

A young boy with brown hair is smiling and looking down at a red McDonald's Happy Meal toy he is holding with both hands. He is also holding a McDonald's drink cup with a straw. The background is a blurred indoor setting, likely a restaurant.

Analyse af kundetilfredshedsundersøgelser via kumulative logit-modeller indeholdende data fra McDonald's

Af: Morten Schou Jørgensen
Afleveret: 13. December 2011
Vejleder: Hans Kurt Kvist

Copenhagen Business School
Erhvervsøkonomi/Matematik-studiet, cand.merc.(mat.)
Institut for Finansiering, Center for Statistik
Kandidatafhandling

Executive summary

Throughout the last decade there has been an increased interest in the topic of customer satisfaction. Although it has been a well known word for most leaders in the service industry for several decades, it is first recently it has become a buzzword in corporations.

The competition in all markets has increased, and it is no longer enough to have a good and solid product in the portfolio. In the service industry, and especially in the restaurant industry, the supply of a meal is high, and the quality of the food is no longer a competitive edge. Today, the consumers want something extra, whether it is an exceptional service or a nice place to enjoy the meal, the restaurant needs to perform well on these parameters. If not, the consumer will find a different place to buy the food next time.

This report will analyze the customer satisfaction at McDonald's in Denmark. The first part of the report will give a brief introduction into the restaurant market in Denmark, then it will look at the definition of customer satisfaction and finally look into a theory, which justifies the increased interest into the topic.

The second part of the report will analyze a customer satisfaction survey, which has been made at McDonald's in Denmark in autumn 2010. I have data available from 14 different restaurants, and these will make the background of the analysis. Are the satisfaction level the same at all restaurants, or are there any differences between them? Is it possible to explain these differences and would it be possible to develop a model that can explain the customer satisfaction? These are some of the questions that will be answered in this report.

Indholdsfortegnelse

1.0 INDLEDNING	3
<i>1.1 BAGGRUND FOR PROBLEMSTILLING</i>	<i>3</i>
2.0 PROBLEMFORMULERING	5
3.0 AFGRÆNSNING	7
4.0 BAGGRUND	8
<i>4.1 RESTAURATIONSBRANCHEN</i>	<i>8</i>
<i>3.2 McDONALD'S I DANMARK</i>	<i>14</i>
<i>3.3 DEFINITION AF KUNDETILFREDSHED</i>	<i>15</i>
<i>3.4 HVORFOR KUNDETILFREDSHED ER EN VIGTIG PARAMETER</i>	<i>16</i>
5.0 DATAINDSAMLING	21
<i>5.1 ANALYSEMETODER</i>	<i>21</i>
<i>5.2 SPØRGESKEMADESIGN</i>	<i>22</i>
<i>5.3 VALIDITET OG RELIABILITET</i>	<i>24</i>
<i>5.4 MISSING VALUES</i>	<i>25</i>
<i>5.5 SPØRGESKEMAETS SPØRGSMÅL</i>	<i>26</i>
6.0 DATASÆTTETS OPBYGNING	28
7.0 DATAGENNEMGANG	34
<i>7.1 HISTOGRAMMER, ALLE RESTAURANTER SAMLET</i>	<i>36</i>
7.1.1 KØN	38
7.1.2 ALDERSFORDELING	38
7.1.3 TIDSRUM	39
7.1.4 UGEDAG	40
7.1.5 BESØGSFREKVENNS	40
7.1.6 BESØGSSTED	41
7.1.7 FAMILIEFORHOLD	41
7.1.8 PRIORITERING	42
<i>7.2 HISTOGRAMMER, BRØNDBY</i>	<i>45</i>
7.2.1 KØN	46
7.2.2 ALDERSFORDELING	47
7.2.3 TIDSRUM	47
7.2.4 UGEDAG	48
7.2.5 BESØGSFREKVENNS	49
7.2.6 BESØGSSTED	49
7.2.7 FAMILIEFORHOLD	50
7.2.8 PRIORITERING	50
7.6 KRYDSTABELLER	52
7.7 KRYDSTABELLER FOR McDONALDS	52
7.7.1 ALDER OG TILFREDSHED/VIGTIG	52

7.7.2 KØN OG TILFREDSHED/VIGTIGT	54
7.7.3 BESØGSFREKVENNS OG TILFREDSHED/VIGTIG	55
7.8 KRYDSTABELLER FOR McDONALD'S BRØNDBY	56
7.8.3 TILFREDSHED OG BESØGSFREKVENNS	57
7.8.4 TIDSRUM OG TILFREDSHED:	58
7.8.5 BØRN OG TILFREDSHED:	58
7.8.6 PLACERING OG TILFREDSHED	59
8.0 TEORI	61
8.1. KONTINGENSTABELLER	61
8.2 ODDS RATIO	62
8.3 TEST FOR UAFHÆNGIGHED	63
8.4 LOGLINEÆRE MODELLER	64
8.5 LOGISTISK REGRESSION	66
8.6 DUMMYVARIABLE	67
8.7 MULTIPEL LOGISTISK REGRESSION	67
8.8 KUMMULATIVE LOGIT MODELLER	68
8.9 PROPORTIONAL ODDS-MODELLEN	69
8.9.1 ODDS RATIO	71
8.9.2 KORRELATIONSKOEFFICIENTER	71
9.0 SAS-PROGRAMMET (DATABASEHANDLING)	73
10.0 ANALYSE AF TILFREDSHEDEN: MODELLER OG RESULTATER	75
10. 1 PROPORTIONAL ODDS MODELLER FOR TILFREDSHED	75
10.2 McDONALD'S DANMARK	75
10.3 McDONALD'S BRØNDBY	88
PARAMETERESTIMATER	93
ODDS RATIO ESTIMATER	94
11.0 KONKLUSION	97
12.0 PERSPEKTIVERING	98
14.0 BILAG	100

1.0 Indledning

1.1 Baggrund for problemstilling

Kundetilfredshed er ikke noget nyt koncept. Faktisk har det mange år på bagen – det er mindst 200 år gammelt. Så langt tilbage som i det 18. Århundrede klarlagde Adam Smith de fundamentale præmisser, på hvilket det frie marked opererer. I bogen ”The Wealth of Nations”, som blev udgivet i 1776, fastlagde han, at idet mennesket konstant higer efter at maksimere deres nytte (utility) – opnå de bedste fordele for den mindste omkostning – flytter de gradvist men ubønhørligt, til udbydere, der kommer tættest på at opfylde dette. Det vil altså med andre ord sige, at kunderne søger efter virksomheder, som performer bedst på det der betyder mest for dem. Kundetilfredshed er det ord, der ofte bruges til at indkapsle dette fænomen.

Med denne betragtning in mente, er det ikke så underligt, at kundetilfredshed er et område, der igennem de seneste år har fået meget opmærksomhed i virksomhederne.

