelsesrisici, vil vi med denne afhandling undersøge, om Data Mining vil kunne skabe et så finmasket net i virksomhedernes interne kontrolmiljø, at besvigelser vil kunne forebygges eller begrænses væsentligt. I forlængelse heraf vil vi undersøge, om det er muligt at udbrede teknologien og implementere den i virksomhedernes interne kontrolmiljø. Med udgangspunkt i problemfeltet kan afhandlingens problemformulering opsummeres til:

Hvordan kan virksomheder ved hjælp af Data Mining styrke deres interne kontroller tilstrækkeligt til at kunne afdække risikoen for besvigelsesrelaterede tab?

1.2.1 Undersøgelsesspørgsmål

For at undersøge problemformuleringen vil vi i afhandlingen besvare nedenstående undersøgelsesspørgsmål. Spørgsmålene er udarbejdet for struktureret at kunne udforme et grundlag, der kan anvendes til at konkludere på problemformuleringen. Undersøgelsesspørgsmålene er delt op i henholdsvis et teori- og et analyseafsnit.

I afhandlingens kapitel 2 vil vi belyse teorien inden for besvigelser, intern kontrol og dataanalyse. De tre afsnit vil stå alene, men vil i sammenhæng give læseren den fornødne forståelse af de problemstillinger og centrale begreber, som afhandlingens undersøgelsesområde omhandler.

I afsnittet om besvigelsesteori udledes de besvigelsesrisici, som har størst konsekvens for virksomhederne. Risiciene danner rammerne for undersøgelsen af Data Mining-teknologien som et værktøj til at forebygge eller opdage besvigelser. Desuden vil vi foretage en gennemgang af begreber og principper inden for intern kontrol. Dette skal bidrage til en vurdering af, om Data Mining kan indgå som en intern kontrol og i givet fald, hvordan en sådan aktivitet metodisk kan implementeres i et kontrolmiljø.

I teoriafsnittet vil vi desuden foretage en redegørelse for, hvordan der i dag arbejdes med dataanalyse, og hvordan Data Mining kan effektivisere denne proces. Afslutningsvist gennemgås forskellige teknikker inden for Data Mining for at give læseren den fornødne viden inden for et meget komplekst område, inden teknikkerne anvendes til analyseafsnittet. Nedenstående tre undersøgelsesspørgsmål vil således blive besvaret i dette kapitel.

- *Hvad karakteriserer de besvigelser, der har haft størst konsekvens for virksomhederne?*
- *Hvad er interne kontroller, og hvordan anvendes de i dag til at forebygge besvigelsers forekomst?*
- *Hvad er dataanalyse, og hvilke anvendelsesmuligheder har teknologien Data Mining.*

I afhandlingens tredje kapitel vil vi analysere anvendelse af Data Mining som et værktøj til opdagelse af besvigelser. Analysen vil omfatte en gennemgang af karakteristika for udvalgte besvigelsestyper og herefter hvilke datapunkter, som kan være relevante analysepunkter ved anvendelse af Data Mining-teknikker. Herefter vil vi foretage en vurdering af værktøjet som et integreret element i de interne kontroller. Afslutningsvis diskuteres fordele og ulemper ved Data Mining i forhold til klassiske interne kontroller. Analyseafsnittet vil sammen med teoriafsnittet skabe grundlaget for en besvarelse af afhandlingens problemformulering. I analysen vil vi besvare nedenstående undersøgelsesspørgsmål.

- *Hvordan kan Data Mining anvendes i virksomhedernes interne kontrolmiljø til imødegåelse af besvigelser?*
- *Hvilke udfordringer/begrænsninger er forbundet med anvendelse af Data Mining?*

1.3 Afgrænsning og begrebsfastlæggelse

Det er forventningen, at læseren af denne afhandling har et velfunderet grundlag inden for regnskab, revision og selskabslovgivning. Dertil er det forventningen, at læser har kendskab til de gængse begreber inden for intern kontrolteori. Der vil således ikke blive foretaget en dybdegående redegørelse for frameworks, lovtekster og lignende i afhandlingens teoriafsnit. Dybdegående redegørelser vil alene finde sted, hvor det forudsættes nødvendigt og er med til at danne rammerne for det videre arbejde. Som følge heraf foretages der løbende afgrænsning i takt med, at der redegøres for teorien, samt når der drages konklusioner. Læser bør altså være opmærksom på den løbende afgrænsning, som indgår i afhandlingens enkelte afsnit. Den løbende afgrænsning medvirker til at begrænse anvendelse af teori, som ikke vurderes relevant for den videre bearbejdning i analysen.

Nedenfor er præciseret, hvilke centrale afgrænsninger inden for afhandlingens delelementer, som læser bør være opmærksom på, ud over den løbende afgrænsning.

Virksomheder – Afhandlingen er afgrænset til alene at omhandle store virksomheder i form af klasse C Stor og opefter. Afgrænsningen foretages med henblik på kompetence- og ressourcebehovet for at kunne anvende dataanalyse som en del af kontrolmiljøet. Det er vores vurdering, at kontrolindsatsen for virksomheder, som ikke opfylder denne klassifikation, ikke vil stå mål med omkostningen herved. Således skal begrebet ”virksomheder” i afhandlingen læses som virksomheder i klasse C Stor og opefter. Dertil er det besluttet, at frasortere virksomhedsformer, hvor ejeren selv agerer øverste ledelse, NGO’er mv., som pga. deres struktur ikke anses for at være målgruppe for afhandlingens problemfelt.

Besvigelsestyper - Afhandlingen er afgrænset til alene at omhandle besvigelsestyperne ”faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen” og ”indregning af fiktiv omsætning”. Der vil blive redegjort for afgrænsningens relevans i teoriafsnittet for besvigelser. Se endvidere skematisering af denne afgrænsning i ”Bilag 2 – Skematiseret afgrænsning inden for Wells’ besvigelsestræ”.

Interne kontroller - Afhandlingens fokus er på anvendelse af dataanalyse og teknologi som en del af det interne kontrolmiljø - nærmere defineret i risikovurderingsprocessen - og som kontrolaktiviteter. Endvidere fokuseres der på opdagelse af besvigelser ved brug af interne kontroller. Med denne begrundelse har vi valgt at afgrænse afhandlingen fra brugen af det generelle framework for interne kontroller udarbejdet af COSO. Endvidere vurderes dataanalyse ikke at kunne skabe tilstrækkelige resultater i det overordnede interne kontrolmiljø, hvorfor overordnede interne kontroller kun behandles overfladisk. I afsnittet for interne kontroller foretages der løbende en specifik afgræsning af delområder.

Periode - For at afhandlingen skal være tidssvarende, foretages der en afgrænsning, således at der alene analyseres på besvigelser begået i perioden 2016-2017. Årsagen til valget af periode skal findes i, at perioden danner grundlaget for den seneste rapportering fra ACFE (Report To The Nations – 2018). Denne afgræsning er med til at sikre, at der udelukkende fokuseres på tidssvarende besvigelser, da dette element er afgørende for valg af teknologi.

Teknologi – I Afhandlingen vil vi udelukkende fokusere på den praktiske og teoretiske anvendelse af Data Mining. Der vil således ikke blive redegjort for matematikken og programmeringen bag de forskellige teknikker.

1.4 Begrundelse for valg af teori

Ved redegørelsen for besvigelsesteorien har vi aktivt forsøgt at belyse relevant teori, som klarlægger, hvad der kendetegner henholdsvis besvigere og besvigelser som én handling. Derfor indeholder teorien en redegørelse for udviklingen i besvigelsesteorien. Besvigelsestrekanten er medtaget for at belyse, at der skal eksistere en mulighed, før en besvigelse kan forekomme. For at uddybe selve besvigelsen redegøres der for en besvigelse som en gerning. Dette gøres for at belyse, at enhver besvigelse vil blive forsøgt skjult i højere eller mindre grad. Afslutningsvis foretages en gennemgang af MICE-modellen og The Fraud Diamond for at identificere, hvilke personer i en virksomhed der reelt kan begå en besvigelse.

Som en del af besvigelsesgennemgangen er anerkendte modeller som fx ”the fraud scale” fravalgt. Dette fravalg er baseret på, at vi i denne afhandling vil fokusere på dataanalyse i et perspektiv af transaktionsdata. Derfor vil integritetsmålinger og adfærdsanalyser ikke være relevante. Disse områder fravælges derfor i teorigennemgangen. Endvidere vurderer vi ikke, at disse områder vil tilvejebringe viden, som ikke allerede indgår i de tidligere omtalte modeller.

Teorigennemgangen er foretaget med det formål at identificere karakteristika inden for besvigelsesteorien, som kan danne grundlaget for afhandlingens analyse. I kombination med et studie af besvigelsers konsekvens for virksomhederne er afhandlingens problemfelt afgrænset til udvalgte besvigelsestyper. Herudover vil redegørelsen klarlægge den betydelige udvikling, som besvigelser er karakteriseret ved, hvilket vil danne udgangspunkt for relevansen af brug af mere agile dataanalyseværktøjer.

I anden del af teoriafsnittet redegøres der for intern kontrol og dennes anvendelsesformål. Denne redegørelse foretages for at danne rammerne og definere centrale begreber inden for intern kontrol, som senere anvendes i afhandlingens analyse. Valget af COSO's besvigelsesframework, modsat det generelle COSO-framework, begrundes i, at anvendelsesformålet ved frameworket og afhandlingen er det samme.

I den sidste del af teoriafsnittet behandles dataanalyse. Dataanalyseteorien er, som beskrevet i afsnittet om kildekritik, forbundet med selvlæring, selvlærte eksperter og et utal af forskellige teknikker. Redegørelsen for teorien inden for dataanalyse, og derigennem Data Mining og Machine Learning, foretages for at give læseren de fornødne forudsætninger for forståelse af et komplekst område, som behandles i afhandlingens analyseafsnit. Valget af teknologierne inden for Data Mining begrundes i Data Minings relativt tværfaglige anvendelsesområde, agilitet og stigende anvendelse. Modsat mere komplekse statistiske værktøjer eller de klassiske regelbaserede værktøjer belyser teorigennemgangen, at Data Mining-teknikkerne vil være bedre rustede til at håndtere den omskiftelighed, der karakteriserer besvigelser. I gennemgangen af teorien omkring Data Mining teknikker omtales association. Denne medtages udelukkende for at give læseren et komplet billede af anvendelsesområderne for teknikkerne. Selve teknikken anvendes ikke i analysen, da associationer mellem transaktioner allerede er givet ved litteraturstudiet omkring faresignaler ved de to besvigelsestyper.

1.5 Videnskabsteori

Den videnskabelige tilgang til afhandlingen bygger på en eksplorativ tilgang. Den eksplorative eller operationelle tilgang har til formål at undersøge sammenhænge og problemstillinger, som er mindre kendte. Tilgangen anvendes på områder, hvorpå der er begrænset viden og litteratur til bekræftelse af den slags mindre kendte sammenhænge. I afhandlingen søger vi at skabe sammenhænge mellem tre fagområder for derigennem at danne nye konklusioner. Tilgangen beskæftiger sig med logiske sammenhænge, og der skabes igennem disse sammenhænge ny viden.

I praksis betyder det, at vi i afhandlingen søger at skabe en sammenhæng mellem dataanalyse og opdagelse af besvigelser samt sammenhængens anvendelse inden for intern kontrol. Da tilgangen har været en kombination af induktion og deduktion, samt at problemfeltet for denne afhandling ikke er velundersøgt i den generelle litteratur, har afhandlingen ikke sit udgangspunkt eller base i en enkelt videnskabsteori. Afhandlingen udvikler sig således eksplorativt for at belyse observationer og bevæger sig herved mellem de forskellige videnskabsteorier. Tilgangen er anvendt for at kunne skabe nye sammenhænge uden at være fastlåst i den generelle teori eller i de enkelte hændelser.

Anvendelse af den eksplorative tilgang har medført, at problemfeltet for afhandlingen flere gange er blevet ændret, når nye og mere relevante vinkler på emnet er opstået. Et væsentligt eksempel herpå er, at vi i afhandlingen har arbejdet med en bred tilgang til dataanalyse. I takt med at der er fundet ny viden og sammenhænge, er emnet blevet tilpasset, således at Data Mining er blevet et centralt element i undersøgelsen. Et sådant tilvalg er udelukkende baseret på en rationalisering og betyder, at andre vinkler og elementer udelukkes. Et andet eksempel er inden for Data Mining teknikker. Ved gennemgangen af litteraturen var associations algoritmer et åbenlyst eksempel på en teknik som kunne identificere besvigelser. I takt med analysen blev udfoldet stod det dog klart, at algoritmen blev erstattet af litteraturstudiet omkring faresignaler ved besvigelser.

Når tilvalg foretages ud fra henholdsvis deduktion og induktion, betyder det, at tilvalget fremadrettet er den sandhed eller hypotese, som forsøges be- eller afkræftet i undersøgelsen. Et eksempel på anvendelse af en induktiv tilgang til undersøgelsen er fx en observation af besvigelseshændelser, hvor vi herefter antager, at karakteristika er gældende for lignende besvigelser. Deduktion har været anvendt, når vi har taget generelle konklusioner fra fx besvigelseslitteraturen og konkluderer, at disse vil være gældende ved nye besvigelser. På den måde kombineres induktion og deduktion for derved at kunne drage logiske konklusioner baseret på både generel teori og enkeltstående observationer.

1.6 Undersøgelsesdesign

For at undersøge problemfeltet har vi i overvejende grad baseret afhandlingen på et litteraturstudie. Dette undersøgelsesdesign er således fremgangsmåden for, hvordan data indsamles, analyseres og fortolkes (Andersen, 2009 s. 99).

Litteraturstudiet er valgt, da det undersøgte emne består af flere delemner. Delemnerne består af besvigelser, interne kontroller og dataanalyse med fokus på Data Mining. De tre emner er isoleret set

forbundet med omfattende litteratur, men det er sjældent, at alle tre emner knyttes sammen. Derfor vurderer vi, at kombinationen af en undersøgelse med et eksplorativt ståsted og anvendelse af et litteratur-review vil være det mest oplagte design til at skabe ny viden på området. Dette bekræftes af, at de tre emner hver for sig kræver meget omfattende kompetencer, hvorfor det vil være svært at finde tilstrækkelig empiri ved brug af interviews eller spørgeskemaer. Besvigelseselementet for undersøgelsen betyder desuden, at der ofte vil være store barrierer forbundet med dataindsamlingen. Virksomheder, som har været udsat for besvigelser, vil sjældent stå frem, hvorfor det ikke er muligt at foretage reelle casestudier.

1.7 Dataindsamling

Til undersøgelsen af problemfeltet har vi primært anvendt sekundær data. Vi har i forbindelse med vores eksplorative tilgang afholdt indledende interviews med ledende medarbejdere inden for Forensics, Risk Advisory, Machine Learning og Internal Controls fra et Big Four revisionshus. Endvidere har vi afholdt indledende interviews med næstformanden i den danske afdeling i ACFE og haft løbende korrespondance med Global Director of Research i ACFE. Fælles for vores indledende interviews var, at der ikke var nogle af eksperterne, der kunne udtale sig specifikt nok til, at de indsamlede primære data bidrog med viden, som ikke allerede var omfattet af vores sekundære dataindsamling i vores litteraturstudier. Således er de indledende interviews alene anvendt til at bekræfte vores indledende hypoteser om, at besvigelser er et problem, som ikke afdækkes tilstrækkeligt i en traditionel intern kontrol. Endvidere blev vi bekræftet i, at virksomhederne i dag arbejder med regelbaserede analyser frontend, og at der til stadighed er problemer med implementeringer af de traditionelle metoder pga. manglende kompetencer. Ved afslutningen af undersøgelsen har vi desuden sparret med de relevante personer for at sikre validiteten i de opnåede konklusioner.

Sekundær data anvendt i undersøgelsen omfatter artikler, rapporter, videoer, websider, bøger og metodiske frameworks. Sekundær data karakteriseres ved, at informationen er indhentet til et andet formål, end det vi undersøger. Da vi ikke har fundet rapporter eller afhandlinger, som undersøger korrelationen mellem de tre delemner, er afhandlingens kilder i overvejende grad af sekundær karakter.

Til brug for analysen har vi konstrueret et datasæt. Datasættet er baseret på de observationer, vi har gjort i forbindelse med et omfangsrigt litteraturstudie af besvigelser, karakteristika ved disse besvigelser og en vurdering af, hvordan besvigelserne fremtræder i virksomhedernes data. Til brug for udarbejdelsen har vi anvendt en teknik fra Data Mining, som hedder oversampling. Oversampling medfører, at man kreerer en overrepræsentation af datapunkter, som man ønsker at undersøge (Baesens, Vlasselaer and Verbeke,

2015). Det betyder, at der i datasættet er flere besvigelser, end der normalt ville være hos en enkelt virksomhed. Denne tilgang anvendes, da det er en af de mest anerkendte træningsmetoder til en algoritme. Endvidere anvendes tilgangen for at sikre, at læseren opnår en forståelse af, hvordan teknikkerne vil genkende mønstre i det undersøgte. Da der ikke er offentligt tilgængelige databaser over besvigelser, har det været nødvendigt at konstruere datasættet for at kunne analysere på de forskellige variables sammenhænge og resultatet af anvendelsen af Data Mining-teknikkerne. Da selve konstruktionen er foretaget på baggrund af de omfangsrige studier, vurderer vi, at datasættet kan anvendes til formålet. Desuden mener vi ikke, at selve datasættet har en påvirkning på konklusionerne, idet det er faresignalerne og teknikkerne bag, som er i fokus for analysen og dermed giver resultaterne.

1.8 Kildekritik

Da besvigelser er et følsomt emne, som typisk afklares internt i virksomhederne og uden politianmeldelse, har det ikke været muligt at indsamle primære data, som kunne bidrage til vidensproduktionen i undersøgelsen. Årsagen hertil er, at eksperter og ledende medarbejdere, der arbejder med disse sager, er omfattet af tavshedspligt og derfor ikke har kunnet udtale sig specifikt nok til, at vi kunne uddrage konklusioner af samtalerne.

Den manglende anvendelse af primær data betyder, at vi ikke direkte har deltaget i vidensproduktionen, der danner grundlaget for undersøgelsen. Derfor er kildekritikken relateret til sekundær data et vigtigt element i vurderingen af undersøgelsens validitet.

I vores undersøgelse er "Report to the Nations" af Association Of Fraud Examinators, herefter benævnt ACFE, et centralt element. Begrundelsen for anvendelsen af ACFE's rapport er, at ACFE er verdens største organisation inden for besvigelsesbekæmpelse med over 85.000 aktive medlemmer globalt. Organisationen har til formål at uddanne og undervise dens medlemmer, særligt virksomhedsledere, i, hvordan man kommer besvigelser til livs. Organisationen har arbejdet med besvigelser siden 1988 og anses således for at have oparbejdet en unik ekspertise inden for feltet. Grundet den unikke ekspertise og globale tilstedeværelse anses ACFE's arbejde med besvigelser som værende et validt grundlag for den videre bearbejdning af afhandlingens problemformulering.

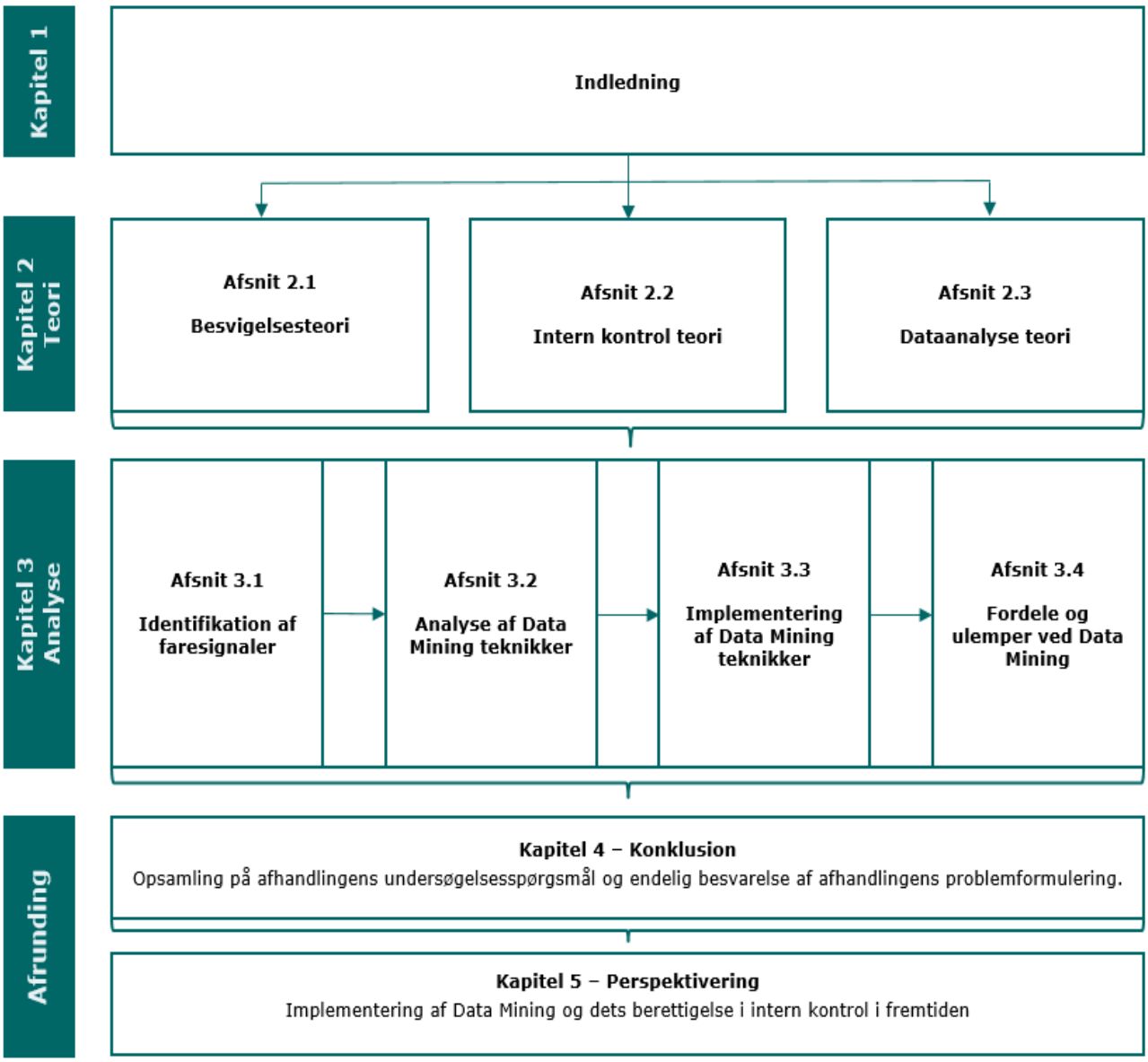
Ved anvendelse af de sekundære data har vi forsøgt at krydsvalidere resultater og observationer mellem forskellige kilder. Særligt "Report to the Nations" er et vigtigt element i vores undersøgelse. Vi har i forbindelse med anvendelse af resultaterne fra seneste udgave af "Report to the Nations - 2018"

sammenholdt resultaterne med 2017-udgaven af samme. Endvidere har vi sammenholdt denne med PWC's Global Economic Crime Survey 2018 (PWC, 2018), og sammenholdt konklusionerne i denne med ”Reports to the Nations” for at sikre validiteten af ”Reports to the Nations”-rapporten. Endvidere har vi gennemgået forudsætningerne bag undersøgelsen for at sikre, at der ikke er tale om overfladiske konklusioner. Vi har i den forbindelse vurderet, at når anerkendte organisationer anvender rapporten som grundlag for flere af deres studier, må den være af en sådan beskaffenhed, at den også kan danne et validt grundlag for vores videre undersøgelse inden for denne afhandlings problemfelt.

Øvrige kilder anvendt i undersøgelsen er baseret på forskellige mere eller mindre anerkendte artikler og publikationer. For at sikre validiteten af disse har vi søgt at anvende anerkendte organisationers og informationsudgiveres udgivelser som grundlag for undersøgelsen. Inden for dataanalyse, og særligt inden for Data Mining, har dette dog været svært, idet emnet er forbundet med selvlæring, selvudnævnte eksperter og praktisk anvendelse. Af den grund er mange kilder af en mere uofficiel karakter. For at sikre validitet af disse kilder har vi sammenholdt informationerne derfra med viden, som vi har opnået gennem undervisningsmateriale fra anerkendte universiteter, såsom MIT og Stanford University. Vi har således søgt at efterprøve, om kildernes indhold har en logisk sammenhæng og dermed bidrager til en valid vidensskabelse i afhandlingen. Afslutningsvis har vi krydsvalideret flere af vores observationer og konklusioner med udgivelser fra SAS Institute, da de er en af de førende vidensproducenter inden for dataanalyse.

1.9 Opgavens struktur

Nærværende afsnit viser indholdet i opgaven og har til formål at skabe overblik for læseren.



figur 1 - Afhandlingens struktur - egen tilvirkning

Kapitel 2 – Teori

AFSNIT 2.1 – BESVIGELSER

Dette teoriafsnit har til formål at give læseren en forståelse af, hvad en besvigelse er, og hvilke betingelser, der i teorien skal være opfyldt, for at denne kan forekomme. Endvidere vil det med afsæt i besvigelsesevolutionen blive belyst, hvor stor omskifteligheden er inden for tilsløringen og udførelsen af en besvigelse. Afslutningsvis vil læseren blive gjort bekendt med de besvigelser, der har størst konsekvens for virksomhederne, og hvordan de i teorien kan udføres. Hensigten med afsnittet er således sammen med de øvrige teoriafsnit at danne det teoretiske fundament for afhandlingens analyseafsnit, samt at give læseren kendskab til de to besvigelsestyper, der er genstandsfeltet den for videre analyse i afhandlingens analyseafsnit.

2.1 Besvigelsesteori

"*The use of one's occupation for personal enrichment through the deliberate misuse or misapplication of the employing organization's resources or assets*" (Wells, 2013 s. 8).

ACFE definerer en besvigelse som en handling, hvor en betroet person igennem sit ansættelsesforhold eller sin ledelsesmæssige position tilegner sig en personlig fordel ved misbrug af virksomhedens aktiver eller ressourcer ved at misbruge den tillid, personen er vist. Denne definition favner bredt og omhandler besvigelsestyperne misbrug af aktiver, korruption og regnskabsmanipulation. I dette afsnit vil vi redegøre for besvigelsesteori på udvalgte områder indenfor misbrug af aktiver og regnskabsmanipulation.

2.1.1 Besvigelsesmodeller – hvilke forhold skal være til stede, før en besvigelse kan forekomme

For at kunne skematisere og forstå, hvilke faktorer der skal være til stede, for at besvigelser kan forekomme, har der gennem de sidste mange årtier været foretaget studier i, hvad der kendetegner besvigere, og hvilke forhold der har medført, at besvigelser har kunnet forekomme.

Besvigelsestrekanten

Det første studie rettet mod kriminalitet begået i erhvervslivet blev foretaget af kriminolog og sociolog Edwin H. Sutherland. Han var den første til at studere økonomisk kriminalitet begået af autoritære og ledende medarbejdere i erhvervslivet. Den type kriminalitet blev i studiet beskrevet som "white-collar crimes", som er et udtryk, Sutherland 80 år efter stadig krediteres for (Dorminey *et al.*, 2012). Sutherlands

studie dannede grundlag for Donald R. Cresseys ph.d.-studie inden for netop samme emne i starten af 1970. Dette studie bestod i interviews af indsatte i Illinois State Penitentiary at Joliet, som afsonede domme for "white-collar crimes". På baggrund af Cresseys observationer opstillede han en hypotese om, at følgende forhold skal være til stede, førend en betroet medarbejder misbruger den tillid, som andre har vist ham:

- Der skal for besvigeren være opstået et økonomisk problem, som vedkommende ikke ønsker at dele med andre af frygt for at miste social status.
- Der skal for besvigeren være en åbning, som muliggør en afhjælpning af problemet i hemmelighed.
- Besvigeren skal være i stand til at rationalisere sine handlinger på en måde, der gør, at vedkommende kan retfærdiggøre det overfor sig selv.

Studiet med de opstillede hypoteser har dannet grundlag for det, der i dag inden for besvigelsesteorien betegnes som besvigelsestrekanten. Besvigelsestrekanten indeholder elementerne incitament, pres og mulighed som illustreret i figur 2.

figur 2 - Besvigelsestrekanten – egen tilvirkning – inspiration (Dorminey et al., 2012)

Besvigelseshandlingstrekanten (besvigelseselementerne)

Den kriminelle handling skal ses som et resultat af, at de tre elementer i besvigelsestrekanten er opfyldt. For at besvigelsen kan realiseres, skal der udføres bestemte handlinger som angivet i besvigelseshandlingstrekanten.

Følgende besvigelseselementer skal være opfyldt, førend besvigelsen kan blive en realitet:

Handling - repræsenterer udførelsen og metoden for besvigelsen, fx underslæb, forfalskning og regnskabsmanipulation.
Tilsløring - består i at skjule handlingen ved fx at forfalske fakturaer og bankafstemninger eller at destruere regnskabsmateriale.
Konvertering - er processen med at omsætte de uretmæssigt erhvervede midler til midler, som besvigeren kan anvende til legitime formål uden risiko for afsløring. Et eksempel herpå kan omfatte fx hvidvaskning af penge.

figur 3 - Besvigelseselementerne - egen tilvirkning med inspiration fra (Dorminey et al, 2012)

Besvigelseselementerne skaber værdi, da de repræsenterer specifikke handlinger, der kan dokumenteres med beviser, og dermed indeholder målbare kontrolpunkter, hvormed svig eller potentielt bedrageri kan forebygges, opdages eller afhjælpes (Dorminey *et al.*, 2012).

Som følge af den teknologiske udvikling og stigende samhandel på det globale marked er kompleksiteten og kreativiteten af besvigelser øget igennem tiden. Besvigelsestrekanten har derfor på trods af den internationale anerkendelse også sine svagheder. Særligt kan besvigelsestrekanten kritiseres for to svagheder:

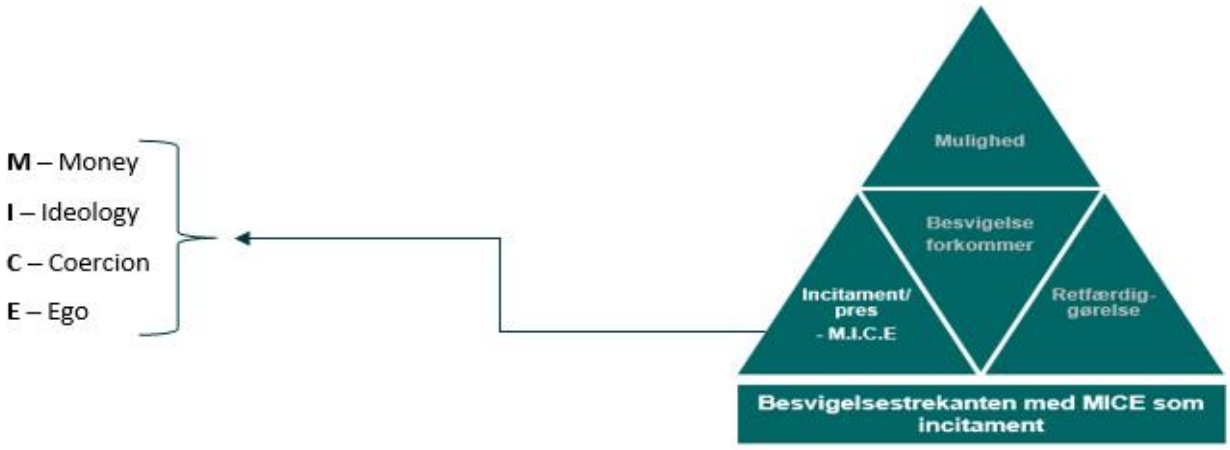
- Den er udviklet til at forklare besvigelser, efter de er sket, da særligt incitament og retfærdiggørelse ikke er målbare størrelser.
- Den er udviklet til at forklare besvigelser begået af mennesker med almindelige moralske værdier og ikke amoralske besvigere.

Besvigelsesteoriens udvikling har således medført en videreudvikling af besvigelsestrekanten. Den har derfor sit nuværende udgangspunkt i nedenstående modeller, der supplerer grundprincipperne i den oprindelige besvigelsestrekant:

- The Fraud Scale
- MICE-modellen
- The Fraud Diamond
- The New Fraud Diamond - som opfølgning på besvigere

Der vil i det efterfølgende ikke blive redegjort for "The Fraud Scale" eller "The New Fraud Diamond", da disse modeller omhandler persontræk og psykologiske aspekter hos besvigeren, der ikke anses for relevante for undersøgelsesområdet i denne afhandling.

MICE-modellen – senere i besvigelsesteoriens historie er elementet "økonomisk pres" i besvigelsestrekanten blevet kritiseret for at være ufyldestgørende, da der eksisterede utallige besvigelsessager, hvor besvigerne viste sig at være ressourcestærke og dermed ikke havde et økonomisk pres som incitament. Derfor er der i starten af årtusindskiftet gennemført flere studier, som har forsøgt at klarlægge, hvad der kendetegner disse besvigeres incitamenter. Disse studier resulterede i 2011 i M.I.C.E-modellens tilblivelse (Dorminey *et al.*, 2012). Modellen forklarer besvigers incitament ud fra følgende fire delelementer:



figur 4 - Besvigelsestrekanten med MICE som incitament - egen tilvirkning med inspiration fra (Dorminey et al, 2012)

Studier har i de seneste år vist, at incitamentet til at begå en besvigelse i udbredt grad skyldes grådighed og ego, hvorfor besvigere ikke længere kun kan beskrives som værende drevet af mulighed, retfærdiggørelse og økonomisk pres, kendt fra den klassiske besvigelsestrekant. I nogle tilfælde kan griskhed, tvang, moralsk habitus og ideologi, fx skatteunddragelse, terrorfinansiering eller besvigerens retfærdighedsfølelse også danne et incitament. MICE-modelen udvider altså definitionen af elementet pres i besvigelsestrekanten med delelementerne, penge, ideologi, tvang og ego, som illustreret i ovenstående figur (Dorminey *et al.*, 2012).

The Fraud Diamond - sætter et fjerde element på besvigelsestrekanten, da studier viser, at visse betingelser skal være opfyldt, for at en besvigelse kan forekomme: Besviger skal have psykiske ressourcer og faglige kvalifikationer til at udføre besvigelsen, bl.a. skal besvigeren være i stand til at håndtere den store stressfaktor, det er at hemmeligholde en besvigelse, have et stort ego, udvise en høj grad af selvsikkerhed, samt have de rette kompetencer til fx at kunne gennemskue forretningsgange og omgå kontrolsystemer.

Gennemgangen af ovenstående besvigelsesmodeller og udviklingen af besvigelsesteorien belyser, hvor mange forskellige faktorer, der spiller ind i en besvigelses tilblivelse. Udviklingen af de forskellige elementer understreger, hvorfor der for virksomheden er flere og flere faktorer, som den skal have for øje i forbindelse med dens risikovurderingsproces, som omtales yderligere i afsnit 2.2 om interne kontroller. For at kunne følge med udviklingen af besvigelsernes karakter bør virksomhederne derfor søge nye løsninger og modeller for interne kontroller for at være foran besvigerne.

Forebyggelse af besvigelser

Ovenstående redegørelse for udviklingen af besvigelsers kompleksitet viser, at alle virksomheder i større eller mindre grad er udsat for truslen om, at der er mulige besvigere i deres virksomhed. Denne observation underbygges af den anerkendte 10-80-10-regel udviklet af NASACT[1]. Reglen udspringer af adskillige besvigelsesstudier, hvor NASACT på baggrund af studierne estimerer, at (AGA Intergov, 2019b):

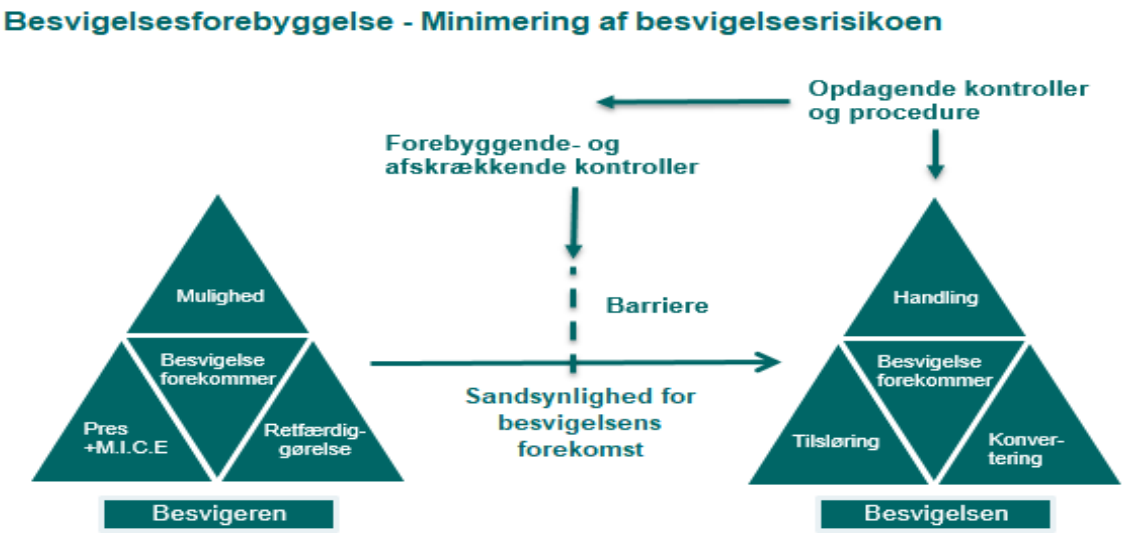
- 10 % af befolkningen aldrig vil begå en besvigelse, og at disse er personer er kendetegnet ved, at de vil gøre alt, hvad der står i deres magt for at tilbagelevere tabte ejendele til rette ejermand.
- 80 % af befolkningen vil begå besvigelser hvis den rette kombination af muligheder, pres og retfærdiggørelse er til stede.

[1] NASACT (National Association of State Auditors, Controllers and Treasurers) er en anerkendt amerikansk organisation for statsautoriserede revisorer, controllere og medarbejdere inden for risikostyring.

- 10 % af befolkningen aktivt kigger efter svagheder i systemer for at få muligheden for at udøve en besvigelse.

Således underbygger 10-80-10-reglen, at op mod 90% af befolkningen er potentielle besvigere, hvorfor det for enhver ledelse er vigtigt, at den rette kombination af muligheder, retfærdiggørelse og pres ikke opstår.

Potentielle besvigere kan dog være svære at identificere, idet besvigelsestrekantens punkter er svært mål- og observerbare. Som følge heraf er det virksomhedens opgave at opsætte barrierer mellem en mulig besviger og selve besvigelseshandlingen. Det er således kontrolorganets opgave at iværksætte forbyggende, opdagende og afskrækkende elementer, som kan minimere sandsynligheden for, at besvigelseselementerne i figur 3 kan opfyldes. Ligeledes har kontrolorganet en opgave i at designe procedurer, som kan opdage en allerede udført besvigelse. Der er således tale om, at kontrolorganet bør designe både forebyggende og opdagende kontroller for at mindske risikoen for, at en potentiel besviger begår besvigelsen.



figur 5 - Besvigelsesforebyggelse - egen tilvirkning med inspiration fra (Dorminey et al, 2012)

Elementer til forebyggelse af besvigelser – Som det fremgår af figur 5 "besvigelsesforebyggelse", er der tre elementer, som bør overvejes for effektivt at forebygge risikoen for, at besvigelseshandlingen kan udføres:

- Forebyggende kontroller
- Afskrækkelse
- Opdagende kontroller/procedurer

Forebyggende kontroller – for at forebygge at en besvigelse kan gennemføres, kan der designes interne kontroller, som medvirker til forebyggelse og afskrækkelse. En klassisk kontrol, som er hyppigt anvendt, er funktionsadskillelse, som kræver mere end én medarbejder til at godkende en særlig risikofyldt handling (Dorminey *et al.,* 2012). Interne kontroller vil blive behandlet yderligere i afsnit 2.2. Endvidere kan følsomhed over for besvigelser og etisk forretningskultur medvirke til, at en besviger i rationaliseringsfasen ikke kan rationalisere sig frem til, hvorfor det er moralsk acceptabelt at begå besvigelsen (Dorminey et al., 2012). Således kan læring og bevidsthed om etisk adfærd også have en forebyggende effekt.

Afskrækkelse – i nær tilknytning til etisk forretningskultur kan tiltag som whistleblower-ordninger, overvågning og code of conduct have en afskrækkende effekt over for en eventuel besviger (Dorminey et al., 2012). De to mest effektive afskrækkelsesformer anses at være frygten for at blive opdaget og frygten for den straf, som forbrydelsen kan medføre (Dorminey et al., 2012). Som nævnt i afhandlingens afgrænsning har afhandlingen ikke fokus på overordnede interne kontroller, hvorfor dette element ikke omtales yderligere i dette afsnit. Overordnede interne kontroller er således alene kort omtalt som et forebyggende element i COSO-afsnittet i kapitlet om interne kontroller.

Opdagende kontroller og procedurer – Som supplement til forebyggende kontroller kan virksomheden designe og implementere opdagende kontroller. Svagheden ved denne type kontrol er imidlertid, at den først kan opdage en besvigelse, efter at den er begået. Opdagende kontroller er typisk mandskabstunge og stikprøvebaserede. Således kan en opdagende kontrol være udformet som stikprøvevis gennemgang af udbetalinger mv. Nogle typer besvigelser har en sådan karakter, at forebyggende interne kontroller ikke kan afdække dem. Eksempler herpå er management override og besvigelser begået af flere ansatte i forening. Således er det vigtigt at have designet nogle opdagende interne kontroller, som kan identificere den type besvigelser. I forlængelse heraf anbefales det blandt andet, at kontrolorganet mfl. sørger for at (Dorminey et al., 2012):

- bevare deres professionelle skepsis
- øge kontrolorganets eller andre relevante parters forståelse af virksomheden
- løbende at foretage en risikovurdering af besvigelsesrisici i virksomheden

Hele processen, som illustreret i figur 5, skal ses som en iterativ-, dvs. gentagen, cirkulær, proces, hvor man på baggrund af erfaringer fra de opdagende kontroller styrker de forebyggende kontroller og dermed også afskrækkelsesmomentet. En kombination af ovenstående elementer kan på den måde bidrage til, at der sættes stærke barrierer op mellem forholdene i besvigelsestrekanten og besvigelseselementerne. Med

stærke barrierer er der større sandsynlighed for, at elementerne i besvigelsestrekanten ikke kan opfyldes, og at selve gerningen, dvs. besvigelsen, derfor ikke kan udføres(Dorminey et al., 2012).

COSO[2]-frameworkets grundprincipper og tilblivelse udspringer af besvigelsesudviklingen og er blevet det mest udbredte framework inden for intern kontrol og derigennem i kampen mod tilsigtede såvel som utilsigtede fejl. Meta-modellen i figur 5 illustrerer grundtankerne og fundamentet for COSO-frameworket. COSO-frameworket redegør metodisk for, hvordan en virksomhed bør designe og implementere et effektivt internt kontrolmiljø til blandt andet at afdække sandsynligheden- og dermed risikoen for, at besvigelseselementerne kan opfyldes. Der vil blive redegjort for relevante frameworks udarbejdet af COSO i afsnittet om interne kontroller.

2.1.2 Besvigelser med størst konsekvens

Kapitlets andet afsnit har til formål at belyse, hvilke former for besvigelser, der oftest forekommer globalt, og hvilken konsekvens de har for virksomhederne. De besvigelsesformer, der har største konsekvens for virksomhederne, vil blive identificeret ud fra den seneste rapportering fra ACFE og vil efterfølgende være genstandsfelt for videre analyse og bearbejdning i afhandlingens efterfølgende kapitler.

ACFE udgiver årligt "Report to the Nations", som beskriver faktuelle besvigelsessager fra det forgangne år, samt hvilken konsekvens de har haft på virksomhederne globalt. Rapporten er offentligt tilgængelig og er et globalt studie i erhvervsmæssige besvigelser, der giver et øjebliksbillede af, hvilke udfordringer virksomheder, regeringer, NGO'er mv. står over for.

"Report to the Nations" har til formål at give detaljeret information inden for fem følgende kritiske områder:

- Metoder, der anvendes til at begå besvigelserne
- Midler, der er anvendt til at opdage besvigelserne
- Karaktertræk ved de organisationer, der er ofre for besvigelser
- Karakteristika ved de personer, som udfører besvigelserne

[2] COSO er en forkortelse for The Committee of Sponsoring Organizations og blev grundlagt for at assistere den nationale kommission i USA med viden om besvigelser inden for finansiel rapportering og faktorerne bag (The Committee of Sponsoring Organizations' (COSO), 2018).