McDonald’s opererer i restaurationsbranchen, som er kendetegnet ved at have mange udbydere. Der er stor konkurrence i markedet omkring kunderne og det er ikke småpenge man konkurrerer om i denne branche, hvorfor en analyse af kundetilfredsheden hos McDonald’s er mere end berettiget.

Hvor restaurationsvirksomheder tidligere anså kunderne som en uudtømmelig kilde, er der i dag et stort udbud af spisesteder, hvilket har ført til en benhård konkurrence om kunderne. I dag er kundernes krav til restauranterne blevet større – hvor man tidligere anså det at få serveret noget spiseligt som det vigtigste, er der i dag flere andre parametre, som også spiller en stor rolle for kunden. Alt dette har kunne ladet sig gøre på grund af det store udbud af spisesteder.

En analyse af kundetilfredsheden kan være med til at give et overblik over kundernes syn på virksomheden. Det er dog vigtigt at holde sig for øje, at en kundetilfredshedsundersøgelse bygger på historiske data og dermed fortæller om virksomhedens tidligere performance. Der

kan sagtens opstilles en model for den fremtidige kundetilfredshed, men man kan ikke være sikker på, at den predikterer fuldstændig korrekt. Den vil dog simulere fremtiden bedre, end hvis man slet ikke havde noget data at bygge på.

2.0 Problemformulering

Jeg vil med denne opgave foretage en analyse af gæstetilfredsheden hos McDonald's.

På baggrund af data fra en kundetilfredshedsundersøgelse i 2010 vil jeg undersøge om tilfredsheden er den samme i de udvalgte restauranter.

Min analyse tager udgangspunkt i spørgeskemaet som er gengivet nedenfor:

Din mening om dit besøg på McDonald's® i dag
Vi arbejder hele tiden på at gøre dit besøg hos os til en god og positiv oplevelse. Vi vil derfor bede dig bruge et par minutter på at besvare nedenstående spørgsmål om dette besøg. Når du afleverer det udfyldte skema i denne restaurant inden udløbsdato, får du en gratis Hamburger eller Cheeseburger som tak for hjælpen.

 i'm lovin' it®

DK00001

Besvar ved at udfylde hele cirklen:

Overhovedet ikke tilfreds

Særlig tilfreds

Vesterbrogade

Se spm. 12
Mest vigtigt

1. Er du overordnet tilfreds med dit besøg i dag?.....

Hvor tilfreds er du med:

2. Tempoet af betjeningen før du fik bestilt.

3. Tempoet af betjeningen efter du fik bestilt.

4. Høflig og venlig betjening.

5. Renligheden (i restaurant el. McDrive).

6. Madens temperatur.

7. Smagen af maden.

8. Friskheden af maden.

9. Min ordre var korrekt. Nej ☐ Ja ☐

Kun hvis du har brugt restauranten...

10. Atmosfæren i restauranten.

Kun hvis du har brugt McDrive...

11. Kommunikationen med vores medarbejder.

Spm. 12 - 25 besvares af alle...

12. Hvis du ser tilbage på spørgsmål 2 - 11, hvad er så mest vigtigt for dig (udfyld kun én cirkel). →

13. I hvilken grad står helhedsindtrykket af besøget mål med prisen.

14. Hvordan vurderer du helhedsindtrykket hos McDonald's i forhold til lignende steder.

15. Har dit besøg i dag været en special god/sjov oplevelse? ... Ja ☐ Nej ☐ (Gå til spm. 17)

16. Hvis ja - Hvad var det, der var specielt godt ved dette besøg? (Fortsæt evt. på bagsiden)

17. Jeg spiser på McDonald's (antal gange pr. måned): ☐ 4+ ☐ 1 - 3 ☐ Sjældnere

18. Ved dette besøg brugte jeg: ☐ Restauranten ☐ McDrive

19. Køn: ☐ Mand ☐ Kvinde

20. Har du børn under 12 år med på dette besøg? ☐ Ja ☐ Nej

21. Alder: ☐ 13 - 19 år ☐ 20 - 29 år ☐ 30 - 39 år ☐ 40 år +

22. Jeg besøgte restauranten: ☐ åbn. - 14.00 ☐ 14.00 - 17.00 ☐ 17.00 - 20.00 ☐ 20.00 - luk

23. I dag er det: ☐ Tor ☐ Fre ☐ Lør ☐ Søn

24. Hvilket postnummer bor du i?

25. Eventuelle kommentarer/forslag: (Du er velkommen til at bruge bagsiden hvis du ønsker mere plads).

©Copyright 2010, Industrial Relations International, Inc. All Rights Reserved. 972-231-8817, DI03

Skal afleveres senest lørdag den 21. november 2010

Det grundlæggende datamateriale er udleveret af IRI Consulting, der har bearbejdet de udfyldte spørgeskemaer og med tilladelse fra McDonald's Danmark. Datamaterialet er

indsamlet fra 14 forskellige restauranter fordelt rundt om i landet og er indsamlet som udfyldte spørgeskemaer. De 14 forskellige restauranter som alle er med i analysen er placeret i følgende byer: Viby, Viborg, Hjørring, Vejle, Svendborg, Frederiksværk, Thisted, Brøndby, Nyborg, Esbjerg, Hillerød, Køge, Valby og Kastrup.

For at forstå, hvorfor en sådan analyse er interessant, vil jeg først beskrive restaurationsbranchen – hvad er der karakteristisk for denne branche, og hvordan opererer de forskellige agenter på dette marked, hvorefter jeg zoomer ind på McDonald's og ser hvilke tiltag de gør for at have en høj kundetilfredshed.

For at kunne afdække eventuelle niveauforskelle i tilfredsheden for de forskellige kunder opstilles generaliserede lineære modeller med tilfredshed som responsvariabel og en række demografiske variable som forklarende variable.

Hvis det viser sig, at der er forskel på tilfredsheden, er målet herefter at kunne beskrive denne forskel. Kan niveauforskellen forklares ud fra de oplysninger jeg har til rådighed?

3.0 Afgrænsning

For at kunne give en dybdegående besvarelse af problemformuleringen, er det nødvendigt med en del afgrænsninger.

Der findes et stort udvalg af teorier omkring kundetilfredsheds. Jeg har i denne opgave valgt at tage en af dem ud og gennemgå den, da den passer godt ind i det jeg har valgt at fokusere på. Man skal bare være opmærksom på, at der findes andre teorier, der kommer med andre bud på kundetilfredshed. Derudover har jeg kun taget en lille brøkdel af de definitioner, der findes omkring tilfredshed og kundetilfredshed. Der ligger ikke nogen større videnskab bag mit udvalg af definitioner, min pointe er blot, at der netop findes mange tolkninger af dette begreb.