- Konsekvens af de begåede besvigelser, og hvilke besvigere der er blevet identificeret

Til grundlag for udarbejdelsen af "Report to the Nations" er der i den seneste rapport foretaget analyse af 2.690 konstaterede besvigelser efterforsket i perioden januar 2016 til oktober 2017. Analysen omfatter hændelser i 125 lande og 23 industrier. Rapporten er den 10. udgave og anses af ACFE for at være en repræsentativ analyse, der giver et øjebliksbillede af omfanget af besvigelser globalt.

Fremhævelser fra ACFE's - Report to the nations 2018

Dette afsnit har til formål at fremhæve væsentlige konklusioner fra "Report to the Nations" for at kunne danne et overblik over:

- Økonomiske konsekvenser, som besvigelser medfører globalt
- Hvilke typer virksomheder, der rammes af besvigelser og konsekvensen heraf
- Konsekvensen af de besvigelser, der forekommer i Vesteuropa
- Hvem det er, der begår besvigelserne
- Hvilke besvigelser, der er hyppigst forekommende og har størst økonomisk konsekvens

Hensigten med fremhævelserne er at underbygge påstanden om, at besvigelser er et reelt problem og at identificere, hvilke besvigelser der på nuværende tidspunkt udgør det største problem. Dermed vil de danne genstandsfelt for afhandlingens analyseafsnit.

Økonomiske konsekvenser

I perioden er der konstateret besvigelsesrelaterede tab for over 45,6 milliarder DKK. Mediantabet pr. sag udgør 848.000 DKK (ACFE, 2018). I ACFE's analyse har man frasorteret den øverste og nederste percentil af datasættet for at rense det for ekstremer og dermed bedre at kunne tegne et mere generelt billede af besvigelsernes omfang. Medtager man ekstremerne, udgør det gennemsnitlige tab i analysen 18 mio. DKK pr. virksomhed (ACFE, 2018).

ACFE udtaler, at det totale omfang af besvigelser i analyseperioden er ukendt, da studiet alene omfatter de kendte sager i analyseperioden. ACFE antager dog, at de 45,6 milliarder DKK alene udgør en lille brøkdel af de tab, som besvigelser har medført i analyseperioden (ACFE, 2018). Det er organisationens vurdering, at omfanget af besvigelser årligt beløber sig til 5% af det globale bruttoprodukt, svarende til 26 billioner DKK.

I 45% af de konstaterede sager udgør det besvigelsesrelaterede mediantab over 1,3 mio. DKK. Hertil kommer, at der er en klar overvægt af sager, som har medført et mediantab på over 6,5 mio. DKK svarende til 22% af sagerne (ACFE, 2018, pp. 4 & pp. 9).

Ser man på varigheden af de perioder, som besvigelsessagerne løber over, varierer disse meget alt efter, hvilken type besvigelse der er begået, og hvor meget besvigeren har gjort for at skjule besvigelsen.

Af analysen fremgår det, at besvigelserne typisk står på i 16 måneder før at de opdages. På baggrund af omfang og varighed må det altså konstateres, at besvigelser er genstand for store tab i virksomhederne og er et reelt problem, som bør have en større, hvis ikke den største bevågenhed fra den øverste ledelses side.

Besvigelsestyper med størst konsekvens – Den hyppigst forekommende besvigelse er misbrug af aktiver - den er repræsenteret i 89% af alle besvigelsessagerne[3]. Mediantabet for misbrug af aktiver udgør 743.000 DKK. Den mindst forekommende besvigelsestype derimod er regnskabsmanipulation, som er repræsenteret i 10% af alle sagerne[4]. Mediantabet for regnskabsmanipulation udgør 5,2 mio. DKK (ACFE, 2018). På trods af at misbrug af aktiver er den uden sammenligning hyppigst forekommende besvigelse, er det også den, der udgør det mindste tab for virksomhederne. Således forekommer regnskabsmanipulation i mere enkeltstående tilfælde, men når de indtræder, udgør de altså et væsentlig større tab for virksomhederne.

Fordeling mellem geografiske områder – Studiet viser, at USA står for 48% af alle de konstaterede besvigelsessager i perioden og har således signifikant flere besvigelsessager end de øvrige geografiske områder. Det er dog værd at bemærke, at Vesteuropa, på trods af at området alene står for 6% af de konstaterede sager (178 ud af 2.960), har et mediantab, der er dobbelt så højt som USA. Det må således konstateres, at selvom besvigelsessagerne er mindre hyppigt forekommende i Vesteuropa, så har de en væsentlig større konsekvens, når de forekommer (se grafik i Bilag 3 – Geografisk fordeling af konstaterede besvigelser).

[3] Denne statistik omfatter besvigelser, hvor misbrug af aktiver har været en del af besvigelsen i kombination med andre typer besvigelser.

[4] Denne statistik omfatter besvigelser, hvor regnskabsmanipulation har været en del af besvigelsen i kombination med andre typer besvigelser.

Da afhandlingen primært henvender sig til den danske læser, vil afhandlingens videre fokus være at analysere besvigelser i Vesteuropa - og særligt i Danmark, da det antages, at læseren vil have en tilknytning til- og interesse i dette geografiske område.

Personerne bag besvigelserne i undersøgelsen – ACFE har fundet, at 19% af alle de konstaterede besvigelser er begået af ledelsen. Der tegner sig altså et billede af, at størstedelen af besvigelserne er begået på medarbejderniveau. Mediantabet for besvigelser begået af ejere/ledelse udgør 5,5 mio. DKK. Det må antages, at mediantabet er så stort som følge af, at den besvigelse, der typisk forekommer på ledelsesniveau, bærer præg af regnskabsmanipulation og sammensværgelse. Regnskabsmanipulation er den besvigelsestype med de største enkeltstående tilsigtede tab, da der typisk skal forekomme større økonomiske ændringer, for at de kan have indflydelse på og fremgå af årsregnskabet og derfor ændre regnskabslæserens opfattelse af regnskabet.

Besvigelser begået i sammensværgelser medfører et langt større tab end de besvigelsestyper, hvor besviger har arbejdet alene (ACFE, 2018). Det fremgår af rapporten, at mediantabet i sager, hvor besviger har arbejdet alene, udgør 483.000 DKK, hvorimod besvigelsessager begået af to i forening har medført et mediantab på 978.000 DKK. Endelig er mediantabet 2,2 mio. DKK i de besvigelsessager, hvor besvigelserne er foretaget af en sammensværgelse mellem tre eller flere medarbejdere. Der er således en signifikant forskel i besvigelsens omfang alt efter antallet af medskyldige i den begåede besvigelse. Årsagen hertil må alt andet lige antages at være, at besvigelserne står på i en længere periode, og at der sandsynligvis vil være ledende medarbejdere indblandet i en sammensværgelse mellem flere medarbejdere. Endvidere kan der være en sammenhæng mellem, hvilken kontrolfunktion de forskellige kontrolindehavere har i forhold til muligheden for at omgå de interne kontroller.

Oftest forekommende besvigelser – I de 2.690 sager, som ACFE har efterforsket i deres seneste udgave af "Report to the Nations", fremgår det, at flere besvigelsessager består af et sammensurium af flere typer besvigelser. Af undersøgelsen fremgår det, at misbrug af aktiver som enkeltstående besvigelse udgør 57% af alle besvigelsestyper, hvorimod regnskabsmanipulation alene udgør 1% (Bilag 4 – Forekomstfordeling imellem besvigelseshovedkategorierne). Således udgør næsten en tredjedel af alle de efterforskede besvigelser en kombination af to eller tre af besvigelseskategorierne: korruption, regnskabsmanipulation eller misbrug af aktiver.

Klassifikation af besvigelsestyper inden for misbrug af aktiver – Da misbrug af aktiver udgør så stor en del af de samlede besvigelsessager, har ACFE i rapporten analyseret, hvilke underkategorier af besvigelser, der bliver begået, herunder hyppigheden og omfanget af disse. I "Bilag 5 – Heatmap - Mest risikobetonet besvigelsestype" er de forskellige underkategorier kategoriseret for at visualisere, hvor stor en risiko de enkelte besvigelsesformer udgør for virksomhederne. Heraf fremgår det, at de hyppigst forekommende besvigelser er:

- Tyveri af aktiver (21%)
- Fakturasvindel i forbindelse med indkøb (20%)
- Tyveri af kontanter (15%)

Samtidig fremgår det af analysen, hvilke besvigelser, der har det største mediantab pr. besvigelse. Heraf udgør nedenstående 3 besvigelsestyper dem med størst mediantab:

- Check- og pengeoverførselsbesvigelser (978.000 DKK)
- Fakturasvindel i forbindelse med indkøb (652.000 DKK)
- Tyveri af aktiver (639.000 DKK)

Svindel med checks og pengeoverførsler er således den besvigelsestype, der medfører størst økonomisk tab, men samtidig forekommer i en væsentlig lavere grad end henholdsvis svindel med fakturaer i forbindelse med indkøbsprocesser og tyveri af aktiver. Ergo må det konkluderes, at "faktureringssvindel i forbindelse med indkøbsprocessen" og "tyveri af aktiver" er de to typer besvigelser, der har størst betydning for virksomhederne.

Ovenstående er forekomsten på globalt plan. I Vesteuropa tegner der sig et billede svarende til ovenstående. Her konkluderer ACFE, at den hyppigst forekommende besvigelse i Vesteuropa er "faktureringssvindel i forbindelse med indkøbsprocessen", som udgør 28% af alle sager (ACFE, 2018).

Forekomst af besvigelser inden for regnskabsmanipulation - Statistikkerne bag "Report to the Nations - 2018" viser, at den hyppigst forekommende form for besvigelser inden for regnskabsmanipulation er indregning af fiktiv omsætning. For besvigelser inden for regnskabsmanipulation udgør fiktiv omsætning ifølge den seneste rapport 34% af alle de konstaterede sager (McNeal, 2019). I den anden ende af spektret ligger periodiseringssvindel, som udgør 16% af de konstaterede besvigelser. Da denne type svindel er den

mindst forekommende, vælger vi ikke at beskæftige os yderligere med den, men at have fokus på de øvrige typer.

Er besvigelser et reelt problem, der bør investeres i?

På baggrund af seneste udgave af "Report to the Nations - 2018" kan det konkluderes, at besvigelser er et stort økonomisk problem for den globale økonomi såvel som for organisationerne. ACFE anslår, at besvigelser årligt koster den globale økonomi 47 billioner DKK svarende til 5% af det globale bruttoprodukt (GWP) i 2017.

I rapporten fra ACFE fremgår det, at misbrug af aktiver er den hyppigste forekommende besvigelsesform. Rapporten konkluderer, at 89% af alle besvigelsessager relaterer sig til misbrug af aktiver, og at disse har et mediantab på 743.000 DKK. Den mindst forekommende besvigelsesform er regnskabsmanipulation, som alene udgør 10% af besvigelsessagerne, hvoraf mediantabet dog udgør 5,5 mio. DKK. Det må derfor antages, at selvom mediantabet for misbrug af aktiver alene udgør 743.000 DKK, så er denne altså ligeså bekostelig som regnskabsmanipulation, da denne er ni gange oftere forekommende.

Af statistikkerne bag "Report to the Nations" fremgår det, at fakturasvindel er den hyppigst forekommende besvigelse inden for misbrug af aktiver og udgør 20% af alle sager relateret hertil. Indregning af fiktiv omsætning udgør 34% af de konstaterede besvigelser relateret til regnskabsmanipulation og er dermed den mest forekommende inden for denne type af besvigelser.

Således kan det konkluderes, at besvigelser har store økonomiske konsekvenser og som resultat heraf bør have ledelsesorganernes opmærksomhed. På baggrund af ovenstående resultater vil "fakturasvindel i forbindelse med indkøbsprocessen" og "indregning af fiktiv omsætning" være genstandsfelt for afhandlingens videre analyse og anbefaling vedrørende anvendelse af ny teknologi til forebyggelse af netop disse to besvigelsestyper.

2.1.3 Typer af besvigelser

Vi vil i dette afsnit redegøre for teorien bag de typer besvigelser, der, efter vores gennemgang af "Reports to the Nations - 2018", har vist sig at have størst betydning og forekommer hyppigst.

De forskellige besvigelsesformer, der historisk set er forekommet i erhvervslivet, er kategoriseret i Joseph T. Wells besvigelsestræ, som angivet i "Bilag 2 – Skematiseret afgrænsning inden for Wells' besvigelsestræ". I de følgende afsnit vil der blive redegjort for besvigelsestyperne "Misbrug af aktiver" og "Regnskabsmanipulation".

2.1.3.1 Misbrug af aktiver

Den første del af Joseph T. Wells' besvigelsestræ består af den type af besvigelser, der klassificeres som misbrug af aktiver.

Inden for misbrug af aktiver forekommer der to undergrupper af besvigelser, som benævnes som misbrug af kontanter og misbrug af andre aktiver. Under disse grupperinger specificeres de forskellige besvigelsesformer. Vi vil som tidligere benævnt analysere og belyse de mest forekommende besvigelser med størst konsekvens for virksomhederne. Under misbrug af aktiver vil der således blive redegjort for begrebet "faktureringssvindel i forbindelse med indkøbsprocessen" i kategorien svigagtige udbetalinger.

Svigagtige udbetalinger, herunder faktureringssvindel i forbindelse med indkøbsprocessen

Svigagtige udbetalinger er en undergruppe til misbrug af aktiver og er kendetegnet ved, at besviger bevirker, at den forurettede virksomhed foretager en betaling af en svigagtigt faktura for fx fiktive varer og ydelser, personlige indkøb eller oppustede fakturaer.

Fælles for dem alle er, at der forekommer en svigagtig handling fra en person i virksomheden, som søger egen vinding på bekostning af virksomheden.

Besvigelser i forbindelse med indkøb kan (faktureringssvindel) kategoriseres i tre typer af besvigelser (Wells, 2013):

1) skyggeselskaber
2) ikke-medskyldige leverandører
3) personlige indkøb på virksomhedens regning. Denne type kommenteres ikke yderligere, da den beløbsmæssige konsekvens i henhold til ACFE rapport ikke er væsentlig.

Ad. 1) Skyggeselskaber

Et skyggeselskab er et selskab, hvis formål er at maskere transaktioner og ikke i sig selv have reel drift. Skyggeselskaber er alene oprettet til det formål at begå besvigelser. Således vil besviger typisk oprette et

fiktivt navn for virksomheden med en tilknyttet postboksadresse. Da virksomheden vil være oprettet med det formål at opsamle svigagtige udbetalinger, vil der også være oprettet en bankkonto i virksomhedens navn med besviger som fuldmagtshaver hertil.

Sådan dannes et skyggeselskab – Som tidligere omtalt er eksistensgrundlaget for et skyggeselskab i forbindelse med svigagtige udbetalinger, at der er oprettet en bankkonto i virksomhedens navn, hvortil den svigagtige udbetaling kan ske til.

Besvigeren vil være nødsaget til at registrere sin virksomhed i den stat, hvor virksomheden er hjemmehørende. I gældende dansk lov foreskrives det, at virksomheder, jævnfør henholdsvis selskabslovens § 9 (Folketinget, 2019b) og loven om visse erhvervsdrivende virksomheder § 8 (Folketinget, 2019a), registreres hos erhvervsstyrelsen. Ved stiftelse skal virksomheden blandt andet oplyse om:

- Navn, bopæl og cpr-nr. på virksomhedens stiftere
- Hvem der er tegningsberettiget
- Virksomhedens adresse
- Formål mv.

Det er værd at bemærke, at alle lande ikke har lige så skærpede registreringskrav som Danmark. Således vil det være muligt for besviger at oprette skyggeselskabet i udlandet, hvor der ikke er samme registreringskrav ved stiftelse af selskab, fx med en postboksadresse for at mindske opdagelsesrisikoen (Wells, 2013). I Danmark og Europa er der indført krav om registrering af reelle ejere for at skabe transparens omkring selskabers konstruktioner og disses ejere.

Udstedelse af falske fakturaer – Når man som besviger har oprettet et skyggeselskab og en dertilhørende bankkonto, skal der fabrikeres falske fakturaer for at kunne fuldføre besvigelsen. Besvigeren udformer den svigagtige faktura og adresserer den til den virksomhed, der er målet for besvigelsen. Herefter kommer fakturaen til godkendelse og endelig betaling. For at fakturaen skal kunne godkendes, gælder det, at besviger har kendskab til det interne kontrolsystem for at kunne omgå det og dermed undgå at blive opdaget.

Godkendelse af fakturaer fra skyggeselskaber – Godkendelsen af disse fakturaer kan ske efter følgende metoder:

Selv-godkendelse – Ved denne type besvigelse er besviger kontrolindehaver og har mulighed for at godkende indkøbet, uden at andre attestationsberettigede skal godkende fakturaen. For at kunne udføre denne form for godkendelse vil det alt andet lige kræve, at besviger er placeret i en ledende stilling og således er øverst i godkendelseshierarkiet (Wells, 2013). Godkendelsen kan også forekomme, hvis virksomhedens godkendelseskontrol er forkert figureret.

Omgåelse af funktionsadskillelse (sammensværgelser) – Denne fakturagodkendelse forekommer i interne kontrolsystemer, hvor et indkøb kræver dobbeltgodkendelse. En sådan godkendelse kendetegner et stærkt internt kontrolsystem. Således vil besvigelser begået af medsammensvorne være svære at opdage, da et sådan kontrolsetup giver ledelse og andre kontrolinstanser anledning til at tro, at medarbejderbesvigelser sjældent forekommer. Et sådan setup fordrer nemlig, at besvigeren har en medsammensvoren, som besvigeren kan stole på, da det kræver, at der er flere, som skal dække over besvigelsen (Wells, 2013).

"Gummistempel"-supervisors – Denne type godkendelse er kendetegnet ved, at kontrolindehaveren ikke er kompetent eller ikke tager sin kontrolindehaver-rolle alvorligt. Har man som besviger kendskab til en attestationsberettiget medarbejder, som godkender fakturaer ukritisk, kan besviger altså udnytte denne kontrolsvaghed og adressere sine fakturaer til denne kontrolindehaver i håbet om, at denne ikke forholder sig til fakturaen (Wells, 2013).

Normalgodkendelse – Denne form for godkendelse må alt andet lige antages at være den mest risikable besvigelsesudførelse. Udførelsen heraf består i, at besviger blot sender fakturaen til godkendelse i det normale godkendelsesflow og håber på, at fakturaen ikke bliver opdaget. Sådan en besvigelse vil alt andet lige fordre, at fakturaen er af mindre, ubetydelige beløb eller leverancer, således at den drukner i mængden af øvrige fakturaer, som ligger til godkendelse (Wells, 2013).

Typiske leverancer fra skyggeselskaber – ACFE angiver, at den oftest forekommende "indkøbssvindel" består af indkøb af immaterielle ydelser, da den type mindsker sandsynligheden for, at besvigelsen opdages. De typiske karakteristika ved skyggeselskabsbesvigelser er, at der ikke foreligger en egentlig leverance/ydelse/vare bag den fremsendte faktura. Således vil der mangle en lagerbeholdning i forbindelse med fakturering af en råvare eller et produkt. ACFE har konstateret, at det typisk er konsulentydelser, der faktureres fra skyggeselskaberne, da det er svært eller umuligt at måle, om virksomheden har modtaget en ikke-fysisk vare (Wells, 2013). I tilfælde, hvor der faktureres for en reel vare, ligger besvigelsen typisk i, at

faktureringen sker til overpriser. Den form for besvigelse er kendt som "pass-through schemes" og foregår ved, at skyggeselskabet indkøber en reel vare fra en underleverandør og videresælger den til den forurettede virksomhed til overpris, hvorved besvigeren opnår en uberettiget gevinst (Wells, 2013).

Hvordan besvigelser, begået ved udbetaling til et skyggeselskab i indkøbsprocessen, kan forebygges, og hvilke advarselssignaler der giver anledning til inspektion, vil blive videre behandlet i afsnit 2.2.3 om interne kontroller.

Ad 2) Ikke-medskyldige leverandører

Ikke-medskyldige leverandører er en anden form for svigagtig udbetaling i forbindelse med fakturering. Denne form for besvigelse kan forekomme i forskellige afskygninger. Fælles for dem er, at den pågældende faktura bevidst bliver fejlbehandlet af besviger. Fejlbehandlingen kan ske på følgende måder:

Dobbeltbetaling af samme faktura – Denne type besvigelse er udformet ved, at besviger bevidst betaler samme faktura to gange og efterfølgende retter henvendelse til leverandøren og beder denne om at tilbageføre fejlindbetalingen til besvigerens eget kontonummer.
Bevidst overbetaling af leverandør – Denne besvigelse følger samme metodik som ovenstående besvigelsestype blot med den modifikation, at i stedet for at udføre en dobbeltbetaling, betales der blot et beløb, der er en lidt højere end fakturabeløbet, og derefter søges overbetalingen udbetalt til besvigers eget kontonummer.
Betaling til forkert leverandør – Denne besvigelse har samme formål som ovenstående. Forskellen her er bare, at besviger foretager en betaling til en forkert leverandør for efterfølgende at søge få betalingen retur til eget kontonummer.
Betaling af gammel eller forfalsket faktura med nyt kontonummer – Denne besvigelse adskiller sig fra de øvrige ved, at besviger finder en gammel faktura og forfalsker den med eget eller en medskyldiges kontonummer.

2.1.3.2 Regnskabsmanipulation

Ved regnskabsmanipulation fragår der, i modsætning til misbrug af aktiver, ikke nogen reel værdi fra virksomheden (Bloch Christensen, Kristensen and Warming-Rasmussen, 2015, p. 15). Der kan dog, som en afledt effekt af regnskabsmanipulation, fragå en reel værdi. Eksempler herpå kunne være, at målet for besvigelsen er, at besviger kan opnå en resultatafhængig løn eller lignende.

Regnskabsmanipulation forekommer i sin hyppigste form som følge af fiktive, manipulerede og/eller manglende posteringer eller oplysninger i virksomhedens regnskab. Regnskabsmanipulation forekommer typisk på ledelsesniveau og har til formål at vildlede regnskabsbruger (Bloch Christensen, Kristensen and Warming-Rasmussen, 2015).

Der kan forekomme forskellige former for incitamenter til at begå regnskabsmanipulation. Regnskabsmanipulation kan have til formål både at overdrive og at underdrive virksomhedens performance. Nedenfor er angivet eksempler på incitamenter til over- såvel som underdrivelse:

Ledelsens incitament til at overdrive performance (Wells, 2013, pp. 275–276)

- Imødekomme aktionærers og ejeres forventninger
- Imødekomme personlige performance-kriterier
- Personlige motiver såsom ego, social status mv.
- Resultatafhængig bonusaflønning
- Opnå lån eller overholde lånevilkår
- Kurspleje af virksomhedens aktier, som i sagen vedrørende PARKEN Sport & Entertainment A/S
- Få virksomheden til at fremstå mere attraktiv, end den er i forbindelse med frasalg

Ledelsens interesser i at underdrive performance (Wells, 2013, pp. 275–276)

- Skattesvig
- Udskyde overskud til næste regnskabsperiode
- Resultatudjævning
- Ledelsesskift midt i regnskabsperioden
- Reducere forventninger, således at fremtidig vækst vil blive bedre påskønnet og belønnet
- Reducere værdien af en ejerledet virksomhed i forbindelse med skilsmisse
- Reducere værdien af en virksomhed ifm. et management buy out

Således er regnskabsmanipulation også den besvigelsestype, der ligger til grund for nogle af de mest kendte besvigelsessager. Eksempler herpå er Nordisk Fjer, IT-Factory, Enron m.fl.

På baggrund af konklusionerne i Report to the Nations fra ACFE vil de efterfølgende afsnit vedrørende regnskabsmanipulation alene redegøre for indregning af fiktiv omsætning, da den vurderes at have størst konsekvens for virksomheden. De øvrige typer regnskabsmanipulation vil ikke blive omtalt yderligere.

Fiktiv omsætning:

Fiktiv omsætning omfatter omsætning, der er indregnet, men ikke forekommet. Fiktiv omsætning kan forekomme til både ikke-eksisterende- og reelle kunder.

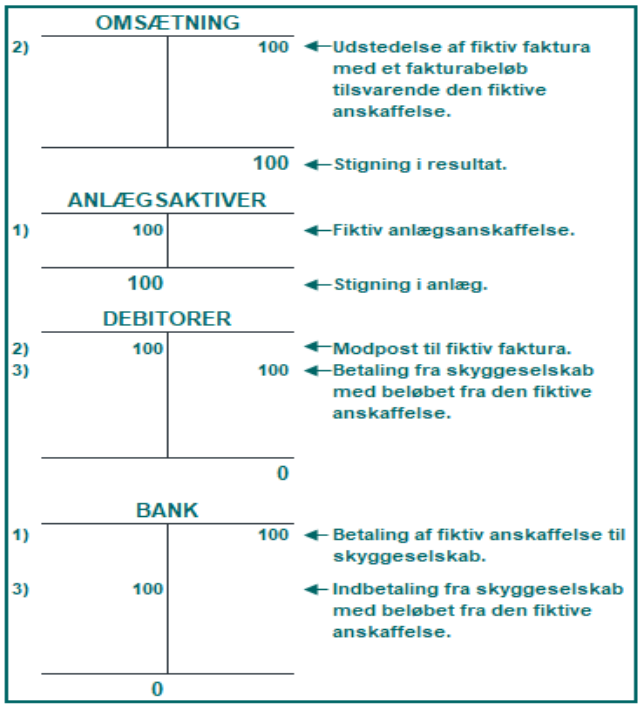
Fiktivt salg til reelle kunder – Kan forekomme ved, at der bliver udskrevet en fiktiv faktura, som danner bogføringsgrundlag for indregning af omsætningen. For at denne type fiktive omsætning ikke skal opdages, sendes fakturaen aldrig til kunden. Således indregnes transaktionen som omsætning og som tilgodehavender fra salg og tjenesteydelser som modpost. Da fakturaen ikke er sendt til kunden, og der ikke er sket levering af en egentlig ydelse, er der risiko for, at denne type besvigelse opdages, da tilgodehavendet med tiden vil blive overforfaldent pga. manglende betaling. Besvigeren vil typisk dække over denne form for besvigelse ved at udstede en kreditnota i det efterfølgende år og dermed udligne det overforfaldne tilgodehavende. Ved at skjule besvigelsen på den måde medfører det modregning i omsætningen det efterfølgende regnskabsår, hvorfor der vil være behov for at lave et større fiktivt salg efterfølgende for at indhente den tabte omsætning. Effekten heraf er, at den type besvigelse typisk akkumulerer til store besvigelser, hvis ikke de opdages i tide (Wells, 2013).

Fiktivt salg til ikke-eksisterende kunder og skyggeselskaber – Endvidere kan fiktiv omsætning forekomme til ikke-eksisterende kunder eller skyggeselskaber. Til forskel for salg til reelle kunder fremsendes der typisk en fiktiv faktura til kunden uden en bagvedliggende leverance. I sådanne situationer vil besvigeren typisk selv have kontrol over den bagvedliggende kunde (Wells, 2013).

Manipulation af salg til eksisterende kunder – Denne type besvigelse omfatter, at man manipulerer med en legitim faktura, sådan at fakturabeløbet eller antallet af varer er pustet op og dermed overstiger den faktiske værdi af salget (Wells, 2013).

Det dobbelte bogholderi medfører, at denne type besvigelse vil medføre en kortsigtet vækst, fordi tilgodehavenderne vil akkumulere over tid og dermed også over tid vil blive overforfaldne. For at dække over denne type besvigelse er man derfor nødt til i den efterfølgende periode at hensætte til tab eller udstede en kreditnota for at udligne det fiktive tilgodehavende. Denne handling medfører således en negativ vækst i den efterfølgende periode. Denne type besvigelse kan med tiden akkumulere til betydelige summer, hvorfor denne type besvigelse som følge af dens natur kan være den type besvigelse, der kan skabe størst skade for virksomheden.

Eksempel på, hvordan denne type besvigelse skjules – Der er set adskillige eksempler på, hvordan besvigere har forsøgt at skjule fiktiv omsætning på en mere sofistikeret måde end ved blot at nedskrive tilgodehavendet i efterfølgende periode. En af disse måder er at finde en anden modpost til tilgodehavendet end regnskabsposten "tilgodehavender fra salg og tjenesteydelser". Et af de eksempler, der er set forekommet, er hvor ledelsen fra et skyggeselskab udsteder en falsk faktura på en anlægsanskaffelse. Besvigeren sørger for, at fakturaen bliver betalt til skyggeselskabet, så den synes som en reel anskaffelse med bagvedliggende dokumentation med undtagelse af det faktiske aktiv. Efterfølgende udsteder besviger en fiktiv faktura med et fakturabeløb tilsvarende den fiktive anskaffelse. Herefter indregnes den fiktive omsætning og tilgodehavendet efter normal forretningsprocedure. Når fakturaen forfalder til betaling, sørger besvigeren for, at den, der har kontrol med skyggeselskabet, betaler med det beløb, der oprindeligt blev modtaget i forbindelse med den fiktive anlægsanskaffelse. Resultatet heraf er, at selskabet hverken er blevet rigere eller fattigere, men derimod har fået maskeret en fiktiv debitor som en fiktiv anlægsinvestering. Pengesporet vil medføre, at debitoren dermed ikke undersøges i efterfølgende periode. Eksemplet er illustreret i figur 6.



figur 6 - Indregning af fiktiv omsætning – egen tilvirkning.

Ovenstående er blot et eksempel på, hvordan besviger kan skjule sin besvigelse. Der findes utallige eksempler på, hvordan fiktiv omsætning kan skjules, og ovenstående er således ikke udtømmende for, hvordan det kan gøres.

Betinget salg – Under kategorien fiktiv omsætning forekommer endvidere betinget salg, som har nær tilknytning til periodiseringssvindel, og som derfor falder ind under begge kategorier. Kategorien kan bestå af omsætning, hvor man har en forpligtelse overfor kunden i forbindelse med fuld returret. Endvidere er det et krav, at der bliver passiveret en andel af salget til at imødegå risikoen for, at kunden returnerer varen.

Endelig kan besvigelsen bestå i for tidlig indregning af omsætning. Det sker, når alle indregningskriterier ikke er opfyldt, således at rettigheder og risikoen ikke er overgået til modtageren. Ved balancedagen kan ledelsen således have incitament til for tidlig indregning af omsætning, der ikke overholder indregningskriterierne, da ledelsen kan have en forventning om, at indregningskriterierne vil blive overholdt i efterfølgende periode. På den måde kan ledelsen fremvise et bedre resultat end det realiserede (Wells, 2013).

Faresignaler, som kan indikere, at der er forekommet fiktiv omsætning (Wells, 2013)**:**

- Eksplosiv vækst eller usædvanlig rentabilitet sammenlignet med resten af branchen
- Negativ eller manglende evne til at generere pengestrømme fra driften trods positiv udvikling i årets resultat
- Væsentlige transaktioner med nærtstående parter
- Usædvanlige eller unødvendigt komplekse transaktioner
- Usædvanligt eller uforklarligt fald i debitorernes omsætningshastighed
- Betydelige salg til kunder, hvor substansen bag eller ejerskabet er ukendt
- Usædvanlig vækst i enkelte afdelinger eller omsætning bogført på virksomhedens hovedkontor
- Posteringer, der ikke følger de normale forretningsgange, fx manuelle posteringer

Disse indikatorer bør alt andet lige anvendes som risikokatalog af kontrolorganet til at designe interne kontroller, der kan forebygge, at fiktiv omsætning forekommer.

2.1.4 Opsummering på besvigelsesteorien

Efter gennemgangen af dette afsnit er læser nu blevet gjort bekendt med, at besvigelsesteorien tegner et billede af en hurtigt omskiftelig omverden, hvor besvigere er hurtige til at tilpasse sig, hvilket medfører, at den bagvedliggende forbrydelse typisk er på forkant med virksomhedernes kontroller. Endvidere belyste

afsnittet, at alle i en virksomhed kan begå besvigelser, såfremt de har ressourcer, kvalifikationer og mulighed.

Det er blevet belyst for læser, at besvigelser årligt koster virksomhederne 26 billioner DKK, og at "faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen" og "indregning af fiktiv omsætning" er de to besvigelser, der har størst konsekvens for virksomhederne.

Centrale forhold og karakteristika fra besvigelsesteorien er skematiseret i figuren nedenfor for at skabe overblik over de centrale områder, der vil danne udgangspunkt for afhandlingens analyseafsnit:

Målbare karakteristika i besvigelsestrekanten og elementer i gerningen

Målbare	Ikke målbare
+ Mulighed (besvigelsestrekanten)	÷ Pres (besvigelsestrekanten)
+ Tilsløring (besvigelseselement)	÷ Retfærdiggørelse (besvigelsestrekanten)
+ Handlingen (besvigelseselement)	
+ Konvertering (besvigelseselement)	

Relevante karakteristika hos besviger

Supplerende element til besvigelsestrekanten	Model og karakteristika
• Incitament	MICE + Grådighed + Ego + Tvang + Ideologier ***Understreger at alle medarbejder er potentielle besvigere***
• Evner	The Fraud Diamond + Håndtere stress + Evnen til at kunne se hvordan evt. kontroller kan omgås ***Ovenstående karakteriserer ansatte på ledelsesniveau***

Besvigelser med størst konsekvens for virksomhederne

Typer af besvigelser	Forekomst og konsekvens
Misbrug af aktiver • Faktureringssvindel i forbindelse med indkøbsprocessen	***57 % af alle besvigelser (mediantab 743.000 DKK)*** • 20 % af alle besvigelser indenfor misbrug af aktiver
Regnskabsmanipulation • Indregning af fiktiv omsætning	***1 % af alle besvigelser (mediantab 5.200.000 DKK)*** • 34 % af alle besvigelser indenfor regnskabsmanipulation

Målbare karateristika ved de to afgrænsede besvigelsestyper:

Faktureringssvindel igennem skyggeselskaber	Indregning af fiktiv omsætning
• Leverandørens reelle ejer er en medarbejder eller nærtstående hertil. • Godkendelsesprocedure omgås i forbindelse. - Omkostningsgodkendelse - Kreditoroprettelse - Ændringer i kreditorstamdata • Ofte immaterielle ydelser • Manglende leverance fra skyggeselskabet	• Tilsløring sker via balancen • Posteringer følger ikke de normale forretningsgange • Manuelle posteringer foretaget på ledelsesniveau • Ændring i samhandelsmønstre • Manglende cashflow trods positive resultater • Management override • Sammenfald mellem leverandører og kunder • Manglende lagertræk og leverance knyttet til omsætningen • Andre karakteristika i form af: usædvanlige beløb og timing mv.

figur 7 - Opsummering af besvigelsesteori skematiseret - egen tilvirkning

AFSNIT 2.2 – INTERN KONTROL

Dette teoriafsnit har til formål at give læseren et kendskab til hvem, der har ansvaret for, at både tilsigtede og utilsigtede fejl ikke forekommer. Herunder hvad interne kontroller er, og hvorfor implementeringen heraf er vigtig i kampen mod besvigelser. Med afsæt i COSO's besvigelsesframework vil vi give læseren indblik i de centrale elementer, der bør overvejes, når man skal sikre sig mod besvigelser ved hjælp af intern kontrol. Afslutningsvis vil der blive redegjort for dataanalysens berettigelse i intern kontrol. Hensigten med afsnittet er derfor sammen med de øvrige teoriafsnit at kunne danne det teoretiske fundament for forståelsen af afhandlingens analyseafsnit.

2.2 Intern kontrol-teori

2.2.1 Hvad er intern kontrol

Interne kontroller er et af virksomhedernes bedste værktøjer til at forsvare sig mod tilsigtede såvel som utilsigtede fejl. I den danske selskabslovgivning fremgår det, at det påhviler ledelsen at drive virksomheden forsvarligt, herunder at der eksisterer de fornødne processer og procedurer for risikovurdering og interne kontroller (Folketinget, 2010).

Når ledelsen skal etablere de fornødne procedurer for risikostyring og intern kontrol, er COSO-frameworket et af de mest anerkendte værktøjer hertil. COSO's overordnede mission er:

"Our mission is to provide thought leadership through the development of comprehensive frameworks and guidance on enterprise risk management, internal control and fraud deterrence designed to improve organizational performance and governance and to reduce the extent of fraud in organizations"(The Committee of Sponsoring Organizations' (COSO), 2018).

COSO-frameworket omfatter 5 overordnede kontrolkomponenter, som ifølge COSO's forskning kan skabe de fornødne rammer for et velfungerende kontrolmiljø. Disse komponenter er af COSO defineret som: kontrolmiljø, risikovurdering, kontrolaktiviteter, informationssystemer og kommunikation samt overvågende aktiviteter (COSO, 2013, p. 5).

Frameworket finder anvendelse på alle dele af virksomhedens drift og de forskellige afdelinger, der understøtter den. Frameworket er opbygget sådan, at virksomheden i sit interne kontrolmiljø indledningsvist anbefales at designe og implementere de overordnede interne kontroller. Overordnede

interne kontroller består af "tone at the top", uafhængighed, integritet og etik mv. (COSO, 2013). Herefter præciserer frameworket, hvordan virksomheden igennem risikovurdering bør identificere risici for tilsigtede såvel som utilsigtede fejl, og hvordan disse bør afdækkes ved hjælp af kontrolaktiviteter. Efterfølgende definerer frameworket, hvordan information bør behandles og kommunikeres af virksomheden. Afslutningsvis understreges det, at kontrolapparatet bør overvåges for at sikre, at kontrollerne faktisk efterkommer det tilsigtede.

De fem komponenter i COSO-frameworket består af:

- Kontrolmiljø
- Risikovurdering
- Kontrolaktiviteter
- Information og kommunikation
- Overvågning og evaluering

Disse fem komponenter defineres af COSO ved hjælp af 17 underliggende principper, som samlet definerer rammerne for en succesfuld implementering i praksis. Sammenhængen mellem de fem kontrolelementer, og de 17 principper for en succesfuld implementering fremgår af "Bilag 6 – COSO-principperne og disses sammenhæng til besvigelses-frameworket".

COSO-frameworket er meget omfattende og danner rammerne for et velfungerende kontrolmiljø. Udfordringen ved frameworket er, at der er tale om generelle principper, og hvis der ses bort fra det enkelte princip om besvigelser beskrevet under risikovurderingen, søger frameworket en generel tilgang til det at drive interne kontroller. Som følge heraf har COSO i 2016 udgivet et framework relateret til besvigelser. Frameworket, som hedder COSO The Fraud Risk Management Guide, herefter omtalt besvigelsesframeworket, angiver fem nye principper, der direkte omhandler besvigelser. Dette framework vil i det efterfølgende afsnit blive uddybet og undersøgt.

2.2.2 COSO-frameworket i et besvigelsesperspektiv

Som resultaterne viste i afsnit 2.1.2, er besvigelser et omfattende globalt problem i mange virksomheder, hvorfor COSO har udgivet det føromtalte besvigelsesframework. De fem nye principper udspringer i høj grad af det oprindelige framework, men præciserer i særlig grad, hvordan det kan anvendes i en besvigelseskontekst. Sammenhængen mellem det oprindelige COSO-framework og de fem nye principper er illustreret i "Bilag 6 – COSO-principperne og disses sammenhæng til besvigelses-frameworket".

Som det fremgår af "Bilag 6 – COSO-principperne og disses sammenhæng til besvigelses-frameworket", er der en direkte kobling mellem det oprindelige framework og COSO's besvigelsesframework. De fem nye principper vil blive beskrevet nedenfor med henblik på at belyse, hvordan interne kontroller kan være med til at forebygge og opdage besvigelser.

Princip nr. 1 - Etablering af governance-politikker for håndtering af besvigelsesrisici - Den første del af COSO-modellen om besvigelser omhandler de overordnede interne kontroller og er centreret i virksomhedens etik, tone at the top og governance, herunder definering af, hvem der er ansvarlig for et besvigelsesprogram. Et åbent besvigelsesprogram kan, som omtalt i afsnittet om besvigelser, have en afskrækkende effekt overfor en potentiel besviger og vil i sin natur medvirke til, at der opbygges en barriere mellem besvigelsestrekantens tre elementer og selve besvigelseshandlingen, som illustreret i figur 5. Hertil kommer etablering af retningslinjer og etablering af direkte adgang til de ansvarlige. De ansvarlige for et eventuelt besvigelsesprogram vil både være den daglige ledelse og de ansvarlige for programmet, men også det øverste ledelsesorgan og eventuelle underudvalg hertil. Virksomhedens ledelse skal derfor sørge for at udarbejde klare målsætninger for virksomhedens håndtering af besvigelser, fx hvem der har ansvaret for håndtering og eksekvering af konsekvenser af dem.

Som beskrevet i afhandlingens afgrænsning søger afhandlingen ikke at undersøge de overordnede interne kontrollers effekt på besvigelser, men derimod hvordan teknologi kan anvendes i de øvrige dele af det interne kontrolmiljø. Konklusionerne fra COSO's Fraud Risk Mangement-guide anses at være gældende herfor, og derfor udgør de overordnede interne kontroller en central rolle i det interne kontrolmiljø for at modvirke eller opdage besvigelser. De overordnede interne kontroller er derfor nødvendige, for at en virksomhed effektivt vil være i stand til at identificere eller modstå besvigelser (L. Cotton, Johnigan and Givarz, 2016, p. 8).

Princip nr. 2 - **Udførelse af risikovurdering** - I anden del af guiden til besvigelsesframeworket omtales risikovurderingen, og hvordan den, hvis udført korrekt, kan være afgørende for, om de øvrige niveauer i kontrolmiljøet vil være i stand til at modvirke eller identificere besvigelser.

Som første led i processen anbefales det virksomheden at etablere et besvigelsesrisikovurderingsteam[5]. Teamet bør sammensættes af forskellige ressourcer og kompetencer, være en blanding af forskellige niveauer og indeholde tilpas med senior-ressourcer for at kunne operere uden begrænsninger og således have et passende mandat (L. Cotton, Johnigan and Givarz, 2016).

Umiddelbart kræver det en del ressourcer, men det er samtidig nøglen til en effektiv tilgang til de efterfølgende dele af kontrolmiljøet. Frameworket påpeger vigtigheden af, at der samles forskellige dele af virksomheden ved at understrege, at driftsansvarlige kender kunder, leverandører, medarbejdere, og de daglige kontroller mv., mens regnskabsafdelingen fx kender til de økonomiske transaktioner, samhandel, likviditet mv. Teamets diversitet medvirker således til, at der bidrages med forskellige vinkler og informationer om de samme områder, hvilket kan styrke risikovurderingsprocessen betydeligt. Hertil vil specialkompetencer inden for it mv. bidrage til, at der fx kan anvendes ny teknologi til processen.

Når virksomheden har etableret et stærkt risikovurderingsteam, påbegyndes selve risikovurderingen. Her er det essentielt, at der foretages en grundig gennemgang af strategi, forretningsmodeller, processer, kontroller, medarbejdere, herunder gamle besvigelsessager, samt enheder og divisioner i virksomheden. Risikovurderingen danner efterfølgende grundlaget for en brainstorm om, hvad der kan gå galt i relation til besvigelser. Her laves en decideret optegning af besvigelsesrisici i virksomheden. Et led i denne optegning er samtidig en vurdering af eksisterende besvigelsesreducerende aktiviteter og kontroller, og hvordan de imødegår eller ikke imødegår forskellige besvigelsesrisici. Når de identificerede risici skal optegnes, er det vigtigt, at teamet foretager en vurdering af de forskellige besvigelsestyper, fx regnskabsmanipulation, misbrug af aktiver mv. Et eksempel herpå, inden for regnskabsmanipulation og misbrug af aktiver, er risikoen for manipulation med eller udstedelse af fiktive fakturaer, og hvordan denne risiko spiller sammen med de to typer af besvigelser. Når denne skal mappes, skal der identificeres enten en eksisterende kontrolaktivitet eller designes en ny. En eksisterende aktivitet kunne være ”to-i-forening” ved udstedelse af fakturaer udenom fakturamoduler. En ny kontrol, såfremt den eksisterende ikke vurderes tilstrækkelig, kunne være stikprøvevis undersøgelse eller dataanalyse af transaktioner med for høj dækningsgrad. Dertil skal interne og eksterne faktorer vurderes (L. Cotton, Johnigan and Givarz, 2016). Fx kan ændringer i konjunkturer medføre øgede risici for misbrug af aktiver eller regnskabsmanipulation.