Jeg har i opgaven valgt at kigge på tilfredsheden for alle restauranter samlet set, men også at se på tilfredsheden for en enkelt restaurant. At valget er faldet på restauranten i Brøndby skyldes udelukkende at det er min arbejdsplads, og derfor kunne det være interessant for mig at grave lidt dybere ned i tilfredsheden for denne restaurant. Det kunne ligeså godt have været for en af de andre restauranter jeg havde valgt ud.

I afsnittet med datagennemgangen har jeg valgt at gennemgå histogrammer og jeg ser også på diverse krydstabeller. Jeg har kun taget et lille udpluk med af alle de histogrammer og krydstabeller, der kunne laves. Jeg har valgt dem ud, som jeg synes tjener opgavens formål bedst muligt.

I analyseafsnittet har jeg valgt at se bort fra vekselvirkninger i min model til beskrivelse af tilfredsheden. Dette skyldes til dels at det vil blive alt for omfattende at tage disse vekselvirkninger med, men tolkningen heraf, når man kommer op på en vekselvirkning med tre variable eller flere, er også meget vanskelig.

Derudover har jeg også valgt ikke at se på de data, der findes for spørgsmål 2-11 i spørgeskemaet i mit analyseafsnit. Det var en mulighed at se, om der også kunne opskrives en model for disse tilfredshedsparametre, men igen vil det komme ud på et omfang, der langt overstiger denne opgaves begrænsninger.

4.0 Baggrund

Jeg vil i dette kapitel give læseren et indblik i det marked, hvorpå McDonald's opererer. Først vil jeg se på hele restaurationsbranchen – en kort beskrivelse af branchen og herunder se på nogle nøgletal, hvorefter jeg vil fortælle lidt om McDonald's, særligt med henblik på, hvilke initiativer de gør for at forøge kundetilfredsheden. Kapitlet rundes af med en teori, der retfærdiggør den øgede fokus på kundetilfredsheden fra virksomhedens side.

4.1 Restaurationsbranchen

Når man ser på restaurationsbranchen i dag, er det kendetegnet ved et stort udbud af steder, hvor man kan få stillet sin sult. Udbuddet af restauranter er stort, og der findes noget for enhver smag og pris. Spektret rækker lige fra den to-stjernede Michelin-restaurant Noma, som har den nykårede verdensmester i VM for kokke, Rasmus Kofoed, ansat, over til den anden yderpol, hvor den lokale bistro, som ofte findes i forbindelse med supermarkedet, befinder sig.

I Danmark har det samlede antal restaurantvirksomheder de senere år ligget stabilt på lidt over 13.000, men dette dækker over at der årligt åbner og lukker op imod 2.500 virksomheder¹. Derudover er der sket en forskydning mellem de forskellige virksomhedskategorier², således at antallet af virksomheder i kategorierne "event catering" og "caféer mv." har været stigende, mens der er blevet lidt færre fastfood-restauranter og virksomheder i kategorien "anden restaurant-virksomhed".

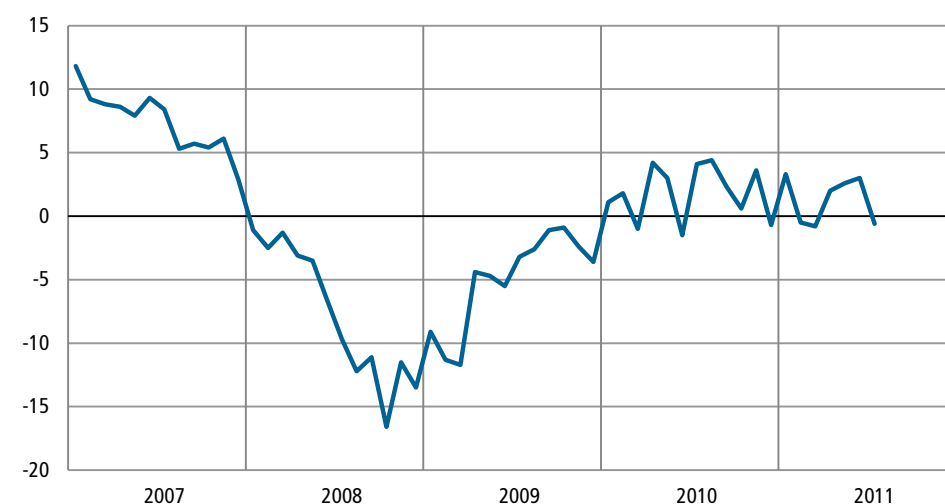
Restaurationsbranchen er en stærk konjunkturfølsom branche – der sælges produkter og ydelser, som både private gæster og forretningskunder relativt let kan skære ned på, hvis der skal spares. Det betyder, at omsætningen på restauranterne følger den generelle udvikling i danskernes økonomi. Dette har branchen også kunne mærke efter finanskrisen og den efterfølgende økonomiske afmatning, som har haft en stor negativ indvirkning på efterspørgslen i restauranterhvervet. Som situationen ser ud i dag, med Grækenland, der er på randen af bankerot og et Italien som heller ikke har det for godt økonomisk, tyder det på at der er lange udsigter til en opblomstring i branchen. Dog viser de allerseneste tal fra

¹ Kilde www.cvr.dk

² Restauranterhvervet er inddelt i følgende fem kategorier: Restauranter, Pizzeriaer, grillbarer, isbarer mv., Event catering, Anden restaurationsvirksomhed og Caféer, værtshuse, diskoteker mv.

Danmarks Statistik, at forbrugertilliden fortsat er stabil³. Forbrugertillidsindikatoren var i juli -0,6 og ligger dermed fortsat på et stabilt niveau. I de seneste seks måneder har niveauet svinget mellem -0,8 og 3,0 med et gennemsnit på 1,0. Indikatoren belyser befolkningens syn på den aktuelle og den fremtidige økonomiske situation.

Forbrugertillidsindikatoren



Umiddelbart kan det virke underligt, at andre landes økonomiske situation har en påvirkning på det danske restaurationsmarked, men her spiller psykologien et pus. Kunderne holder på deres penge og sparer op i tilfælde af, at en finansiell krise også skal skylle ind over Danmark og ramme netop dem. Dette er med til at starte en negativ spiral, da økonomien kunne sættes i gang igen ved at kunderne øger deres forbrug. Herved ville man kunne få gang i hjulene igen i den danske økonomi, men forbrugerne vælger i stedet at spare op til dårligere tider, som kun rykker nærmere ved netop denne handling. Der ligger en stor opgave hos politikerne og den finansielle sektor for at oprette forbrugernes tillid til markedet igen.

Omsætning

Efter en årrække med højkonjunktur og stigende omsætning, faldt omsætningen i restauranterhvervet med 5,3 pct. fra 2008 til 2009. Som det fremgår af nedenstående tabel blev restaurantkategorierne ikke ligeligt påvirket af den økonomiske krise. Det største

³ NYT fra Danmarks Statistik, nr. 338 20. Juli 2011

omsætningsfald skete i kategorien 'anden restaurantvirksomhed', som er en kategori, der bl.a. dækker over kantiner og selskabslokaler. Virksomhederne i denne kategori tabte i 2009 10,7 pct. af omsætningen sammenlignet med året før. Blandt de traditionelle restauranter og fastfood-restauranterne var omsætningsnedgangen på henholdsvis 5,3 og 5,4 pct.