[5] Fremadrettet benævnt risikovurderingsteam

Som værktøj til vurdering af besvigelsesrisici kan dataanalyse og ny teknologi være med til at identificere potentielle risici. COSO fremhæver i forbindelse hermed:

"Organizations often utilize employee surveys, facilitated sessions, and other data-gathering techniques to gain a more reflective perspective on fraud risks. They then use data analytics techniques to compile, display, and analyze the results" (L. Cotton, Johnigan and Givarz, 2016).

I praksis betyder det, at virksomhederne har mulighed for både at opnå en dybere viden om regnskabsmæssige posteringer, men også en større forståelse af mønstre i samhandlen med leverandører og kunder, medarbejdertilfredshed mv. Metoder til risikovurdering ved brug af data kan være[6] (L. Cotton, Johnigan and Givarz, 2016):

- ***Datastratificering eller regelbaseret analyse***: Kategorisering af data ud fra regler om fx beløbsstørrelse, tidspunkt mv.
- ***Riskscoring***: Vægtning af risiko på transaktioner, således der er størst fokus på risikofyldte områder.
- ***Trendanalyse***: Sammenligning af udvikling på tværs af perioder, afdelinger, personer eller forskellige regnskabsposter. Som fx omsætning ift. debitorudviklingen enten på aggregeret eller disaggregeret niveau, ud fra fx afdelinger, personer eller kunder.
- ***Visualiseringsteknikker***: Visualisering som værktøj. Fx en analyse med omsætning holdt op mod vareforbrug, hvor dækningsgraden altså vises. Posteringer med unormale dækningsgrader danner herefter grundlag for risici.
- ***Data Mining***: Analyse af mere statistisk eller intelligent karakter. Her er tale om løbende overvågning ud fra statistik, normaler og forventninger. Der vil blive redegjort for begrebet Data Mining i afsnit 2.3.2.1.

Dataanalyse giver et optimeret grundlag, sammenlignet med manglende anvendelse af dataanalyse, for at foretage en mere risikorettet vurdering af besvigelsesrisiciene i virksomhederne. I afsnit 2.3 arbejdes der videre med mulighederne inden for dataanalyse.

Når risiciene i virksomheden er optegnet og klassificeret ud fra en væsentligheds- og sandsynlighedsbetragtning, udarbejder teamet et oplæg til det øverste ledelsesorgan (L. Cotton, Johnigan

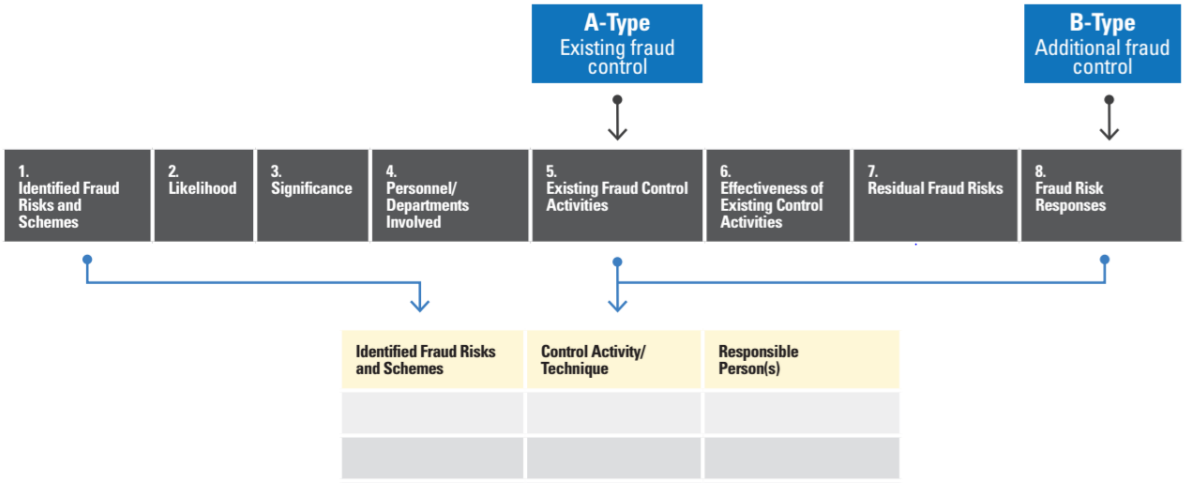
[6] Ej udtømmende

and Givarz, 2016). I oplægget bør det endvidere fremgå, om der bør ageres på den pågældende risiko og i hvor høj grad. En metode er udarbejdelse af et "heatmap". I afsnit 2.2.4 vises et eksempel på, hvordan det kan anvendes i risikovurderingsprocessen. Resultaterne fra risikovurderingen skal herefter forankres i virksomhedens målsætninger, og de nødvendige ressourcer skal allokeres til det videre arbejde for at afdække de identificerede risici.

Afslutningsvis er det afgørende for effekten af risikovurderingen, at den løbende revurderes ud fra interne og eksterne faktorer, herunder resultaterne af det udførte arbejde, som en iterativ, dvs. cirkulær og gentagen, proces.

Princip nr. 3 - Design og implementering af forebyggende og opdagende besvigelseskontroller

Med udgangspunkt i risikovurderingsplanen skal der designes og implementeres nye kontroller eller identificeres eksisterende kontroller, der kan afdække de identificerede risici. Processen vedrørende risikovurdering og design af kontrolaktiviteter er illustreret i nedenstående figur, hvor punkt 1-3 består af selve risikovurderingsprocessen, hvor den enkelte besvigelsesrisiko defineres. Efterfølgende angives det i punkt 4, hvor i virksomheden besvigelsesrisikoen er placeret. På baggrund heraf undersøges det, om der allerede eksisterer en kontrol, som effektivt afdækker den identificerede besvigelsesrisiko i punkt 5-7. Såfremt der ikke eksisterer en intern kontrol, som effektivt afdækker besvigelsesrisikoen, designes der en ny kontrol hertil i punkt 8. Outputtet fra processen afføder således en optegnelse af de identificerede besvigelsesrisici og af de interne kontroller, der er designet til at afdække disse, samt en placering af ansvaret for implementeringen af den nye kontrol.



figur 8 - Documenting Fraud Risk Assesment Summary Matrix, control activities areas (L. Cotton, Johnigan and Givarz, 2016)

I intern kontrol arbejdes der med 2 typer af kontroller, henholdsvis præventive og opdagende kontroller, som vil blive uddybet nedenfor.

Forebyggende kontroller - er kontroller, som har til formål at sikre, at en besvigelse ikke kan forekomme. Forebyggende kontroller kan være i form af effektiv funktionsadskillelse, prædefinerede undersøgelser ved forskellige typer transaktioner, fx transaktioner med nærtstående parter eller transaktioner, initieret af ledelsen. Endvidere kan opsætning af godkendelseshierarkier og fysiske inspektioner være velegnede eksempler på præventive kontroller. De præventive kontroller vil ofte være dyre at implementere, da de i høj grad kræver personaleressourcer både ved manuelle og automatiske kontroller. Årsagen til, at der overhovedet implementeres præventive kontroller, skyldes i høj grad, at det kan være omkostningstungt at skulle inddrive aktiver, som allerede er fragået virksomheden som konsekvens af en besvigelse, også benævnt ”pay and chase”. Forebyggende kontroller er ofte kendt af medarbejderne i virksomheden (L. Cotton, Johnigan and Givarz, 2016).

Opdagende kontroller er derimod kontroller, hvor formålet er at opdage eventuelle besvigelser, efter at en transaktion er sket. Der kan være tale om analytiske undersøgelser af mønstre i fx samhandel, betalinger, transaktioner af unormal karakter mv. Dertil kan der være tale om stikprøvevise test af udvalgte områder eller reviews i relation til en månedsrapportering. Opdagende kontroller er som oftest skjulte fra den normale virksomhedsdrift og baserer sig på både interne og eksterne informationer. Kontrollerne anvendes i høj grad til optimering af både de forebyggende kontroller, men også selve risikovurderingen (L. Cotton, Johnigan and Givarz, 2016).

Virksomheder bør anvende en blanding af både forebyggende og opdagende kontroller. De forebyggende kontroller kan have en reducerende effekt i sig selv, da virksomhedens ansatte har kendskab til dem, og de danner således også del af en åben Fraud Governance-politik, der har en intern afskrækkende effekt. Kontrollerne er som regel kendte i virksomheden, hvilket også betyder, at de er mulige at omgå. Det er ikke muligt, hvis der også skal tages højde for omkostninger, at eliminere alle besvigelser ved hjælp af forebyggende kontroller. Det ville kræve enorme ressourcer, både i form af mandskabskraft og it mv. Kombinationen af forebyggende og opdagende kontroller skal derfor sikre, at områder, hvor der ikke kan etableres forebyggende kontroller, afdækkes af opdagende kontroller (L. Cotton, Johnigan and Givarz, 2016). Et eksempel på en kombineret kontrol er fx bankbetalinger, hvor der etableres en to-i-forening-kontrol, og hvor det dermed sikres, at der er funktionsadskillelse på området. Dog er det for omkostningstungt for virksomheden at kontrollere stamdata på samtlige betalinger. Som konsekvens heraf

implementeres en præventiv kontrol rettet mod ændringer i stamdata, hvor alle ændringer i stamdata skal godkendes, før de kan gennemføres. I forlængelse af dette foretages der stikprøver på udvalgte bilag, efter at betalingen er gennemført. Den sidste kontrol vil have karakter af en opdagende kontrol.

Når virksomhederne skal vælge mellem de forskellige kontroltyper, kan dataanalyse være nøglen til et effektivt kontrolmiljø, som både fungerer hensigtsmæssigt og også kan forblive omkostningsoptimalt. Dette skyldes, at dataanalyse kan identificere og "trigge" anomalier eller ændrede mønstre i samhandlen.

Med udgangspunkt i risikovurderingen og en vurdering af ovenstående kontroltyper bør virksomheden designe og implementere de nødvendige kontroller. Ved design forstås (COSO, 2013, p. 87):

- At kontrollen afdækker det, som den har til hensigt
- At medarbejderen har passende kompetencer
- At der anvendes et passende niveau for acceptable afvigelser
- At der bliver foretaget passende undersøgelser af afvigelser eller fejl
- At der bliver foretaget en identifikation af kontrolrisikoen
- At der bliver foretaget en analyse af, om kontrollen er afhængig af andre kontroller

Dette betyder i praksis, at for optimalt at afdække en eventuel risiko relateret til regnskabsmanipulation bør virksomheden begrænse brugerrettigheder og opsætte en "to i forening-godkendelse"-procedure ved manuelle posteringer. Ved en manipuleret transaktion vil det derfor ikke give mening at bede ledelsen om at skrive under på transaktionen, hvis besvigeren selv har initieret denne. Derfor er designet vigtigt.

Efterfølgende er det helt nødvendigt, at kontrollen rent faktisk bliver implementeret, hvilket groft sagt kan oversættes til, at de ansvarlige for kontrollen udfører de punkter/elementer, som kontrollen er designet til, samt at de udfører kontrollen kontinuerligt. I forlængelse heraf omhandler implementeringsdelen den løbende opfølgning og de korrigerende tiltag, der er knyttet til kontrollen. I praksis er der tale om en revurdering af kontrollen ud fra erfaringer og resultater (L. Cotton, Johnigan and Givarz, 2016).

Ved valg af kontroller er det centralt, at der anvendes forskellige typer af kontrolaktiviteter. De kan være:

- Fraud governance (overordnede interne kontroller)
- Funktionsadskillelse
- Transaktionsbaserede kontroller
- HR-kontroller fx ved ansættelse, aflønning, performance, medarbejder-surveys, Fraud Risk-træning

- Whistleblower-ordninger
- Kontroller, som imødegår management override
- Dataanalyse og kontinuerlig overvågning, nærmere omtalt i afsnit 2.3

Når virksomheden har opsat de nødvendige kontroller, og der er skabt en kobling mellem disse og de identificerede risici, har virksomheden en plan for, hvordan den vil afdække besvigelsesrisikoen. I den forbindelse er det nødvendigt at kontrolaktiviteter, samt selve planen dokumenteres, da dokumentationen både danner rammer for de efterfølgende processer i kontrolmiljøet, men samtidig også for den iterative proces, som intern kontrol også er.

Afslutningsvis bør virksomheden løbende rapportere om resultater og eventuelle mangler i kontrolmiljøet til risikovurderingsteamet. Dette skal gøres således, at teamet har adgang til faktisk information og derigennem kan foretage en rettidig opdatering af risikovurderingen.

Princip 4 - Udførelse af besvigelsesundersøgelser - Den fjerde del af besvigelsesframeworket omhandler, hvordan virksomheden omsætter resultaterne af dens interne kontroller til konkrete undersøgelser af scenarier, hvor der er risiko for besvigelser. Afsnittet tager udgangspunkt i anmeldelser af besvigelser, som er identificeret enten gennem kontrolaktiviteterne eller gennem anmeldelser. Dette område er dog ikke fokus for afhandlingen, hvorfor der, jævnfør afgrænsningen, udelukkende foretages en kort opridsning af området.

Besvigelsesundersøgelser er reelle undersøgelse af hændelser, hvori man har en formodning om, at der blev begået besvigelser. Når disse undersøgelser udføres, er der tale om detektivlignende arbejde, hvor alt materiale vedrørende den pågældende hændelse bliver indhentet og undersøgt. Endvidere anvender virksomhederne i efterhånden større grad dataanalyse, herunder forventningsbaserede analyser, for at kunne undersøge hændelserne. Undersøgelserne anvendes til at tage juridiske skridt mod en eventuel besviger. I relation til afhandlingen er det primært resultaterne af en undersøgelse, der har interesse. En undersøgelse kan vise svagheder i kontrolmiljøet eller eventuelle behov for yderligere kontrol. Endelig kan undersøgelserne klarlægge risikozoner eller områder, hvori fx nye kontroller baseret på analyse kan have en opdagende eller forebyggende effekt (L. Cotton, Johnigan and Givarz, 2016).

Princip 5 - Overvågning og evaluering af programmet for håndtering af besvigelsesrisici

Den sidste del af COSO-frameworket omkring besvigelser er den løbende overvågning og evaluering af hele besvigelsesprogrammet. Dette omhandler periodiske rapporteringer til det øverste ledelsesorgan samt særskilt evaluering på delområder. Den periodiske rapportering, som er resultatet af alle de udførte aktiviteter, omfatter antallet af besvigelsessager, som er kendte, både afværgede eller realiserede, samt læring fra de udførte undersøgelser. Dertil kommer den løbende evaluering af delområder, fx ved virksomhedsopkøb og rapportering af resultater fra besvigelsesprogrammet for denne enhed.

Rapporteringen bør sendes til det øverste ledelsesorgan, og herefter bør også risikovurderingsrådet have adgang hertil. Denne rapportering er nemlig central i forhold til vurderingen af, om der bør ske revaluering af kontrolaktiviteter og risici mv. I forlængelse af dette bør revalueringen også omfatte eventuelle ændringer i virksomhedens nær- og fjernmiljø, fx kan konjunkturændringer eller lovmæssige forhold have en væsentlig indvirkning på besvigelsesrisici i virksomheden (L. Cotton, Johnigan and Givarz, 2016). Hele processen omkring intern kontrol er således en iterativ proces, som hele tiden kræver revurdering, fornyelse og tilpasning.

2.2.3 Konkrete kontrolaktiviteter rettet mod besvigelser

I kapitlet om besvigelser blev det konstateret, at "faktureringssvindel i forbindelse med indkøb" og "indregning af fiktiv omsætning" er de to typer af besvigelser, der har størst konsekvens for virksomhederne. Dette afsnit har redegjort for COSO-frameworket og dettes medvirken til at opbygge en barriere mellem besvigelsestrekanten og besvigelseshandlingen, som omtalt i afsnit 2.1.1. Nedenfor vil der blive redegjort for traditionelle kontrolaktiviteter og processer, som kan anvendes til at forebygge og opdage besvigelser indenfor de to ovennævnte besvigelsestyper.

Indledningsvist vil ledelsen have defineret et mål som en del af deres besvigelsesprogram om, at de ønsker at minimere besvigelsesrisikoen inden for faktureringssvindel og indregning af fiktiv omsætning. Som følge heraf skal det af ledelsen udpegede risikovurderingsteam identificere risici relateret til disse objektiver, designe passende kontroller og udpege kompetente kontrolindehavere til at varetage den løbende kontrolaktivitet til imødegåelse af de identificerede risici.

Faktureringssvindel – For eksemplets skyld afgrænses gennemgangen til alene at omhandle "svigagtige udbetalinger igennem skyggeselskaber i forbindelse med indkøbsprocessen", som omtalt i afsnit

2.1.3.12.1.3. Indledningsvis bør risikovurderingsteamet foretage en grundig gennemgang af virksomhedens strategier, processer og forretningsområder mfl.

På baggrund af gennemgangen bør risikovurderingsteamet være i stand til at opnå et indgående kendskab til virksomheden og derfor være i stand til at påbegynde sin brainstorming om mulige risici inden for faktureringssvindel, samt hvilke konsekvenser det kan have for virksomheden. I en virksomhed, som sælger tjenesteydelser, antages det, at risikovurderingsteamet vil identificere følgende risici [7]:

- Der er risiko for, at der foretages betaling af fiktive fakturaer **(1)**
- Der er risiko for, at der foretages indkøb til overpris **(2)**

Når ovenstående risici er identificeret, anbefales det, at risikovurderingsteamet klassificerer de identificerede risici efter en vurdering af både sandsynligheden for besvigelsens forekomst og en vurdering af, hvor væsentlig denne vil være. Et relevant værktøj til dette er fx et heatmap. På baggrund af ovenstående rapportering til ledelsen angiver ledelsen overfor risikovurderingsteamet, hvilke risici de ønsker afdækket. Det antages i eksemplet, at ledelsen, på baggrund af de foregående undersøgelser og heatmap-rapporteringen, ønsker begge risici afdækket. Som følge heraf er det risikovurderingsteamets opgave at kontrollere, om der er designet og implementeret velfungerende kontroller, som imødegår disse risici. Hvis ikke dette er tilfældet, er det risikovurderingsteamets opgave at designe og implementere sådanne interne kontroller. Når risikovurderingsteamet skal designe sådanne kontroller, anbefales COSO-frameworket.

Som beskrevet i afsnit 2.1.3.1 kan svigagtige udbetalinger igennem skyggeselskaber i forbindelse med indkøbsprocesser ske som følge af manglende effektive interne kontroller på godkendelsesområdet. Udførelsen af ovenstående besvigelse begås således ved, at der udstedes en ikke-gyldig faktura af et skyggeselskab, hvorefter godkendelseskontrollerne omgås. Et klassisk kontroldesign til at imødegå en sådan besvigelse vil typisk bestå af forebyggende og opdagende kontroller:

Besvigelsesrisiko i forbindelse med leverandøroprettelse samt ændring af stamdata – For at afdække denne besvigelsesrisiko vil en effektiv forebyggende kontrol kunne bestå i, at der alene kan ske betaling af en faktura, som kan knyttes til en leverandør, der allerede er oprettet i virksomhedens kreditormodul. For at imødegå risikoen for at der foretages udbetalinger til skyggeselskaber, kan en effektiv intern kontrol

[7] Ej udtømmende, blot eksempler

endvidere være, at kreditoroprettelser skal ske to-i-forening, og at leverandørernes reelle ejer dokumenteres i forbindelse med oprettelsen. Ved ændringer i kreditorstamdata kan en effektiv forebyggende kontrol være, at ændringer ligeledes skal ske to-i-forening. Alternativt kan der implementeres en opdagende kontrol, hvor en kontrolindehaver løbende skal foretage gennemgang af en log, der dannes, hver gang der sker ændringer i stamdata (Wells, 2013).

Besvigelsesrisiko i forbindelse med manglende funktionsadskillelse – En helt essentiel og klassisk forebyggende intern kontrol, som imødegår ovenstående risiko, er godkendelse to-i-forening. Således kan en effektiv, forebyggende kontrol være, at der skal ske fakturagodkendelse to-i-forening på baggrund af et opsat godkendelseshierarki, hvor det kontrolleres, at fakturaen fremtræder som legitim, før den godkendes og lægges til betaling. Fakturaer fra skyggeselskaber vil typisk være mangelfulde og mangle CVR-numre, telefonnumre mv. (Wells, 2013).

For ovenstående kontroller er det vigtigt, at kontrolindehaverne har de rette kompetencer og således er i stand til at opdage de besvigelser, som kontrollen er designet til fange.
De førnævnte kontroldesigns er ikke udtømmende og er blot eksempler på, hvordan den adresserede risiko traditionelt afdækkes i et internt kontrolmiljø.

Besvigelsesrisiko i forbindelse med indregning af fiktiv omsætning **-** Som beskrevet i afsnit 2.1.3.2 er der tale om en type besvigelse, der har store økonomiske og tillidsmæssige konsekvenser, når den forekommer. Ligesom i eksemplet med faktureringssvindel antages det, at risikovurderingsteamet har udført de fornødne procedurer og er nået frem til, at risikoen for fiktiv omsætning er til stede, og at den er relateret til fakturering til skyggeselskaber. Det er vurderet, at der er tale om en risiko, som forekommer med rimelig høj sandsynlighed og med stor betydning.

Risikovurderingsteamet har vurderet, at risikoen i høj grad bør adresseres ud fra en væsentligheds- og sandsynlighedsbetragtning. Som led i risikovurderingsprocessen er der analyseret på eksterne og offentligt tilgængelige sager. I lighed med eksemplet fra afsnit 2.1.3.1 er det konstateret, at en besvigelse af denne type kan skjules ved indbetaling fra et skyggeselskab, hvorefter der vil være et tilsvarende køb hos selskabet. Købet vil ofte blive balanceført og efterfølgende afskrevet løbende. Rækkefølgen af transaktionen har teamet ikke vurderet til at være afgørende, da transaktionen ofte vil være kontrolleret af samme part.

Ved design af kontrolaktiviteter tages informationer om samhandlen med en part, som både er kunde og leverandør, i betragtning. Risikovurderingsteamet designer først en forebyggende kontrol, der systemmæssigt kræver, at alle leverandører, som samtidig er kunder, kræver 2-i-forening-godkendelse ved oprettelsen. Kontrollen opsættes, så den også gælder modsatrettet.

Samtidig vurderer risikovurderingsteamet, at der er risiko for omgåelse af denne kontrol ved en eventuel sammensværgelse. Derfor vil risikovurderingsteamet designe en kontrol til opdagelse af posteringer, hvor kunde og leverandør er den samme, og periode eller beløb er sammenfaldende. Så snart sådanne posteringer foretages, registreres de i en logfil, som teamet har adgang til. Teamet planlægger, som en del af kontrolaktiviteten, at alle mistænkelige transaktioner undersøges så langt som til underliggende dokumentation for både køb og salg.

Ovenstående er typiske eksempler på kontrolaktiviteter rettet mod konkrete risici, men eksemplerne er ikke udtømmende. I det næste afsnit redegøres der for, hvordan arbejdet med interne kontroller kan understøttes af dataanalyse, og hvordan dataanalyse metodisk kan anvendes som kontrolaktivitet.

2.2.4 Hvordan kan dataanalyse anvendes metodisk i den interne kontrol

Når virksomheden er i gang med at planlægge kontrolaktiviteter eller risikovurderingshandlinger, kan de som før nævnt anvende forskellige typer dataanalyse for at styrke kontrolmiljøet. ACFE understreger i ”Reports to the Nations, 2018” mulighederne ved anvendelse af dataanalyse som værktøj til opdagelse af besvigelser (ACFE, 2018).

Rapporten viser, at der er en høj opdagelsesrate forbundet med brug af dataanalyse eller anden datamonitorering. Endvidere påpeger rapporten, at kun 37% af de undersøgte virksomheder har implementeret kontroller af denne karakter. Der ses umiddelbart en stor forskel mellem resultater, som opnås ved dataanalyse, og den faktiske og ringe anvendelse af dataanalyse. Dette afsnit vil derfor belyse, hvordan dataanalyse kan anvendes i intern kontrol.

Dataanalyse kan blandt andet anvendes til at bekræfte, at kontroller eksisterer og fungerer effektivt. Fx kan dataanalyser programmeres til at identificere eventuelle brud på funktionsadskillelse i en nedskrivningsproces på et varelager, hvor der ikke må være personsammenfald mellem den, der dokumenterer nedskrivningsbehovet, og den, der godkender det. Afslutningsvis kan dataanalyse virke

afskrækkende over for en besviger. Mulighed og rationalisering bliver reduceret, og derfor mindskes risikoen for besvigelser (L. Cotton, Johnigan and Givarz, 2016).

COSO har i forlængelse af besvigelses-frameworket givet et bud på, hvordan virksomheder kan implementere dataanalyse som en del af det interne kontrolmiljø som illustreret i figur 9. Der gives en kort introduktion for at demonstrere den metodiske anvendelse i intern kontrol. Redegørelse for dataanalyse fremgår af afsnit 2.3.

figur 9 - Data Analytics and Fraud Risk Management (L. Cotton, Johnigan and Givarz, 2016)

Første del af guiden omhandler således **analysens design.** Når der skal udarbejdes en dataanalyse, skal risikoen, som forsøges afdækket, klarlægges. Herefter designes analysen, så den er i stand til at imødegå risikoen. Når designet er på plads, identificerer man potentielle datakilder og adgangen til disse, hvorefter det er muligt at udvikle en plan og definere de konkrete handlinger (L. Cotton, Johnigan and Givarz, 2016). Området indeholder arbejde med valg af analysedesign, som kan være valg af teknologier, regelsæt, datateknikker, anvendte værktøjer mv. En stor del af det at arbejde med dataanalyse omhandler derfor de valg, som træffes i denne fase. Valgene er afgørende for resultaterne af analysen og kræver både faglige, operationelle og tekniske ressourcer. Dette skal ses i lyset af, at besvigelser ofte er komplekse og svære at identificere. Kompleksiteten afspejles i valget af analyser og kan kræve kombination af mange parametre og datakilder. I Afsnit 2.2.4.1 gives et eksempel på arbejdet med en konkret risiko i en hypotetisk virksomhed.

Anden del omhandler **indsamling af data** til analysen. Når analysen er designet, skal den udarbejdes i samarbejde med relevante it-folk. Der skal endvidere opsættes adgang til datakilderne, og i den forbindelse skal det sikres, at dataindsamlingen sker fuldstændigt og nøjagtigt. Til brug for analysen kan der fx opsættes et data warehouse, der kan anvendes som grundlag for virksomhedens dataanalyseplatform. I nogle tilfælde kan analysen programmeres direkte i virksomhedens it-infrastruktur. Denne proces uddybes yderligere i afsnit 2.3.2. Når dataadgange er sat op, skal der testes på udvalgte perioder for at validere, om data er brugbare til formålet (L. Cotton, Johnigan and Givarz, 2016).

I tredje del opsættes selve analysen, og **data renses og bearbejdes** ud fra de nødvendige, opsatte parametre. De ansvarlige for analysen samarbejder her med eksperter om at designe de nødvendige regler for at kunne udføre analysen. Her er tale om matematiske regler, og i nogle tilfælde kan det være nødvendigt med mere avancerede teknikker. Fx kan virksomheder med mange indtægtskilder have svært ved at oprette regler, der omfavner alle posteringstyper. Her kan mønster- eller linkanalyse anvendes til at kunne skabe realistiske forventninger til de forskellige transaktionstyper. Mønster- eller linkanalyse er en teknik, som søger kendte mønstre og afvigelser fra disse mønstre, samt identificerer nye mønstre (Techopedia, 2019). Virksomheden bør herefter teste analysen og validere, at resultaterne er korrekte, og at der ikke opstår for mange negativer eller falske positiver. Resultaterne analyseres, og analysen tilpasses herefter på denne baggrund (L. Cotton, Johnigan and Givarz, 2016).

Fjerde del omhandler selve **dataanalysen**. Her evalueres resultaterne fra analysen. Såfremt der er mange afvigelser, eller mængden af anomalier er for stor, kan der anvendes teknikker til at risikorette de fremadrettede undersøgelser, fx i form af ”risk scoring”. Her opdeles populationer ud fra karakteristika, som kan hjælpe teamet med at prioritere, hvad der skal undersøges yderligere. Denne risikoprioritering bør herefter indarbejdes i analysen, så der kun fokuseres på relevante posteringer. Et eksempel kunne være, at der er mange små posteringer, som ikke er relevante for undersøgelsen, og at man derfor etablerer et threshold (L. Cotton, Johnigan and Givarz, 2016).

Sidste del af guiden omhandler håndteringen af **observationer, og hvordan der reageres herpå.** Teamet bør have en strategi for indhentning af dokumentation og validering heraf. Dertil kommer selve stikprøvestørrelsen og grundigheden af undersøgelsen af disse. Afslutningsvis defineres processen for eskalering af fundne fejl og et system for, hvordan disse registreres (L. Cotton, Johnigan and Givarz, 2016).

Frameworket danner altså grundlag for, hvordan der kan arbejdes med planlægning og eksekvering af dataanalyse i de interne kontroller. De ovenstående tiltag vil naturligvis være forskellige, afhængigt af om der arbejdes med risikovurdering eller kontrolaktivitet, da formålene med de 2 områder er forskellige. Såfremt der arbejdes med risikovurdering, vil analyserne have en mere klarlæggende karakter, mens arbejdet med kontrolaktiviteter oftest vil være af opdagende eller forebyggende karakter.

I nedenstående er anvendelse af frameworket eksemplificeret ved risikoen for indregning af fiktiv omsætning, der i afsnit 2.1.2 blev identificeret som de besvigelser, der er af størst økonomisk betydning.

2.2.4.1 Eksempel på anvendelse af dataanalyse i intern kontrol

I eksemplet tages der udgangspunkt i en handelsvirksomhed, hvor der er etableret et risikovurderingsteam. Risikovurderingsteamet ønsker at anvende dataanalyse som en del af deres kontrolaktiviteter rettet mod besvigelser. Risikovurderingsteamet har igennem en historisk analyse identificeret, at risikoen for indregning af fiktiv omsætning er til stede. I forbindelse med risikovurderingen er det blevet vurderet, at fiktiv omsætning vil fremgå som unormale posteringer på omsætningskonti, særligt omkring månedsafslutningen.

Design - Risikovurderingsteamet opstiller en hypotese om, at unormale posteringer på omsætningskonti vil have en unormalt høj dækningsgrad. Risikovurderingsteamet identificerer relevante datakilder til at være lagersystemet og ERP-systemet. I disse kilder vurderer risikovurderingsteamet, at tabellerne for salgsposteringer og tilhørende vareforbrug samt lagerafgange er relevante. Datakilderne er fuldt tilgængelige for teamet. Til brug for analysen planlægges det, at der skal anvendes en regelbaseret analyse med efterfølgende visualisering. En regelbaseret analyse er en analyse, der er baseret på underliggende regler. En regel i dette tilfælde er, at der skal være sammenhæng mellem 2 datapunkter på et udvalgt parameter, fx ordrenummer. Analysen skal identificere salgsposteringer, som afviger fra den forventede dækning. Resultaterne visualiseres i et dashboard til efterfølgende undersøgelse. Derudover planlægges det, at der foretages analyse og undersøgelse af alle posteringer, hvor der ikke er registreret et underliggende vareforbrug, eller hvor der er et vareforbrug, men ikke en tilhørende lagerafgang samt kreditnotaer i den efterfølgende periode.

Dataindsamling - Risikovurderingsteamet opnår adgang til alle salgstransaktioner, lagertransaktioner, vareforbrug og finansposteringer. Risikovurderingsteamet validerer parametrene ved at afstemme disse med deres oprindelige datakilde, fx en råbalance eller et salgsmodul, hvorefter der bygges et link fra disse kilder til analyseplatformen. Analyseværktøjer kan være værktøjer som fx Business Intelligence, hvilket vil blive uddybet i afsnit 2.3.1. Herefter tester teamet analysen ved at validere data ved afstemning til grundlaget.

Dataopbygning og beregning - Risikovurderingsteamet opsætter en regelbaseret analysemodel ud fra en forventning om korrelation mellem omsætning, vareforbrug og lagerafgange baseret på ordrenummer. Der opsættes i dette tilfælde simple regler, analysen eksekveres, og der analyseres på resultaterne gennem et visualiseringsværktøj, fx Tableau.

Dataanalyse - Risikovurderingsteamet får nu resultaterne af analysen. Her vises en graf over posteringer, hvis dækningsgrader afviger markant fra gennemsnittet. I visualiseringen heraf er anomalierne sorteret efter størrelse, og det giver dermed risikovurderingsteamet mulighed for at klikke sig igennem data og se grundlaget herfor. Risikovurderingsteamet vil få mange afvigelser i dækningsgrader, men fælles for mange af disse er, at de er tilknyttet fragt. Risikovurderingsteamet genovervejer herefter analysen, da anomalier som følge af udsving i fragtomkostninger ikke er væsentlige, og teamet filtrerer dem derfor fra. Risikovurderingsteamet får resultaterne tilbage, og der er nu en overkommelig mængde af anomalier. Disse opdeles efter risiko, hvor posteringer uden vareforbrug og lagerafgang har størst risiko. Risikovurderingsteamet tildeler risikokategorierne en vægtning til brug for efterfølgende undersøgelse.

Observation og rapportering - Resultaterne fra analysen er nu gennemgået, og der er identificeret 20 posteringer af særlig interesse. Risikovurderingsteamet beslutter at gennemgå alle posteringer og indhente underliggende dokumentation for de gennemførte transaktioner. Materialet bliver gennemgået, og det viser sig, at størstedelen af posteringerne skyldes fejlregistreringer på ordrenummeret, som således er uden for risiko. Dog er der 5 posteringer, som alle er udført af samme sælger og herefter godkendt af samme chef op til månedsafslutningen. Teamet kan se, at relevante kontroller på området er omgået. Risikovurderingsteamet rapporterer herefter til besvigelsesudvalget, at en dybdegående undersøgelse bør iværksættes, da det vurderes, at der er indikationer på, at sælgeren og hans chef har manipuleret med salg op til månedsafslutningen. Som ovenstående eksempel viser, er det muligt at tilgå en konkret risiko ud fra det metodiske framework, som blev præsenteret i det foregående afsnit.

2.2.5 Opsummering på interne kontroller

Efter gennemgangen af dette afsnit er læser nu blevet gjort bekendt med, at det er den øverste ledelses ansvar, at tilsigtede såvel som utilsigtede fejl ikke forekommer. Et af de stærkeste værktøjer til at sikre, at disse fejl ikke forekommer, er interne kontroller. Interne kontroller kan implementeres med COSO's udførlige guides til, hvordan et effektivt kontrolmiljø kan designes. Læser er ligeledes gjort bekendt med, at intern kontrol består af aktiviteter, som virksomheden iværksætter med henblik på at imødegå de risici, virksomheden igennem sine aktiviteter eksponeres for. Grundet stor efterspørgsel har COSO publiceret et besvigelsesframework, som er rettet direkte mod bekæmpelsen af besvigelser ved hjælp af intern kontrol. Besvigelsesframeworket fremhæver, at interne kontroller i høj grad kan supporteres af dataanalyse, hvilket også bekræftes i ACFE's seneste rapportering "Report To The Nations, 2018".

Redegørelsen i dette afsnit danner det teoretiske fundament og begrebsrammen for den videre analyse i afhandlingens analyseafsnit. Endvidere er afsnittet anvendt til at bekræfte den indledende hypotese om, at dataanalyse effektivt kan anvendes til at opdage og forebygge besvigelser igennem intern kontrol, hvilket igen fører til videre analyse af analyseteknikker. Teorien vil blive anvendt ved vurdering af om Data Mining kan anvendes i virksomhedernes interne kontrolmiljø.

AFSNIT 2.3 – DATAANALYSE

Afsnittets hensigt er, sammen med de øvrige teoriafsnit, at danne det teoretiske fundament for forståelsen af afhandlingens analyseafsnit. Endvidere vil læseren blive gjort bekendt med, hvordan dataanalyse traditionelt anvendes, og hvordan databearbejdning ved hjælp af Data Mining kan være med til at effektivisere processen. Herunder vil læseren få et overordnet kendskab til et udvalg af Data Mining-teknikker, som vil blive anvendt og analyseret i besvigelseskontekst i afhandlingens analyseafsnit.

2.3 Teori om Dataanalyse

I daglig tale er dataanalyse et løst begreb, som omfatter at analysere på forskellige dataformer ved brug af forskellige teknikker. For at komme tættere på hvad dataanalyse er, har vi fremhævet et par forskellige definitioner på området:

"The process of evaluating data using analytical and logical reasoning to examine each component of the data provided. This form of analysis is just one of the many steps that must be completed when conducting a research experiment. Data from various sources is gathered, reviewed, and then analyzed to form some sort of finding or conclusion. There are a variety of specific data analysis methods, some of which include Data Mining, text business intelligence, and data visualizations" (BusinessDictionary.com, 2019).

"The Process of inspecting, cleaning, transforming, and modeling data with the objective of discovering useful information, arriving at conclusions, and supporting the decision making process is called Data Analysis" (Predictiveanalyticstoday.com, 2019).

"Data Analytics is the practice of using data to drive business strategy and performance. It includes a range of approaches and solutions, from looking backward to evaluate what happened in the past to looking forward to do scenario planning and predictive modelling" (Deloitte ZA, 2019).

"Data analytics, as it applies to fraud examination, refers to the use of analytics software to identify trends, patterns, anomalies, and exceptions within data" (Zack and ACFE, 2015, p. 6).

Ovenstående definitioner viser, at der er stor forskel på, hvordan dataanalyse defineres, og at definitionen afhænger af anvendelsesområdet. Fælles for dem er analyse af data ved brug af teknologi og software for derved at opnå indsigt i og ny viden om det undersøgte. Afhandlingen definerer herefter dataanalyse som

en proces, hvori data analyseres for at drage konklusioner om historiske begivenheder ved brug af teknikker til databearbejdning.

Da dataanalyse har en bred anvendelse, findes der et utal af forskellige teknologier og teknikker. For at få belyst de mest anvendelige inden for bekæmpelse af besvigelser vil der først blive redegjort for de mest almindelige aktuelle teknikker, hvorefter teorien vedrørende de nyeste og mest innovative og potente teknikker vil blive behandlet.

2.3.1 Business intelligence og klassiske dataanalyser

I intern kontrol har der, som beskrevet i afsnit 2.2.2, været anvendt klassiske kontroller i form af manuelle kontroller ved stikprøvevis gennemgang af transaktioner, godkendelseskontroller mv. eller automatiserede kontroller, hvorimod anvendelsen af dataanalyse har været begrænset. I intern kontrol vinder forskellige teknikker frem, hvilket hos mange virksomheder er begyndt med anvendelsen af "Business Intelligence", også kaldet BI.

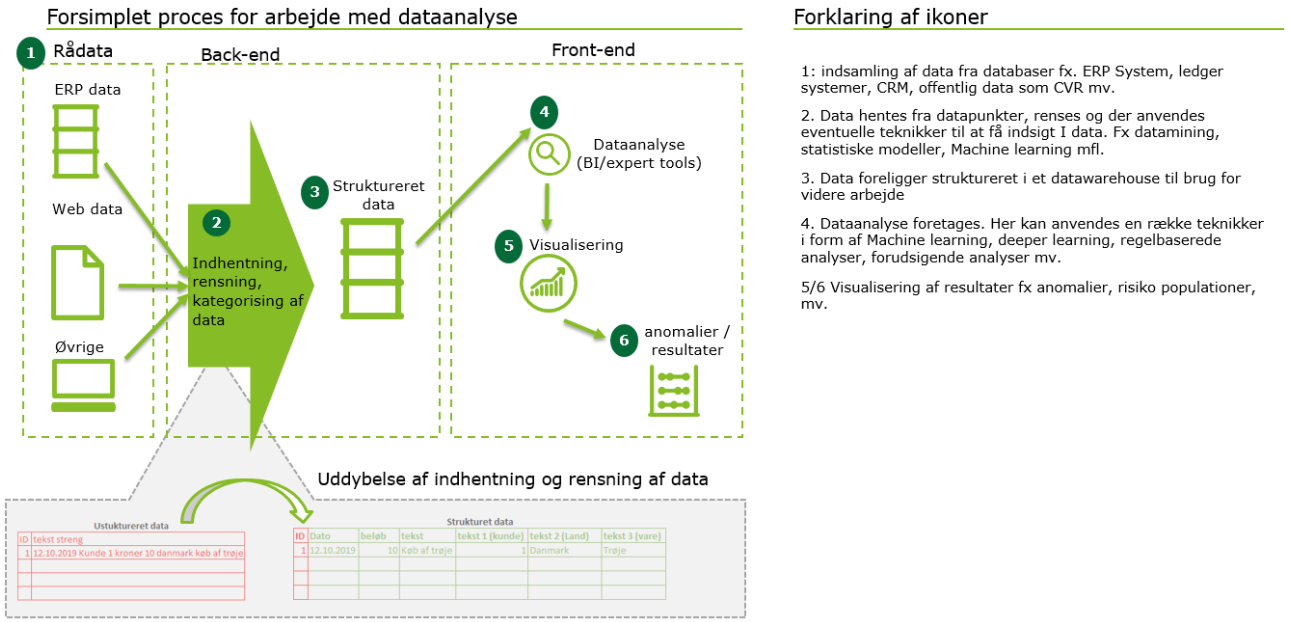
BI er et præsentations- og analyseværktøj, som typisk implementeres på toppen af virksomhedens ERP-system eller på toppen af et data warehouse, der opsamler og gør data læsbare til et givent formål.

Hos nogle virksomheder implementeres BI som en samlet del af virksomhedens it-infrastruktur og har derfor mange datakilder. Værktøjet anvendes til at præsentere data ud fra fastesatte parametre (Tableau, 2019a). En af de mest anerkendte organisationer inden for generel teknologi, Gartner, præciserer, at arbejdet med BI omfatter Data Mining, tekstbearbejdning, predictive analytics, applied analytics og statistik (Gartner, 2019). Hertil kommer den underliggende optimering af dataindsamling, it-infrastruktur og databaseopbygning.

Anvendelse af BI har været central i udviklingen inden for virksomhedsstyring for at kunne træffe begrundede realtidsbeslutninger baseret på data for at kunne respondere rettidigt på ændringer i virksomhedens forretningsområder. Denne udvikling har samtidig skabt grundlaget for, at der arbejdes struktureret med data til fejlsøgning og problemidentifikation hos virksomhederne (Tableau, 2019a).

Rent procesmæssigt betyder det altså, at det er et værktøj, der implementeres i virksomhedernes frontend-systemer og derfor anvendes som et ekspertværktøj over de strukturerede data, der leveres fra backend-

systemerne i virksomheden. For at give en forståelse af processen fra rå-data til slutresultatet i form af anomalier eller lignende er processen visualiseret i nedenstående procesdiagram.



figur 10 - Egen tilvirkning, dataanalyseprocessen

Backend-processen:

(1) Som det fremgå af procesdiagrammet, starter processen ved, at relevante datakilder identificeres. Disse datakilder er virksomhedens rå-data. **(2)** Rå-data i virksomhedens ERP- og CRM-systemer mv. foreligger ofte i en form, der kan defineres som ustruktureret data. Ustruktureret data vil typisk være meget omfangsrigt, uoverskueligt og svært at analysere. **(3)** For at kunne arbejde videre med virksomhedens rå-data, er man nødsaget til at forarbejde rå-data. Således vil det næste skridt i processen være at indsamle rå-data i en database i form af et data warehouse eller lignede, hvor rå-data bliver renset og dermed bliver gjort læsbare. Når de ustrukturerede data er blevet behandlet, defineres de som strukturerede data og kan nu anvendes til videre analyse.

Frontend-processen:

(4) For at kunne arbejde videre med det strukturerede data, anvender man typisk et expert tool i form af BI på toppen af det strukturerede data. Dette værktøj arbejder typisk regelbaseret eller efter en af de andre gængse anvendte analyseteknikker. Kendetegnet ved denne del af processen er, at den ligger frontend, kræver manuel behandling og derfor kræver løbende vedligeholdelse, så den følger virksomhedens udvikling og dens situation. **(5)** Som følge af den foregående analyse og visualisering heraf vil der typisk

være nogle anomalier, der skal undersøges manuelt. **(6)** Disse danner herefter grundlag for indhentning af underliggende dokumentation for at undersøge, om der evt. kan være tale om en besvigelse eller lignende.

Som beskrevet i afsnittet om frontend-processen har man historisk anvendt mere eller mindre manuelle analyseteknologier inden for BI. Nogle af disse teknikker blev listet i starten af dette afsnit. De enkleste teknikker er uddrag af både regelbaserede og statistiske analyser samt regressionsanalyser. En forsimplet redegørelse af disse teknikker følger i nedenstående afsnit.

Regelbaserede analyser er fundamentalt set analyser, der er baseret på input i form af regler, der herefter skaber et forventet resultat. I afsnit 2.2.4.1 blev en regelbaseret analyse gennemgået. Regler i en analyse til brug for intern kontrol kan illustrere og redegøre for sammenhænge mellem datakilder, beløbsmæssige specifikationer, oplysninger om land, afdeling, bruger, type m.fl. Regler har været og er fortsat det mest anvendte værktøj inden for dataanalyse, da de er nemme at sætte op og implementere. Regler er baseret på menneskelige erfaringer og input og regelbaserede analyser supporteres ofte af en visuel præsentation, hvorved brugeren kan foretage sin gennemgang og herved vurdere eventuelle risici samt eventuelt foretage yderligere undersøgelser.

Statistisk analyse har gennem historien haft stor anvendelse inden for intern kontrol og besvigelsesundersøgelser. Et eksempel på en anerkendt teknik er bl.a. Benfords lov. Denne teknik anvendes til at analysere en mængde data ud fra statistiske forventninger til en talrækkefølge. Hypotesen bag denne teknik er, at mennesker vil anvende unaturligt forekommende værdier oftere end naturligt forekommende værdier. Ved hjælp af dette kan man undersøge anomalier (Nigrini, 1999). Andre teknikker omfatter anvendelse af indikatorer til kategorisering af data ud fra fx beløbsmæssig størrelse, tidsmæssig placering, medianer, gennemsnit, kvartiler, standardafvigelser mfl. En anden statistisk analyse, der ofte har været anvendt, er regressionsanalyse.

Regressionsanalyse omhandler analyse af en eller flere variabler ud fra en mængde af uafhængige variabler. Regressionsanalyse eksisterer i forskellige kompleksiteter - fra simpel lineær regression til de mere avancerede modeller i form af logistisk regression (Danske, 2017). Formålet med en sådan analyse er at be- eller afkræfte den undersøgte variabels sammenhæng til den uafhængige population. Fx vil en samlet periodes ordrer kunne danne grundlaget for de uafhængige variabler, hvor en delperiodes ordrer sammenlignes. Ordrer, som falder udenfor normalen, vil herefter være at betragte som anomalier. Regressionsanalyse uddybes yderligere i afsnit 2.3.2.3.

2.3.1.1 Udviklingen inden for BI

Udviklingen inden for BI har været markant de seneste år. Traditionel BI er baseret på rapporter fra it-afdelingen, som er udviklet i samarbejde med ledelsen. Analyserne er typisk meget statiske og derfor også omkostningsfulde og bliver ofte ikke videreudviklet eller ændret efter brugerens behov (Baesens, Vlasselaer and Verbeke, 2015, p. 10). Ændringer kræver omfattende arbejde og ressourcer, hvilket har medvirket til, at traditionel BI opleves som et værktøj med begrænset anvendelse og efterfølgende stor irritation og frustration. Traditionel BI anvendes fortsat til statiske forespørgsler og rapporteringer som klassiske månedsrapporteringer eller udvalgte målepunkter, fx salgs- og segmentpræstationer.

I de seneste år har BI udviklet sig i et meget højt tempo og er gået fra en statisk løsning til at være et brugerbaseret værktøj med stor agilitet og dynamik. Anvendelsesområderne er blevet markant bredere, da værktøjerne nu kan modelleres efter brugerens ønske, der som udgangspunkt har adgang til enorme mængder data (Tableau, 2019a). Udviklingen skyldes, at udviklere i højere grad er lykkedes med at automatisere processen omkring databearbejdningen og dermed implementeringen af denne i backend-processen.

Inden for BI sker der fortsat en signifikant udvikling, og Tableau, som er en stor spiller inden for dataanalyseplatforme og er et produkt af forskning hos Stanford University, forudser bl.a., at kunstig intelligens vil finde vej ind i BI-værktøjer i form af Machine Learning og Deeper Learning. Dertil forudser Tableau, at teknologier som Natural Language Proofing[8] efterhånden bliver mere avancerede og anvendelige, hvilket skaber yderligere indsigt i data. I samme publicering forventer Tableau, at virksomhederne bliver bedre til at anvende dataanalyse (Tableau, 2019b).

2.3.2 Databearbejdning ved Data Mining og Machine Learning

I forbindelse med at virksomhederne har fået større indblik i deres data, udvikles og investeres der i højere grad i automatisering og smartere løsninger. I forlængelse heraf søges det ved hjælp af Data Mining og Machine Learning at komprimere hele frontend-processen og implementere denne som en del af backend-processen, fordi den manuelle bearbejdning og manglende agilitet i frontend-baserede værktøjer kan medføre manglende tilpasning af analyser og derfor give brugeren forkerte resultater. Endvidere kan det være omkostningstungt at vedligeholde og videreudvikle disse analyser, jf. afsnit 2.3.1.1.

[8] Teknologi til at blandt andet at tolke rådata og dermed automatisk rense og forberede disse til analysebrug.

2.3.2.1 Data Mining-processen

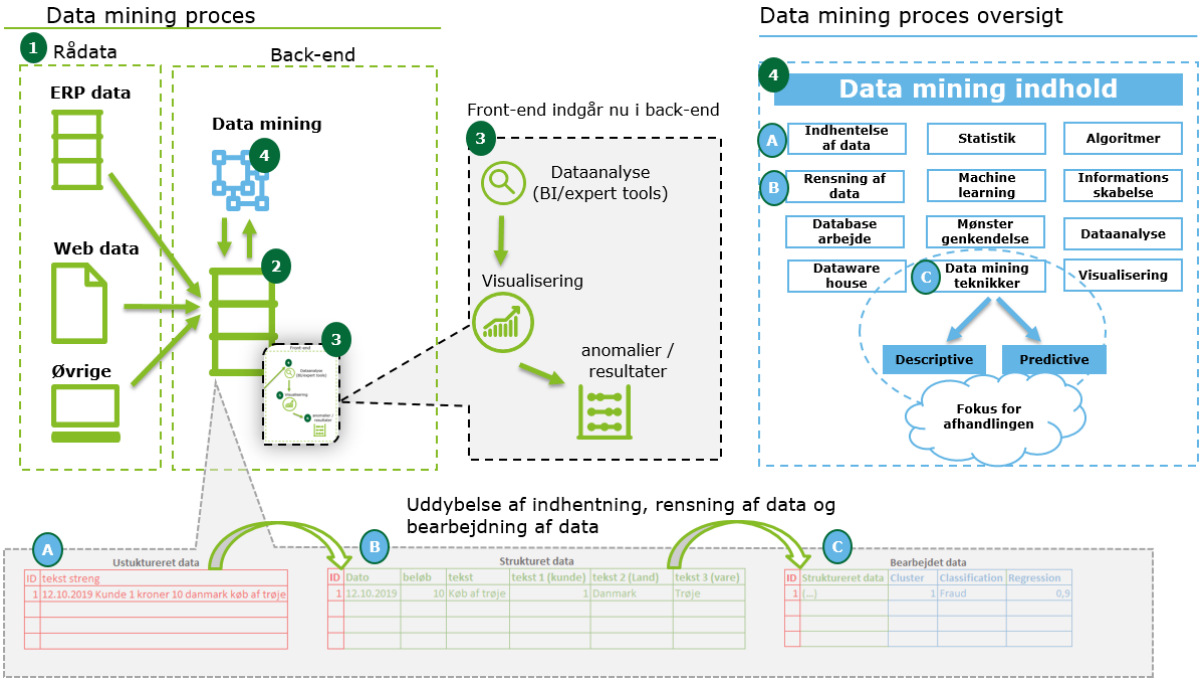
For at kunne komprimere hele frontend-processen og implementere denne som en del af backend-processen, kan virksomheder med fordel gøre brug af Data Mining for at automatisere processen og gøre den mere effektiv. I modsætning til den traditionelle brug af expert tools, dvs. ekspertværktøj og BI i frontend-processen, integreres hele bearbejdningen, analysen og visualiseringen af backend-data ved hjælp af Data Mining. Data Mining er en samlebetegnelse for det at indsamle og bearbejde rådata. Data Mining består derfor bl.a. af følgende elementer:

- Indhentning af data
- Rensning af data
- Bearbejdning af data mv.

Ovenstående er ikke udtømmende, men blot et lille udpluk af alle de kategorier, der falder ind under samlebetegnelsen "Data Mining". I figur 11, på næste side, er det illustreret, hvordan man ved hjælp af Data Mining kan strukturere rå-data.

(A) I selve data-indsamlingsprocessen vil man typisk få et datasæt, der er svært læsbart. **(B)** For at kunne arbejde videre med det, er man derfor nødsaget til at implementere en proces, der sikrer, at datasættet renses og struktureres, for at man kan arbejde videre med det. Data vil normalt ved udtræk fra SQL-databaser mv. bestå af tekststrenge, der omfatter numeriske værdier, datoer, lande og kunder mv. Som tidligere omtalt kan man anvende teknologier som Natural Language Proofing til at rense et sådant datasæt, så tekststrenge automatisk opbrydes og defineres i særskilte kolonner. indhentnings- og rensningsprocessen er blot få delelementer af Data Mining. Teorien vedrørende denne del af processen omtales ikke yderligere, da afhandlingens fokus er bearbejdning af data, hvorfor den videre redegørelse for Data Mining-teorien alene vil omhandle Data Mining-teknikker.

For at skabe et overblik over, hvilke dele af den traditionelle proces omkring dataanalyse der optimeres ved hjælp af Data Mining, er processen visualiseret i nedenstående procesdiagram figur 11.



figur 11 - Egen tilvirkning, Data Mining-processen

Som tidligere omtalt består dataanalyseprocessen traditionelt i 3 overordnede trin i form af **(1)** identifikation af relevante datakilder, **(2)** indhentning og rensning af rå-data i backend og **(3)** bearbejdning af struktureret data i frontend-funktionen. Processen, hvad angår input af rå-data, indhentning og rensning af rå-data, er uændret i denne del af processen.

Til forskel fra den traditionelle proces omkring dataanalyse, så indarbejdes, jf. punkt **(3)**, bearbejdning, analyse, visualisering, samt håndtering af anomalier direkte i et data warehouse. Således arbejder man ved hjælp af Data Mining-teknikker videre på det rensede datasæt. **(C)** Det vil sige, at man ved hjælp af automatiseret klassifikation, clustering og lignende teknikker strukturerer data til et givent formål, hvorpå en given model kan konkludere på anomalier eller præcist påpege, hvilke anomalier der bør undersøges nærmere. Der henvises til afsnit 2.3.2.3 for yderligere redegørelse af Data Mining-teknikker.

For at demonstrere den procesmæssige tilgang til dataanalyse ved anvendelse af Data Mining redegøres der for illustrationen i figur 11 nedenfor.

(1) Identifikation af relevante datakilder i form af ERP- og CRM-systermer eller andre relevante kilder. **(2)** Efterfølgende indhentes **(A)** og renses **(B)** de ustrukturerede data fra virksomhedens databaser som ved den traditionelle proces. **(3)** For at undgå den manuelle og potentielt omkostningstunge proces, der ligger i

at opsætte analyser og udlede konklusioner heraf, komprimerer man hele frontend-processen og indarbejder denne i backend-databasen. **(4)** Således bearbejdes datasættet ved at designe algoritmer på baggrund af anvendelige Data Mining-teknikker, som efterfølgende arbejder selvstændigt med datasættet og lærer heraf direkte i backend-databasen. **(C)** Bearbejdningsprocessen ved hjælp af Data Mining-teknikker medfører således, at der tilføjes nye informationer i de strukturerede data, som kan anvendes til analyseformål.

2.3.2.2 Machine learning-teorien

Som angivet i figur 11 er Machine Learning en integreret del af Data Mining. Der vil derfor blive redegjort for Machine Learnings relevans i forhold til Data Mining i dette afsnit.

Machine Learning er en underkategori til AI og formår selv, eller med begrænset indblanding, at skabe et resultat ud fra et datasæt eller tidligere erfaringer fra lignende analyser. Således medfører Machine Learning, at computere er i stand til at frembringe resultater tilnærmelsesvis på samme måde som et menneske. Målsætningen er således at få en maskine til at efterligne den kognitive intelligens, som et menneske besidder. Inden for Data Mining er Machine Learning således et sæt af metoder, som tillader en computer at genkende mønstre og trends, og på baggrund heraf at forudsige fremtiden eller frembringe konklusioner.

Machine Learning er en bred betegnelse for automatisk bearbejdning af data ved brug af algoritmer, matematik og programmering. Teknologien er anvendt inden for mange forskellige områder og består forsimplet af, at en maskine, i dette tilfælde en computer, bearbejder data og finder mønstre heri. Når maskinen genkender mønstre i et datasæt, er den programmeret til at vide, hvordan den skal bearbejde det. Teknologien arbejder dermed ud fra mønstre i historiske data og bearbejder nye data gennem tidligere opnået erfaring (EDUCBA, 2019c).

Eksempler på Machine Learning er Siri, Apples virtuelle assistent, som er trænet til at reagere, når den kan genkende en kommando. Metodikken er, at softwaren bag algoritmen har indsamlet en masse historisk data og kan genkende mønstre heri. Når Siri genkender et mønster i en kommando, agerer den på baggrund af dens læring ved algoritmen og returnerer korrekte output til brugeren. Som følge af de større og større akkumulerede datamængder er Machine Learning et effektivt værktøj til at bearbejde data og opdage mønstre, der skal reageres på.

Som tidligere nævn består Machine Learning af anvendelsen af algoritmer. Algoritmerne består af et sæt regler, der skal følges for at kunne løse en konkret problemstilling. Det kan fx både være komplekse matematiske ligninger og særlige instruktioner, som modellen opsættes til at følge. Modellen baseres på enten anerkendte matematiske modeller eller kombinationen af et test-datasæt og et foruddefineret output til at lære, hvordan den i en besvigelseskontekst kan forudsige eventuelle anomalier.

I Machine Learning arbejdes der inden for tre forskellige kategorier: superviseret læring, ikke-superviseret læring og forstærket læring (EDUCBA, 2019c). Forstærket læring er ikke i fokus for resten af afhandlingen.

Ikke-superviseret læring – handler om, at der anvendes anerkendte metoder og teknikker på et eksisterende datasæt. Der arbejdes uden en definition på outputtet, ligesom der ikke anvendes testdata til hjælp for algoritmen (Koen and Towards data science, 2019). Forskellige metoder hertil er oplistet i "Bilag 8 – Data Mining og underkategorier" og vil blive gennemgået i næste afsnit. Ikke-superviseret læring kan sidestilles med, at en elev skal løse en matematisk ligning, men kun kan anvende den læring, som eleven er i besiddelse af. Eleven har adgang til sine lærebøger og kan derfor slå op og genfinde den information, som eleven tidligere har været i besiddelse af. Eleven kommer herefter med sit bud på løsningen, hvorefter læreren vurderer, om eleven har løst opgaven, eller om eleven bør anvende en ny teknik til opgaveløsningen. I ikke-superviseret læring blander læreren sig således ikke i selve opgaveløsningen, men nøjes med at vurdere og evaluere resultaterne heraf.

Superviseret læring – handler i modsætning til ikke-superviseret læring om at definere et output samt at træne en algoritme ved hjælp af testdata (Koen and Towards data science, 2019).

Inden for superviseret læring skal algoritmen indledningsvis trænes for at kunne analysere på de data, som modellen bliver fodret med. Således er første inputdata et såkaldt test-datasæt, hvor brugeren fodrer algoritmen med historiske data. Jo flere anomalier, der er forekommende i datasættet, jo bedre, da algoritmen således har flere eksempler at lære fra.

Hvis eksemplet fra ikke-superviseret læring genanvendes, vil det betyde, at eleven nu får at vide, at han skal opdele nogle billeder efter, om de ligner katte eller hunde. Som testdata får eleven en bunke meget tydelige billeder af katte og hunde og laver derefter en opdeling af billederne. Læreren gennemgår efterfølgende resultaterne og evaluerer herpå. Når træningen har medført et tilfredsstillende resultat, får eleven efterfølgende det samlede datasæt. I dette tilfælde vil datasættet bestå af en stor bunke billeder af

hunde og katte i forskellige afskygninger. Eleven opdeler nu på baggrund af sin læring billederne ud fra sin viden om hunde og katte.

Machine learning kan derfor kort beskrives som automatiseret læring af en computer for at efterligne menneskelige kompetencer og derigennem automatisere store dele af arbejdet vedrørende databearbejdning og analyse. Inden for Machine Learning eksisterer der en teknologi, som hedder Deeper Learning. Denne teknologi omhandler det at skabe algoritmer, der efterligner den menneskelige hjernes opbygning for at kunne løse meget komplekse problemstillinger. Da besvigelser konstant ændrer sig, jf. afsnit 2.1.1, og på grund af de enorme krav til anvendelse af Deeper Learning-teknologien, vil den ikke blive belyst yderligere i afhandlingen.

Machine Learning er langt mere effektiv end manuelle kontroller og kan på længere sigt betyde, at der kan optimeres på omkostningerne forbundet med besvigelsesanalyser (PWC, 2019, p. 15). Machine Learning svarer til at have hundredvis af besvigelsesanalytikere til at fortage analyser. Fordelen ved Machine Learning er, at denne teknologi er i stand til at udføre analyserne i realtid og kan analysere på data på en brøkdel af den tid, et menneske vil bruge. Endvidere er det en fordel for denne teknologi med øgede datamængder, da den derfor har flere hændelser at analysere på (Ravelin, 2019).

Omkostningerne for virksomheden ved øgede datamængder ville stige forholdsvist, hvis det var personale, der skulle yde den større indsats, som det ville kræve at kunne kontrollere omfanget af den stigende mængde data (Ravelin, 2019).

2.3.2.3 Data Mining-teknikker

Som beskrevet i de foregående afsnit indeholder Data Mining forskellige teknikker til at bearbejde data. Disse defineres som Data Mining-teknikker. Teknikkerne er en overordnet betegnelse for bearbejdning af data ved hjælp af forskellige teknikker til at identificere eller skabe nye sammenhænge i data. Data Mining-teknikkerne kan således udvinde nye informationer, identificere mønstre og sammenhænge, som et menneske ellers ikke ville kunne gennemskue. Teknikkerne anvendes til både at danne et overblik over store mængder data, for herefter at gøre det læsbart for en bruger og at klargøre data til yderligere analyse. Brugeren kan på baggrund af denne information drage konklusioner, foretage undersøgelser eller træffe beslutninger (EDUCBA, 2019a).

Da Data Mining er en overordnet betegnelse for et utal af forskellige datateknikker, fremgår der i "Bilag 8 – Data Mining og underkategorier" en oversigt over de mest gængse teknikker inden for nedenstående anvendelsesområder.

figur 12 - Data Mining, uddrag fra Bilag 8 – Data Mining og underkategorier

Descriptive Data Mining - også kaldet ikke-superviseret Data Mining - anvendes til at finde mønstre og informationer i et datasæt baseret på historiske eller nylige hændelser. Denne type databearbejdning anvender altså historiske data til at skabe sammenhænge eller grupperinger i data. Analyseformen er reaktiv, da den kræver, at noget er sket, før hændelsen kan analyseres. Ved brug af denne teknik vil brugeren have besvaret, hvilken og hvorfor en hændelse er sket.

Eksempler på anvendelsesformer er: Trendanalyser, hvor datapunkter sammenholdes over en eller flere perioder, associationsskabelser, fx skjulte sammenhænge, og undergrupperinger mfl. Descriptive Data Minings primære formål er på den måde at sige noget om historikken og om at identificere, hvad der er "normalen", og hvad der falder uden for "normalen". Teknikken baserer sig udelukkende på karakteristika i de undersøgte data. Analyserne kræver minimal menneskelig indblanding, idet de ofte er baseret på anerkendte statistiske modeller og Machine Learning.

Dette kaldes inden for datavidenskab for ikke-superviseret Data Mining (Techdifferences, 2019) eller ikke-superviseret Machine Learning. Ved anvendelse af ikke-superviseret læring indlæses data i modellen, og modellen bearbejder data. Der anvendes ikke testdata til støtte for algoritmen til at finde output. Ikke-superviseret læring handler derfor om at skabe sammenhæng og mønstre baseret på karakteristika i data ved brug af statistik og Machine Learning. Machine learning vil blive uddybet senere i dette afsnit.

Analyseteknikken har, jf. ovenstående figur, to undergrupper bestående af association og clustering. I "Bilag 8 – Data Mining og underkategorier" fremgår forskellige underteknikker til disse. I nedenstående afsnit vil udvalgte teknikker blive uddybet.

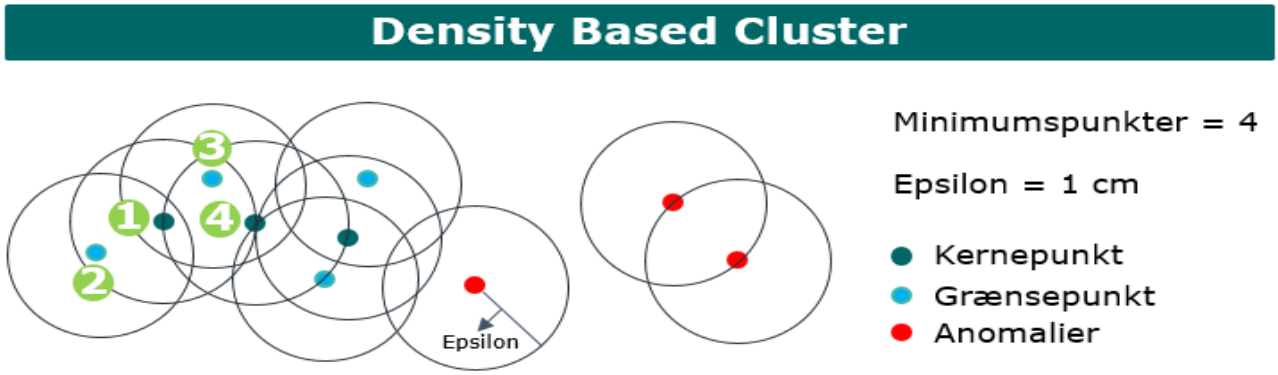
Association er en undergruppe til Descriptive Data Mining. Denne undergruppe indeholder teknikker til at skabe associationer eller sammenhænge mellem forskellige variabler i et datasæt. Der er tale om transaktionsbaserede teknikker. Et eksempel herpå er en teknik, der kaldes for Market Basket Model eller mere teknisk betegnet Apriority Algorithm, som i praksis er simpel og let at forstå. Modellen anvendes primært til at skabe sammenhæng mellem produkter i et supermarked. Fx vil der ofte være en sammenhæng mellem køb af mælk og køb af morgenmadsprodukter. Teknikken virker ved at analysere store datasæt for at finde sammenhæng mellem forskellige variabler. I førnævnte model vil indkøbskurve være betegnet som baskets eller kurve, og produkter vil være items. Analysen sammenligner herefter items med baskets for at finde sammenhænge i indkøbsvaner (Leskovec, 2016). Den variabel, som har færrest numeriske punkter, vil altid være basket i analysen. Fx vil der være langt færre indkøbskurve end den samlede mængde produkter i et supermarked.

Associationer kan skabes ved regler, hvor data analyseres via algoritmen, der herefter selv foreslår associationer i det undersøgte. Disse mønstre kan herefter danne grundlaget for yderligere analyse. En bruger kan fx have en forventning om en sammenhæng, og hvis der i resultaterne ikke er denne sammenhæng, kan det herefter danne grundlag for yderligere undersøgelse. I afsnit 2.2.4.1 blev der gennemgået et eksempel på sammenhæng mellem omsætning og lagertræk. Algoritmen vil i dette eksempel forventes at finde en sammenhæng mellem omsætningstransaktionerne og lagertransaktionerne. For transaktioner uden sammenhæng vil manglen herpå fremgå tydeligt, når forventningen var, at der burde være en sammenhæng. Resultaterne vil herefter kunne danne grundlag for yderligere undersøgelse.

Clustering er en anden undergruppe inden for Descriptive Data Mining. Clustering omhandler dannelsen af grupper eller clustre i et datasæt baseret på forskellige teknikker. Teknikkerne omfatter bl.a. hierarkiskabelse eller densitet. Hierarki er en metode til identificering af enten den største eller mindste variabel i et datasæt og på baggrund af dennes karakteristika at skabe et hierarki. De øvrige variabler placeres herefter i de forskellige hierarkier, indtil der er skabt hierarkier blandt alle variabler i det undersøgte datasæt (EDUCBA, 2019b).

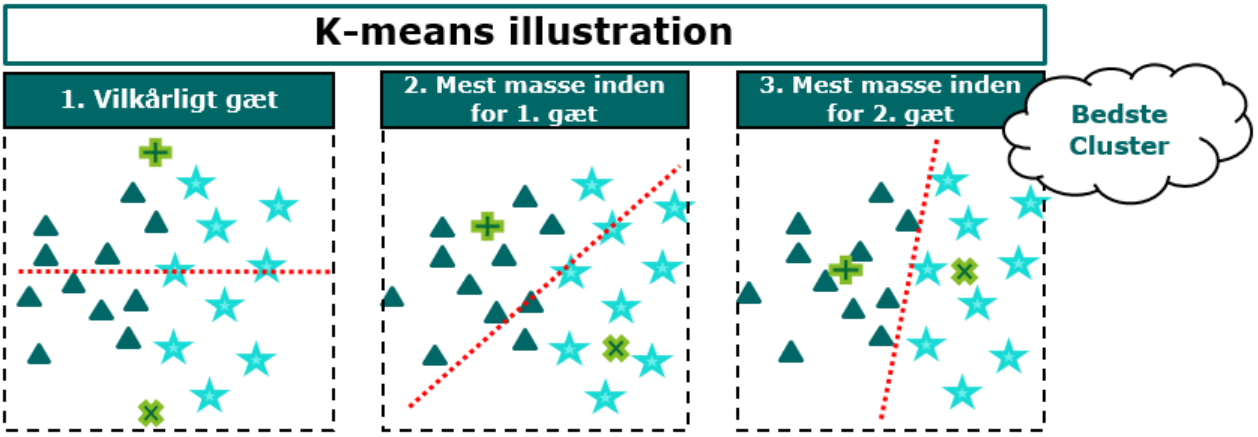
En anden metode er Density Based Clustering også kendt som DBSCAN. Denne metode går ud på at danne clustre ved at identificere kernepunkter og grænsepunkter. De punkter i datasættet, der ikke er tilknyttet clustre indeholdende kernepunkter og grænsepunkter, kategoriseres som anomalier. Forud for en Density Based Cluster-analyse skal variabler i algoritmen defineres. Disse variabler omfatter epsilon og

minimumspunkter. Epsilon angiver afstanden mellem datapunkterne i et datasæt, for at disse indgår i et del-cluster. Minimumspunkter definerer, hvor mange datapunkter der skal indgå i et del-cluster, før der er tale om et cluster. Når disse variabler er fastsat, identificerer algoritmen alle kernepunkter og grænsepunkter i datasættet. Kernepunkter er punkter, som opfylder begge fastsatte variabler. Et grænsepunkt er et datapunkt, som ikke opfylder variablens minimumspunkt, men som inden for den foruddefinerede epsilonværdi er tilknyttet et kernepunkt. Processen gentages for hvert eneste grænsepunkt, indtil alle punkter indgår i en klynge eller er blevet defineret som anomalier (IBM Analytics, 2017). Dette er illustreret i nedenstående figur, hvor datapunktet 1 defineres som et kernepunkt, da der indenfor en radius på 1 cm. (epsilonværdien) er fire andre datapunkter (minimumspunkter). Datapunktet 2 defineres som et græsepunkt, da der ikke forekommer fire andre datapunkter indenfor epsilonværdien. Dette datapunkt udspringer ikke som en anomalie, da der indenfor epsilon findes et kernepunkt (datapunkt 1). Sådan gentages processen, indtil alle datapunkter er clustred i del-clustre og anomalier er identificeret. I nedenstående figur indgår alle kerne- og grænsepunkter i samme del-cluster, hvorefter anomalier identificeres som illustreret med rød.



figur 13 - Density Based Clustering - egen tilvirkning

En af de mere anvendte teknikker inden for clustering er K-means. K-means forsøger at skabe clustre baseret på en eller flere lineære opdelinger. I nedenstående er dette illustreret:



figur 14 - Egen tilvirkning, K-means illustreret (Kutz and Washington University, 2016b)

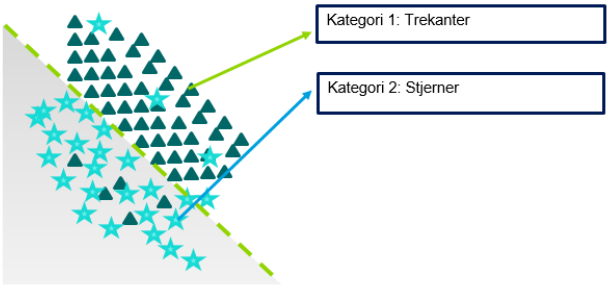
K-means fungerer ved, at der indsættes to vilkårlige punkter i et koordinatsystem bestående af de datapunkter, der analyses på. Disse punkter danner herefter modpoler, og mellem disse punkter tegnes en lineær streg. Hver side af stregen er nu en delpopulation. Herefter bearbejder algoritmen data for at lede efter masse i denne population. Når punktet for mest masse er fundet, flyttes de oprindelige punkter nu til de nye områder med mest masse og dermed de mest karakteristiske for populationen, og der tegnes en ny streg. Denne proces gentager algoritmen, indtil der fås to eller flere clustre med størst mulig sammenhæng (Kutz and Washington University, 2016b).

Clustering er en effektiv metode til at identificere anomalier i et datasæt. Når der er tilpas store mængder data, vil clustre kunne skabes mellem transaktioner, som har lignende karakteristika. Ved eksemplet fra før vil det i praksis betyde en gruppering af ordrer inden for fx segmenter eller indtægtskilder. Alternativt vil andre variabler kunne anvendes, såsom dækningsgrader, beløb, tidspunkt for postering, enhedspris mv. Posteringer, som falder uden for clustrene, de såkaldte anomalier, vil fremgå tydeligt for brugeren, som herefter vil have mulighed for at arbejde videre med dem, enten ved brug af en klassisk undersøgelse, jf. afsnit om intern kontrol, eller anvendelse af informationen til yderligere modellering. Ikke-superviseret learning kan være et effektivt grundlag, når der skal arbejdes med Predictive Data Mining. De mønstre, som kan skabes inden for Predictive Data Mining, kan danne rammerne for testdata til brug for de superviserede teknikker. Dette vil blive uddybet under klassificering.

Predictive Data Mining er den anden gren i Data Mining og omhandler arbejdet med superviseret læring. Superviseret læring er teknikker, hvor der arbejdes med outputsøgning. Algoritmerne trænes med testdata, og forventet output defineres. Testdata er et uddrag af et datasæt, der indeholder forskellige variabler, men det er afgørende, at det indeholder en tilpas mængde af de observationer, som brugeren ønsker at

identificere. Vil brugeren træne en model til at identificere besvigelser, skal datasættet som minimum indeholde et par posteringer, som kan kategoriseres som anomalier eller besvigelser. Et effektivt værktøj til dette er de føromtalte Descriptive Data Mining-teknikker, som kan identificere anomalier. Det kan være grundlaget for Predictive Data Mining, og således kan uddrag fra et datasæt fra en clustering-teknik danne grundlag for testdata. Endvidere kan informationen, som tilvejebringes gennem denne teknik, lette processen med modellering ved anvendelse af Predictive Data Mining. Når der arbejdes med superviseret læring, gentages øvelsen, såfremt resultaterne ikke har den ønskede præcision. Der er således tale om en trial and error-baseret programmering. Predictive Data Mining omfatter både klassificering af data, regression og det at forudsige resultater baseret på karakteristika i datainput samt vægtningen heraf. Der er derfor ikke udelukkende tale om teknikker, hvis formål er at forudsige fremtiden, men især også teknikker, der kan anvendes til at søge hændelser, som er defineret af brugeren. Inden for den type Data Mining arbejdes der ofte med to undergrupper, klassifikation og regression. Disse vil blive uddybet nedenfor (Techdifferences, 2019).

Klassifikation er en gruppe af teknikker, der anvendes til at klassificere data ud fra prædefinerede kategorier. Kategorier er oversat fra ordet "labels" og vil sige, at noget fx defineres som enten besvigelse eller ikke-besvigelse. Kategorierne opsættes af brugeren af algoritmen, og disse skal prædefineres, før klassificeringen kan påbegyndes. Eksempler på kategorisering er fx: ja eller nej, høj, lav eller middel. Der arbejdes enten med binær kategorisering, hvilket vil sige to variable, fx 1 eller 0, eller ja eller nej, eller multipelkategorisering ved mere end to kategorier. Når kategorierne er sat op, anvendes der forskellige teknikker til at klassificere data i opsatte kategorier. Nedenfor vises et eksempel på kategorisering af data (Kutz and Washington University, 2016a).

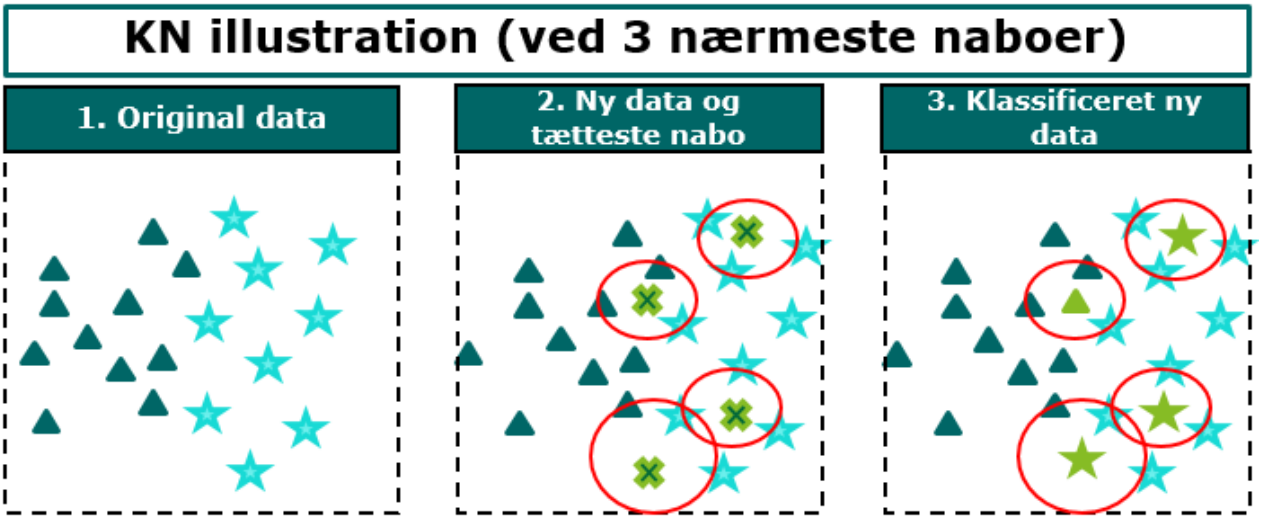


figur 15 – Klassifikation af stjerner og firkanter, egen tilvirkning

I eksemplet ovenfor vil brugeren klassificere sine data som enten stjerner eller trekanter. For eksemplets skyld er data illustreret ud fra deres reelle karakteristika ved anvendelse af stjerner og trekanter. Brugeren

klassificerer sit datasæt ved brug af en algoritme og en binær kategorisering, 1 = Trekanter, 2 = Stjerner. Algoritmen bearbejder data og tegner herefter den grønne streg, der illustrerer, at alt over stregen kategoriseres som en trekant, og alt under stregen betegnes som en stjerne. Brugeren er interesseret i at undersøge sine stjerner yderligere. Som illustrationen viser, sker der i denne klassificering en fejlklassificering af hhv. de 4 stjerner, som bliver til trekanter, og de 5 trekanter, som bliver til stjerner. Dog har brugeren ved hjælp af klassificering reduceret sit datasæt betydeligt og kan nu i højere grad undersøge sine observationer.

Klassifikation er, som ovenstående eksempel viser, en meget brugbar teknik til at opdele data i mindre undergrupper og dermed muliggøre en risikorettet undersøgelse af de datapunkter, der kategoriseres efter brugerens ønske. Overstående viser endvidere, at klassificerings modeller bygger på struktureret læring, da brugeren her har defineret et output. Inden for klassificering er der mange forskellige teknikker - de spænder fra ”normal” Machine Learning til de mere avancerede deeper learning-modeller. I ”Bilag 8 – Data Mining og underkategorier” er forskellige teknikker oplistet. En af de mere anvendte klassifikationsteknikker inden for superviseret læring er ”K-nearest neighbor”. Denne teknik forsøger at skabe klassifikation ved at lede efter et datapunkts nærmeste tilhørsforhold og er illustreret nedenfor.

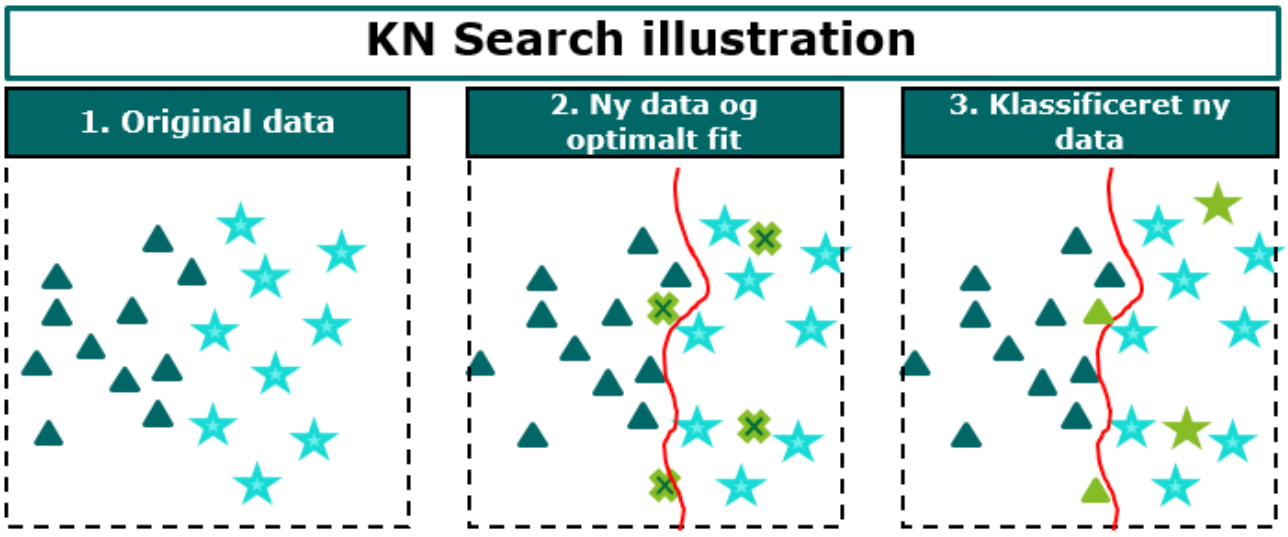


figur 16 – Egen tilvirkning, KNN-illustration

Illustrationen viser først test-data. Herefter indsættes nye data, her illustreret som X’er. Når testdata er indsat, søger algoritmen, i dette eksempel, efter de tre nærmeste naboer da K i eksemplet er defineret som tre. Der anvendes således tre naboer for at minimere risikoen for fejlklassificering. Herefter lader algoritmen de tre nærmeste naboer ”stemme” om, hvorvidt datapunktet er en stjerne eller en trekant. Trin

tre viser, hvordan de nye datapunkter er blevet transformeret til stjerner eller trekanter. **K-nearest neighbor** eller k-nærmeste nabo er således en effektiv måde at bruge sit test datasæt til at klassificere nye datapunkter. Det baserer sig på en logik om, at data placerer sig efter sine karakteristika, og at disse karakteristika kan forklare noget om data. Dog kan ovenstående også medføre, at enkelte punkter bliver klassificeret forkert. I billede to ses et punkt, der ligger mod venstre og over imod trekanterne, men i billede tre bliver dette til en stjerne, pga. de nærmeste naboer. Hvis dette ikke er korrekt, sker der en fejlklassificering (Kutz and Washington University, 2016c).

En metode, der kan anvendes som alternativ eller på toppen af ovenstående teknik, er **K-nearest neighbor search**. Denne metode er baseret på en tankegang om optimale tilhørsforhold. I stedet for at anvende nærmeste nabo søger den derfor efter en samlet logik mellem alle datapunkterne. I nedenstående er dette illustreret.



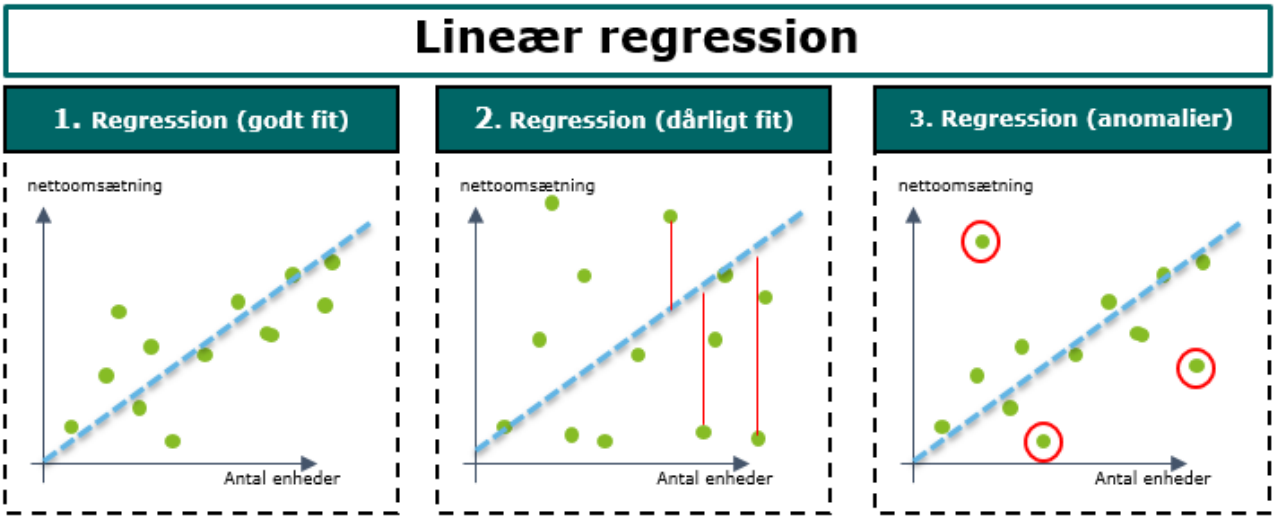
Figur 17 – Egen tilvirkning KNS

Denne metode kan således opnå et andet resultat, som i nogle tilfælde kan være mere retvisende for de undersøgte data. Oftest vil disse skulle krydsvalideres ved brug af forskellige metoder. Derfor er det essentielt, at resultaterne efterprøves med forskellige testdata for at sikre, at der ikke sker en fejlklassificering ved et overfit (Kutz and Washington University, 2016c).

Regression er det sidste ben inden for Predictive Data Mining. Regression kan opdeles i henholdsvis lineær og ikke-lineær regression. Regression anvendes til at klassificere komplekse problemstillinger eller til at forudsige fremtidige hændelser. Regression er som udgangspunkt et værktøj til at forklare sammenhænge mellem forskellige datapunkter ud fra en forklarende og afhængig variabel. Da regression befinder sig

inden for feltet for superviseret Machine Learning, betyder det, at variablerne i en regressionsanalyse vil være defineret af brugeren(Baesens, Vlasselaer and Verbeke, 2015). Outputtet i en sådan analyse vil derfor være defineret på forhånd, ligesom anvendelsen enten vil være rettet mod et klassifikations- eller et forudsigelsesspørgsmål. Ikke-lineær regression vil ikke blive anvendt yderligere, hvorfor denne teknik ikke vil blive uddybet. Vi vil i afhandlingen ikke fokusere yderligere på forudsigende analyser.