Kategorien 'cafeer, værtshuse, diskoteker mv.' er den eneste, som i hvert fald omsætningsmæssigt, ikke er blevet synderligt påvirket af den generelle nedgang i økonomien, idet omsætningen blot faldt med 0,5 pct. sammenlignet med 2008.

År	2001	2002	2003	2004	2005	2006	2007	2008	2009	Ændring i pct. (08-09)
Restauranter	9.340	9.404	8.215	8.817	9.844	10.899	12.504	13.236	12.536	-5,3
Pizzariaer, grillbarer, isbarer mv.	5.675	5.838	5.905	5.968	6.192	6.415	6.641	6.665	6.302	-5,4
Event catering	1.709	1.170	1.374	1.565	1.695	1.807	1.855	1.868	1.794	-4,0
Anden restaurationsvirksomhed	2.398	2.919	2.939	3.013	3.423	3.531	3.879	4.262	3.805	-10,7
Caféer, værtshuse, diskoteker mv.	3.201	2.984	4.088	4.171	4.327	4.564	4.615	4.530	4.506	-0,5
I alt	22.323	22.315	22.521	23.534	25.481	27.216	29.494	30.561	28.943	-5,3

Tabel 1 - Restauranterhvervets omsætningsfordeling i mio. kr.⁴

Danskernes udespisevaner

Den faldende omsætning skyldes dels, at danskerne går mindre ud og spiser, og dels at gennemsnitsforbruget pr. besøg er faldet. I 2008 foretog danskerne 241 mio. restaurantbesøg på tværs af de forskellige restauranttyper, men i 2009 faldt antallet af besøg til 234 mio. besøg, svarende til et fald på 2,9 pct.

Færre restaurantbesøg kan således ikke forklare hele omsætningstabt, hvilket betyder at omsætningen pr. restaurantbesøg i 2009 for første gang i mange år er faldet. Gæsterne bruger ganske enkelt færre penge, når de går ud og spiser. Dette kan eksempelvis være sket ved, at de har købt færre eller billigere drikkevarer eller i højere grad fravælger forretten eller desserten. Nogle gæster vælger en billigere restaurant end tidligere.

En del af forklaringen på det faldende forbrug pr. restaurantbesøg er også, at erhvervs-virksomhederne generelt har skåret kraftigt ned på deres repræsentationsbudget, og mange forretningsmøder er flyttet tilbage i egne mødelokaler. Restaurantgæster, som går på

⁴ Normtalsanalyse 2009/10 – Horesta s. 105

restaurant i erhvervsøjemed er typisk kendetegnet ved et højt forbrug. Når de reducerer antallet af restaurantbesøg mere end de private restaurantgæster, medfører dette også et fald i den gennemsnitlige omsætning pr. restaurantbesøg.

Det er ikke alle restauranttyper, som har haft færre gæster i 2009. Størst tilbagegang ses blandt pizza- og shawarmabarerne, som har oplevet et fald i besøgstallet på hele 9,1 pct. i forhold til 2008. De traditionelle restauranter, som står for ca. halvdelen af den samlede restaurantomsætning, har haft en tilbagegang i besøgstallet på 4,1 pct., mens cafeerne – som de eneste – har øget deres besøgstal, om end fremgangen udgjorde beskedne 1,3 pct. Besøgstillene understøtter således omsætningstillene, hvor cafeerne også er den virksomhedskategori, som har haft det mindste omsætningsfald i 2009.

De senere års udvikling, hvor fastfood-restauranterne har tabt markedsandele til cafeer og restauranter ser ud til at være fortsat i 2009. Dette tyder på, at det ikke kun er den økonomiske afmatning, som har indflydelse på danskernes restaurantbesøg, men også at der er tale om en generel ændring i forbrugeradfærd, hvor besøg på fastfood-restauranter i stigende grad bliver fravalgt. Det skal dog pointeres, at det primært er pizzeriaerne, som har oplevet en kraftig nedgang, mens kategorierne burgerbarer/burgerkæder samt grillbarer har mærket et væsentligt mindre fald i besøgstallet fra 2008-2009.

Faldet i det samlede antal fastfoodbesøg kan hænge sammen med, at de yngste aldersgrupper går væsentligt mindre på restaurant end tidligere. Antallet af restaurantbesøg pr. person pr. år blandt de 12-19-årige er således faldet med 29,1 pct. Der har ligeledes været et mindre fald i antal restaurantbesøg foretaget af de 20-29-årige, hvorimod alle aldersgrupper over 30 år faktisk går hyppigere på restaurant end i 2001. Antallet af restaurantbesøg varierer også i høj grad mellem landsdelene. København By er den landsdel, hvor befolkningen går mest ud og spiser. I København By aflægger hver person gennemsnitligt 73 restaurantbesøg om året. Dette tal inkluderer både besøg på traditionelle restauranter, cafeer, værtshuse og fastfood-restauranter. Udover København By ligger Fyn og Københavns Omegn en smule over landsgennemsnittet med henholdsvis 53 og 52 restaurantbesøg pr. år. Årsagen til at netop disse landsdele ligger over gennemsnittet er, at der bor relativt flere unge, og det er dem, der hyppigst går ud og spiser.

Dette skal også i høj grad ses i lyset af danskernes udespisevaner. Finanskrisen har haft sin indvirkning på danskernes tilrådighedsbeløb, det er for manges vedkommende skrumpet en hel del ind, og dette har påvirket købekraften hos kunderne.

Tager man fat i den seneste økonomiske opgørelse fra hotel- og restaurationsbranchen (HORESTA), kan man få et indblik i den økonomiske udvikling der har været i restauranterhvervet.

Faktum er, at danskerne årligt bruger i omegnen af 28 milliarder til at spise ude (se tabel 1), og hvis man som virksom kan få en lille del af denne kage, vil det ikke være dårligt.

Kæderestauranter

I de seneste år er der kommet flere kæderestaurantkoncepter i Danmark, ligesom de fleste restauranter har ekspanderet med flere enheder. Kæderne har således øget deres markedsandel og denne tendens lader til at fortsætte i de kommende år. Kæderestauranterne hører til de største restauranter på det danske marked. Trods økonomisk krise og generel omsætningsnedgang på det danske restaurationsmarked, har kæderestauranterne fastholdt omsætningsniveauet. Der er intens kamp om kunderne og lancering af attraktive tilbud på det danske fastfoodmarked.⁵

En gennemsnitlig kæderestaurantfilial i årets Normtalsanalyse opnår en omsætning på 14,6 mio. kr., og dermed hører kæderestauranterne til de største restauranter på det danske marked.