Lineær regression er en teknik, der anvendes på en forklarende og en afhængig variabel, hvor variablerne har en lineær sammenhæng. Det betyder, at såfremt en variabel stiger med 1, vil den have en tilnærmelsesvis lineær effekt, hvis den tilsvarende variabel stiger med 2. Fx vil variabel 1 kunne være solgte enheder og variabel 2 nettoomsætning. Såfremt der sælges 1 enhed, vil nettoomsætningen være 100 kr., og hvis der sælges 2, vil omsætningen formentlig være 200 kr. Dette betyder en lineær korrelation. Ved brug af lineær regression vil der blive beregnet en optimal linje mellem datapunkterne i analysen. Hvis der fx er 10 salgsobservationer, beregnes en regressionskurve, der udtrykker de projicerede indtægter. I nedenstående illustration vises forskellige eksempler på lineær regression.



figur 18 – Egen tilvirkning, lineær regression

I illustrationen fremgår tre forskellige grafer. Den første graf viser forskellige salgsobservationer og en regressionslinje. Regressionslinjen er den forventede nettoomsætning, baseret på de solgte enheder ud fra observationspunkterne. Der er tale om en regression med et godt fit, hvilket betyder, at observationerne ligger tæt på den beregnede regressionslinje. I graf to er det modsatte illustreret. Her ses store variationer mellem regressionslinjen og observationerne. Det betyder, at der vil være stor usikkerhed forbundet med forventningen. I det sidste billede er anomalier identificeret ved at sammenholde det, som kaldes error

eller fejl, hvilket er afstanden fra observationen til regressionslinjen, med de øvrige observationers fejl. Eksemplet her viser, hvordan regression, i lighed med teknikkerne under klassifikation, kan klassificere datapunkter og dermed tildele dem til en kategori.

Redegørelsen for Data Mining viser hermed, at der kan arbejdes med forskellige teknikker til at bearbejde data og finde mønstre heri. En vigtig del af arbejdet handler om at finde de rigtige teknikker til at behandle problemstillingen. Tilgangen hertil er trial and error-baseret og handler om at afprøve forskellige teknikker på problemstillingen ud fra problemstillingens- og teknikkernes kendetegn. Det betyder, at arbejdet med Data Mining ofte vil være en iterativ proces, hvor der sker en naturlig rotation i principperne, ligesom konklusioner bør efterprøves ved anvendelse af andre teknikker eller med nye data. Såfremt krydsvalideringen ikke foretages, er risikoen for falske positiver et ikke ubetydeligt parameter. En falsk positiv er fx, når et datapunkt kategoriseres som værende A, men burde være B. Det betyder fx, at en besvigelsestransaktion kategoriseres som en normal transaktion og dermed mister undersøgelsen sin validitet. For mange falske positiver eller negativer kan gøre arbejdet med efterfølgende undersøgelser uoverskuelige (Statistics How To, 2015).

Der er store fordele ved at bruge en kombination af teknikkerne både på tværs af Descriptive og Pedictive Data Mining, men også imellem teknikkernes undergrupper. Fordele og ulemper vil blive analyseret og diskuteret i næste kapitel.

Ved redegørelsen for Data Mining blev Machine Learning præsenteret. For at få en dybere forståelse af arbejdet med Machine Learning henvises der til næste kapitel.

2.3.3 Opsummering på dataanalyse

Efter gennemgangen af dette afsnit er læser nu blevet gjort bekendt med, hvad dataanalyse er, og hvordan det traditionelt anvendes. Det er blevet belyst, at den traditionelle tilgang kan effektiviseres ved at implementere dataanalyse i virksomhedens backend-systemer ved hjælp af Data Mining, herunder Machine Learning. Der er ligeledes redegjort for de mest anvendelige teknikker med afsæt i afgrænsende besvigelsestyper. De mest anvendelige teknikker inden for henholdsvis ikke-superviseret og superviseret læring er ”Density Based Clustering”, ”K-Means”, ”K-Nearest Neighbor”, ”K-Nearest Search” og ”Lineær Regression”.

Hensigten med dette teoriafsnit har været at skabe den fornødne teoretiske forståelse for læseren af anvendelse af dataanalyse, samt hvilke Data Mining-teknikker, der kan anvendes til opdagelse af besvigelser. Det teoretiske fundament og kendskabet til de nye Data Mining-teknikker vil blive anvendt og analyseret i besvigelseskontekst i afhandlingens analyseafsnit.

2.4 Delkonklusion på teorien

I delkonklusionen vil vi konkludere på undersøgelsesspørgsmålene inden for besvigelsesteori, intern kontrol og dataanalyseteorien. Konklusionerne vil danne rammerne for det efterfølgende analyseafsnit.

- ***Hvad karakteriserer de besvigelser, der har haft størst konsekvens for virksomhederne?***

Konsekvensen af besvigelser er, at virksomheder på globalt plan årligt taber 26 billioner DKK. Dette tal er estimeret af ACFE i deres seneste rapportering "Report to the Nations". Bag estimatet forekommer der et utal af forskellige besvigelser. I rapporten fremgår det, at regnskabsmanipulation er den type besvigelse, som har den største enkeltstående konsekvens. Herudover er misbrug af aktiver den hyppigst forekommende besvigelsestype. De to besvigelsestyper har sammenlagt omtrent den samme beløbsmæssige konsekvens for virksomhederne. Indenfor disse to kategorier er "indregning af fiktiv omsætning" og "faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen" de besvigelsestyper, som har haft de største konsekvenser for virksomhederne. Med afsæt i denne rapport konkluderes det, at vi ved at afdække disse besvigelser vil opnå den største værdi med analysen.

En af de besvigelser, der har størst konsekvens for virksomhederne er "faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocesser". Denne type besvigelse er kendetegnet ved, at en betroet medarbejder stjæler udbetalinger i indkøbsprocessen i forbindelse med fakturering. Centralt for udførslen af denne besvigelse er, at besviger skal være i stand til at omgå godkendelsesprocedurer i indkøbsprocessen. Denne type besvigelse fordrer, at en fiktiv/manipuleret faktura eller betaling bliver godkendt, hvorfor godkendelses- og review-elementet er det helt essentielle kontrolpunkt i denne besvigelsestype.

Indregning af fiktiv omsætning er kendetegnet ved, at den begås på højt ledelsesniveau. For at kunne lykkes med den type besvigelse kræver det, at besvigeren udfører opgaven fare for nemt at blive opdaget. Denne form for besvigelse anses dog for at være af en mere avanceret og sofistikeret karakter end andre typer besvigelser, hvorfor denne type besvigelse typisk vil foregå med medsammensvorne eller ved hjælp

af ledelsens tilsidesættelse af kontroller. Kompleksiteten i tilsløringen af sådan en besvigelse vil således være mere kompleks end for øvrige besvigelser, da besvigeren antageligvis må have evner og kompetencer, der gør det muligt.

Centrale forhold og karakteristika fra besvigelsesteorien er skematiseret i figur 19. Resultaterne i nedenstående skema anvendes også som grundlag i analysens identifikation af besvigelseskarakteristika.

Besvigelser med størst konsekvens for virksomhederne

Typer af besvigelser	Forekomst og konsekvens
Misbrug af aktiver	***57 % af alle besvigelser (mediantab 743.000 DKK)***
• Faktureringssvindel i forbindelse med indkøbsprocessen	• 20 % af alle besvigelser indenfor misbrug af aktiver
Regnskabsmanipulation	***1 % af alle besvigelser (mediantab 5.200.000 DKK)***
• Indregning af fiktiv omsætning	• 34 % af alle besvigelser indenfor regnskabsmanipulation

Målbare karateristika ved de to afgrænsede besvigelsestyper:

Faktureringssvindel igennem skyggeselskaber	Indregning af fiktiv omsætning
• Leverandørens reelle ejer er en medarbejder eller nærtstående hertil.	• Tilsløring sker via balancen
• Godkendelsesprocedure omgås i forbindelse.	• Posteringer følger ikke de normale forretningsgange
- Omkostningsgodkendelse	• Manuelle posteringer foretaget på ledelsesniveau
- Kreditoroprettelse	• Ændring i samhandelsmønstre
- Ændringer i kreditorstamdata	• Manglende cashflow trods positive resultater
• Ofte immaterielle ydelser	• Management override
	• Sammenfald mellem leverandører og kunder
	• Manglende lagertræk og leverance knyttet til omsætningen
• Manglende leverance fra skyggeselskabet	• Andre karakteristika i form af: usædvanlige beløb og timing mv.

figur 19 - Skematisering af centrale forhold og karakteristik fra besvigelsesteorien til analyseafsnit - egen tilvirkning

Anden del af teoriafsnittet omhandlede interne kontroller, hvor nedenstående undersøgelsesspørgsmål blev forsøgt besvaret.

- Hvad er interne kontroller, og hvordan anvendes disse i dag til at forebygge besvigelsers forekomst?

Intern kontrol er påkrævet af den øverste ledelse for at sikre forsvarligt opsyn med virksomheden. I mange tilfælde vælger den øverste ledelse at uddelegere arbejdet til den daglige ledelse, men ansvaret påhviler fortsat den øverste ledelse. Interne kontroller består af aktiviteter, som virksomheden iværksætter med henblik på at imødegå de risici, som virksomheden igennem sine aktiviteter eksponerer sig overfor. COSO-Frameworket om interne kontroller fra 2013 danner rammerne for de interne kontroller i mange virksomheder rundt om i verden. I COSO-Frameworket opdeles interne kontroller i 5 komponenter.

Komponenterne består af: kontrolmiljø, risikovurdering, kontrolaktiviteter, it og kommunikation, samt overvågende og monitorerende aktiviteter.

Ud fra et besvigelsesperspektiv har der været behov for præcisering og dermed også øget fokus på området som konsekvens af det store problem, som besvigelser er for mange virksomheder. I forlængelse heraf har COSO udgivet en udførlig guide til håndtering af besvigelser. I afhandlingen er det primært besvigelsesframeworkets forhold ved risikovurdering og kontrolaktivitet, som er i fokus. Ved gennemgang af dette framework blev det identificeret, at interne kontroller i høj grad kan supporteres af dataanalyse. I publikationen om dataanalyse er der indarbejdet en metodisk tilgang til dataanalyse som et effektivt værktøj til bekæmpelse af besvigelser, hvilket ligeledes fremhæves af ACFE i deres seneste rapport.

I teoriafsnittets tredje del blev teorien om dataanalyse behandlet, herunder formålet om at besvare nedenstående undersøgelsesspørgsmål.

- *Hvad er dataanalyse, og hvilke anvendelsesmuligheder har teknologien Data Mining.*

Dataanalyse kan defineres som en proces, hvori der analyseres på data for at drage konklusioner om historiske begivenheder ved brug af teknikker inden for databearbejdning. De seneste år har virksomhederne arbejdet mere intensivt med BI og expert-tools som frontend-værktøjer, hvor bearbejdning og analyse bliver foretaget på toppen af strukturerede data fra virksomhedens backend-systemer.

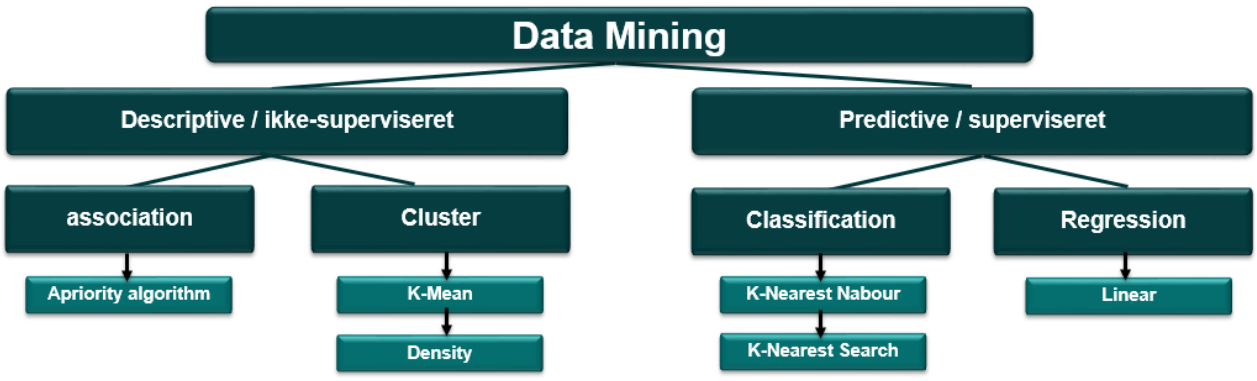
Da virksomheder, som følge af globaliseringen og den digitale transformation, løbende står over for en særdeles omskiftelig omverden, kan frontend-databearbejdning være omkostningstungt og svært at udvikle og vedligeholde bl.a. pga. stort forbrug af personaleressourcer. Ved hjælp af ny teknologi i form af Data Mining og Machine Learning kan virksomhederne implementere deres frontend-systemer i form af bearbejdning og analyse direkte i deres backend-systemer. Backend-systemer består af databaser, data warehouses og integrerede analyseværktøjer. Ved hjælp af Data Mining og derigennem Machine Learning samt disses bagvedliggende algoritmer, kan hele eller dele af databearbejdningen overlades til maskiner. Maskinerne kan bearbejde og analysere på data på en brøkdel af den tid, et menneske vil bruge, og sandsynligvis komme med mere præcise resultater. Årsagen hertil er, at Data Mining-teknikker er i stand til automatisk at tilpasse bearbejdningen af rå-data til algoritmen, i takt med at teknologierne lærer af de

data, som algoritmen bearbejder. Endvidere kan nye teknikker nemt tilføjes til virksomhedens analyseapparat.

Data Mining er en bred samlebetegnelse for processering af rå-data. Data Mining omfatter blandt andet indhentning af data fra forskellige kilder som fx ERP- og CRM-systemer, rensning af rå-data, visualisering mv. Yderligere kan Data Mining-teknikker anvendes til at frembringe ny viden. Med udgangspunkt i afhandlingens problemfelt og afgrænsning fandt vi, at de relevante dele af processen udelukkende vedrørte anvendelse af Data Mining-teknikker, hvorfor de øvrige dele af processen udelukkende danner de forståelsesmæssige rammer for læseren.

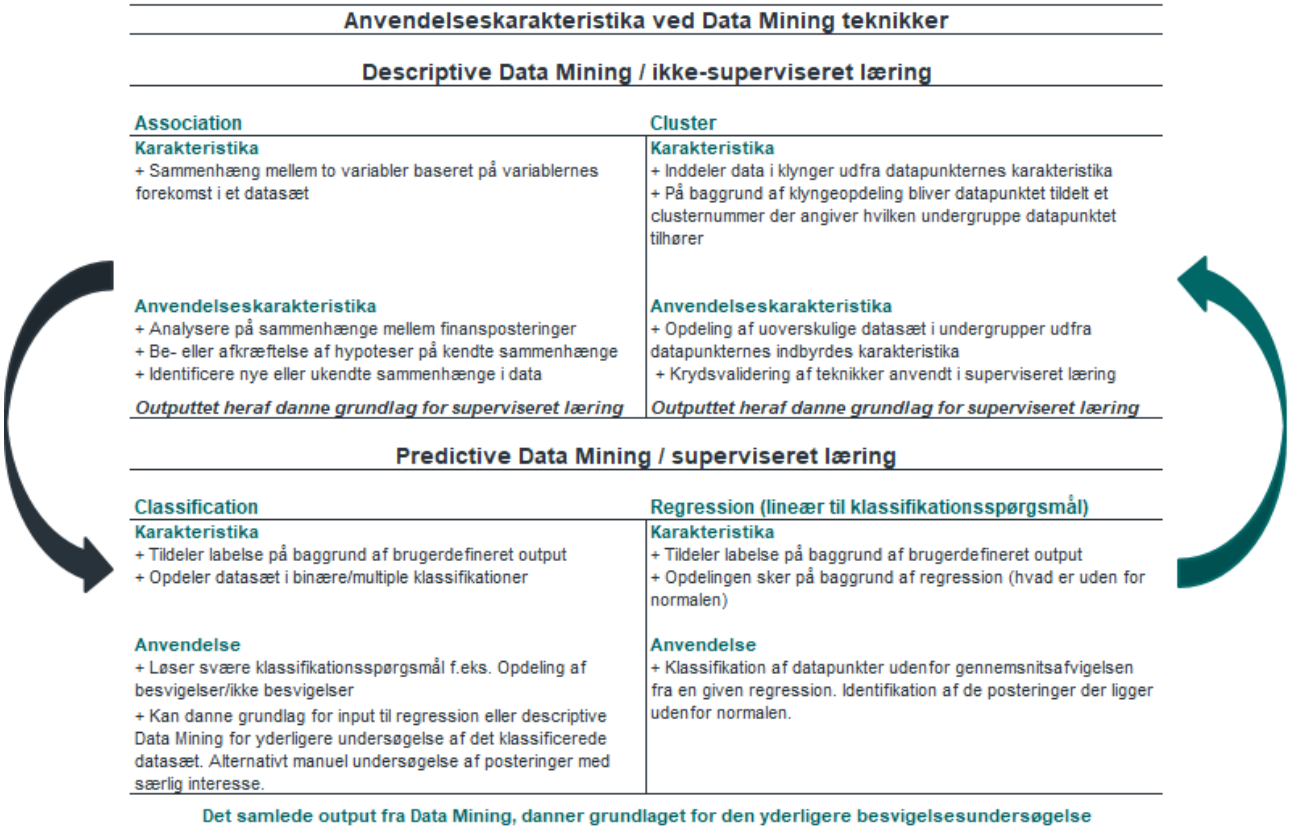
Data Mining-teknikker opdeles i superviseret og ikke-superviseret læring. Superviseret læring omfatter, at en algoritme bliver trænet til at komme med et foruddefineret output. Algoritmen trænes herefter ved brug af testdata indeholdende de observationer, som brugeren vil have algoritmen til at finde. Superviseret læring består af to underkategorier, ”Classification” og ”Regression”, som med hver deres tilgang har til formål at bearbejde data ud fra et særligt sæt regler. Ikke-superviseret læring består i anvendelse af anerkendte matematiske eller statistiske modeller og teknikker på eksisterende datasæt, hvor brugeren alene forholder sig til og vurderer resultaterne fra modellen. Ikke-superviseret læring er ligesom superviseret læring delt op i to underkategorier. Underkategorierne til ikke-superviseret læring er ”Association” og ”Cluster”.

Inden for de nævnte underkategorier eksisterer der et utal af forskellige teknikker, som hver især har deres berettigelse inden for hver sit område. Inden for bekæmpelse af besvigelser må det konstateres, at en håndfuld teknikker er særligt relevante til bearbejdning af data. Disse teknikker og tilhørsforhold inden for Data Mining er illustreret i nedenstående figur.



figur 20 - Relevante Data Mining-teknikker til analyse - egen tilvirkning

I figur 21 er ovenstående datateknikkers karakteristika og anvendelsesområder skematiseret for at danne overblik over centrale kendetegn udvundet af dataanalyseteorien til besvarelse af det tredje undersøgelsesspørgsmål. De centrale kendetegn er endvidere en del af grundlaget for valg af Data Mining-teknikker i analysen.



[9]

figur 21 - Skematisering af Data Mining-teknikker til analyseafsnit - egen tilvirkning

[9] Pilene illustrerer at der er tale om en iterativ proces. Grundlaget for predictive Data Mining anvendes således igen i descriptive Data Mining og modsat.

Kapitel 3 – Analyse og diskussion

3 Analyse

Dette kapitel har til formål at analysere faresignaler, der indikerer besvigelser inden for afhandlingens afgrænsede besvigelsestyper. Herunder, hvilke målbare karakteristika der ligger bag de identificerede faresignaler. I forlængelse heraf vil det blive udledt af analysen, hvilke datakilder der indeholder informationer om de identificerede karakteristika. Når faresignaler, karakteristika og datakilder er identificeret, vil der ske en analyse af, hvordan disse observationer kan anvendes til at designe Data Mining-analyser, og af, hvilke teknikker der er mest anvendelige. Slutteligt vil det blive analyseret, om de valgte teknikker og analyser bør implementeres som en del af virksomhedernes interne kontroller, og om disse løsninger er bedre end dem, der anvendes i dag.

3.1 Identifikation af faresignaler

Vi har ved et større litteraturstudie besøgt relevante databaser inden for besvigelsesundersøgelser for at kortlægge kendte faresignaler ved henholdsvis besvigelser inden for ”faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocesser” og ”indregning af fiktiv omsætning”.

Kilderne henviser i overvejende grad til ACFE’s studier af besvigelser begået på verdensplan og bekræfter herigennem observationerne i ”Report To The Nations”. Litteraturstudiet medførte, at der blev identificeret adskillige faresignaler, som hver i sær har vist sig at være til stede i forbindelse med afsløringen af en bagvedliggende besvigelse. Analysen affødte en bruttoliste over typiske faresignaler, som kan danne grundlag for opdagelsen af nye besvigelser og derigennem arbejdet med Data Mining. Faresignalerne er af generel karakter pga. disses følsomhed over for offentligheden og derfor ikke er konkretiseret i reelle casestudier.

Bruttolisten over faresignaler om faktureringssvindel i skyggeselskaber i forbindelse med indkøbsprocesser fremgår af ”Bilag 9 – Bruttoliste for faresignaler i forbindelse med faktureringssvindel”. Den tilsvarende bruttoliste for fiktiv omsætning fremgår af ”

Bilag 10 – Bruttoliste faresignaler, indregning af fiktiv omsætning". De nævnte bilag indeholder det samlede grundlag for nedenstående analyse af, hvilke faresignaler og de afledte datapunkter, der har relevans for opdagelsen af besvigelser, og som derigennem kan anvendes i Data Mining-teknikkerne.

3.1.1 Faresignaler, variable og relevante datakilder inden for faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen

Den omtalte bruttoliste vurderes for omfattende til videre analyse, hvorfor vi har bearbejdet bruttolisten og søgt sammenhænge imellem de enkelte faresignaler. Som følge heraf er bruttolisten indsnævret til de mest centrale faresignaler inden for faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen.

Vi fandt i forbindelse med litteraturstudiet, at flere af de identificerede faresignaler falder ind under den kategorisering, som blev foretaget i teorigennemgangen. Studiet understreger således, at faresignaler inden for de tre kategorier er nogle, som virksomheden bør identificere i forbindelse med design og implementering af Data Mining-teknikker. Rette design og implementering heraf vil medføre, at virksomheden effektivt vil kunne forebygge og opdage mulige besvigelser inden for denne kategori af besvigelser. De tre omtalte kategorier er opstillet nedenfor og vil danne grundlag for den strukturelle analyse heraf.

- Reelle ejere
- Omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange
- Typen af leverancer og manglende leverancer

På baggrund af kategoriseringen af de identificerede faresignaler, har vi rationaliseret os frem til relevante variable og datakilder, som vurderes anvendelige i forbindelse udarbejdelsen af en effektiv Data Mining-model til at forebygge og opdage besvigelser inden for faktureringssvindel. Sammenhængen mellem kategorier, faresignaler, variabler og datakilder vil blive analyseret i nedenstående afsnit for at demonstrere rationalet bag valget af variabler og datakilder til design af de relevante analyser. For læseren er det endvidere illustreret i "Bilag 11 – Reelle ejere, udvalgte faresignaler, variabler og datakilder".

Reelle ejere - er som tidligere beskrevet en effektiv tilsløringsmetode i forbindelse med faktureringssvindel. Således vil besvigeren, som tidligere beskrevet, forsøge at tilsløre besvigelsen ved at foretage fakturering til et skyggeselskab, som besvigeren selv kontrollerer. Der er adskillige faresignaler, som kan indikere, at der

forekommer faktureringssvindel. Der vil naturligt være et sammenfald mellem de enkelte faresignaler. Særligt kan mønstre i fakturanumre indikere, om der er tale om en besvigelse. Indikationer på den type besvigelser kan være, at der forekommer usædvanlige mønstre i fortløbende fakturanumre fra en bestemt leverandør, fx uafbrudt fakturanummerering fra en bestemt leverandør, hvilket kan tyde på, at leverandørens eneste kunde er den pågældende virksomhed (Vona, 2017). Dette forekommer typisk i et skyggeselskab, som alene er konstrueret til at frarøve den forurettede virksomhed dens aktiver. Relevante variabler til at identificere besvigelsen ved fakturanumre vil være: **(a)** seneste fakturanummer på leverandørniveau sammenholdt med **(b)** foregående fakturanummer på leverandørniveau. Hvis afvigelser mellem hvert fakturanummer kun udgør 1 eller en anden unormalt forekommende værdi, kan det være udtryk for fakturering fra et skyggeselskab, og at der derfor er risiko for en besvigelsesforekomst. Relevante datakilder til opdagelse af besvigelse her vil være ERP-systemer i form af kreditormoduler og underliggende kreditorposteringer, da disse indeholder informationer om fakturanumre og leverandører og ikke kræver den store arbejdsindsats.

Andre eksempler på faresignaler, der kan indikerer, at der er tale om en besvigelse inden for kategorien ”reelle ejere”, er fx datering af faktura indenfor 90 dage efter virksomhedsregistreringen af en relevant leverandør, adressesammenfald eller indkøb fra en ny samhandelspartner af produkter eller ydelser, hvor virksomheden ellers har en stamleverandør. For at kunne identificere besvigelser ved ovenstående faresignaler, vil relevante variabler være fakturanumre, gennemsnitsafvigelser i fakturanumre på leverandørniveau, fakturadato såvel som stamdata på leverandører og medarbejdere. Ved at analysere på tværs af disse variable vil tilsløringen af reelle ejere i de fleste tilfælde blive identificeret som anomalier. Det vil således blive analyseret, hvordan disse variabler i samspil med de rette Data Mining-teknikker kan identificere potentielle besvigelser i relation til faresignaler tilknyttet reelle ejere. Datapunkterne vurderes at kunne findes i virksomhedens ERP-systemer for kreditormoduler og de bagvedliggende kreditorposteringer, såvel som i virksomhedsregistre og HR-moduler.

Omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange - Det må antages, at der kan forekomme eksempler, hvor tilsløringen i forbindelse med reelle ejere er så gennemført, at besvigelsen ikke kan identificeres ud fra de faresignaler, som relaterer sig hertil. Derfor kan faresignaler i relation til omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange være med til at afsløre besvigelsen. En indikator på at godkendelsesprocedurer eller sædvanlige forretningsgange er blevet omgået for at udføre besvigelsen, kan være, at der sker indkøb under tærsklen for review eller godkendelse. Det vil sige, at besviger bevidst fabrikerer fiktive fakturaer i et skyggeselskab, som ligger under grænsen for

godkendelse eller review, hvilket kan medføre, at det beløbsmæssige omfang af besvigelsen med tiden akkumulerer på enkelte leverandører (AGA Intergov, 2019a). Andre faresignaler kan være, at der sker indkøb af varer i afdelinger, som ikke er i nær tilknytning til den indkøbte leverance, eller at godkendelse af små omkostninger sker på managerniveau eller højere, hvilket ikke følger de normale forretningsgange. Dertil kommer anvendelse af Benfords lov mod den medarbejder, der initierer godkendelsen. Det vil sige, at hvis en given medarbejder har overvægt at godkendte fakturaer, hvor beløbet strider mod Benfords lov, så kan det være et tegn på en besvigelse. For at kunne identificere besvigelsen igennem ovenstående faresignaler, vil en kobling af treshold, antal tidligere indkøb over/under treshold, indkøber, godkender og fakturabeløb være interessante variabler, som kan være udslagsgivende i forhold til normalen inden for godkendelsesprocedurer og sædvanlige forretningsgange. Relevante datapunkter vurderes at kunne udledes af datakilder såsom ERP-systemer i form af kreditormoduler og bagvedliggende kreditorposteringer såvel som HR-/lønsystemer og omkostningsgodkendelsessystemer.

Typen af leverancer og manglende leverancer - Er tilsløringen af besvigelsen så sofistikeret, at den ikke har kunnet afsløres igennem reelle ejere eller omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange, kan "typen af leverancer" medføre, at besvigelsen afsløres. Et faresignal, der kan indikere, at der er begået en besvigelse, er, at der er overvægt af serviceydelser på enkelte leverandører (Wolf, 2016). Årsagen hertil er, at besviger typisk vil forsøge at tilsløre besvigelsen ved at foretage indkøb af serviceydelser, grundet deres uhåndgribelige karakter. Leverancen af ydelsen er således svær at måle i modsætning til et materielt indkøb. Relevante variabler herpå er produkttyper på posteringsniveau for at kunne udrede, om der er tale om en immateriel leverance eller ej, samt leverandørnavn/-nummer. For leverandører, hvor der foretaget materielle indkøb, kan besvigelsen afsløres ved manglende leverancer. Et faresignal kan være, at leverandørfakturaen ikke kan spores til efterfølgende fragtdokument eller varemodtagelse (AGA Intergov, 2019a). Endvidere kan store svindprocenter på enkelte varenumre tyde på, at der ikke er sket en levering svarende til det pågældende indkøb, og at der dermed kan være tale om en besvigelse. Relevante variabler herpå vil være fragtdokument/varemodtagelse, fakturanummer og svind på varenumre. Disse variabler vil således blive videreført i analysen. Ovennævnte variabler vurderes at kunne udledes af datakilder som fx ERP-systemer i form af kreditormoduler, bagvedliggende kreditorposteringer og lagermoduler.

3.1.2 Faresignaler inden for indregning af fiktiv omsætning

Den anden besvigelse, som blev identificeret, med størst konsekvens for virksomhederne er indregning af fiktiv omsætning. Indregning af fiktiv omsætning forekommer i forskellige variationer, herunder med forskellige niveauer af tilsløring og kompleksitet. I "Bilag 10 – Bruttoliste faresignaler, indregning af fiktiv omsætning" fremgår en oversigt over kendte faresignaler og karakteristika, der er identificeret, når indregning af fiktiv omsætning er blevet opdaget. Vi har med udgangspunkt i bruttolisten foretaget en udvælgelsesproces baseret på en vurdering af, om de enkelte faresignaler overlappes af andre faresignaler, og om de forskellige faresignaler vil være til stede i virksomhedens data og derfor kan anvendes til analyse.

De faresignaler, som er fundet relevante, vil blive gennemgået i nedenstående afsnit inden for følgende tre kategorier: Tilsløring via balancen, posteringer som ikke følger normale forretningsgange og ændring i samhandelsmønstre. Resultatet af analysen vil danne grundlag for valget af Data Mining-teknikker. Faresignalerne er identificeret ved en rationalisering og vurdering af, hvor i virksomhedernes data faresignalerne vil fremgå ved at vurdere, hvor besvigelserne sandsynligvis vil fremgå tydeligst. Vi har løbende sikret, at datapunkter kan beregnes eller fremfindes ved at sammenholde datapunkterne med anerkendte økonomisystemers tabeldokumentation[10]. I "Bilag 14 – Fiktiv omsætning - udvalgte faresignaler, variabler og datakilder" har vi desuden skematiseret dette.

Tilsløring via balance er som beskrevet en effektiv tilsløringsmetode til at skjule besvigelser. Karakteristika vil på et aggregeret niveau være uforklarlige stigninger i aktivsummer samt manglende pengestrømme. Hvis der analyseres på det laveste niveau, vil det betyde, at selve transaktionen ikke skal være pengestrømsgenererende eller være forsøgt tilsløret et sted, hvor det ikke er forventningen, at der skal ske en løbende udligning. Det vil derfor være at foretrække, at virksomhederne analyserer på transaktionernes modkonto. Poster, som ikke bogføres som en debitor eller tilsvarende, bør herefter tilkendegives som en postering indeholdende en risiko for besvigelse. Efterfølgende vil brugerinformationen være et vigtigt parameter for den videre analyse. De to punkter vil, i forlængelse af den øvrige information om fx dato, beløb mv., skabe grundlaget for den videre analyse. Datapunkterne vil kunne genfindes i ERP-systemet, hvorefter brugeroplysninger kan beriges med information fra et HR-system.

[10] Tabeldokumentationen omfatter information om forskellige tabellers felter. Fx indeholder en debitorliste felterne "kunde", "kundenummer" mfl. Udgangspunktet for vurderingen er Microsoft Dynamics Navision tabeldokumentation.

Posteringer, som ikke følger normale forretningsgange, omfatter mange forskellige faresignaler, der i høj grad afhænger af tilsløringsformen. For enkelte af disse faresignaler vil der være et naturligt sammenfald. Fx vil "bogføring foretaget af personer, som sjældent foretager bogføring", "brug af manuelle posteringer" og "salg uden vareforsendelse" formentlig have sammenfaldende variabler i højere eller mindre grad. Det mest sofistikerede vil være at forsøge at skjule besvigelsen i det reelle faktureringssystem, således at posteringen ikke automatisk tilkendegives som manuel. For at identificere denne type besvigelse, vil forholdet mellem omsætningsposteringen og det tilhørende vareforbrug i form af en dækningsgrad være anvendelig. Hertil vil oplysningen om, om der er sket en reel lagerafgang, kunne være i stand til at verificere salget. Hvis dækningsgraden er unormalt høj kan det i nogle tilfælde indikere, at der er noget galt. Dækningsgraden som variabel vil for andre typer af tilsløring inden for denne kategori også være udslagsgivende. Fx vil "ændring af regnskabsposteringer" også medføre en unaturlig høj dækning. Derfor tages disse variabler med i det videre arbejde. Informationen vil som udgangspunkt stamme fra virksomhedernes ERP-system. Herefter vil data vedrørende ordrer kombineres med oplysninger fra et CRM-system. For lageroplysninger vil disse enten hentes i undertabel i ERP-systemet eller i et særskilt lagersystem.

Ændring i samhandelsmønstre omhandler besvigelser, hvor de normale forretningsgange omgås for at tilsløre besvigelser. Eksempler på faresignaler ved denne type tilsløringsform er brugen af falske kunder eller unormale transaktioner med nærtstående parter. For besvigelser, som er tilsløret ved anvendelse af falske kunder, vil det være stamoplysningerne på kunderne, som har relevans. Såfremt en kunde er fiktiv, vil der være forskellige punkter, som kan valideres og dermed have interesse i relation til opdagelsen eller forebyggelsen af besvigelser. Navn og adresse vil være de første og mest nærliggende variabler at validere. Her kan oplysningerne valideres op imod offentlige registre som fx CVR. Såfremt oplysninger i de offentlige registre og virksomhedens stamdata er overensstemmende, tilskrives der en variabel, som tyder på, at kunderne reelt er registreret hos en myndighed. En anden relevant validering kan være, at de normale forretningsgange er overholdt. Her kan en validering op imod virksomhedens CRM-system være interessant. Såfremt der ikke er oprettet en ordre på den pågældende kunde, beriges datasættet med et nyt felt, hvori det fremgår, at der ikke er oprettet en ordre. Øvrige variabler, vi har fundet relevante, er, om det er første salg til kunden, eller om der er sammenfald mellem den person, som både initierer salget og har oprettet kunden. De forskellige informationer kan kombineres og skabes ud fra selskabets ERP-posteringer og ændringslogs, databasen bag fx Virk.dk. og CRM-systemet.

En anden besvigelseshandling inden for kategorien" ændring i samhandelsmønstre" omhandler tilsløring via samhandlen mellem nærtstående parter. Her er posteringerne ofte interessante, hvis de afviger fra normalen. For at analysere denne indikator opdeles posteringer, alt efter om der er tale om handel med nærtstående parter eller ej. Her søges at berige posteringsinformationen med et ja eller nej. I denne proces kan oplysninger fra bogholderiet, i form af modkontoen, eller oplysning om, hvorvidt kunden er registreret som nærtstående, finde anvendelse. En anden tilgang er et check af virksomhedens navn mod kundens navn eller kontrol af, om der sammenfaldende adresser. Er der sammenfald, flages denne som nærtstående. Andre måder at opnå informationen vil være: Opslag mod et datasæt indeholdende alle nærtstående parters navne, eller opslag mod centrale registre, som indeholder diagrammer over nærtstående fx BiQ.dk. Såfremt der er tale om en nærtstående part, kan øvrige relevante variabler bestå af dagens nummer i den måned[11], posteringen er foretaget på eller antal gange, posteringen overstiger gennemsnittet for den normale samhandel med kunden. Data til brug for disse datapunkter vil stamme fra virksomhedens ERP-system, Virk.dk, lignende databaser eller oplysninger fra CRM-systemet.

I det næste afsnit vil ovenstående variabler danne grundlaget for de forskellige Data Mining-teknikker, som kan anvendes til at afdække de identificerede besvigelsesrisici. Overblik fremgår i "Bilag 14 – Fiktiv omsætning - udvalgte faresignaler, variabler og datakilder".

3.2 Analyse af teknikker

I det foregående afsnit blev forskellige faresignaler og deres karakteristika gennemgået. Disse karakteristika blev herefter omskrevet til datapunkter og derigennem til variabler, som kan anvendes til analyseformål. I dette afsnit vil vi vise, hvordan kendte karakteristika kan danne grundlag for arbejdet med Data Mining som et værktøj til opdagelse af besvigelser.

Første del af arbejdet er at identificere to variabler, som i fællesskab kan sige noget om en population af datapunkter. Ved den rette kombination af variabler vil der opstå mønstre i data, som ellers ikke ville være mulige at identificere. Denne del af analysen er således kritisk for, om resultaterne er brugbare til opdagelse eller forebyggelse af besvigelser. I nedenstående afsnit vil vi analysere de forskellige datapunkter inden for de to risici, som blev identificeret i afsnit 2.1.2. Når datapunkterne er valgt, vil vi vise, hvordan

[11] Ved dagens nummer i måneden menes, at den 12. oktober er dag 12, hvilket også vil være gældende for den 12. januar. Der ses dermed bort fra måneder.

forskellige teknikker inden for Data Mining kan bearbejde data og derigennem danne grundlag for videre undersøgelse.

3.2.1 Data Mining-teknikker rettet mod faktureringssvindel i forbindelse med indkøbsprocessen

Som der blev redegjort for i teoriafsnittet om besvigelser, er disse meget varierende og typisk på forkant med virksomhederne. Årsagen til, at besvigerne er på forkant med virksomhederne, er, at de hele tiden tænker i mere sofistikerede metodikker for at tilsløre besvigelsen. For at være foran besvigerne, må virksomhederne være mindst lige så kreative som besvigerne og tænke i strukturer, omstændigheder og forhold, som ikke vil blive fanget i de traditionelle interne kontroller. Inden for faktureringssvindel er det således vores vurdering, at der eksisterer tre typer af faktureringssvindel inden for brugen af skyggeselskaber. De tre typer kombinerer forskellige områder inden for besvigelsesteorien og kan tilsløres på så sofistikerede og komplicerede måder, at de er svære at fange i traditionelle interne kontroller:

- Direkte fakturering igennem skyggeselskaber på baggrund af fiktive fakturaer
- Fakturering igennem skyggeselskaber via ikke-medskyldige leverandører, hvormed der ikke længere er samhandel
- Salg af varer til overpriser igennem skyggeselskaber (viderefaktureringssvindel)

Det vurderes, at alle tre konstruktioner kan afsløres ved at analysere de i afsnit 3.1.1 identificerede karakteristika bestående af:

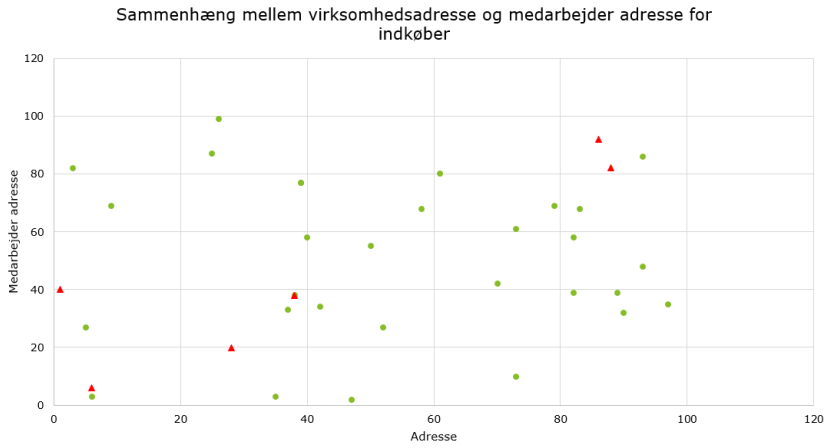
- Reelle ejere
- Omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange
- Typen af leverancer og manglende leverancer

Vi vil i nedenstående afsnit foretage en analyse af, hvilke variabler fra afsnit 3.1.1 der, i samspil med de udvalgte Data Mining-teknikker, kan være med til at afsløre ovenstående besvigelser. Analysen tager udgangspunkt i et datasæt, som er konstrueret ud fra litteraturstudiet af faresignaler og karakteristikaene i disse. Datasættes over besvigelser kombineres herefter med posteringer, som udtrykker normale transaktioner. Datasættes anvendes udelukkende for at kunne analysere variablernes indbyrdes sammenhænge og for hvordan Data Mining teknikkerne kan identificere besvigelser. I datasættet er der anvendt oversampling, som er en teknik inden for datascience, som kort fortalt går ud på at

overrepræsenterer det, som der analyses efter (Baesens, Vlasselaer and Verbeke, 2015) - dette tilfælde besvigelseslignende posteringer. I datasættet er besvigelseslignende posteringer illustreret med røde trekanter og normale posteringer med grønne.

3.2.1.1 Reelle ejere

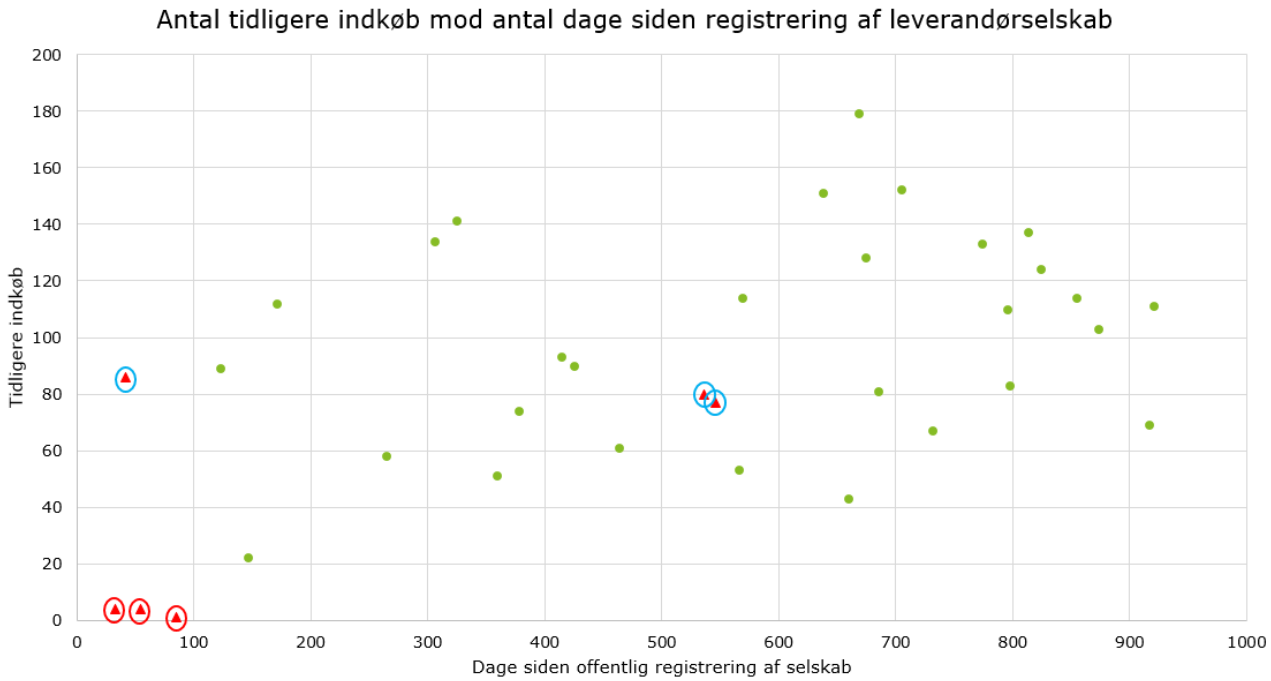
Som benævnt i afsnit 3.1.1 er der adskillige faresignaler forbundet med reelle ejere i forbindelse med fakturering igennem skyggeselskaber i forbindelse med indkøbsprocessen, som kan indikere, at der er sket en besvigelse. Således bør virksomheden analysere de variabler, som har nær tilknytning til de pågældende faresignaler. Det antages, at en af de første analyser, der vil blive foretaget, vil være at se, om der er sammenfald mellem medarbejder- og leverandøradresser. Inden der kan analyseres herpå, skal data bearbejdes, således at de strukturerede data definerer adresser med unikke talværdier, som således symboliserer en given adresse. På den baggrund sammenholdes variablerne for at se, om der er sammenfald. Dette eksempel er illustreret i figur 22.



figur 22 - Analyse af sammenhæng mellem virksomhedsadresse og adresser for indkøber - egen tilvirkning

Som det fremgår af illustrationen, har besviger tilsløret besvigelserne i en sådan grad, at der alene er sammenfald mellem medarbejdere og leverandøradresser på én af posteringerne. Tilsløringen fremgår effektiv, da de røde (besvigelsesposteringerne) er kamufleret blandt de grønne (normale) posteringer. Således fanger denne analyse kun 1/6 af de begåede besvigelser og tegner således ikke et generelt mønster for, hvor Data Mining-teknikker vil berige undersøgelsen med yderligere information.

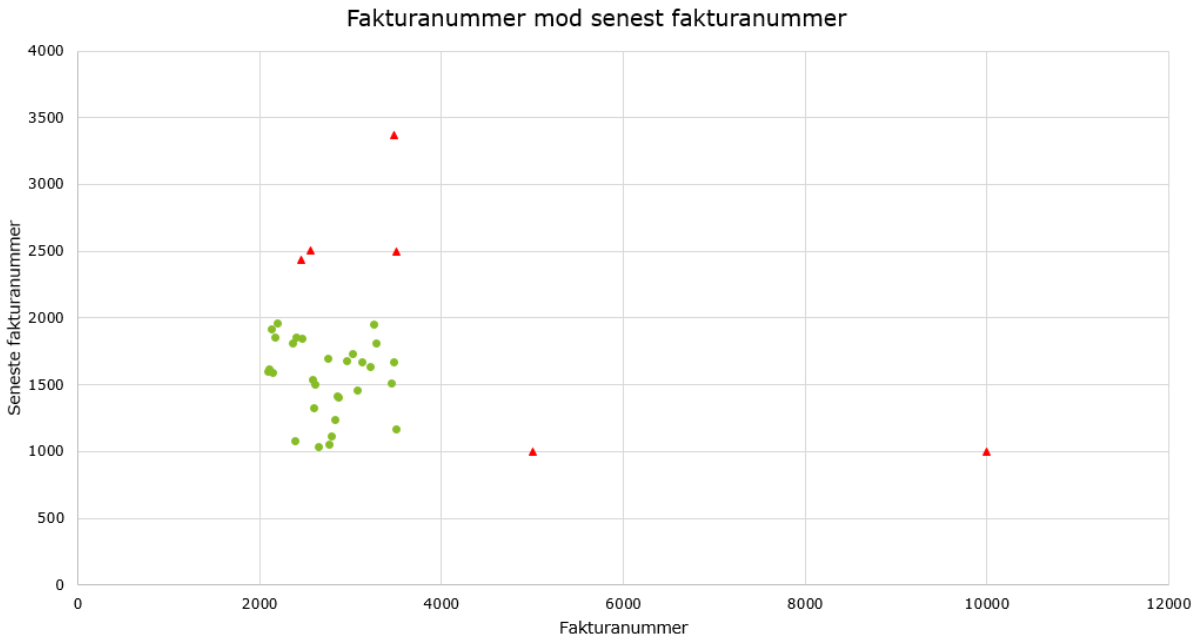
Som følge af at ovenstående analyse ikke gav det ønskede resultat, vil det være interessant at se på, om der er sammenhæng mellem det tidligere antal indkøb og registreringshistorikken på leverandørniveau. Hypotesen vil her være, at en besvigelse igennem et skyggeselskab vil ske kort efter, at skyggeselskabet er oprettet. Denne analyse er illustreret i figuren nedenfor.



figur 23 - Analyse af sammenhængen mellem tidligere indkøb mod antal dage siden registrering af leverandørselskab - egen tilvirkning

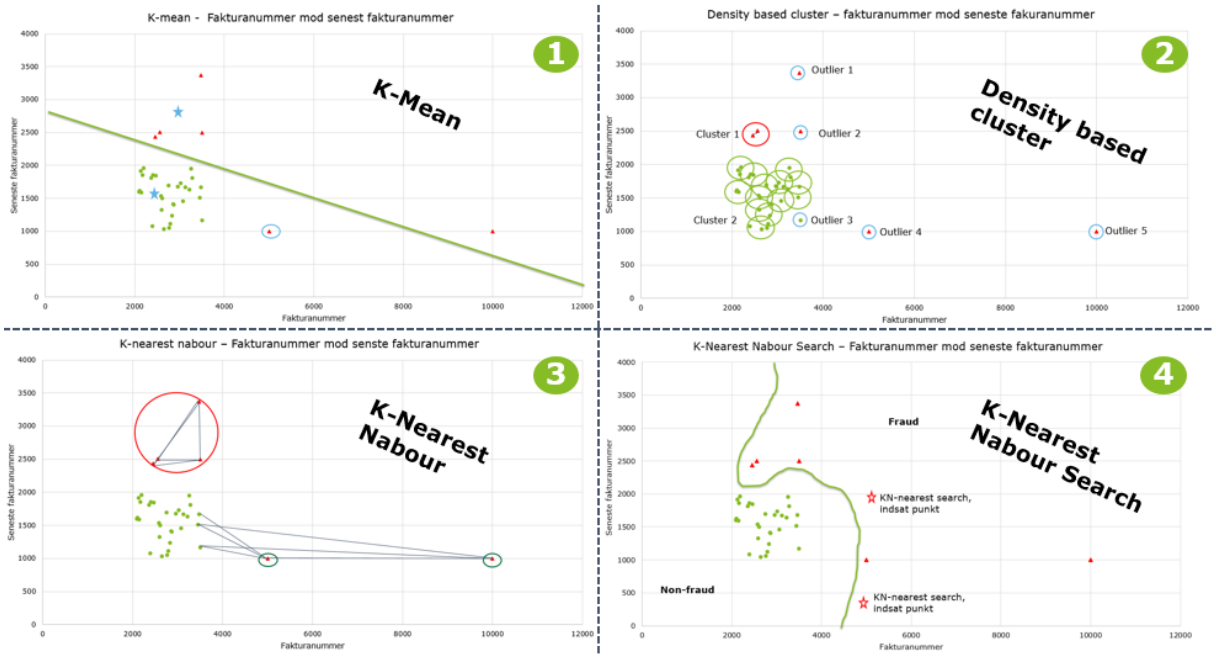
Som det fremgår af illustrationen, er det særligt tre posteringer, der springer i øjnene (røde cirkler). Disse datapunkter må anses som anomalier, fordi de bærer de karakteristika, der kan indikere en besvigelse, da der er sket indkøb kort tid efter, at selskabet er registreret i det centrale virksomhedsregister. De øvrige datapunkter følger dog ikke disse karakteristika, da det antages, at besviger i to tilfælde har købt et eksisterende selskab, der nu anvendes som skyggeselskab, hvilket gør, at disse følger mønstret på normale posteringer. Således er der risiko for, at Data Mining-teknikker, som fx K-Nearest Neighbor, vil medføre to falske positiver, som illustreret med blå markeringer i illustrationen. Således er det alene 3/6 af besvigelserne, der identificeres i denne analyse.

Som supplement til ovenstående analyser vurderes det, at en ustruktureret udvikling i fakturanumre kan identificere besvigelserne, da det er svært ved manipulation at ramme en naturlig udvikling i fakturanumre. Således sammenholdes fakturanummer med seneste fakturanummer for at se, om der tegner sig et særligt mønster. Analysen er illustreret i figur 24.



figur 24 - Analyse af sammenhængen mellem fakturanummer og seneste fakturanummer - egen tilvirkning

Som det fremgår af ovenstående illustration, tegner der sig et mønster, hvor de posteringer, som ikke er omfattet af besvigelser, samler sig i en klynge med de øvrige besvigelser, hvorimod posteringer omfattet af besvigelser stikker ud. Således vurderes alle 6 besvigelser at kunne identificeres som anomalier i denne analyse. På baggrund heraf foretages en videre analyse af, hvordan Data Mining-teknikker kan anvendes til at identificere besvigelserne. Anvendelse af forskellige Data Mining-teknikker er illustreret i figur 25, hvor hver kvadrant viser, hvilke anomalier der identificeres afhængig af anvendelsen af de forskellige teknikker.



figur 25 - Anvendelse af Data Mining-teknikker på faresignaler relateret til reelle ejere - egen tilvirkning

Som det fremgår af ovenstående analyse, er der ingen lineær tendens i datapunkterne, hvorfor valget af Data Mining-teknikker bør omfatte cluster- og klassifikationsteknikker.

I billede 1 i illustrationen er anomalier forsøgt identificeret ved anvendelse af cluster teknikken K-Means, som i dette eksempel vil være i stand til at identificere 5/6 af besvigelserne. Det er givet, at K=2 og algoritmen således er programmeret til at danne to klynger. Som nævnt i teoriafsnittet vil denne algoritme sætte to tilfældige fiktive datapunkter i datasættet og tegne en linje mellem disse to punkter for at danne de antal klynger, som algoritmen er programmeret til. Når linjen er sat, vil algoritmen sætte et nyt vilkårligt punkt i det dataområde, hvor gennemsnitsafstanden fra det fiktive punkt og punkterne i klyngen er kortest. Herefter dannes en ny klynge og et nyt punkt med kortest gennemsnitsafstand til datapunkterne i klyngen og så fremdeles. Resultatet heraf er, at algoritmen identificerer 5/6 af besvigelserne, men det medfører samtidig, at én af besvigelserne ikke identificeres og dermed klynges som en falsk negativ sammen med de øvrige posteringer, der følger en normal udvikling i fakturanummereringen.

I billede 2 er der anvendt Density Based Clustering, hvor følgende parametre er givet; at Epsilon er 1 cm og minimumspunkter i clustrene er 2.

Algoritmen søger at klyngedanne alle kerne- og grænsepunkter. Datapunkter, der ikke kan klyngedannes, identificeres således som anomalier eller outliers. Som følge heraf er algoritmen i stand til at danne en

klynge (Cluster 2) af de normale posteringer, hvor udviklingen i fakturanumre følger normalen. Klyngen dannes, da alle disse har to naboer inden for en radius på 1 cm (kernepunkter), eller har et kernepunkt inden for en radius på 1 cm (grænsepunkter). På baggrund af de regler, som den ikke-superviserede algoritme er fodret med, danner den således også to falske positive i Cluster 1 (rød markering) og én falsk negativ i form af outlier 3. Denne teknik er således i stand til at identificere 4/6 af besvigelserne. Svagheden ved teknikken i det anvendte eksempel er dog, at den vil medføre, at to besvigelser ikke vil blive undersøgt, og at en normal postering vil blive udtaget til test.

I billede 3 i illustration er der anvendt K-Nearest Neighbor, hvor der søges efter de tre nærmeste naboer for at identificere posteringer, der afviger fra normalen. Ved anvendelse heraf fremgår det, at modellen er i stand til at identificere 4/6 af besvigelserne (markeret med en rød cirkel). De to resterende besvigelser, som ikke følger samme mønster som de øvrige besvigelser, bliver her fejlklassificeret som falsk negative (markeret med grønne ringe). Årsagen er, at disse to besvigelser er enkeltstående og således tilnærmelsesvis minder om normalen, og pga. manglende nabopunkter i form af besvigelser bliver de klassificeret som falsk negative. Således er denne teknik ikke i stand til at identificere alle besvigelserne på egen hånd, hvorfor den bør suppleres af yderligere en teknik.

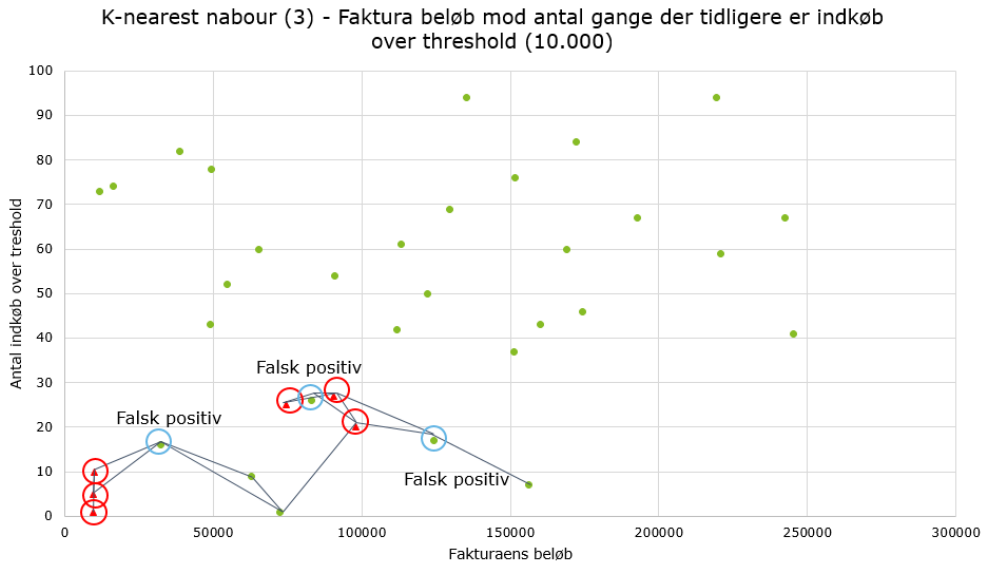
I billede 4 er klassifikationsteknikken KN-Nearest Search anvendt for med et nyt forsøg at identificere de seks besvigelser i datasættet. Her kan algoritmen ved at indsætte nye og fiktive datapunkter, som har de samme karakteristika som de oprindelige punkter, dog med en lille variation, tegne et nyt billede af posteringerne. Algoritmen tegner ud fra de nye punkter den mest ”rigtige” kurve, hvorefter posteringerne klassificeres.

Ved at kombinere alle fire teknikker, som beskrevet ovenfor, vil det således være muligt at krydsvalidere resultaterne og dermed også de identificerede anomalier for at sandsynliggøre, hvor stor risikoen er for, at de identificerede anomalier er besvigelser, som dermed kræver nærmere undersøgelse. I ovenstående eksempel bliver de fire datapunkter øverst til venstre bekræftet i tre ud af fire analyser som anomalier.

3.2.1.2 Omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange

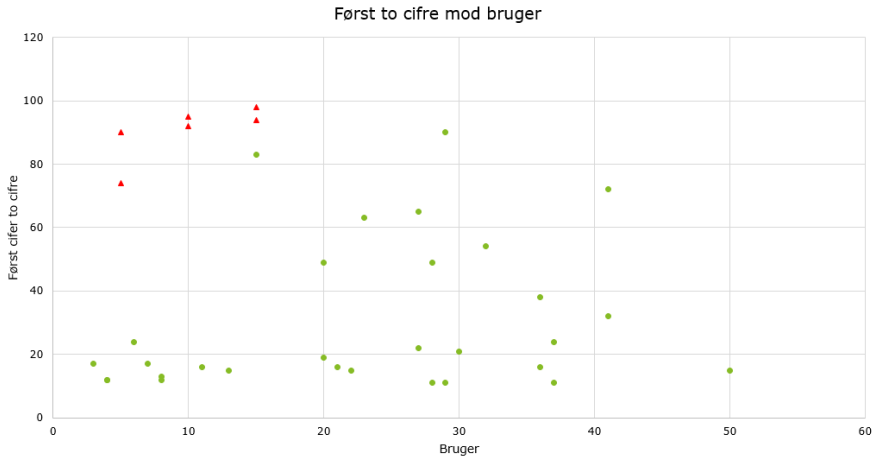
I det givne eksempel antages det, at tilsløringen ved reelle ejere er så sofistikeret, at besvigelserne ikke giver udslag i de første analyser, og således stadig er tilsløret for virksomheden. Således udføres der, parallelt med de øvrige analyser, analyser på de identificerede faresignaler relateret til omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange.

Som nævnt i afsnit 3.1.1 så er der særligt tre faresignaler forbundet med omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange, som kan indikere, at der er sket en besvigelse. Risikovurderingsteamet bør i deres risikovurdering analysere de variabler, som har nær tilknytning til de pågældende faresignaler. Interessante variabler at analysere ud fra de identificerede faresignaler er historiske antal indkøb på leverandøren over treshold sammenholdt med seneste fakturabeløb. Ved visualisering af denne analyse kan det konstateres, at besvigelserne ikke er så udslagsgivne her, og analysen vil umiddelbart ikke sige så meget. Ved anvendelse af Data Mining-teknikker, som fx K-Nearest Neighbor, vil 33% af anomalierne bestå af falske positiver. Analysen er illustreret i figur 26.



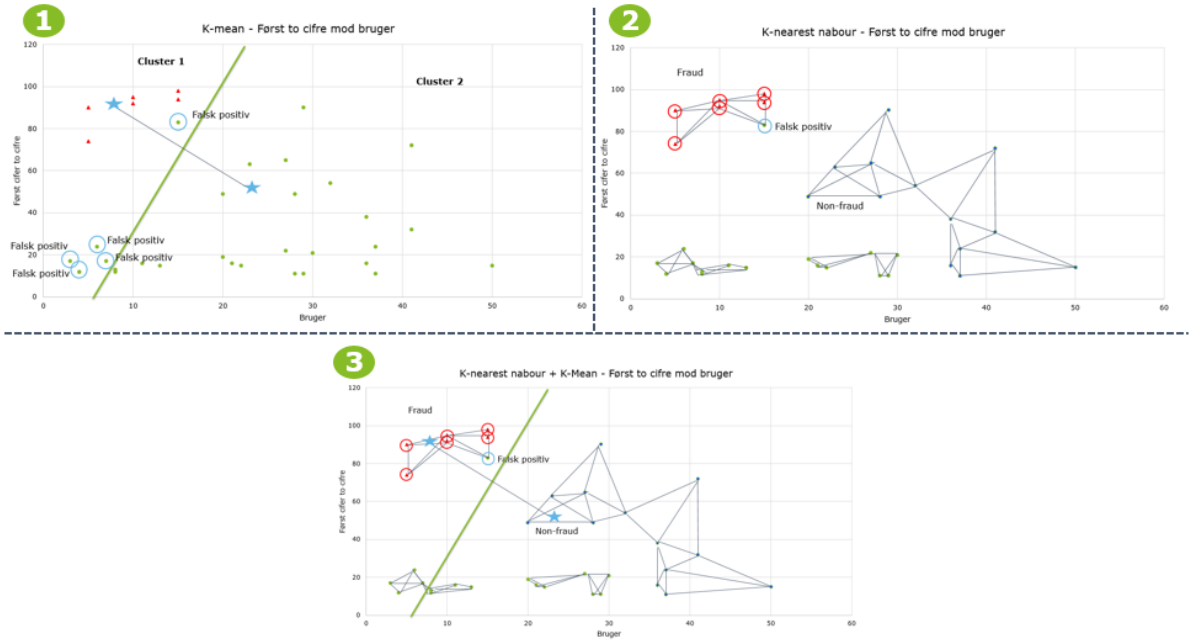
figur 26 - K-Nearest Neighbor-analyse - fakturabeløb med indkøb over treshold

Da analysen ikke gav det ønskede resultat, kan variabler som de to første cifre i fakturabeløbet sammenholdt med godkendelsespersonen være med til at identificere eventuelle besvigelser. Årsagen hertil vil være, at fakturabeløb i Benfords lov bør variere på en bestemt måde. Så er der overvægt af eksempelvis tallet 9, 8 og 7 i de første cifre i fakturabeløbet på bestemte brugere, kan det indikere, at en besviger bevidst køber ind lige under treshold og på den måde søger at tilsløre besvigelsen ved at omgå godkendelseskontrollen. I illustrationen nedenfor er denne analyse eksemplificeret på baggrund af et fiktivt datasæt.



figur 27 - Analyse af to første cifre mod bruger – egen tilvirkning

Af denne analyse fremgår det, at posteringer relateret til besvigelserne adskiller sig væsentligt i forhold til de normale posteringer. Det vurderes, at denne analyse er den mest optimale til at opdage besvigelser, som er tilsløret ved omgåelse af godkendelsesprocedurer. Det vil i forlængelse heraf blive demonstreret, hvordan kombinationen af udvalgte Data Mining-teknikker vil være i stand til at identificere besvigelserne uden menneskelig indblanding.



figur 28 - Data Mining-teknikker anvendt på de to første cifre mod bruger - egen tilvirkning

Som det var tilfældet i eksemplet på reelle ejere, er der ingen lineær tendens i datasættet i ovenstående analyse, hvorfor regressionsanalyser hurtigt bør forkastes. Således bør besvigelsesteamet anvende Data Mining-teknikker såsom cluster- og klassifikationsteknikker.

I billede 1 i figur 28 er der inden for ikke-superviseret læring anvendt en cluster-teknik i form af K-Means til at forsøge at identificere de 6 besvigelser, som forekommer i datasættet. Algoritmen er instrueret i at danne to klynger i det givne datasæt. Teknikken vil i dette eksempel være i stand til at identificere alle 6 besvigelser. Teknikken har dog den svaghed, at den vil identificere 5 falske positiver i analysen. Således vil der ved anvendelse af denne analyse blive identificeret næsten det dobbelte antal anomalier, end hvad der bør have besvigelsesteamets opmærksomhed. Det vil således kræve flere ressourcer end nødvendigt.

Som følge af et ikke-optimalt resultat ved anvendelse af K-Means er der i billede 2 i figur 28 inden for superviseret læring anvendt en klassifikationsteknik i form af K-Nearest Neighbor med tre naboer. Heraf fremgår det, at algoritmen er i stand til at klassificere alle 6 besvigelser i en klasse for sig, dog med en falsk positiv. De øvrige posteringer af normal karakter ligger dog så spredt, at teknikken vil klassificere datasættet i 4 særskilte grupper, hvorfor denne teknik i sig selv ikke vil give besvigelsesteamet et brugbart output til videre undersøgelse.

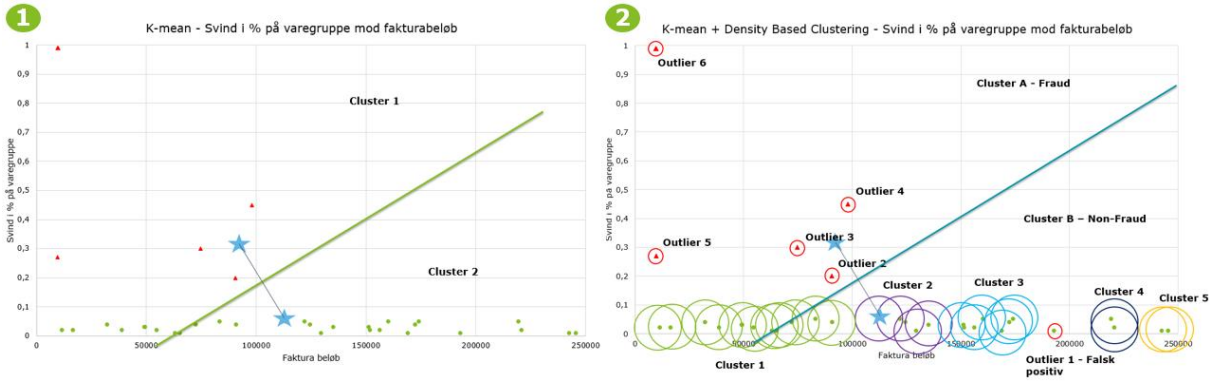
Som følge af de to modellers individuelle svagheder i det givne datasæt vil en kombination af de to resultater kunne bidrage til, at det alene er relevante anomalier, der identificeres. I billede 3 er denne krydsvalidering illustreret. Som det fremgår af illustrationen, vil 3/4 klynger af posteringer tydeligt kunne inddeles automatisk i henholdsvis Cluster 1 og Cluster 2. Den sidste klynge fordeler sig dog stadig i hver sin klynge i K-Means modellen. Ved krydsvalideringen bør modellen ved K-Nearest Neighbor komme frem til, at de fire falske positive nederst til venstre har så nær tilknytning til hinanden, at disse kan afkræftes som anomalier, og at de dermed tilhører Cluster 2 sammen med de øvrige normale posteringer.

Ved anvendelse af ovenstående teknikker vil algoritmerne være i stand til at identificere 6/6 af besvigelser som anomalier og én falsk positiv anomali til videre undersøgelse af besvigelsesteamet. Dette illustrerer således, at teknikkerne ikke er fejlfrie, og forkert anvendelse og forkert kombination heraf vil medføre forkerte resultater og dermed ikke have nogen værdi for besvigelsesteamet.

3.2.1.3 Typen af leverancer og manglende leverancer

Det antages videre i det tænkte eksempel, at besvigelserne er så sofistikeret tilsløret, at de ikke har givet udslag i analyserne relateret til reelle ejere og relateret til omgåelse af godkendelsesprocedurer og sædvanlige forretningsgange. Da det blev identificeret i afsnit 3.1.1, at faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen typisk kan identificeres ved typen af leverance eller mangel på samme, køres der parallelt med de øvrige analyser en analyse på disse faresignaler.

Som benævnt i afsnit 3.1.1 er der særligt to faresignaler forbundet med typen af leverancer og manglende leverancer, som kan indikere, at der er sket en besvigelse. Således bør risikovurderingsteamet i deres risikovurdering analysere de variabler, der har nær tilknytning til de pågældende faresignaler. I dette analyseafsnit tages der udgangspunkt i manglende leverancer, da typen af leverancer vurderes svær analyserbar ud fra numeriske data uden brug af Text Mining-teknikker. Relevante variabler til analyse om manglende leverancer kan være historisk svind på bestemte varegrupper/varenumre og fakturabeløb. Valget af disse variabler begrundes i, at de manglende leverancer søges tilsløret af besviger ved foretagelse af nedskrivninger i form af svind og kassationer på de varer, der ikke er leveret. Således vil svindprocenter på disse varegrupper afvige fra normalen, hvorfor besvigelserne bør fremstå som anomalier i denne analyse. Analysen er illustreret i figur 29 på baggrund af et fiktivt datasæt indeholdende 6 besvigelser.



figur 29 - Data Mining-teknikker anvendt på analyse om svind mod fakturabeløb - egen tilvirkning

I eksemplet har risikovurderingsteamet vurderet, at cluster-teknikker inden for ikke-superviseret læring er de mest effektive teknikker til på egen hånd at identificere besvigelser i et givent datasæt.

I billede 1 i figur 29 er der anvendt K-Means, hvor det er givet, at K = 2, og algoritmen skal således opdele datasættet i to klynger (besvigelser/ikke-besvigelser). Ved anvendelse af denne teknik er algoritmen i stand

til at identificere alle besvigelser selvstændigt. Teknikken har dog den svaghed, at den i datasættet sammen med besvigelserne identificerer 7 falske positiver, hvorfor outputtet af denne model som enkeltstående teknik vil medføre et uhensigtsmæssigt resultat.

I billede 2 er analysen suppleret med en Density Based Clustering-teknik, hvor risikovurderingsteamet har defineret Epsilon som 1 cm og minimumpunktet som 2. Denne teknik er i stand til at identificere 5 klynger, som den således vurderer følger normalen i datasættet. Således er denne teknik i stand til på egen hånd, at identificere 6 anomalier i datasættet, hvoraf en i dette tilfælde vil være en falsk positiv.

Kombinationen af disse to teknikker og krydsvalideringen af resultaterne herfra vil således være i stand til på egen hånd at identificere alle besvigelser, afkræfte den ene falske negative og dermed give et korrekt svar, som medfører, at det alene vil være besvigelserne, der vil blive identificeret som anomalier til videre undersøgelse af besvigelsesundersøgelsesteamet.

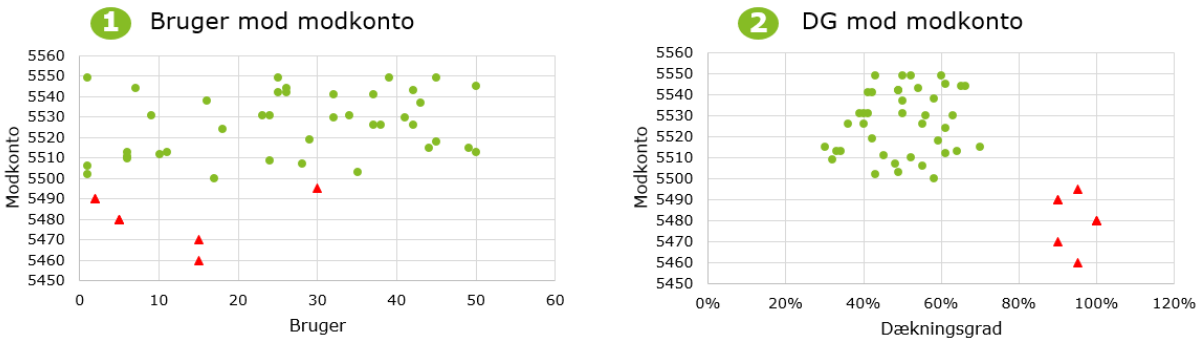
På baggrund af ovenstående analyser kan det udledes, at analyserne i samspil med hinanden, hvis rigtigt anvendt ud fra virksomhedens karakter, er i stand til i realtid at identificere mulige besvigelser i forbindelse med fakturering igennem skyggeselskaber i forbindelse med indkøbsprocessen. Ved krydsvalidering mellem relevante Data Mining-teknikker kan der automatisk sorteres i falske positive og negative, således at det alene er relevante anomalier, der bliver udslagsgivne og sendes videre til besvigelsesundersøgelsesteamet. Således kan virksomhederne ved programmering af algoritmer backend ved en begrundet professionel udvælgelse af analyser være i stand til at konstruere et så fintmasket net, at en potentiel besvigelse vil være svær at tilsløre for en besviger, uden at det bliver opdaget i realtid eller relativt hurtigt. Det er endvidere værd at bemærke, at disse analyser kan have den afledte effekt, at de også afdækker andre besvigelser og eventuelt utilsigtede fejl. Fx vil tilsigtede og utilsigtede dobbeltbetalinger af leverandørfaktura give udslag i analysen i form af ustruktureret udvikling i fakturanumre.

3.2.2 Data Mining-teknikker rettet mod indregning af fiktiv omsætning

Som beskrevet indledningsvist bør første del af processen ved opdagelsen af besvigelser være at identificere relevante variabler, som kan afsløre underliggende mønstre i data og herigennem identificere posteringer af særlig interesse. I afsnit 3.1.2 blev der identificeret tre undergrupper inden for indregning af fiktiv omsætning, baseret på disses målbare karakteristika. Gruppernes karakteristika danner rammerne for bearbejdningen af de identificerede variabler samt valget af Data Mining-teknikker.

3.2.2.1 Tilsløring via balancen

Som beskrevet i afsnit 2.1.3.2 kan besvigelser tilsløres via balancen. Såfremt der anvendes et anlægsaktiv, kan der gå lang tid, før resultatopgørelsen belastes af besvigelsen. Derudover er det nemmere for en besviger at skjule besvigelsen på en balancekonto, hvor fokus er mindre, end det er på en debitorkonto. Overforfaldne debitorer vil ofte blive undersøgt, hvorfor det vil kræve en omfattende tilsløring fra besvigeren for at kunne anvende en debitor som modkonto til sin besvigelse. Med udgangspunkt i denne viden kan en relevant variabel være posteringens modkonto. Der vil være en forventning om, at alle salgsposteringer i højere eller mindre grad følger samme forretningsgang. Derfor vil posteringer, som ikke følger den normale forretningsgang, være af særlig interesse. Derudover er det interessant at finde ud af, hvem der foretager denne type postering. Derfor vil bruger og modkonto være variabler, som kan skabe ny information og samtidig være med til at afsløre potentielle besvigelser. I figur 30 billede 1 er en sammenstilling af disse variabler illustreret.



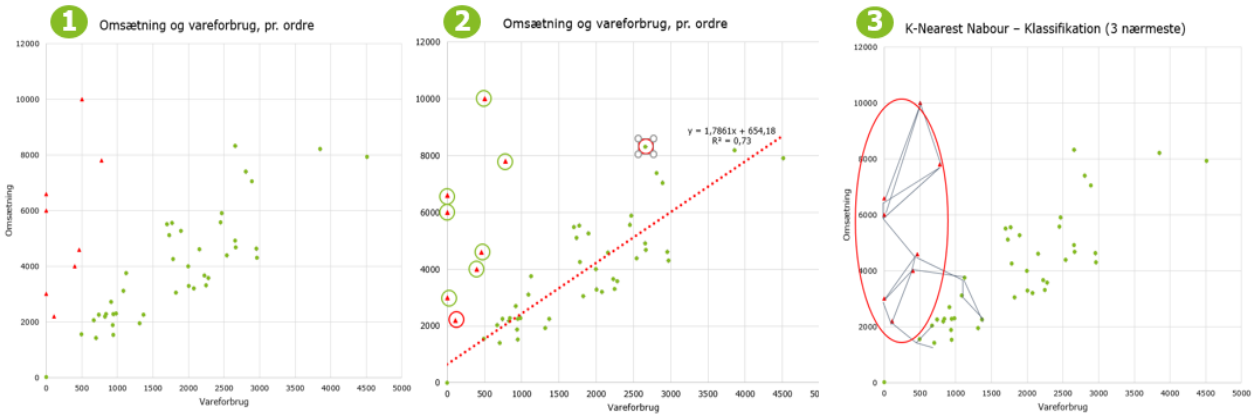
figur 30 - Tilsløring via balancen - egen tilvirkning

I de to billeder er normale transaktioner farvet grønne. Røde transaktioner indikerer besvigelser. I billede 1 ses det, at der er nogle relativt tydelige forskelle mellem de normale posteringer og de besvigelsesrelaterede posteringer. Kontiene, som starter med 55, indikerer debitorkonti. De øvrige indikerer anlægsaktiver. I billede 1 kan der dog fortsat herske tvivl om anvendelse af Data Mining-teknikker, da punkterne ikke har de ønskede karakteristika ved brug af de to variabler. Hvis informationen om modkontoen beholdes, men brugeren udskiftes med dækningsgrad (DG), opnås der et meget tydeligt resultat. De besvigelsesrelaterede posteringer fremstår nu tydeligt som anomalier i det samlede datasæt. Her vil en Data Mining-teknik som K-means nemt finde anvendelse, da der er grundlag for et mønster, og den type data ikke vil kræve, at output defineres. Outputtet vil herefter danne grundlag for yderligere undersøgelse. Da denne type besvigelse vil kunne elimineres ved søgning efter poster, som ikke følger det normale posteringsflow, vil vi ikke uddybe anvendelsen af Data Mining-teknikker på dette område yderligere.

3.2.2.2 Posteringer, som ikke følger normale forretningsgange

Posteringer, som ikke følger de normale forretningsgange, indeholder forskellige former for besvigelser, herunder ledelsens tilsidesættelse af kontroller, indtægtsføring uden vareforbrug mm. Ved vurderingen af karakteristika for denne type besvigelse, vurderer vi, at dette ofte vil komme til udtryk ved en unormalt høj dækningsgrad. En unormalt høj dækningsgrad vil, ved brug af Data Mining-teknikker, være baseret på de øvrige posteringer i populationen.

Vi har med udgangspunkt i et fiktivt udarbejdet datasæt genereret omsætnings- og vareforbrugspunkter på ordreniveau. I billede 1 i nedenstående illustration er normale posteringer, som ikke er omfattet af besvigelser, vist med grøn. Hertil har vi indsat besvigelseslignende posteringer, som har en højere dækning end gennemsnittet. Disse er illustreret med røde trekanter. I billede 2 fremgår en analyse baseret på Predictive Data Mining ved brug af regression. I billede 3 fremgår en illustration af anvendelse af Predictive Data Mining, herunder anvendelse af klassifikationsalgoritmen K-Nearest Neighbor.



figur 31 - Egen tilvirkning, anvendelse af Data Mining på omsætning og vareforbrug

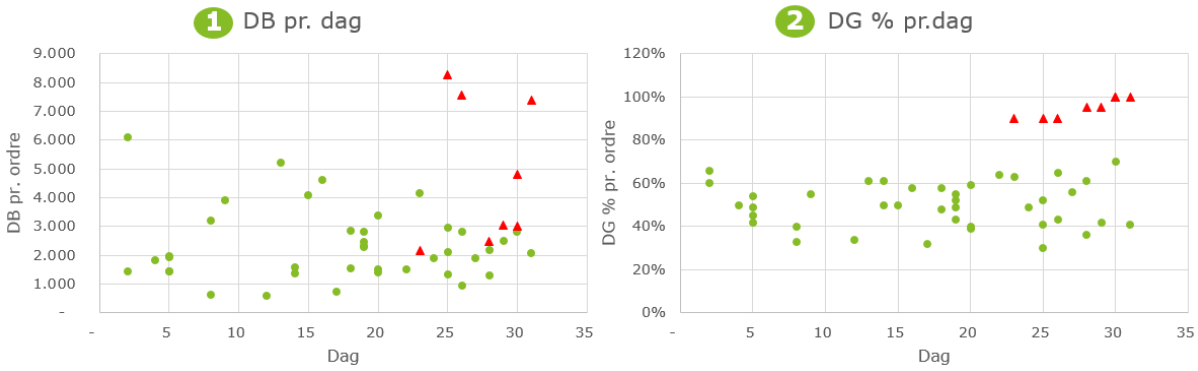
Som det fremgår af billede 1 i ovenstående figur, er der en lineær tendens i datapunkterne. For hver ekstra enhed, som sælges, sker der en lineær udvikling i både omsætning og vareforbrug. Imidlertid kan det visuelt aflæses, at der er enkelte posteringer, som ikke følger denne trend. Når der er tale om lineær tendens, finder regressionsanalyse sin anvendelse. Regressionen kan fortælle, hvilke datapunkter der afviger fra normalen ud fra de enkelte punkters afvigelse til regressionslinjen og den gennemsnitlige afvigelse for alle punkterne til regressionslinjen. I billede 2 er punkter, som afviger fra normalen, illustreret med grønne cirkler. Dertil er der to røde cirkler, hvor punktet mod venstre viser et punkt, som ikke klassificeres som en postering af interesse. Punktet til højre viser en normal omsætningspostering, som fejlklassificeres som værende interessant. Regressionsanalysen viser her, at mange af de

besvigelsesrelaterede posteringer vil blive klassificeret korrekt. Dog har modellen i eksemplet sine svagheder, da der både er en falsk positiv og en falsk negativ. Dette vil i praksis betyde, at kontrolaktiviteten vil overse en besvigelse og samtidig anvende ressourcer på undersøgelse af en normal postering.

For at optimere resultaterne kan andre teknikker anvendes. I billede 3 er anvendelsen af K-Nearest Neighbor illustreret. Her er modellen illustreret ved anvendelse af 3 nærmeste naboer til at klassificere datapunkterne. I modellen er det kun illustreret for punkter, hvor der er en grænseflade mellem besvigelseslignende posteringer og de normale omsætningsposteringer. Modellen medtager her alle datapunkterne, leder efter de 3 nærmeste naboer og lader dem bestemme, om der er tale om en besvigelse eller ej. Algoritmen opnår her et rigtigt resultat, da alle besvigelseslignende posteringer bliver klassificeret korrekt ud fra deres karakteristika. I dette tilfælde viser klassifikationsalgoritmen sig som det bedste valg af de to modeller.

Da modellerne indbyrdes kan lave fejl, kan en kombination af de forskellige teknikker skabe et mere validt resultat. Såfremt resultaterne fra begge modeller opsummeres, vil der være et sammenfald på syv posteringer. Derudover vil der være to posteringer, som afviger i teknikkerne. Da teknikkerne er baseret på statistik, er der derfor en risiko for, at der opstår falske positiver eller falske negativer, og derfor vil en kombination være at foretrække. Til brug for undersøgelsen af posteringerne vil de ni posteringer og informationen om den manglende sammenhæng på to posteringer være en nøgleinformation. Dette mindsker risikoen for, at potentielle besvigelser ikke fanges. Når der arbejdes med opdagelse af besvigelser, bør en højere præcision prioriteres, således at alle besvigelser fanges, selvom dette betyder, at der også sker en undersøgelse af posteringer, som ikke er besvigelsesrelaterede.

For at øge præcisionen af de foretagne analyser kan der foretages analyse på andre variabler, som karakteriserer besvigelser. Fx vil der antageligt blive foretaget mere regnskabsmanipulation i slutningen af en måned end i starten. Hvis denne viden kombineres med, at dækningsbidrag for en sådan transaktion ofte vil være høj, fås nedenstående illustration.



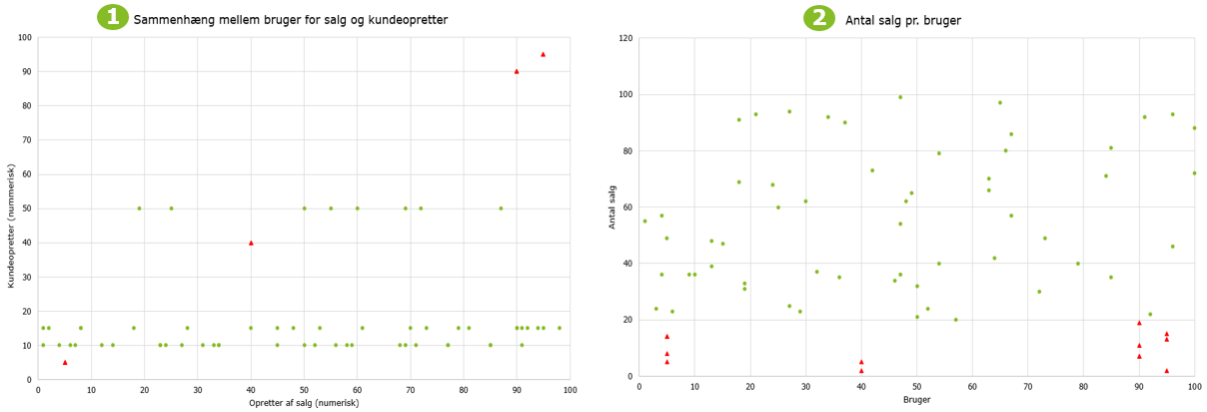
figur 32 - DB og DG pr. dag - egen tilvirkning

I billede 1 vises dækningsbidrag pr. dag pr. ordre. Her illustreres det, at enkelte posteringer adskiller sig, og at nogle posteringer vil være svært identificerbare blandt normale salg. Såfremt dækningsbidrag skiftes ud med dækningsgrad, bliver de besvigelsesrelaterede posteringer nu meget tydelige. Billede 2 kan herefter danne grundlag for yderligere undersøgelse ved brug af de forskellige Data Mining-teknikker. Datapunkterne vil kunne blive analyseret ved brug af forskellige clustering- eller klassifikationsteknikker, hvor K-nearest Search vil være den mest oplagte.

Anvendes disse data i kombination med resultaterne fra den foregående analyse, kan undersøgelsen nu reduceres fra de ni posteringer af interesse til syv posteringer af interesse. Ved brug af krydsvalidering på tværs af forskellige teknikker og variabler inden for samme datasæt, kan sikkerheden af resultaterne øges, og sandsynligheden for falske positive eller negative kan reduceres. Det betyder, at algoritmerne i højere grad medfører effektivitet, da undersøgelserne nu kun vil omfatte relevante posteringer.

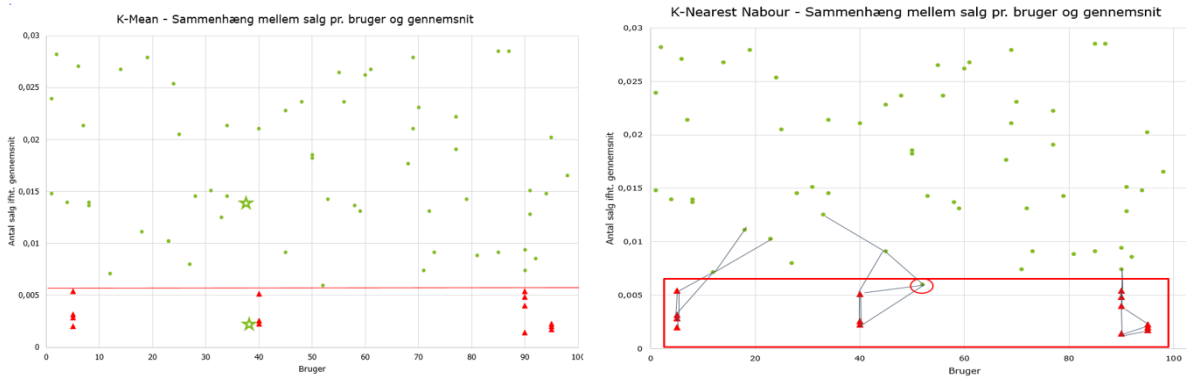
3.2.2.3 Ændrede mønstre i samhandel

En anden type af karakteristika på besvigelser relateret til indregning af fiktiv omsætning er ændrede mønstre i samhandlen. Det kan være i form af falske kunder, ændrede samhandelsmønstre for reelle kunder, handel med nærtstående m.fl. Ved denne type besvigelse kan det være svært at finde oplagte variabler at analysere. Når der arbejdes med Data Mining-teknikker, vil det første skridt kræve, at der arbejdes med de forskellige variabler, som kan have interesse. I nedenstående illustration har vi vist, hvordan der kan arbejdes med de samme faresignaler på forskellige måder.