McDonald's er markedsleder både på det danske marked, men også på globalt plan. Den største konkurrent på det danske marked er Burger King, der kom som de første på markedet i 1977, men de blev hurtigt overhalet af McDonald's som rykkede ind på det danske marked i 1981.

⁵ Normtalsanalyse 2009/10 – Horesta

I dag har kunderne et hav af muligheder hvis de ønsker at spise fast food. Hvor markedet tidligere var kendetegnet ved 2 store kæder, McDonald's og Burger King, er mulighederne i dag mangedoblet, og hvor udvalget tidligere kun var "usunde" burgere, er der kommet mange sundere alternativer til. Både i form af nye restauranter, men også en udvikling af de eksisterende restauraners sortiment.

Det nyeste koncept indenfor kæderestauranter er en modernisering af den traditionelle danske grillbar. Under navnet "Tulip TimeOut" tilbydes danskerne baguette, burger og pølser i omgivelser, hvor den personlige betjening, hygiejne og kvalitetsprodukter er i fokus. Også en modernisering af den traditionsrige pølsevogn er kommet på markedet under navnet "Steff's Place", for at tage konkurrencen op på det lukrative fast-food marked⁶.

⁶ Dansk Handelsblad – 25. Februar 2011.

3.2 McDonald's i Danmark

McDonald's fik sin entre på det danske marked tilbage i 1981 og var en af de første restauranter som i dag er kendte under konceptet fast food-restauranter. Danskerne tog rigtig godt imod dette nye koncept og i dag findes der 83 restauranter rundt om i Danmark, de fleste er af typen McDrive.

McDonald's er markedsleder både på det danske marked, men også på globalt plan. Den største konkurrent på det danske marked er Burger King, der kom som de første på markedet i 1977, men de blev hurtigt overhalet af McDonald's som rykkede ind på det danske marked i 1981.

McDonald's restauranter i Danmark havde i 2010 en samlet omsætning på 1.6 mia. kr. ⁷

Hvis man skal finde en forklaring på, hvorfor det går McDonald's så godt, er der mange faktorer der gør sig gældende. Én af dem er den megen fokus der bliver knyttet til kundetilfredshed. Lige siden Ray Kroc blev ansvarlig for McDonald's har der været en stor interesse for kunderne og dette har været med til at udvikle McDonald's til den type restaurant den er i dag. Der bliver lyttet til kunderne og der tages hensyn til deres synspunkter i et omfang det er muligt. Senest er der introduceret den nye produktionslinie – Made For You – hvor gæstens mad først bliver tilberedt i det øjeblik den bliver bestilt. Dette er sket som et led i at højne kvaliteten af produktet.

Når man taler om kundetilfredshed hos McDonald's er der flere parametre man kan måle på. Jeg har i denne opgave valgt at sætte mit fokus på den årlige KTU (kundetilfredshedsundersøgelse) der finder sted i alle restauranter, hvor der bliver udleveret et skema til alle gæster der besøger restauranten i den pågældende periode. Et andet måleredskab man kunne benytte hertil var at se på antallet af kundeklager, mens antallet af gæster (transaktioner) og omsætningen kan sammenlignes med samme periode sidste år for at give en indikation om man har at gøre med en sund forretning.

Et relativt nyt koncept i kundetilfredshedsundersøgelser – Mystery Shopping – har også fundet sin vej til McDonald's. Hver restaurant bliver besøgt to gange om måneden af en sådan shopper, som evaluere besøget efter nogle parametre, som man er kommet frem til er de vigtigste for at få en god oplevelse som kunde.

⁷ Årsregnskabet for McDonald's i 2010, fundet på deres hjemmeside www.mcdonalds.dk

3.3 Definition af kundetilfredshed

I dette afsnit vil jeg definere kundetilfredshed, som er hele omdrejningspunktet for denne opgave og efterfølgende have en diskussion af begrebet.

At være tilfreds, ved vi alle hvad det vil sige at være, men skal man give en definition af selve begrebet, vil der komme mange forskellige og interessante bud herpå. Dette skyldes at vi hver især forbinder ordet med nogle forskellige værdier, og det er med til at forvirre debatten når man taler om tilfredshed. Det er derfor nødvendigt at man har samme ståsted – den samme referenceramme – når man taler om tilfredshed, da man ellers i realiteten kan tale om to forskellige ting.

Philip Kotler definerer tilfredshed som:

"A person's feelings of pleasure or disappointment resulting from comparing a product's perceived performance (or outcome) in relation to his or her expectation".

Et andet eksempel på en definition af tilfredshed er Oliver (1997) som benytter følgende:

"Satisfaction is the consumer's fulfillment response. It is a judgement that a product or service feature, or the product or service itself, provided (or is providing) a pleasurable level of consumption-related fulfillment, including levels of under- or overfulfillment."

Hermed fokuserer han på, at tilfredshed er et resultat af, om produktet lever op til forbrugerens forventninger. Johnson et al. (1995) argumenterer for kundetilfredshed:

"Customer satisfaction is a cumulative construct that is affected by market expectations and performance perceptions in any given period, and is also affected by past satisfaction from period to period."

Denne definition fokuserer på, at kundetilfredshed dannes af en instrumental konstruktion, som består af kundens forventninger og produktets præstation. Kundetilfredsheden er således basalt en efterkøbsevaluering. Dette understøttes af Fornell (1992), der definerer

kundetilfredshed:

“Customer satisfaction is expressed as a function of prepurchase expectations and postpurchase perceived performance.”

Ud fra dette kan det konstateres, at kundetilfredshed er en funktion af kundens forventninger til produktet før købet samt kundens opfattelse af produktet efter købet⁸.

Når man taler om kundetilfredshed mener jeg specielt at der er to ting man skal lægge mærke til, hvilket ovenstående definition også indikerer. Dette er den opfattede kvalitet og forventningerne. Den opfattede kvalitet hænger nøje sammen med produktets egenskaber sammenholdt med kundens forventninger. Tilfredshed eller utilfredshed kan altså opstå, hvis der er et negativt forhold mellem disse to begreber.

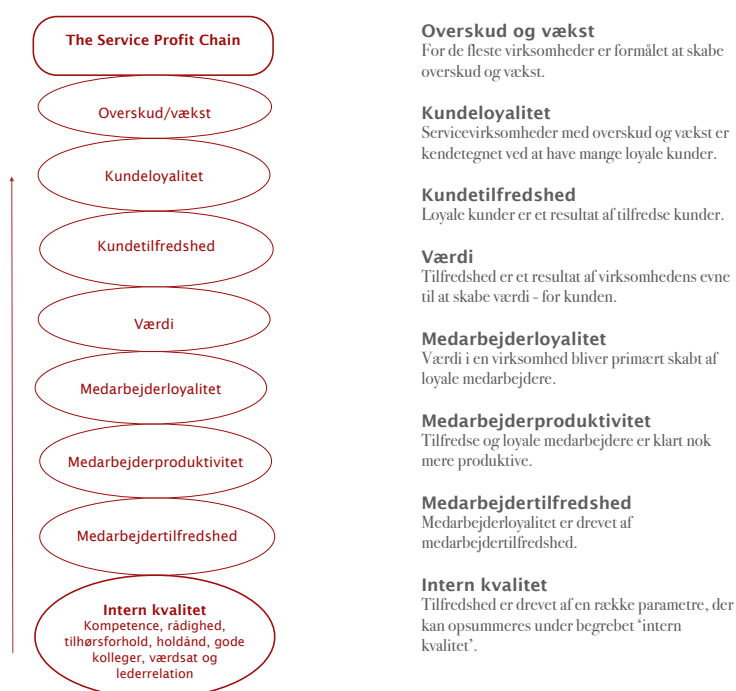
Som man kan se af ovenstående, findes der mange forskellige definitioner af tilfredshed, og dette er blot et lille udsnit blandt mange. Jeg har ikke valgt at ligge mig fast på en definition til brug i denne opgave, men definitionerne er medtaget for at illustrerer, at tilfredshed kan opfattes forskelligt. Hvad én person opfatter som meget tilfredsstillende, kan være normen for en anden!

3.4 Hvorfor kundetilfredshed er en vigtig parameter

Hvis man skal finde en forklaring på, hvorfor kundetilfredshed er blevet en parameter virksomhederne bruger meget tid og energi på, kan man blandt andet hvide fat i service værdikæde-teorien. Teorien om The Service Profit Chain (service værdikæden) er udviklet af forskere fra Harvard University i midten af 1990'erne og går kort fortalt ud på, at profit og vækst i bund og grund afhænger af virksomhedernes interne kvalitet og værdiskabelse for egne medarbejdere. Det vil sige, at tilfredse og loyale medarbejdere, skaber tilfredse og loyale kunder. Forskerne har påvist, at der er en klokkeklar årsagskæde – en trinvis proces – som fører til vækst og ekstraordinære resultater. Hele filosofien bag service værdikæden kan illustreres som i modellen herunder: Vejen til det bedste resultat starter med at skabe den bedst mulige interne kvalitet, som primært opnås ved godt lederskab.

⁸ Kristensen; Martensen; Grønholdt 1998, s. 9

Fra toppen ser årsagskæden sådan ud:



Når et marked har for meget af alt – og det er situationen for de fleste serviceudbydere i dag – så bliver det købers marked. Ny medieteknologi har gjort det nemt at komme til orde i form af markedsføringsbudskaber af den ene eller den anden art, men det er i stigende grad svært at blive hørt. Virksomhederne forsøger ihærdigt at overbevise kunderne om, at lige netop deres produkt er bedre end alle de andre. Men det budskab har man hørt så mange gange nu, at der lukkes af for støjen. Derimod vender man sig imod noget man kender og ved at det fungerer. Hvis man skulle vove sig ud i nye oplevelser søger man gerne råd fra nogen man kender og hører deres anbefalinger. Det er "word of mouth" der er med til at afgøre kundens beslutningsproces. I den forbindelse er loyale kunder væsentligt mere tilbøjelige til at fortælle andre om deres loyalitet end kunder, der blot er tilfredse. Begeistrede kunder fortæller andre om deres oplevelser og skaber ambassadører for virksomheden. De bliver loyale kunder, og de kommer igen og igen. At tiltrække nye kunder, er derfor ikke længere et spørgsmål om at lave mere larm end alle de andre, men at begeistrede brugere fortæller andre om deres oplevelse. Det medfører et skift i fokus fra at være "larmende" i jagten på kunder til at udvikle lidenskabelige ambassadører.

Der er to ting omkring tilfredshed, der er vigtige at holde sig for øje. En kunde, der er tilfreds, er ikke automatisk en loyal kunde. Det er kun *super* tilfredse kunder, der bliver loyale. Derfor

er "tilfreds" ikke godt nok i en verden med for meget af alt. Dernæst er den barske sandhed, at det, der sidste år var fantastisk, jo blot er det forventede niveau i år. Kundens forventninger flytter sig hele tiden. Hvis man vil fastholde loyaliteten, skal man blive bedre og bedre.

Måden hvormed man kommer fra tilfredse kunder til begejstrede kunder er ved at skabe værdi for kunden, som næsten altid har en emotionel komponent, der gør en oplevelse specielt mindeværdig for kunden.

Værdi

Den primære faktor af kundetilfredshed er værdi – at kunden har fået mere for sine penge, end kunden selv synes, det er værd. Tilfredshed er derfor et dybt personligt anliggende og helt forskelligt fra person til person. Heri ligger én af nøglerne til at forstå, hvorfor medarbejder-loyalitet bliver så vigtigt. Hvis to produkter stort set kan det samme og koster det samme, bliver det, der gør forskellen, en emotionel komponent. Beder man en person om detaljeret at

beskrive en fantastisk serviceoplevelse, er det sjældent guldvandhaner, 48'tommer B&O fladskærme eller Børge Mogensen sofaerne, der bliver nævnt. Det er typisk noget emotionelt: Der blev gjort noget ekstra. De fandt en sjov løsning til mit barn, eller vi fik særlig opmærksomhed/omtanke/hensyn osv. Nøglen til at skabe værdi, ud over blot almindelig tilfredshed, er altså evnen til at knytte en emotionel forbindelse med kunden og skabe emotionelle wow oplevelser. Derfor burde det også være indlysende, at en servicemedarbejder, der er ked af at komme på arbejde, har væsentligt sværere ved at knytte en emotionel forbindelse til sin kunde end den medarbejder, der kommer på jobbet med armene over hovedet og en indstilling til, at det her bliver en dejlig dag. Værdi kan i servicevirksomheder sættes på en brøkstreg⁹:

$$V = \frac{R + P}{\$ + E} = \frac{\text{Resultat} + \text{Proces}}{\text{Pris} + \text{Egen indsats}}$$

⁹ Heskett, J., W. E. Sasser Jr., and L. Schlesinger. *The Service Profit Chain*. N.Y.: Free Press, 1997.

Produktive medarbejdere

Værdien er et direkte resultat af produktive medarbejdere. Her skal produktivitet ikke ses i forhold til hvor meget output, man kan producere på en time. Medarbejderens produktivitet i servicesammenhæng har noget at gøre med, i hvor høj grad medarbejderen forstår sit produkt, og i hvor høj grad han/hun forstår at oversætte kundens ønsker og behov til løsninger inden for de rammer. Produktive medarbejdere har både højt produktkendskab og højt kundekendskab. Derfor er medarbejdere, der har været længe i virksomheden, som regel meget mere "produktive" end en helt ny medarbejder. Det er altså ikke ligegyldigt, om man har en svingdørskultur på gulvniveau.