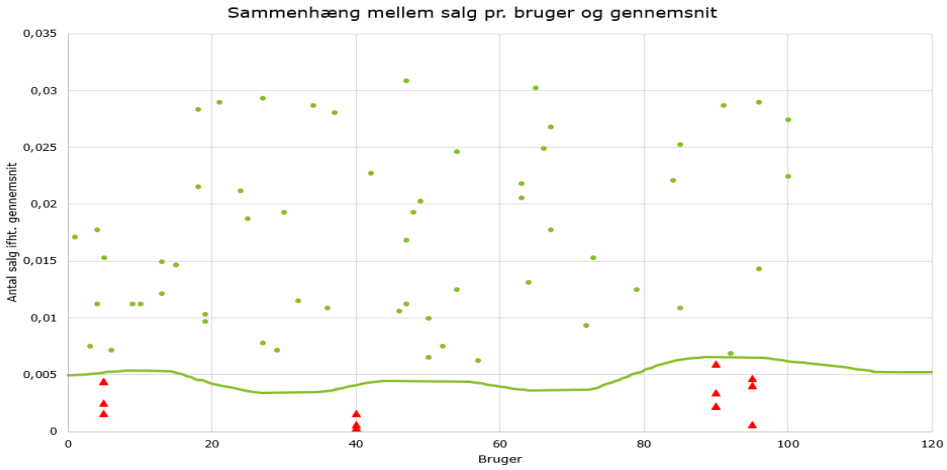
figur 33 - Faresignaler illustreret på forskellige måder, egen tilvirkning

I billede 1 fremgår korrelationen mellem den person, som initierer en ordre, og den person, som har oprettet kunden, hvorefter ordren gennemføres. Forventningen er her, at der ikke bør være sammenfald mellem disse personer. Dog er det ikke ensbetydende med, at der er tale om en besvigelse. Derfor kan der være et behov for at se faresignalerne på en anden måde. Fx kan det være relevant at kombinere denne information med, hvor mange salg hver bruger har foretaget. Dette fremgår af billede 2. I nedenstående er brugen af clustering-teknikken K-means og klassifikationsteknikken K-Nearest Neighbor visualiseret på baggrund af outputtet fra figur 33.



figur 34 - K-Means og K-Nearest neighbor sammenhæng mellem salg pr. bruger og gennemsnit

Som ovenstående viser, er K-means i stand til at kategorisere data korrekt. Derimod sker der en fejlklassificering af et datapunkt ved brug af K-Nearest Neighbor. For at optimere resultaterne kan K-Nearest Search anvendes som et supplement. Dette fremgår i figur 35.



figur 35 - K-Nearest Neighbor search sammenhæng mellem salg pr. bruger og gennemsnit

De tre forskellige Data Mining-teknikker er altså i stand til at klassificere eller gruppere transaktioner ud fra transaktioner, der følger normalen, og dermed identificere mulige besvigelser som anomalier. Ved kombinationen af de forskellige metoder fås der en validering af anomalierne, ligesom den falske positive postering fra K-Nearest Neighbor kan sorteres fra. Algoritmen danner hermed et præcist grundlag for den videre undersøgelse.

Arbejdet med forskellige faresignaler og afledte variabler samt forskellige teknikker viser derfor, at det er muligt at se mønstre og tendenser i data, som ellers ville være meget svære at identificere ved brug af klassiske analyser eller stikprøver. De mere avancerede teknikker kan derfor være med til at identificere eller begrænse besvigelser i virksomhederne. Endvidere viser gennemgangen, at en besviger vil skulle gå til ekstremer for at foretage en tilsløring, som kombinationen af variabler eller teknikker ikke vil være i stand til at afsløre. Dette understreger, at Data Mining-teknikker som værktøjer kan håndtere den omskiftelighed, der er forbundet med besvigelser. Endvidere kan det udledes, at værktøjet er agilt, da forskellige teknikker hurtigt kan anvendes på datasæt, uagtet at der er tale om strukturerede eller ustrukturerede data. Dette betyder i praksis, at algoritmerne kan køre i realtid, da algoritmerne implementeres direkte i virksomhedernes backend it-infrastruktur.

3.3 Hvordan bør det implementeres som interne kontroller

I teoriafsnittet omkring interne kontroller påpegede vi, at anvendelse af dataanalyse eller kontinuerlig overvågning var anbefalelsesværdigt som en del af de interne kontroller. Herefter redegjorde vi for anvendelse af dataanalyse i det interne kontrolmiljø, og for, hvordan dette metodisk kunne indarbejdes både som en del af risikovurderingsprocessen, men også som reel kontrolaktivitet.

Data Mining-anvendelsesområdet spænder bredt og kan, som tidligere beskrevet, variere fra meget komplekse algoritmer i form af regressioner og klassifikationsalgoritmer til mere simple associations- eller cluster-algoritmer. Da Data Minings anvendelsesområder kan variere fra opdeling af data til klassificering af data, betyder det, at teknikkerne vil kunne anvendes i flere dele af det interne kontrolmiljø. Da Data Mining-processen foregår i virksomhedernes backend og kræver forståelse af datapunkter, matematik og programmering, ses der et behov for specialistkompetencer.

SAS Institute, som er en anerkendt udbyder af viden inden for programmering, Machine learning mv., har i bogen" A guide to Data Science for Fraud Detection" fastslået, at en god besvigelses-datavidenskabsmand bør besidde kompetencer inden for programmering, modellering, forretning, kreativitet, visualisering og kommunikation (Baesens, Vlasselaer and Verbeke, 2015, p. 33). Der er i deres definition tale om en meget bred profil med krævende kompetencer og umiddelbart en kombination af kvalifikationer, der mildt sagt ikke er nemme for virksomhederne at finde. Som tidligere nævnt foreslår COSO, at der etableres et besvigelsesrisiko-vurderingsteam bestående af mange forskellige kompetencer, herunder forretningsforståelse. I lys heraf anfægtes det realistiske i, at en person skulle kunne varetage alle de kompetencer, som omtalte bog opremser. I stedet fremhæver vi følgende kompetencer som essentielle: programmering, modellering, kommunikation og samarbejde. Med disse kompetencer vil virksomhederne kunne kombinere deres etablerede forretningsforståelse med de nødvendige specialistkompetencer. Virksomhedernes avertering efter en mindre krævende profil vil formentlig bevirke ansøgninger fra et bredere kandidatfelt.

Når virksomhederne har de fornødne ressourcer i form af kompetente medarbejdere, kan arbejdet med anvendelse af Data Mining som en integreret del af kontrolmiljøet påbegyndes. I afsnit 2.2.4 fremgik COSO's forslag til, hvordan dataanalyse kunne anvendes som metodisk værkstøj både som kontrolaktivitet og til risikovurdering. Da processen for anvendelse af dataanalyse er identisk for kontrol og risikovurdering, vil nedenstående omfatte begge områder.

Det første centrale punkt omhandler at etablere et besvigelsesrisikovurderingsteam bestående af de fornødne kompetente medarbejdere med forretningsmæssig forståelse. Teamet skal derfor indeholde den føromtalte datavidenskabsmand med speciale i besvigelser samt personer fra it, operations, finansafdelingen mv. Herefter foretages optegning af kendte besvigelsessager og deres karakteristika. Kontrolsvagheder og risikofyldte aktiviteter identificeres. Dernæst bør teamet anvende offentligt kendte faresignaler. Som led i risikovurderingsprocessen identificeres de potentielle risici, hvor teamet har behov for yderligere kendskab for at vurdere risicienes tilstedeværelse. Eksempler kunne være forbigåelse af godkendelseskontroller illustreret pr. godkender, manglende korrelation mellem salg og vareforbrug for udvalgte ordrer eller samhandel med selskaber, hvor dets ejerforhold er ukendte.

I andet punkt finder Data Mining sin anvendelse. Som beskrevet i foregående afsnit om design og anvendelse af de forskellige Data Mining-teknikker til imødegåelse af udvalgte risici, er de ikke-superviserede eller deskriptive teknikker effektive til at skabe overblik over sammenhænge, som er svære at identificere for mennesker. Til dette arbejde kan de prædiktive teknikker hjælpe teamet med at klassificere data og derigennem få af- eller bekræftet risici, før der bliver anvendt ressourcer på videre undersøgelse af disse. Data Mining kan ved korrekt anvendelse skabe en stærkere risikovurdering og derved undgå, at der bliver spildt ressourcer på områder, hvor der kun foreligger en hypotetisk risiko. Særligt kan det være ressourcekrævende at foretage store stikprøveundersøgelser af fx omsætningsposteringer, hvor dækningen afviger fra normalen, hvis disse ikke er opdelt i underpopulationer. Der kan fx foreligge serviceaftaler eller andre ydelser på solgte produkter, som ikke altid medfører et direkte vareforbrug. Disse poster kan ved brug af Data Mining grupperes samlet og sorteres fra, før den videre analyse foretages, hvilket vil lette teamets vurdering af risici på området.

Når risikovurderingsteamet har udfærdiget sin risikovurdering og tilknyttet eksisterende opdagende og forebyggende kontroller, kan teamet påbegynde arbejdet med anvendelse af Data Mining som kontrol-aktivitet. Data Mining kan anvendes som både forebyggende og opdagende kontrol.

Forebyggende kontroller er defineret som kontroller, hvis formål var at begrænse besvigelser, før de opstår. Disse kontroller er ofte kendt af medarbejderne i virksomhederne og anvendes på områder, hvor mulighederne for besvigelser ofte ville være store, hvis kontrollerne ikke var til stede. Eksempler herpå kunne være kontrol af bankudbetalinger, omkostningsgodkendelse mm. Da forebyggende kontroller kan omgås, når en besviger har tilpas forståelse af kontrollen og er kvalificeret til at omgå den, kan Data Mining være et supplement, som gør, at kontrolaktiviteten bliver meget vanskelig at forstå og dermed også

sværere at omgå for en besviger. Da Data Mining ikke er baseret på regler, men trends, karakteristika og tendenser i data, vil det kræve, at besvigeren har fuld indsigt i opsætningen af algoritmen og de underliggende data, for at denne kontrol vil kunne omgås. Et eksempel på denne type kontrol er fx, at der ikke kan ske udbetaling til en leverandør, som er flaget af en Data Mining-algoritme, fordi leverandøren har sendt sin første faktura til virksomheden inden for 90 dage fra dets registrering hos en offentlig myndighed. Besvigelsesteamet skal herefter "frigive" fakturaen, før den kan udbetales. Kontrollen kan indarbejdes som et ekstra led i virksomhedens godkendelseshierarki. Når den normale godkendelse er gennemført, bearbejder algoritmen data, og såfremt denne flages, sker der en automatisk flytning af bilaget til en ny godkendelsesgruppe. Denne gruppe varetages fx af besvigelsesteamet. En anden forebyggende kontrol, som kan implementeres, er anvendelse af en regressionsalgoritme på omsætningsposteringer. Før en omsætningspostering kan effektueres, bearbejder regressionsalgoritmen automatisk posteringen og sammenholder den med periodens øvrige omsætningsposteringer ved brug af posteringernes regressionslinje. Hvis den undersøgte posteringsafvigelse fra regressionen overstiger gennemsnittet af afvigelser, flages den som en potentiel besvigelse. Posteringen skal herefter godkendes af besvigelsesteamet, før den kan effektueres.

Opdagende kontroller blev defineret som kontroller, der havde til formål at opdage en besvigelse, efter den var begået. Data Mining vil i dette tilfælde blive anvendt som analyse på posteringer, som allerede er eksekveret i bogføringen og eventuelt også er betalt i banken. Analyserne vil kunne indarbejdes ligesom kontrollerne beskrevet under forebyggende kontroller. Forskellen vil udelukkende bestå i, at posteringerne ikke vil blive sendt til godkendelse hos besvigelsesteamet, men i stedet sendes til undersøgelse hos teamet.

Som ovenstående viser, har virksomhederne mulighed for at indarbejde Data Mining i både deres forebyggende og opdagende kontroller. Denne pointe realiseres ved, at der ved brug af Data Mining arbejdes i virksomhedernes backend it-infrastruktur, hvorfor det er muligt at indarbejde teknikkerne som en integreret del af virksomhedens ERP-system og dermed direkte i det indbyggede kontrolmiljø. Det betyder, at der opstår en mulighed for at flytte flere kontrolaktiviteter fra at være opdagende til at være forebyggende. Dette underbygges både ved, at besviger har sværere ved at omgå kontrollerne, da disse ikke er regelbaserede, og at kontrollerne opererer kontinuerligt og effektivt. Brugeren vil ikke opleve kontrollerne, hvorfor performance på økonomisystemet ikke vil påvirkes negativt. Ved flytningen fra opdagende til forebyggende kontroller kan virksomhederne begrænse det såkaldte "pay and chase", hvorved risikoen for tab vil blive yderligere reduceret. Data Mining som teknologi indeholder således fordele relateret til både begrænsning og opdagelse af besvigelser, men også i relation til de tab, som

efterfølgende er forbundet med indkrævningsprocessen. Data Mining ses derfor som en oplagt del af det interne kontrol miljø og anbefales at blive implementeret som forebyggende kontrol i så vid udstrækning som muligt.

3.4 Fordele og ulemper ved Data Mining

Data Mining fremstår som en potent teknologi i bekæmpelsen af besvigelser, men det fremgår også tydeligt, at der er tale om et område, som ikke har stor udbredelse i virksomhedernes interne kontrolmiljøer. I afsnit 2.2.4 fremgik en konklusion fra ACFE's "Report to the Nation", som fremhævede, hvor stor en effekt dataanalyse havde på opdagelsen af besvigelser, men at den sjældent blev anvendt af virksomhederne. Når dataanalyse sjældent bliver anvendt, kan konklusionen i høj grad videreføres til anvendelse af Data Mining, da denne teknik er mere avanceret end de klassiske værktøjer, hvorfor Data Mining-teknikkers udbredelse må antages at være mere begrænsede

Der er mange fordele relateret til opdagelse og forebyggelse af besvigelser, herunder begrænsningen af tab relateret til midler fragået virksomheden. Endvidere ses flere muligheder, da teknologien har mange anvendelsesområder og kan håndtere kontinuerlige skift i kompleksitet og karakteristika ved besvigelser. De store ulemper skal findes i, at business casen kan være svær at retfærdiggøre, fordi de reelle tab til besvigelser sjældent er kendte af virksomhederne, og at der ofte hersker en udpræget omkostningsminimeringsmentalitet samt en "det sker ikke for os"-mentalitet. Endelig kan kompetencekløften og risikoen for fejlagtig implementering være med til at skabe usikkerhed om beslutningen for anvendelse af teknologien.

Når der skal rationaliseres, kommer det i sidste ende an på en vurdering af omkostninger versus fordele ved investeringen. Såfremt ACFE's hypotese står til troende, bør der allokeres store midler og være et stort incitament til at foretage investeringen. Kombineres hypotesen med besvigelsesteorien og udviklingen af kompleksiteten i besvigelser, vil virksomhederne i den nærmeste fremtid stå i en situation, hvor deres klassiske kontroller ikke vil være tilstrækkelige til at forebygge eller opdage besvigelser. Derfor vil det være et spørgsmål om, hvornår, og ikke hvis, de bliver ramt af besvigelser. Fordelene ved at få kompetencer inden for programmering og modellering ind i kontrolmiljøet må ikke undervurderes. Når rejsen med Data Mining først er begyndt, vil der opstå store muligheder inden for automatisering af de klassiske kontroller, hvilket vil underbygge hypotesen yderligere.

Den store udfordring ved Data Mining er instrumentets begrænsede anvendelse i virksomhederne, særligt inden for intern kontrol. Når anvendelsen er begrænset, betyder det også, at kompetencerne dertil vil være svære at finde. At virksomhederne formentligt vil have meget svært ved at finde de rigtige kompetencer, vil være en stor ulempe for virksomhederne og vil dermed tale imod anvendelse af Data Mining som en del af virksomhedernes inhouse kontrolmiljø. Virksomhederne ville samtidige skulle foretage store investeringer i software og procesarbejde. Disse udfordringer er af en sådan karakter, at det ikke kan forventes, at virksomhederne vil lykkedes med en implementering af teknologien inden for den nærmeste fremtid. Dette underbygges af, at der primært arbejdes med teknologien inden for den finansielle sektor samt inden for sundhedssektoren (Baesens, Vlasselaer and Verbeke, 2015).

Når der imidlertid er så store fordele ved teknologien, og besvigelser er så stort et problem for virksomhederne, bør det vurderes, om kontrolarbejdet med besvigelser kan outsources som en såkaldt managed service. En managed service er, når en virksomhed outsourcer dele af dens processer eller områder af virksomheden til tredje part, fx til en softwareudbyder eller et konsulenthus (Forbes, 2018). Managed services har den fordel, at der kan ske en skalering af ydelserne, samt at der opnås stordriftsfordele. Endvidere giver et større volumen grundlag for udvikling af medarbejdere, innovation, investering i ny software, procesoptimering mv. Det vil nedbryde kompetencebarrierer og til dels også omkostningsudfordringen ved, at virksomhederne skulle foretage store initialinvesteringer. Alle ulemperne, som blev identificeret i starten af afsnittet, vil kunne håndteres igennem stordriftsfordele og samling af ressourcer på tværs af mange virksomheder. Ved managed service-aftaler opstår der dog nye ulemper. Den mest centrale ulempe vil være relateret til sikring af, at der sker en tilstrækkelig tilpasning af analyserne til den pågældende virksomhed, således at der fortsat analyseres på de rigtige risici. Håndteringen af en sådan ulempe kræver, at virksomhederne i deres besvigelsesrisikovurderingsteam inviterer tredjeparten ind. Managed service-udbyderen vil derfor skulle være en integreret del af virksomhedens interne kontrolmiljø og vil uundgåeligt blive en strategisk samarbejdspartner for virksomheden. Tredjepartens primære formål i en sådan serviceaftale vil derfor være at stille processer, kompetencer og teknologi til rådighed, hvor virksomheden selv ville skulle forestå arbejdet med risikovurdering.
I "Bilag 15 – Fordele og ulemper ved implementering af Data Mining" har vi opsummeret yderligere og ulemper, som er identificeret ved gennemgangen af analysen og teorien bag Data Mining.

3.5 Delkonklusion på analyseafsnittet

I delkonklusionen vil vi nu konkludere på undersøgelsesspørgsmålene inden for afhandlingens analyse. Disse konklusioner danner, sammen med teoriafsnittet, rammerne for konklusionen på problemfeltet.

- ***Hvordan kan Data Mining anvendes i virksomhedernes interne kontrolmiljø til imødegåelse af besvigelser?***

I analysen blev de udvalgte besvigelsesrisici relateret til indkøbssvindel og fiktiv omsætning, undersøgt for hvilke karakteristika som indikerede besvigelser, og hvordan disse karakteristika, kunne fremfindes i virksomhedernes datakilder. Karakteristikaene dannede grundlag for forskellige variabler, som i samspil med hinanden kunne give ny information om en transaktion. Den nye viden dannede herefter rammerne for, hvordan Data Mining-teknikkerne, kan anvendes til at imødegå de identificerede besvigelsesrisici. Ved gennemgangen af de forskellige risici viste vi i analysen, at Data Mining-teknikkerne kan frembringe ny viden ved at dele de undersøgte data op i mindre populationer. Opdelingen blev foretaget ved clustering- og klassifikationsteknikker. På enkelte områder kan regressionsanalyse anvendes til at klassificere data, dog kræver disse en lineær sammenhæng mellem de undersøgte variabler. Den lineære sammenhæng var udelukkende relateret til fiktiv omsætning. Analysen viste, at der var stor forskel på resultaterne afhængigt af de undersøgte variabler, og de foretagne Data Mining-analyser. I resultaterne fremgik det, at der opstod en risiko for både falske negativer og falske positiver, men at disse kan begrænses ved krydsvalidering mellem teknikkerne og ved anvendelse af flere forskellige variabler mod den samme risiko. Ved krydsvalideringen blev resultaterne mere pålidelige, og der blev samtidig opnået en reduktion i, hvor mange posteringer, virksomhederne skulle foretage yderligere undersøgelse af. På baggrund af analysen konkluderer vi, at det er muligt at identificere besvigelser ved brug af Data Mining.

Herefter analyserede vi på, hvordan Data Mining kunne implementeres som en del af virksomhedernes interne kontrolmiljø. Vi fandt, at teknikkerne kan anvendes som en del af det metodiske framework fra COSO omhandlende implementering af dataanalyse i de interne kontroller, som gennemgået i teoriafsnittet. Analyserne kan grundet deres karakter blive implementeret som forebyggende kontroller. Dette skyldes, at Data Mining virker realtid og derfor kan indgå som grundlag for virksomhedernes godkendelseskontroller.

Nedenfor vil vi konkludere på afhandlingens sidste undersøgelsesspørgsmål:

- ***Hvilke udfordringer/begrænsninger er forbundet med anvendelse af Data Mining?***

På baggrund af vores analyse kan det konkluderes, at det vil være fordelagtigt, hvis Data Mining implementeres som en forebyggende kontrol i virksomhedens interne kontrolsystem. Fordelen herved vil være, at posteringer af besvigelseslignende karakter skal godkendes af et dedikeret besvigelsesrisikovurderingsteam, før at disse gennemføres i bogføringen eller banken, og dermed også før besvigelsen kan medføre et tab for virksomheden. Vi fandt i vores analyse, at der er store omkostninger forbundet med at inddrive penge, som er fragået virksomheden, hvorfor Data Mining som forebyggende kontrol vurderes som en teknologi, der er værd at investere i. Vi fandt dog i vores analyse, at der er store udfordringer med at finde ressourcer med rette kompetencer inden for besvigelsesrelateret datavidenskab, samt at omkostninger forbundet med software og processer er meget høje. Selvom der er store tab forbundet med besvigelser, kan det være svært at retfærdiggøre en business case for investering i Data Mining, når besvigelsestab for den enkelte virksomhed kan være ukendte, og opstartsinvesteringen i processer og software er stor.

For at imødegå udfordringen omkring omkostninger og kompetencer foretog vi en vurdering af, hvordan dette kunne håndteres, hvis det ikke skulle varetages af virksomhederne selv. Med udgangspunkt i dette blev der foretaget en vurdering af, om besvigelsesopdagelse og anvendelse af Data Mining kunne varetages af en softwareleverandør eller af et konsulenthus som en managed service. Ved en managed service fandt vi, at der vil kunne opnås stordriftsfordele, og dermed kan omkostningerne til kontrolaktiviteterne sænkes. Endvidere vil stordriftsfordelene og de ekstra ressourcer gøre det nemmere at tiltrække de fornødne kompetencer. Afslutningsvis fandt vi, at der kunne være en risiko for, at en standardisering på tværs af virksomheder kunne medføre, at analyserne ikke blev tilstrækkeligt tilpasset den enkelte virksomhed, og dermed ikke nødvendigvis vil være optimal.

Kapitel 4 – Konklusion

4 Konklusion

I dette afsnit vil vi indledningsvis præsentere de konklusioner, der er draget løbende i afhandlingen for til slut at præsentere en konklusion på afhandlingens problemformulering.

I indledningen og i afsnittet om besvigelsesteori konkluderede vi, at der var enorme tab forbundet med besvigelser. Besvigelser kan begås af alle virksomheders ansatte, ligesom alle virksomheder i teorien er udsat for en besvigelsesrisiko. Konklusionen har sit udgangspunkt i besvigelsesteoriens påstand om, at de fleste mennesker vil kunne begå en besvigelse, hvis de har en mulighed, et pres, og handlingen kan retfærdiggøres.

Vi fandt ved gennemgangen af besvigelsesteorien, at besvigelser er meget varierende, og at de bliver mere og mere komplekse. Disse konklusioner bekræfter behovet for, at der etableres stærke interne kontroller til at forebygge og opdage besvigelser, før de indtræffer, og hvis de allerede er indtruffet, at de efterfølgende bliver identificeret hurtigere.

Traditionelt har mange virksomheder arbejdet med interne kontroller. Arbejdet har for mange været foretaget med udgangspunkt i COSO-framework om intern kontrol. Traditionelle interne kontroller består af forebyggende og opdagende kontroller som fx godkendelse, afstemning, stikprøver mv. Da besvigelsesrelaterede tab for virksomhederne fortsat er enorme, publicerede COSO et framework rettet mod håndtering af besvigelsesrisikoen i virksomhedernes kontrolmiljø. Et helt centralt område for arbejdet med besvigelser er etableringen af et besvigelsesrisikovurderingsteam. Teamet bør repræsentere kompetencer fra alle dele af virksomheden. Endvidere fremhæver frameworket dataanalyse som et nøgleelement. ACFE fremhæver i deres seneste rapport, ”Report To The Nations”, at besvigelsesrelaterede tab, identificeret ved hjælp af dataanalyse, medførte 52% mindre tab, og at disse blev opdaget 58% hurtigere end besvigelser, der blev afsløret på andre måder. Således kan det konkluderes, at dataanalyse kan være et effektivt supplement til de klassiske interne kontroller og supportere disse, da størrelsen af tabene ved interne klassiske kontroller bekræfter, at de ikke har været tilstrækkelige, og at besvigelsesrelaterede tab kan reduceres ved anvendelse af dataanalyse.

Dataanalyse eksisterer i mange variationer, hvorfor vi foretog en teoretisk gennemgang af udviklingen. Her redegjorde vi for det intensive arbejde med Business Intelligence, hvilket har medført en strukturering af data i virksomhedernes databaser. Denne strukturering skaber mulighed for nye og mere avancerede teknologier. Det må dog konstateres, at Business Intelligence-værktøjer ofte ikke er særlige agile eller frontend-baserede, hvorfor vi fandt, at disse værktøjer formentligt ikke vil kunne håndtere kompleksiteten og omskifteligheden i besvigelser. Afledt heraf fandt vi, at Data Mining er en teknologi med karakteristika, der både kan håndtere kompleksiteten, de store datamængder og omskifteligheden, som besvigelser indeholder. Data Mining er en Machine Learning-teknologi, som effektivt kan operere i realtid i virksomhedernes it-infrastruktur som en integreret del af databaserne, hvilket i praksis benævnes backend-kodning. Backend-kodningen har en stor fordel i modsætning til frontend-systemer, da backend-kodning er langt mere agil og derudover har direkte adgang til store mængder data. Ved den teoretiske gennemgang af Data Mining identificerede vi forskellige teknikker inden for superviserede og ikke-superviserede teknikker, som kan anvendes mod besvigelser. Disse omfattede forskellige clusterings, klassifikationer og regressionsanalyser.

Med udgangspunkt i de karakteristika, der blev identificeret i forbindelse med gennemgangen af besvigelsesteorien, foretog vi et studie i, hvilke faresignaler der er relevante inden for ”faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen” og ”fiktiv omsætning”. Disse karakteristika kan omdannes til datapunkter og derigennem variabler, som igen kan anvendes til at skabe ny viden. De identificerede variabler kan i forskellige kombinationer danne grundlag for Data Mining-teknikkerne. I vores analyse fandt vi, at kombinationerne af variabler og teknikker vil være i stand til at identificere besvigelseslignende posteringer. Arbejdet med Data Mining-teknikker kan medføre, at der identificeres falske positiver og falske negativer, som kan gøre de efterfølgende undersøgelser mere omfattende.

Der konkluderes i den forbindelse, at det er muligt at reducere falske positiver og falske negativer ved at krydsvalidere mellem forskellige Data Mining-teknikker og variabler. Afledt af muligheden for reduktion af falske positiver og negativer samt identifikation af posteringerne konkluderer vi, at Data Mining kan anvendes som et værktøj til at opdage besvigelser. Endvidere konkluderer vi, at Data Mining vil være mere præcis end en klassisk opdagende kontrol, og at krydsvalideringen vil medføre færre udslagsgivende posteringer, som ikke er i besvigelsesrisikogruppen og dermed vil være mere effektiv end de klassiske kontroller. Data Mining er derfor et oplagt værktøj til bekæmpelse af besvigelser.

Som konsekvens heraf analyserede vi på implementeringen af Data Mining som en integreret del af virksomhedernes kontrolmiljø. Analyserne kan med fordel anvendes som grundlag for godkendelseskontroller og dermed også som forebyggende kontroller, fordi programmeringen sker backend og derfor kan opsættes til at analysere i realtid. Realtidsanalyserne kan identificere transaktioner med besvigelseslignende karakter, før disse bogføres. Transaktionerne kan herefter sendes til godkendelse hos et dedikeret besvigelsesrisikovurderingsteam, som derigennem kan undersøge posteringerne. Da transaktionerne standses, før de bogføres, kan virksomhederne på den måde opnå en besparelse relateret til inddrivelse af værdier, som er fragået virksomheden. Implementeringen kræver dog store ressourcer i form af omfattende kompetencer inden for datavidenskab og besvigelser. Endvidere fandt vi, at det kræver store investeringer i teknologi og processer. Selvom fordelene bl.a. består af store besparelser i form af en begrænsning i besvigelsesrelaterede tab, nye kompetencer, brede anvendelsesområder m.fl., er business casen svær at forsvare. Det skyldes, at tabene for virksomhederne ofte er ukendte, eller at der hersker en "det sker ikke for os"-mentalitet. I forlængelse heraf er det vores konklusion, at det vil være svært for virksomhederne at finde de fornødne kompetencer til arbejdet med Data Mining.

Afledt af ulemperne opstår der en mulighed relateret til outsourcing af besvigelsesbekæmpelsen. Denne type outsourcing kaldes managed services og kan varetages af fx softwareudbydere eller konsulentfirmaer. Ved at samle en ydelse hos en managed service-udbyder, kan udbydere skalere arbejdet på tværs af flere virksomheder. Dette muliggør, at investeringen kan bredes ud blandt flere parter, ligesom det kan være nemmere at tiltrække, udnytte og uddanne de relevante kompetencer. Afledt af disse fordele for udbyderen kan virksomhederne outsource området til leverandøren for et lavere beløb, end de selv skulle have anvendt til at opsætte kontrollerne, og således kan business casen nemmere retfærdiggøres. Outsourcingen kræver, at der opbygges strategiske partnerskaber, da det er essentielt, at arbejdet med Data Mining og besvigelsesopdagelse er baseret på faktiske risici og er tilpasset de enkelte virksomheder. Det konkluderes derfor, at anvendelse af en managed service-udbyder vil kræve, at relevante personer fra leverandøren skal indgå i virksomhedernes besvigelsesrisikovurderingsteam.

På baggrund af ovenstående opsummering af afhandlingens løbende konklusioner vil vi nu, med afsæt i afhandlingens problemformulering, som gengivet på næste side, drage den endelige konklusion.

Hvordan kan virksomheder ved hjælp af Data Mining styrke deres interne kontroller tilstrækkeligt til at kunne afdække risikoen for besvigelsesrelaterede tab?

Data Mining er som teknologi i stand til at forebygge og opdage besvigelser bedre og mere effektivt end de klassiske interne kontroller. Dog skal Data Mining ses som et supplement og ikke en erstatning til de traditionelle interne kontroller. På nuværende tidspunkt er det ikke realistisk, at virksomhederne selv kan opbygge kontroller baseret på Data Mining-teknologien, da business casen ikke kan forsvares, og kompetencerne formentlig heller ikke er til stede. Løsningen herpå kan være strategiske alliancer mellem virksomhederne og en managed service-udbyder, som har ressourcer og kompetencer til arbejdet med Data Mining og derigennem til besvigelsesopdagelsen. Virksomhederne deler viden og risici med serviceudbyderen, som stiller software, processer og kompetencer til rådighed. Herigennem kan virksomhederne anvende Data Mining som et effektivt værktøj i deres interne kontrolmiljø i kampen mod besvigelser og herigennem begrænse besvigelsesrelaterede tab.

Kapitel 5 - Perspektivering

5 Perspektivering

I dette afsnit vil vi berøre, hvordan Data Mining kan anvendes bredere end de to besvigelsesrisici, som har været i fokus for denne afhandling. Vi vil vurdere, hvordan Data Mining kan skabe yderligere værdi for virksomhederne og vurdere Data Minings anvendelse mod andre risikoområder.

Hvis besvigelsesopdagelsen bliver en service, som fx varetages af et konsulenthus, kan der opstå nye måder, hvorpå besvigelser nemmere kan identificeres, forebygges og begrænses. Særligt inden for benchmarking mellem forskellige virksomheder enten geografisk eller industrimæssigt ses der muligheder. Benchmarking er, når en observation sammenlignes med en tilsvarende observation hos en anden virksomhed, ofte inden for samme industri eller geografiske område. Såfremt en managed service-udbyder specialiserer sig inden for en industri, kan benchmarking være et potent værktøj. Hvis der identificeres besvigelser i en virksomhed, vil den nye viden kunne anvendes til at opsætte nye analyser og derigennem nye kontrolaktiviteter for alle de øvrige virksomheder inden for fx den samme industri. Virksomhederne vil igennem den strategiske alliance opnå en fordel mod besvigere. Den kombinerede viden har potentialet til i fremtiden at kunne håndtere den omskiftelighed, som besvigelser er omfattet af, ligesom managed service-udbyderen vil kunne effektivisere dens processer på tværs af geografiske områder eller industrier. Benchmarking har derved potentialet til at begrænse besvigelser ydereligere, ligesom benchmarkingen formentlig også vil tale for effektivisering og dermed omkostningsbesparelser.

Data Mining som værktøj har endvidere mange fremtidige anvendelsesmuligheder. I denne afhandling har vi fokuseret på to besvigelsesrisici og kun inden for private virksomheder, men teknologien har potentiale til at afdække mange flere besvigelsesrisici. Data Mining-teknikkerne kan anvendes på alle områder, hvor det er muligt at observere hændelser i data. Således vil mange af de besvigelsestyper, som besvigelsesteorien indeholder, forventeligt kunne afdækkes ved at kombinere forskellige variable og teknikker, som vi har gjort i analysen.

Et meget specifikt anvendelsesområde er inden for offentlige virksomheder og organisationer, hvor der senest verserer en sag fra Forsvarets Ejendomsstyrelse i medierne, hvor betroede medarbejdere ved hjælp af fiktive indkøb, bestikkelse og omgåelse af godkendelseskontroller har begået besvigelser for offentlige

midler. Rigsrevisionen har fundet, at to medarbejdere har modtaget bestikkelse for 1,8 mio.kr. (Ingvorsen and Vilsbøll, 2019), og at indkøbere i adskillige tilfælde selv har varemodtaget og godkendt fakturaer (Forsvarsministeriet, 2019). Som følge af rigsrevisionens offentliggørelse har sagen fået enorm medieomtale, og Forsvarsministeren har måttet blande sig i sagen. Ministerens indblanding har medført, at sagen skal kulegraves for at undersøge det samlede omfang. Denne besvigelsessag er blot et enkelt eksempel på, hvor store de økonomiske konsekvenser, direkte såvel som indirekte, igennem sagsomkostninger, undersøgelser, fragåede ressourcer fra organisationen mv., kan være for virksomheder og organisationer.

En sådan sag vidner om, at der er tale om mangelfulde interne kontroller. De interne kontroller har højst sandsynligt været til stede, men har ikke fungeret effektivt. Måske skulle der blot ske en efterprøvning og opdatering af de interne kontroller, men vores bud er, at der er et behov for at arbejde med ny teknologi for på den måde at sikre, at kontrollerne ikke fremadrettet kan blive forbigået. Derfor vurderer vi, at Data Mining også vil have sin berettigelse inden for det offentlige inden for de næste fem til ti år. Sagen om Forsvarets Ejendomsministerie kunne ved brug af Data Mining være blevet opdaget, da de forskellige besvigelsesgerninger ville medføre direkte observerbare datapunkter. Fx ville det forhold, at der er sammenfald mellem varemodtager og fakturagodkender hurtigt springe frem i nogle af de teknikker, som vi har gennemgået i afhandlingens analyse. Desuden ville det offentlige have mulighed for at eliminere de udfordringer, vi identificerede som relateret til anvendelse af Data Mining som intern kontrol, idet staten har mulighed for at arbejde med teknologien som en managed service på tværs af forskellige styrelser. Endvidere ville det også være muligt at inddrage kommuner og lignende i denne kategori. Der vil derfor være volumen nok til at foretage de fornødne investeringer.

Endelig har vi i afhandlingen kun demonstreret enkelte af de mange forskellige Data Mining-teknikker, der eksisterer. Fx fremgår der i ”Bilag 8 – Data Mining og underkategorier” mange forskellige teknikker, som vi ikke berørte i analysen. Eksempler var Neural Networks, som anvendte Deeper Learning-teknologien. Denne teknologi afgrænsede vi os fra, da teknologien er meget teknisk krævende, særligt hvis denne skulle implementeres hos en enkelt virksomhed. Endvidere kræver teknologien, at analysens design rettes mod en enkelt risiko. Derfor er det meget omkostningskrævende for den enkelte virksomhed at arbejde med Deeper Learning. Her bliver benchmarking, arbejdet på tværs af sektorer, og derigennem managed services, interessant. Arbejdet på tværs af virksomheder muliggør anvendelse af den langt mere komplekse teknologi, som er tættere på menneskets kognitive intelligens. Deeper Learning-teknologien har et uudforsket potentiale for besvigelsesopdagelse, men teknologiens karakteristika beviser klart, at der

eksisterer muligheder inden for feltet. Neural Networks består af forskellige statistiske funktioner - også kaldet lag. Når en funktion aktiveres af en transaktion, flyttes posteringen med tilføjelse af den nye viden til næste lag. Når alle lag er gennemgået, er netværket i stand til at klassificere en postering med en meget høj præcision. SAS Institute konkluderer, at Neural Network er i stand til at løse stort set alle komplekse problemstilling til den ønskede præcision (Baesens, Vlasselaer and Verbeke, 2015).

Data Mining, herunder Machine Learning, indeholder således et utal af muligheder inden for intern kontrol og besvigelsesopdagelse, når denne kan skaleres. Denne afhandling er derfor kun et første bud på anvendelse af den potente teknologi.

Litteraturliste

ACFE (2018) *Report to the Nation: Occupational Fraud and Abuse. Association of Certified Fraud Examiners (2018)*.

AGA Intergov (2019a) *AGA - Billing Schemes*. Available at: https://www.agacgfm.org/Tools-Resources/intergov/Fraud-Prevention/Tools-by-Fraud-Type/Billing-Schemes.aspx (Accessed: 12 November 2019).

AGA Intergov (2019b) *AGA - Fraud Triangle*. Available at: https://www.agacgfm.org/Tools-Resources/intergov/Fraud-Prevention/Fraud-Awareness-Mitigation/Fraud-Triangle.aspx (Accessed: 10 December 2019).

Andersen, I. (2009) *Den skinbarlige virkelighede*. 5th edn, *Samfundslitteratur*. 5th edn. doi: 10.1016/j.chc.2017.11.015.

Baesens, B., Vlasselaer, V. Van and Verbeke, W. (2015) *Fraud Analytics, Using descriptive, predictive and social network techniques. A guide to data science for fraud detection*. Wiley & SAS Insititute.

Bloch Christensen, M., Kristensen, R. H. and Warming-Rasmussen, B. (2015) *Revisors opklaring af besvigelser*. 2nd edn. Karnov Group.

BusinessDictionary.com (2019) *What is data analysis? definition and meaning - BusinessDictionary.com*.

CFI (2019) *Fraud Red Flags - Learn to Identify Internal Indicators of Fraud*. Available at: https://corporatefinanceinstitute.com/resources/knowledge/other/fraud-red-flags/ (Accessed: 11 November 2019).

COSO (2013) *Internal Control — Integrated Framewo rk Framework and Appendices*.

Danske, G.-D. S. (2017) *regressionsanalyse*.

Deloitte ZA (2019) *Data Analytics | Risk Advisory | Deloitte*.

Dorminey, J. *et al.* (2012) 'The evolution of fraud theory', *Issues in Accounting Education*. doi: 10.2308/iace-50131.

EDUCBA (2019a) *Data Mining Techniques | Top 7 Data Mining Techniques for Best results*. Available at: https://www.educba.com/data-mining-techniques/ (Accessed: 8 January 2020).

EDUCBA (2019b) *What is Clustering in Data Mining? | Application of clustering in data mining*.

EDUCBA (2019c) *What is Machine Learning? | How It Works | Technique & Advantage*.

Folketinget (2010) *Selskabsloven (lov om aktie- og anpartsselskaber) med kommentarer*. Available at: https://www.retsinformation.dk/pdfPrint.aspx?id=209846.

Folketinget (2019a) *LEV-loven - Bekendtgørelse af lov om visse erhvervsdrivende virksomheder - retsinformation.dk*. Available at: https://www.retsinformation.dk/Forms/R0710.aspx?id=209437#id1c301cb3-8153-4fe1-ad7e-d9d87145faa8 (Accessed: 6 September 2019).

Folketinget (2019b) *Selskabsloven - Bekendtgørelse af lov om aktie- og anpartsselskaber (selskabsloven) - retsinformation.dk*. Available at: https://www.retsinformation.dk/Forms/R0710.aspx?id=209846 (Accessed: 6 September 2019).

Forbes (2018) *Council Post: What Is A 'Managed Service' Anyway?*

Forsvarsministeriet (2019) *Redegørelse vedrørende forløbet i forbindelse med Rigsrevisionens undersøgelse af indkøb hos Forsvarsministeriets Ejendomsstyrelse*. Available at: https://fmn.dk/nyheder/Pages/opfoelgning-efter-svindel-i-forsvarsministeriets-ejendomsstyrelse.aspx.

Gartner (2019) *Business Analytics*.

Gravitt, J., Johnson, J. and Horwath, C. (2008) *Recognizing Financial Statement Fraud Red Flags - Criminal Law - United States*. Available at: http://www.mondaq.com/unitedstates/x/56058/White+Collar+Crime+Fraud/Recognizing+Financial+Statement+Fraud+Red+Flags (Accessed: 11 November 2019).

Hubler, M. (2016) *A Guide to Financial Statement Fraud, Red Flags & Prevention Tips*. Available at: https://www.purdueglobal.edu/blog/business/guide-financial-statement-fraud/ (Accessed: 11 November 2019).

IBM Analytics (2017) *(115) Machine Learning - Unsupervised Learning - Density Based Clustering - YouTube*.

Ingvorsen, E. S. and Vilsbøll, S. M. K. (2019) *Sådan foregik mulig svindel: Cykelskur blev til tekøkken og steg 80.000 kroner | Indland | DR*, *Danmarks Radio*. Available at: https://www.dr.dk/nyheder/indland/saadan-foregik-mulig-svindel-cykelskur-blev-til-tekoekken-og-steg-80000-kroner (Accessed: 10 December 2019).

Kassem, R. (2019) *Additional red flags for financial statement fraud | Download Table*. Available at: https://www.researchgate.net/figure/Additional-red-flags-for-financial-statement-fraud_tbl3_319044233 (Accessed: 11 November 2019).

Koen, S. and Towards data science (2019) *Not yet another article on Machine Learning! - Towards Data Science*.

Kutz, N. and Washington University (2016a) *Data Analysis: Clustering and Classification (Lec. 1, part 1) - YouTube*.

Kutz, N. and Washington University (2016b) *Data Analysis: Clustering and Classification (Lec. 1, part 2) - YouTube*.

Kutz, N. and Washington University (2016c) *Data Analysis: Clustering and Classification (Lec. 1, part 3) - YouTube*.

L. Cotton, D., Johnigan, S. and Givarz, L. (2016) *Fraud Risk Mangement Guide*. The Committee of Sponsoring Organizations' (COSO).

Leskovec, R. (2016) *(115) Lecture 20 — Frequent Itemsets | Mining of Massive Datasets | Stanford University - YouTube*.

Lormer, D. (2017) *An Essential Guide to Accounts Payable Fraud | i-Sight*. Available at: https://i-sight.com/resources/accounts-payable-fraud/ (Accessed: 11 November 2019).

McNeal, A. (2019) 'EXT RE Help to a Danish study in the most common fraud committed'. ACFE.

Nigrini, M. J. (1999) 'I ' ve Got Your Number', *JOURNAL of ACCOUNTANCY*, May(5), pp. 79–83.

Oversight Systems Inc (2019) *Purchase Order Fraud: How to Identify Red Flags in Procure-to-Pay*. Available at: https://www.oversightsystems.com/blog/purchase-order-fraud-procure-to-pay-red-flags (Accessed: 11 November 2019).