Loyale medarbejdere

For at kunne skabe den nødvendige begejstring hos kunden i form af den emotionelle oplevelse, er man som nævnt selv nødt til at være begejstret og være i stand til at møde kunden emotionelt. Det forudsætter, at medarbejderne i virksomheden er begejstrede for deres arbejde, for at have det nødvendige overskud til kunderne. For hvis man ikke er specielt glad for sit arbejde, kan det være meget svært at overbevise kunderne om det fantastiske ved virksomheden.

Loyale medarbejdere skal her både ses i den betydning, at de bliver længe, dvs. de får opbygget en stor viden om kunder, processer og virksomhedens servicekultur. Men det er også loyal i den betydning, at medarbejderens indstilling over for virksomheden er, at man passer på omkostninger, gør noget ekstra for kunden osv. Det er en form for loyalitet, der smitter direkte af på bundlinien (slukker man f.eks. lyset på lagret, når man går fredag eftermiddag, eller lader man det brænde til mandag og tænker: "Who cares?"). Det modsatte af en loyal medarbejder er den medarbejder, der svarer gæsten: "Don't ask me, I just work here..."

Tilfredse medarbejdere

Hvad driver en høj medarbejderloyalitet? Det gør logisk nok en høj grad af medarbejdertilfredshed. Men på samme måde, som kundeloyalitet drives af begejstring mere end blot tilfredshed, så drives medarbejderloyalitet på samme måde af begejstring og ikke blot af tilfredshed. Medarbejdertilfredshed er forholdsvis nemt at opnå. At opnå medarbejderens begejstring er derimod en meget sværere og meget mere kompleks opgave. I service værdikæden drives medarbejderbegejstring af begrebet "intern kvalitet", som igen dækker over en række delelementer, der alle skal være til stede for at sikre tilfredse medarbejdere.

Intern kvalitet

Intern kvalitet sikrer motivationen hos medarbejderne. Her er det vigtigt at understrege, at motivation både har noget med jobkontekst og jobindhold at gøre. Jobkonteksten er de ydre rammer – så som personalegoder, lønforhold, fysiske rammer m.m. Og det udgør faktisk ikke det væsentligste, når vi taler om motivation. Konteksten er snarere det, der sikrer, at medarbejderen ikke bliver demotiveret. Det siger sig selv, at hvis man har utilfredsstillende fysiske rammer, så kan man ikke gøre sig håb om at fastholde de bedste medarbejdere. Det er imidlertid jobindholdet, der er den afgørende faktor for motivation og medarbejder-tilfredshed: For motivation opstår først i det øjeblik, man bidrager med noget, som giver mening for os. Selvstændig beslutningskompetence, anerkendelse fra kolleger og ikke mindst ledelsen, ansvar, tillid og mulighed for faglig udvikling, er blandt de vigtigste topscorere, der giver tilfredse og motiverede medarbejdere. Det, der typisk frustrerer de rigtig gode servicemedarbejdere mest, er, når deres kompetencebegrænsninger er skruet sammen på en sådan måde, at de ikke er i stand til at yde den service, de mener, kunden bør have i en given situation.

Der er altså et samspil mellem kundetilfredsheden og virksomhedens overlevelse på markedet. Det er netop denne sammenhæng, som igennem det sidste årti er gået op for virksomhederne, og som dermed har betydet, at der er kommet en markant yderligere fokus på netop kundetilfredshed.

5.0 Dataindsamling

5.1 Analysemetoder

Når McDonalds står overfor at skulle måle kundetilfredsheden, er der flere forskellige metoder, der kan benyttes hertil. Alt efter, hvad formålet med undersøgelsen er, kan man benytte forskellige interview-metoder, både kvalitative og kvantitative. De kvalitative bygger ofte på personlige interviews, hvor man har én eller flere personer til samtale med en kvalificeret konsulent/moderator. Undersøgelsen kan foregå som fokusgruppe, hvor man ved at samle en gruppe personer (ca. 8-12 stk.) og præsentere dem for spørgsmål eller materiale, kan få nuanceret information på et bestemt område¹⁰. En anden mulighed er hukommelsesanalyse, hvor man får personer til at nedskrive deres erindringer i en bestemt situation, hvorefter den enkeltes historie diskuteres i forum. Derved kan man udnytte disse erindringer som ressource til teoriudvikling og forskning¹¹. Kvalitative undersøgelser kan ses som mere subjektive end kvantitative undersøgelser, som anses for at være objektive. Ved kvalitative studier kan man diskutere sig frem til konklusioner, og der behøver ikke være noget endeligt resultat. Kvalitative undersøgelser er dog sjældent repræsentative. De kvalitative studier kan sagtens bruges som selvstændige undersøgelser til f.eks. be- eller afkræftelse af hypoteser om trends, livsstil eller indstillinger til konkrete kampagnekoncepter, men de kan også bruges som baggrund for den kvantitative spørgeramme. Omvendt kan de også være med til at uddybe svarene fra en kvantitativ analyse.

Kvantitative analyser bruges til at fastsætte størrelsen eller omfanget af bestemte handlemåder, forudsætninger, indstillinger til varen osv. Undersøgelsen er ofte repræsentativ, og den statistiske usikkerhed kan beregnes, så man med en vis sandsynlighed kan sige, hvordan hele populationen eller universet ser ud baseret på en analyse af en begrænset stikprøve. Dataindsamling kan foregå på flere måder, f.eks. ved telefoninterview, personlige interview, postale spørgeskemaer eller online analyser på internettet.

¹⁰ Kilde: http://www.dma-research.com/organisationsamf_fokusgrupper.html

¹¹ Kilde: Torben Hansen, Hukommelsesanalyse.

Nedenfor har jeg listet eksempler på fordele og ulemper ved de enkelte analysemetoder:

Interviewmetode	Styrker	Svagheder
Personlige interview	Visuel og mundtlig information Interaktiv/forhandling Uddybende forklaringer af interviewer Spontanitet Personligt	Geografiske ulemper Gruppepres Kemi mellem interviewer og respondent matcher ikke Dyrt og omfattende
Postale interview	Verbal Mere tid til besvarelse	Ingen mulighed for visuel kommunikation Langsom dataindsamling Mindre personligt Mere formelt/seriøst
Telefon interview	Mundtlig information Interaktiv/forhandling Nogen ikke-verbal (følelsesmæssig) kommunikation Spontanitet	Intet visuelt, ingen tæthed på respondenterne Begrænset rapport Kun én-til-én (one-on-one)
Online analyser	Visuel og verbal information Svarene kommer hurtigt ind Åbne spørgsmål besvares bedre Mulighed for chat-interviews Lavere omkostninger	Begrænsede forklaringsmuligheder Upersonlig kontakt (ingen ansigt og stemme) Manglende repræsentativitet Useriøse svar pga. anonymitet

Jeg har i denne opgave kun at gøre med kvantitative undersøgelser, hvor datamaterialet er indsamlet via spørgeskemaer, men som jeg vil komme ind på senere, kunne det være interessant at foretage en kvalitativ analyse også, men det vil blive alt for omfattende at medtage i denne opgave.