Pinkasovitch, A. (2019) *Detecting Financial Statement Fraud*. Available at:

https://www.investopedia.com/articles/financial-theory/11/detecting-financial-fraud.asp (Accessed: 11 November 2019).

Predictiveanalyticstoday.com (2019) *What is Data Analysis ?*

PWC (2018) 'Pulling Fraud out of The Shadows', *PwC's Global Economic Crime and Fraud Survey 2018*. Available at: https://www.pwc.com/gx/en/forensics/global-economic-crime-and-fraud-survey-2018.pdf.

PWC (2019) *Re-inventing Internal Controls in the Digital Age*.

Ravelin (2019) 'Machine learning for fraud detection'. Available at: https://www.ravelin.com/insights/machine-learning-for-fraud-detection?utm_term=fraud machine learning&utm_campaign=Awareness+-+Machine+learning&utm_source=adwords&utm_medium=ppc&hsa_ver=3&hsa_grp=58391109750&hsa_kw=fraud machine learning&hsa_tgt=kwd-30262 (Accessed: 8 January 2020).

Reaseach, S. G. (2019) 'The Challenger Audit Brand : Impossible or inevitable ? Emerging Trends', (January).

Riley, R., Kranacher, M. J. and Wells, J. T. (2010) *Forensic accounting and fraud examination*.

Statistics How To (2015) *False Positive and False Negative: Definition and Examples - Statistics How To*.

Tableau (2019a) *Business intelligence: what it is and why it matters*.

Tableau (2019b) *Top 10 2019 Business Intelligence Trends | Tableau Software*.

Techdifferences (2019) *Difference Between Descriptive and Predictive Data Mining (with Comparison Chart) - Tech Differences*.

Techopedia (2019) 'What is Link Analysis? - Definition from Techopedia', *Techopedia*.

The Committee of Sponsoring Organizations' (COSO) (2016) *Fraud risk management guide - executive summary*.

The Committee of Sponsoring Organizations' (COSO) (2018) *About Us*.

Vona, L. (2017) *Spotting Potentially Fraudulent Shell Companies During Audits*. Available at: https://misti.com/internal-audit-insights/spotting-potentially-fraudulent-shell-companies-during-audits (Accessed: 11 November 2019).

Wells, J. T. (2013) *Principles of Fraud Examination, Fourth Edition*. 4th edn, *John Wiley & Sons, Inc.* 4th edn. Wiley.

Wolf, R. (2016) *Billing Schemes: 3 Ways Fraudsters Can Cash In on Your Nonprofit*. Available at: https://www.gma-cpa.com/blog/billing-schemes-3-ways-fraudsters-can-cash-in-on-your-nonprofit (Accessed: 12 November 2019).

Zack, G. M. and ACFE (2015) 'Using Data Analytics to Detect Fraud', pp. 1–16.

Bilag

Bilag 1 – Mail fra ACFE

Dear Kim,

My sincere apologies in the delay in our response. I was able to pull some statistics on the frequency of various forms of financial statement fraud from our 2018 *Report to the Nations*. The results of that analysis is noted below.

	Percent of FSF cases
Fictitious Revenues	34%
Improper Disclosures	28%
Concealed Liabilities and Expenses	26%
Improper Asset Valuations	19%
Timing Differences	16%

As for your request for case studies, you might find the cases in our "Stories Behind the Numbers" series helpful: https://secure-web.cisco.com/1r7H1WRAwt48IE73In739R5aWzRuCdBmBDBCSRBh5RIQZMKceSAC9QchcaN-sto-Gdb_f8iBbcBFqtl_EOlYlzQ8iOYKhmFpvv4yi2Sc9999f44XHE7NHdHTkJGWEb66Lw1nb5C_VQRx3DuGqKL5Bt4xXbotQC0dxFyO_XdWAiOg-NNSSOi5TLthi9VWPEtNOX1I8acvLxQccfnoEZoV2j9F6E11slnHHdEMxFpvEyrGQ3_eCrmxwCcPLfxq1LsISNHQ63PkIV5D81Trwm7MEztaYI-4Qm3DVr0q4aVrzrhLWwsGAsn2T1nqS0nX_6_fUlBkhRAoBBHjKDloZDp1Bxg/https%3A%2F%2Fwww.acfe.com%2Freport-to-the-nations%2Fbehind-the-numbers%2F

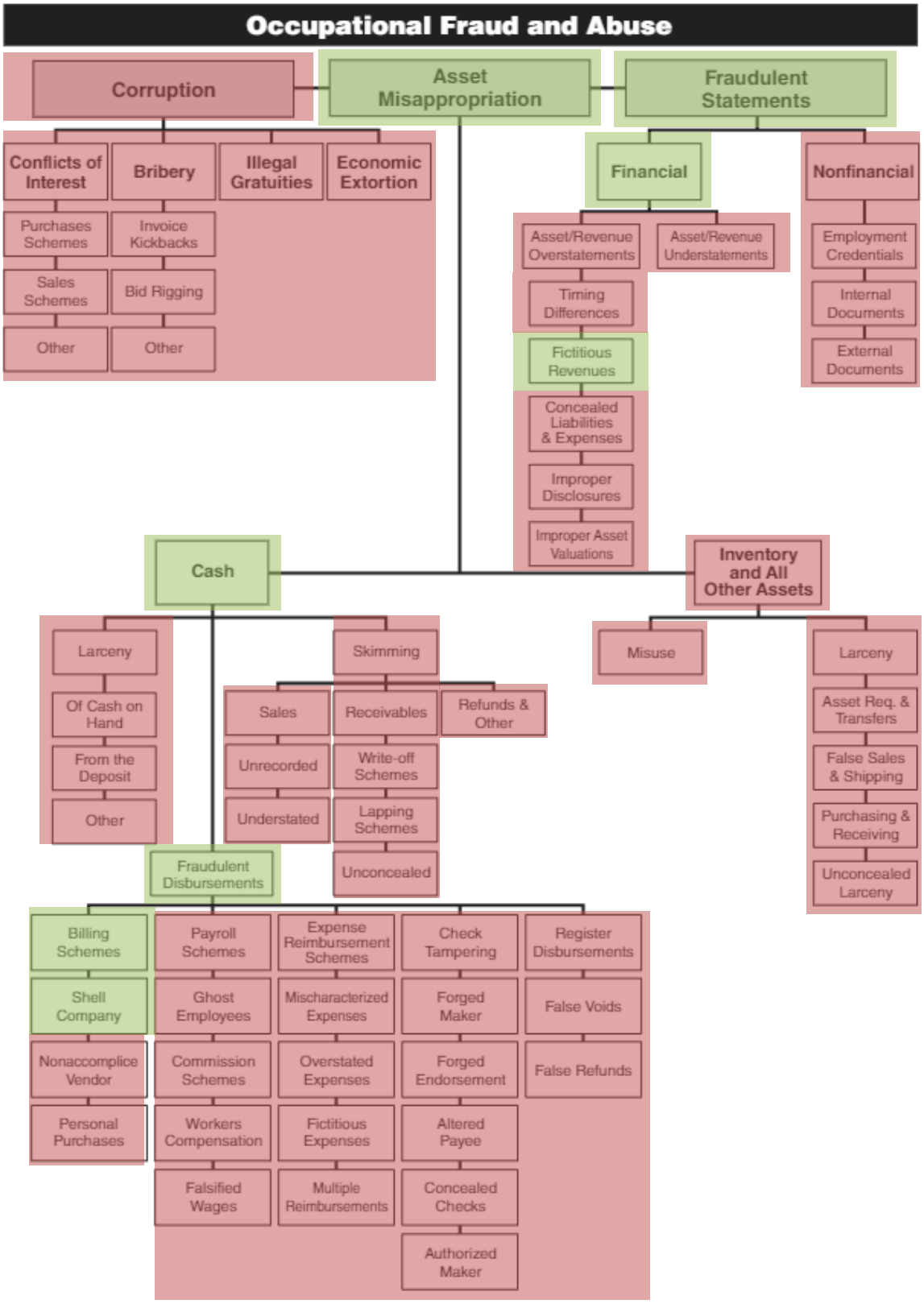
I wish you the best of luck with your studies! Please let me know if there's any other assistance we can provide.

Best regards,

Andi McNeal, CFE, CPA
Director of Research
Association of Certified Fraud Examiners
Global Headquarters • The Gregor Building
716 West Avenue • Austin, TX 78701
Tel: +1 512-478-9000
Fax: +1 512-478-1444
E-mail: amcneal@ACFE.com

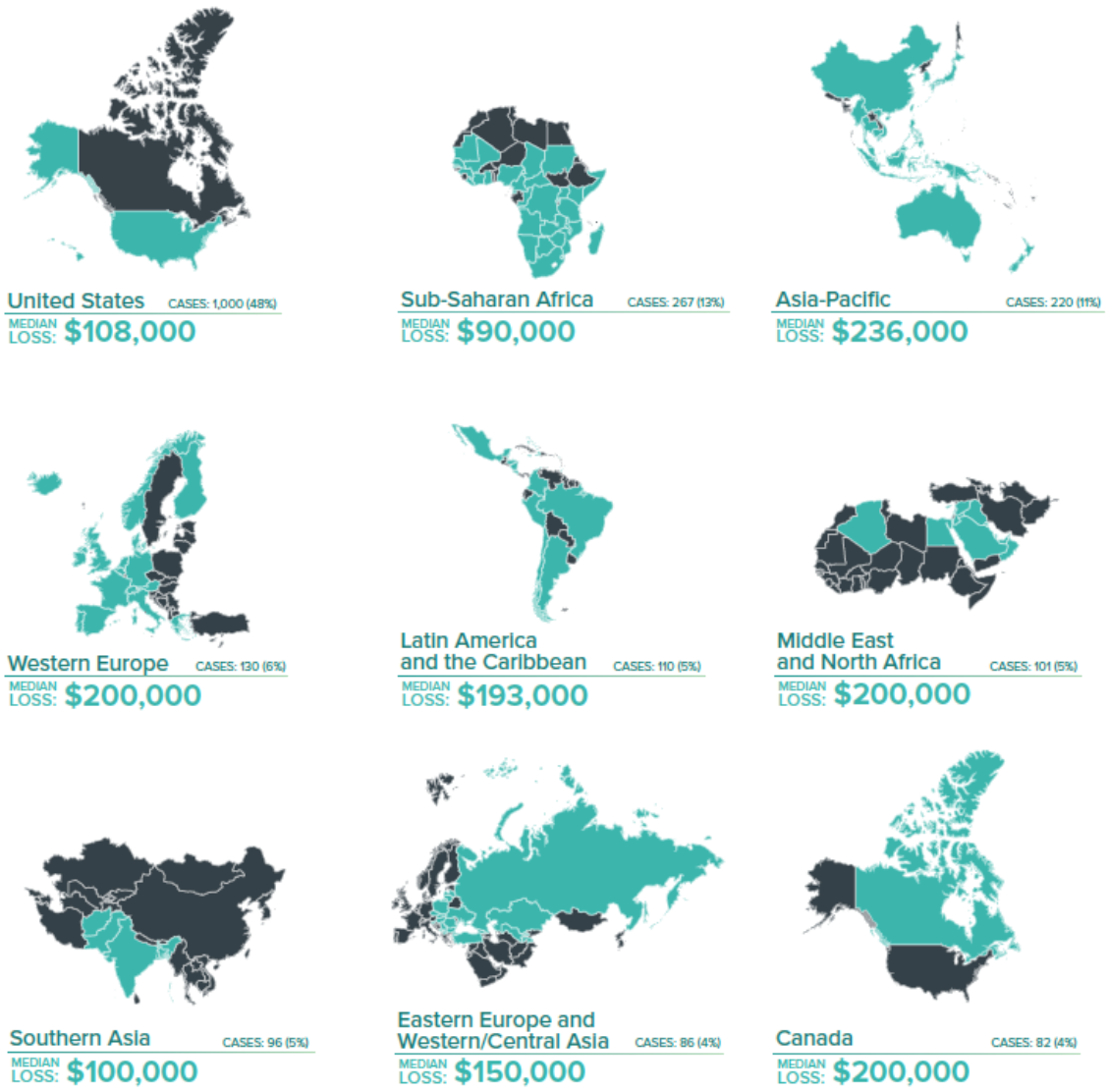
Miss the *ACFE Global Fraud Conference* in Austin? Register for the Virtual Conference and see more than 75 sessions, earn up to 26 CPE.

Bilag 2 – Skematiseret afgrænsning inden for Wells' besvigelsestræ



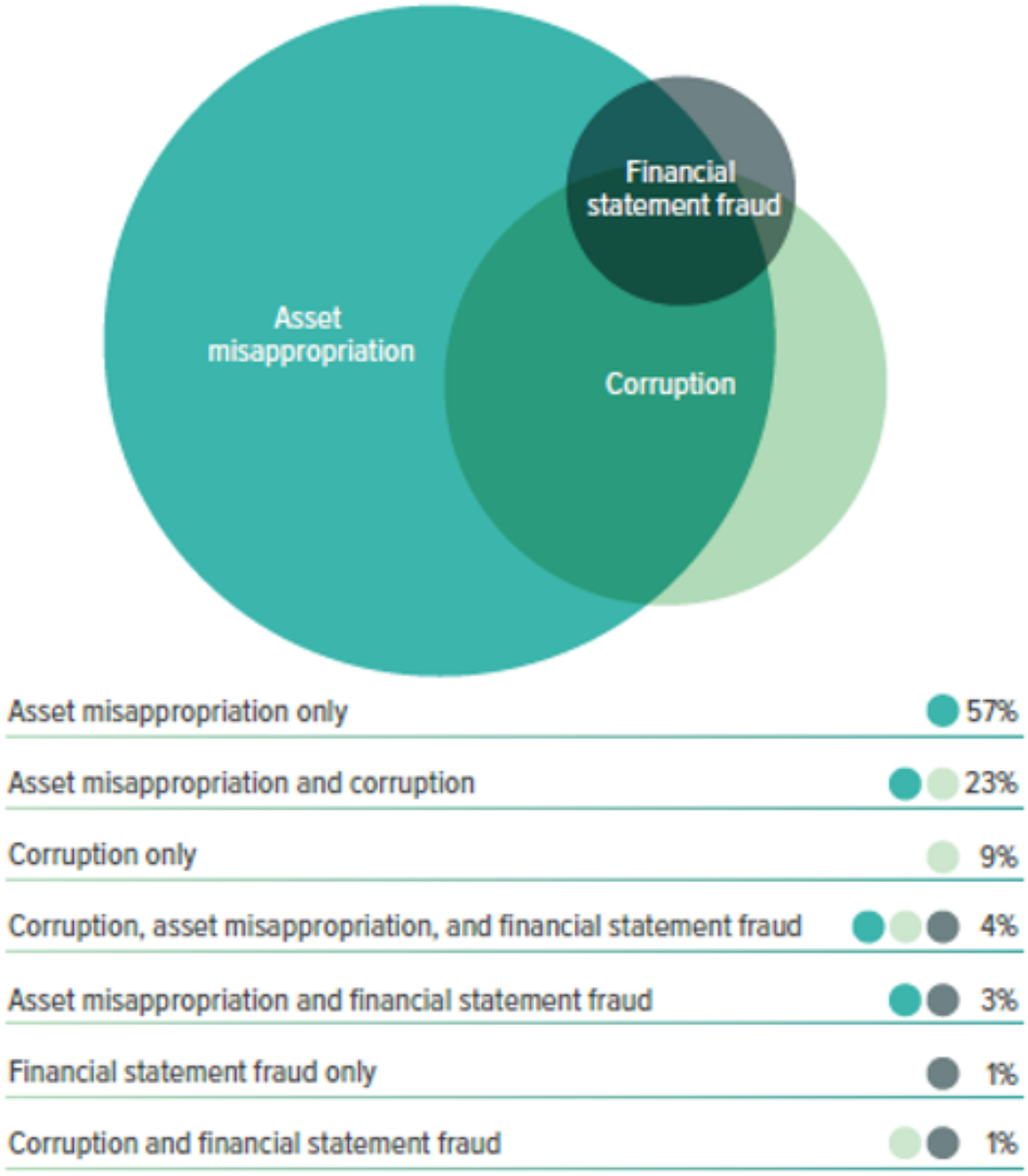
Besvigelsestræet, (Wells, 2013)

Bilag 3 – Geografisk fordeling af konstaterede besvigelser



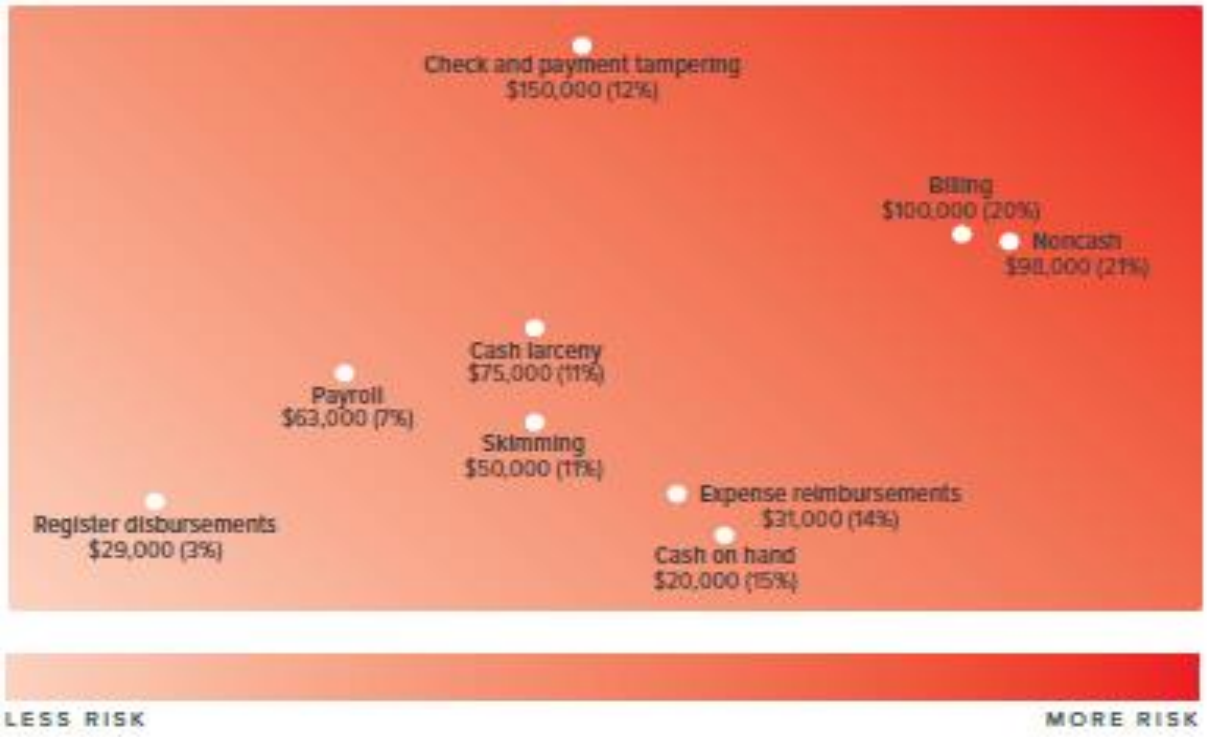
Geografisk fordeling af konstaterede besvigelser (ACFE, 2018, p. 7)

Bilag 4 – Forekomstfordeling imellem besvigelseshovedkategorierne



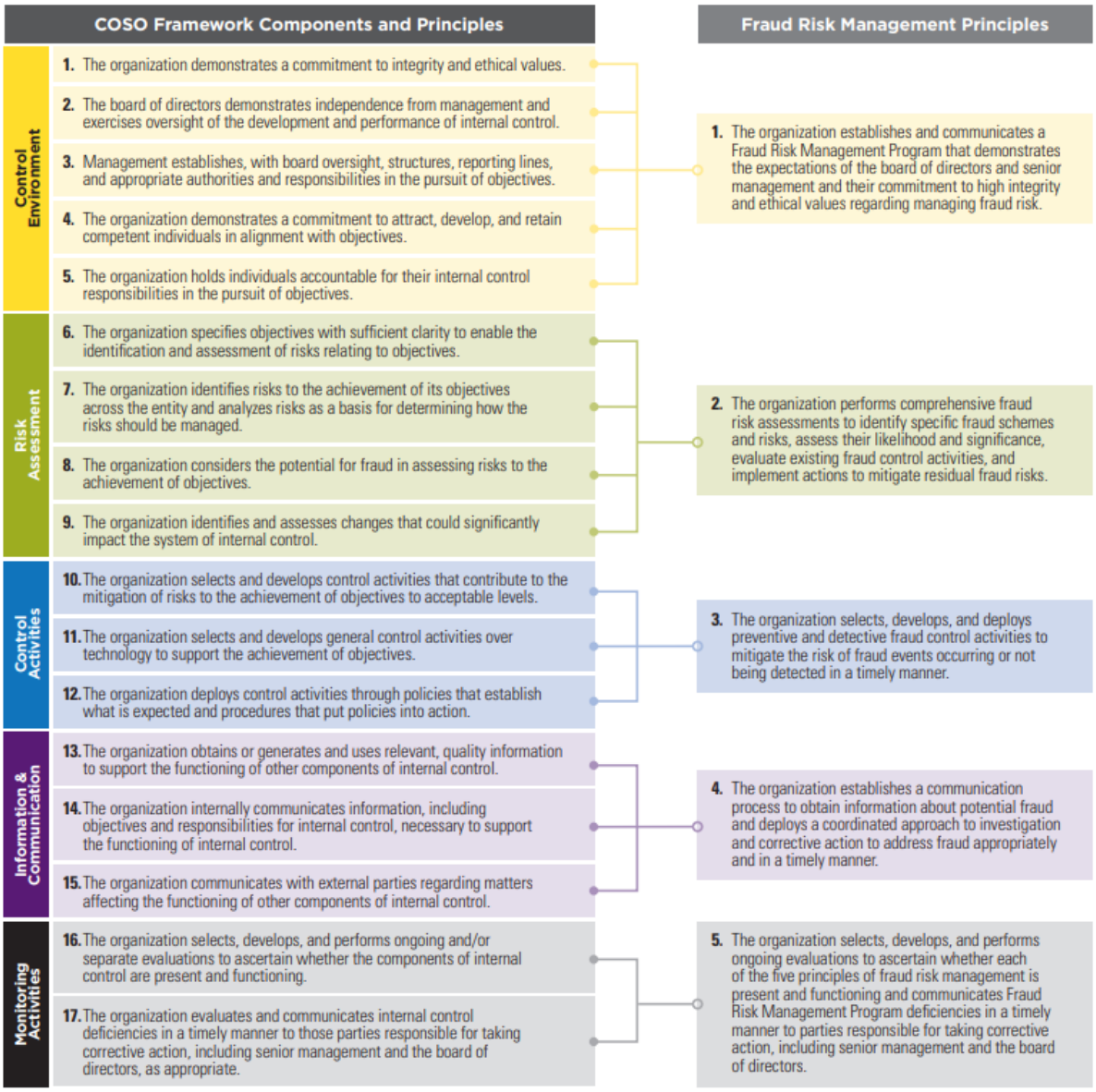
Forekomst inden for besvigelseshovedkategorierne (**ACFE, 2018, p 12)**

Bilag 5 – Heatmap - Mest risikobetonet besvigelsestype



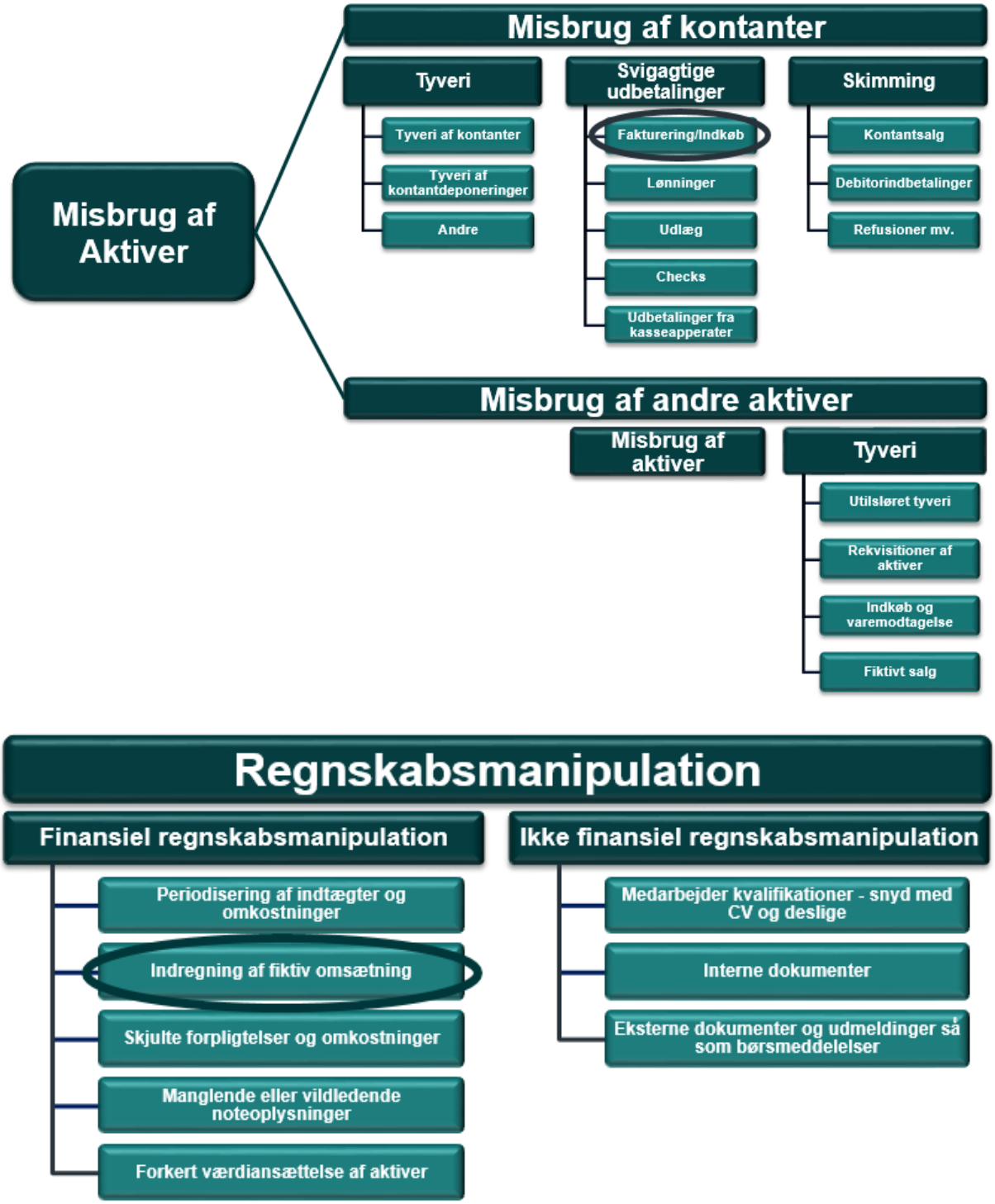
Mest risikobetonet besvigelsestype (ACFE, 2018, p. 12)

Bilag 6 – COSO-principperne og disses sammenhæng til besvigelses-frameworket



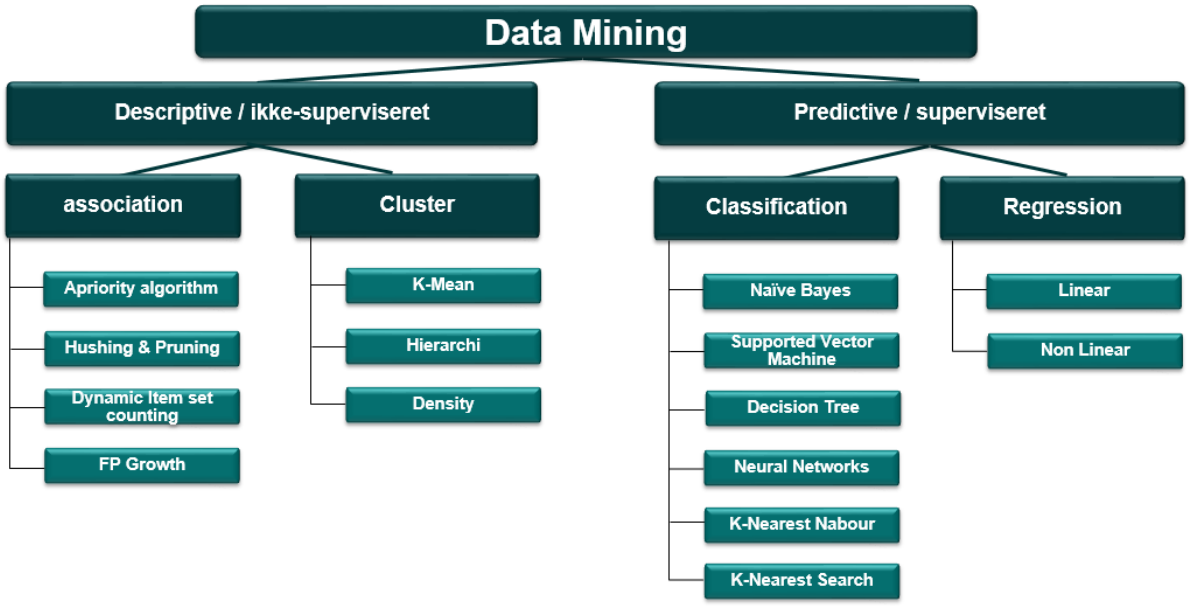
Relationship Between the 2013 COSO Framework's Five Components and 17 Internal Control Principles and this Guide's Five Fraud Risk Management Principles (The Committee of Sponsoring Organizations' (COSO), 2016).

Bilag 7 - Besvigelsestræet inden for afgrænsning:



Kilde: Egen tilvirkning med inspiration fra Wells' besvigelsestræ (Wells, 2013)

Bilag 8 – Data Mining og underkategorier



Kilde: Egen tilvirkning på baggrund af litteraturstudie af Data Mining og Machine Learning

Bilag 9 – Bruttoliste for faresignaler i forbindelse med faktureringssvindel

Faresignal	**Kilde**
• Invoices cannot be traced to shipments • Multiple payments to single vendor on the same date • Pattern of purchases just below review level • Unusually quick turnaround of invoices • Payment to multiple vendors for same product • Extreme inventory shortage • Expenses increase dramatically • Unexplained rise in cost of goods sold • Unexplained decrease in gross/net profits • Excessive materials orders • Goods not purchased at optimal point • High level approval of a low level transaction	(AGA Intergov, 2019a)
• Sequtentiel invoice numbers from one vendor • Illogical invoice number range identicates a higher level for concealment • Date of invoice is within 90 days off the vendores registration date • Dublicate adress, vendor, email etc. • Description of goods or services is not consistent with a real business	(Vona, 2017)
• Excessive number of adjusting entries • Missing documents	(CFI, 2019)
• Buyers processing POs for vendors outside their normal responsibilities • POs for vague or poorly defined services including “blanket orders” • Multiple POs to one vendor in the same or similar amounts from the same requesting group • Identical items purchased in different amounts simultaneously or within short periods of time • POs split by type of work (e.g., one purchase order for labor and another for material) • Recurring purchases that fall just under review/authorization thresholds to avoid the scrutiny required for larger purchases • Invoices that predate Pos • POs that have been amended after invoices have been submitted • Vendors issuing invoices with missing or improper PO numbers, or verbal Pos • Repeated use of a one-time-vendor record for the same address	(Oversight Systems Inc, 2019)
• Rounded amounts • Very large payments to one vendor • Unusually large purchases on an employee’s company-issued credit card • Invoices that are missing key details, such as address and phone number • "Vendor’s prices that seem unusually low or high"	(Lormer, 2017)
• More often services than goods • Unfamiliar vendors or variations on an approved vendor’s name – a fraudster may try to bypass red flags by altering the name of an approved company such as changing ABC, Inc. to ABC, LLC • Invoices for unspecified or poorly defined services – fraudulent invoices may be vague in order to avoid drawing attention	(Wolf, 2016)

• Vendors that only have a P.O. box address – it could be a sign they don't actually exist • Sudden increases in purchases from one vendor – a fraudster may just be cashing in on your organization • "Vendor addresses that match employee addresses – the company invoicing you for five hours of business consulting probably isn't being run out of your employee's living room" • Unusually quick turnaround of invoices – speedy turnaround may mean your fraudster is in a hurry to cash in • Large billings that are broken into multiple smaller invoices that will not attract attention – flying under the radar is a fraudster's top priority, and they will do anything to avoid detection	

Bilag 10 – Bruttoliste faresignaler, indregning af fiktiv omsætning

Faresignaler	**Kilde:**
• Fake costumers • Increasing in revenue, without corresponding growth in cashflow • Discrepancies and unexplained items on accounting reconciliation • Missing/altered documents • Significant or unusual changes in assets or liabilities • High revenue figures during a time in which competitors are in a downturn • Reinvoice past due accounts • Sales without shipment • Sales without orders • Alteration of accounting records	(Hubler, 2016)
• The company maintains consistent gross profit margins while its industry is facing pricing pressure. This can potentially indicate failure to recognize expenses or aggressive revenue recognition • A large buildup of fixed assets. An unexpected accumulation of fixed assets can flag the usage of operating expense capitalization, rather than expense recognition • Outsized frequency of complex related-party or third-party transactions, many of which do not add tangible value (can be used to conceal debt off the balance sheet) • A rapid and unexplainable rise in the number of day's sales in receivables in addition to growing inventories. This suggests obsolete goods for which the firm records fictitious future sales	(Pinkasovitch, 2019)
• Unusual increase in sales from few companies, branches, or people within the compay as whole • High level revenue postings does not follow patterns of other sales in terms of discounts etc. • Improper recording of revenue, using balancesheets accounts that does not correspond to normal cause of business eg. Accrued revenue • Large, complex and unusual transcation close to period-end	(Kassem, 2019)
• Significant volume of sales to entities whose substance and ownership is not known: The red flag may be possibility of inflating the sales • Sales recorded by corporate headquarter: There may be a possibility that the sale has not yet taken place or rights of risk of ownwership has not passed to the purchaser • Outstanding accounts receivable from customers which are difficult to identify	

• Significant transactions with related party entities or special purpose entities not in the ordinary couse of business, where those are entities are not audited or audited by other firms • A significant to entities which substanse or ownership is not known	(Riley, Kranacher and Wells, 2010)
• Entries are made by individuals that rarely or newer performeds financial journal entries • Management override of controls • Use of journal entries	(Gravitt, Johnson and Horwath, 2008)

Bilag 11 – Reelle ejere, udvalgte faresignaler, variabler og datakilder

Faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen
- Reelle ejere

Faresignaler	Variable *	Datakilder
+ Fortløbne fakturanumre fra en leverandør	(a) Seneste fakturanummer på leverandørniveau (b) Foregående fakturanummer på leverandørniveau	- ERP-systemer (kreditormodul og kreditorposteringer)
+ Ustruktureret udvikling i fakturanumre på leverandørniveau	(a) 2x standardafvigelse på summen af fakturanumre pr. leverandør (beregnet variabel), (b) Afvigelse mellem observationspunktet (fakturaen) og foregående faktura (beregnet)	- ERP-systemer (kreditormodul og kreditorposteringer)
+ Datering af leverandør faktura indenfor 90 dage efter virksomhedsregistrering	(a) Fakturadato (b) Registreringsdato i det centrale virksomhedsregister	- ERP-systemer (kreditormodul og kreditorposteringer) - www.virk.dk eller lignende registre
+ Sammenfald af adresse mv. på flere leverandører	(a) Leverandørnavn (b) Adresse, tlf.nr. mv.	- ERP-systemer (kreditormodul og kreditorposteringer)
+ Leverandører er hjemmehørende på en postboksadresse	(a) Leverandørnavn (b) Adresse i form af postboks adresse (ja/nej) - enten direkte i teksten eller ved opslag mod virksomhedsregister (beriget variabel)	- ERP-systemer (kreditormodul og kreditorposteringer) - www.virk.dk eller et adresseregister
+ Sammenfald mellem medarbejder- og leverandøradresser	(a) Leverandøradresse (b) Medarbejder adresse (c) Sammenfald mellem adresser (ja/nej) beregnet variabel	- ERP-systemer (kreditormodul og kreditorposteringer) - HR/lønsystem
+ Indkøb af samme produkt/ydelse fra en ikke kendt samhandelspartner (leverandør)	(a) Ny samarbejdspartner (ja / nej), beregnet variabel (b) Varegruppe / varetype/ varekategori (c) Er der en anden samhandelspartner på variabel b.	- ERP-systemer (kreditormodul og kreditorposteringer) - Lagermodul eller lagersystem

** For alle posteringer indhentes hele faktureringsposteringerne i form af dato, beløb, tekst, konto, modkonto, valuta, løbenummer, kunde-/leverandørnavn, kunde-/leverandørnummer, brugeroplysninger, herunder intialer, afdeling mv.*

Bilag 12 – Godkendelsesprocedurer - udvalgte faresignaler, variabler og datakilder

Faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen
- Omgåelse af godkendelsesprocedure/forretningsgange

Faresignaler	Variable *	Datakilder
+ Indkøb under treshold for review eller godkendelse	(a) Treshold for godkendelse (b) Varians mellem treshold og fakturabeløb (beregnet) (c) antal indkøb pr. leverandør over/under threshold (d) bruger sammenfald	- ERP-systemer (kreditormodul og kreditorposteringer) - Godkendelseshierarki
+ Indkøb af vare i afdelinger der ikke er i nær tilknytning til den indkøbte leverance	(a) Brugerens (indkøber og/eller godkender) kostcenter eller tilhørssted (afdeling) (b) Omkostningen/varens kostcenter eller tilhørssted (afdeling) (c) Manglende sammenfald mellem kostcenter for bruger og indkøb (ja/nej), beregnet variabel	- ERP-systemer (kreditormodul og kreditorposteringer) - HR / lønssystem - Omkostningsgodkendelsessystem
+ Godkendelse af mindre omkostninger foretaget højt i godkendelseshierakiet	(a) Godkender (b) Godkender er manager eller højere (ja/nej) (c) Godkendelse er ikke foretaget logisk i forhold til godkendelse hierarki (ja/nej)	- ERP-systemer (kreditormodul og kreditorposteringer) - HR / lønssystem - Omkostningsgodkendelsessystem

** for alle posteringer indhentes hele faktureringsposteringerne i form af dato, beløb, tekst, konto, modkonto, valuta, løbenummer, kunde-/leverandørnavn, kunde-/leverandørnummer, brugeroplysninger, herunder intialer, afdeling mv.*

Bilag 13 – Typen af-/og manglende leverance - udvalgte faresignaler, variabler og datakilder

Faktureringssvindel igennem skyggeselskaber i forbindelse med indkøbsprocessen
- Typen af leverancer

Faresignaler	Variable *	Datakilder
Typen af leverancer		
+ Overvægt af serviceydelser for enkelte leverandører	(a) Produkttype på posteringsniveau (ydelse/vare) (b) Leverandørnavn/kundenummer	- ERP-systemer (kreditormodul og kreditorposteringer)
Manglende leverancer		
+ Leverandørfaktura kan ikke spores til fragtdokumenter/varemodtagelse	(a) Fragtdokument/varemodtagelse (b) Fakturanummer	- ERP-systemer (kreditormodul og kreditorposteringer) - Lagermodul

** for alle posteringer indhentes hele faktureringsposteringerne i form af dato, beløb, tekst, konto, modkonto, valuta, løbenummer, kunde-/leverandørnavn, kunde-/leverandørnummer, brugeroplysninger, herunder intialer, afdeling mv.*

Bilag 14 – Fiktiv omsætning - udvalgte faresignaler, variabler og datakilder

Indregning af fiktiv omsætning

Red flags	Variable*	Datakilder
Tilsløring via balancen		
+ Anvendelse af unormale balancekonti til indregning af omsætning	(a) modkonto (b) bruger	- ERP posteringer - HR/lønsystem
Posteringer som ikke følger normale forretningsgange		
+ Ændrede regnskabsposteringer	(a) Postering ændret ja/nej (b) Beløb før (c) beløb efter (d) bruger	-ERP posteringer - Logfil Ændringer - HR/lønsystem
+ Genfakturering til debitorer med overfaldne balancer eller hvor kreditmaks er nået	(a) overforfalden saldi (b) Kreditmaks overskredet (ja/nej) (c) Bruger	- ERP posteringer (salg og debitor) - HR/lønsystem - CRM system
+ Bogføringer foretaget af personer som sjældent foretager bogføringer	(a) Bruger (b) Historik, antal posteringer bruger har foretaget (c) Dage siden sidste postering	
+ Brug af finansposteringer (journal entries)	(a) Posteringstype, finanspostering (ja/nej) (b) Modkonto (c) Bruger (d) beløb	-ERP Posteringer, general ledger -HR/Lønsystem
+ Salg uden vareafsendelse (uden vareforbrug eller lagertræk)	(a) Ordrenummer (b) dækningsgrad (c) lagertræk (ja/nej)	- ERP posteringer (salg, debitor, vareforbrug) - Lagerposteringer
Ændring i samhandelsmøntre		
+ Anvendelse af falske kunder	(a) Kundenavn (b) Adresse i form af postboks adresse (ja/nej) – enten direkte i teksten eller ved opslag mod virksomhedsregister (beriget variabel) (c) Salg oprettet i CRM (ja/nej) (d) Første salg til kunde (ja/nej) (e) sammenfald mellem opretter af kunde og salg (ja/nej)	- ERP-systemer (debitormodul og salgsposteringer) - www.virk.dk eller et adresseregister - CRM posteringer - Logfiler på kundeoprettelse
+ Transaktioner med nærstående partner på udsædvanlige tidspunkter eller af usædvanlig størrelse	(a) nærtstående (ja/nej) enten kendt variable eller baseret på navne sammenfald (b) posteringsdag (c) Antal gange postering overstiger den normale samhandel	- ERP-systemer (debitormodul og salgsposteringer) - www.virk.dk ifht. Ejerstruktur - CRM posteringer

** for alle posteringer indhentes hele faktureringsposteringerne i form af dato, beløb, tekst, konto, modkonto, valuta, løbenummer, kunde-/leverandørnavn, kunde-/leverandørnummer, brugeroplysninger, herunder intialer, afdeling mv.*

Egen tilvirkning på baggrund af bruttolisten over redflags i bilag 10. Variabler og datakilder er identificeret ved rationalisering og gennemgang af mulige variabler i ERP-systemers API-dokumentation.

Bilag 15 – Fordele og ulemper ved implementering af Data Mining

Fordele og ulemper ved implementering af Data Mining som en del af kontrolmiljøet

Fordele	Ulemper
+ Forebyggende kontroller: Data Mining kan indbygges som forebyggende	- Kræver nye kompetencer. Programmering, data science, modellering og matematik er ikke i overflod.
+ Forebyggende kontroller kan mindske Pay & Chase og dermed begrænse tab	- Kræver et skift i mindsæt. Tør virksomhederne at bero sig på matematik, statestik og algoritmer. Eller foretages det på toppen af det eksiterende kontrolmiljø.
+ Mere effektivt til opdagelse af besvigelser end klassiske kontroller. Kan teste alle posteringer og er ikke begrænset af menneskelig viden i form af regelopbygning.	- Vil kræve en væsentlig upfront investering. Der skal ansættes nye kompetancer, der skal investeres i software og analyser skal designes, testes og implementeres
+ Tidssvarende og kan håndtere omskifteligheden i besvigelser	- Reele besvigelsestab kender ikke, hvorfor businesscase kan være svær at forsvarere (det sker ikke for os mentalitet)
+ Omkostningseffektivit når implementeret, kan frigive mandskabstimer til risikovurderingsarbejde eller opsætning af nye kontroller	- Kræver kontinuerlig udvikling og kompetenceløft. Data Mining vil løbende kræve revurdering og optimering. Foretages de rigtige analyser
+ Hvis virksomheder efter ACFE hypotese taber 5 % af deres omsætning årligt, er der væsentlige ressourcer til at retfærdiggøre investeringen	- Kræver trial and error, der findes ikke best practice. Det vil derfor kræve mange forsøg og test, før en at ledelsen vil have tiltro til teknikkerne.
+ Teknologien kan anvendes på mange områder og skabe værdi på tværs af virksomheden	- Fejlagtig opsætning kan medføre falsk negativer og falsk positiver, hvilket kan betyder at besvigelser ikke fanges eller at der vil igangsættes store undersøgelser på baggrund af fejl.
+ Kan udbygges i takt med risikobilledet eller det tekniske muligheder ændrer sig. Er først skridt mod højere grad af indsigt og automatisering.	