5.2 Spørgeskemadesign

Når en kvantitativ undersøgelse skal foretages, laves et spørgeskema, som skal opfylde en række krav. Først skal der dannes et overblik over, hvilken form for information, man ønsker at få fra respondenterne. Det er vigtigt, at et repræsentativt udsnit af relevante svarpersoner får de samme spørgsmål, og at de får spørgsmålene i samme rækkefølge. Foretages dataindsamlingen struktureret, bliver det lettere at få svar, som kan give en struktureret databehandling. Spørgsmålene skal formuleres, så problemstillingen bliver belyst objektivt og neutralt. Skal man f.eks. følge en udvikling over tid eller sammenligne to interviewmetoder, så

skal der altid anvendes samme formulering ved spørgsmålene. Der må ikke stilles ledende spørgsmål, og man skal helst undgå svære og uforståelige ord samt upræcise formuleringer.

Ved et interview bruges ofte flere typer af spørgsmål, og her nævnes nogle:

Lukkede spørgsmål	Her er svarmulighederne formuleret på forhånd, og man skal her krydse af ud for ét eller flere svar.
Åbent spørgsmål	Ingen svarmuligheder er angivet, men et helt åbent skrivefelt til besvarelse af spørgsmålet.
Halvåbne spørgsmål	Både svarmuligheder og åbent skrivefelt, som f.eks. kan bruges til at angive et svar, som ikke er med blandt de forudbestemte svarmuligheder.
Numeriske spørgsmål	Her angives et tal, f.eks. alder eller postnummer.
Skalerings-spørgsmål	Svaret angives på en skala, hvor f.eks. 1 er "Overhovedet ikke tilfreds" og 10 er "Særdeles tilfreds".
Frekvens-spørgsmål	Kan bruges ved spørgebatterier, hvor svar til flere spørgsmål skal afgives på samme skala.

I spørgeskemaet kan man indlægge spring, som aktiveres ved at respondenterne angiver bestemte svar, og derved springes enkelte spørgsmål over. I kundetilfredsheds-spørgeskemaet, skal spørgsmål nummer 11 kun besvares af kunder, der har været i McDrive. Dermed kan kunder, der har været instore (indenfor i restauranten) springe dette spørgsmål over.

For at kunne opdele respondenterne i undergrupper efter f.eks. demografi, bør man altid tage en række baggrundsvARIABLE med i undersøgelsen. Det kan f.eks. være at spørge respondenterne om køn, alder, indkomst og uddannelse.

Under problemformuleringen (side 5) kan man se det spørgeskema, som bliver brugt i kundetilfredshedsundersøgelsen hos McDonald's.

Spørgeskemaet indeholder spørgsmål, hvortil der er brugt forskellige skaleringer ved svarmulighederne. Her ses, hvad de forskellige skalaer bruges til:

<i>Nominal skala</i>	Hvert nummer/tal er knyttet til et og kun et objekt. En form for indekssværdi, f.eks. ja/nej-spørgsmål. Demografiske variable såsom køn og bopæl vil ofte være nominale.
<i>Ordinal skala</i>	Rangordning af svarmulighederne. Det er muligt at angive, om noget er bedre end noget andet. Når man f.eks. kan sige, at et mærke er bedre end et andet, men man kan ikke sige, hvor meget bedre.
<i>Interval skala</i>	Måleenhed med arbitrært nulpunkt, dvs. man vælger selv sit udgangspunkt. Samme afstand mellem punkter på skalaen svarer til samme værdi i det målte. Her kan beregnes gennemsnit og standardafvigelse.
<i>Ratio skala</i>	Måleenhed med fast nulpunkt. Denne type skala indeholder alt fra de tre ovenstående typer skalering. Man kan identificere og klassificere objekter, rangordne objekter samt sammenligne intervaller og forskelle.

5.3 Validitet og reliabilitet

For at man kan drage brugbare konklusioner ud fra kundetilfredshedsundersøgelsen, skal resultaterne af undersøgelserne have en vis grad af korrekthed og troværdighed – med andre ord: analysen skal være valid og reliabel. Validitet er spørgsmålet om, hvorvidt der reelt måles det, som der ønskes målt, altså om spørgsmålene reflekterer den egentlige betydning af det undersøgte begreb eller emne. Der må i spørgeskemaet altså ikke være ”ugyldige svar”, det vil sige, at der er blevet spurgt til noget andet, end det man rent faktisk ønskede at spørge til. Gyldighedsproblemer kan opstå, hvis forudsætningerne for at svare på spørgsmålet varierer meget i den gruppe, man spørger.

At et spørgeskema er troværdigt, vil sige, at det, som skemaet måler, ved gentagne målinger på samme forhold skal stemme overens. Disse krav stilles, så man kan træffe vigtige beslutninger på baggrund af resultaterne.

Der er ingen vandtætte metoder til at sikre, at et spørgeskema er gyldigt og troværdigt, men en blanding af fornuft og afprøvning(er) af skemaet kan hjælpe hertil. Det er med andre ord essentielt, at man konstant fastholder et kritisk blik på spørgsmålene, og stiller spørgsmålet: Er det sandsynligt, at respondenterne forstår det samme ved spørgsmålet, som den der har formuleret det?

5.4 Missing values

Når markedsanalyser foretages, kan der forekomme problemer med manglende svar på spørgsmålene. Hvis et stort antal respondenter ikke svarer på spørgsmål, kan bortfaldet få stor betydning for undersøgelsens reliabilitet, specielt hvis antallet af manglende svar er mere end 10%. Bortfald kan f.eks. skyldes omfanget af spørgeskemaet, manglende incitament til at deltage eller at respondenterne ikke kan eller vil svare. I datasættene angives de steder, hvor der ikke er svaret, som missing-values. Der er flere muligheder for at behandle disse missing values:

Metode	Forklaring
Udelukkende enkeltvis	Her tages alle respondenter med manglende svar ud af datasættet. Dette kan dog resultere i en lille stikprøvestørrelse, da mange respondenter vil have et manglende svar et sted i spørgeskemaet.
Udelukkende parvis	I stedet for at tage respondenterne ud, som ikke har svaret på enkelte spørgsmål, inddrages de respondenter, som har svaret på alle relevante spørgsmål. På denne måde kan analyserne for datasættet altså være baseret på forskellige stikprøvestørrelser.
Indsætte neutral værdi	Man kan indsætte en værdi de steder, hvor respondenterne ikke har angivet noget svar. Værdien kan være gennemsnittet af de andre respondents svar.
Indsætte beregnet værdi	Her ser man på de svar, respondenterne har afgivet ved andre spørgsmål, og på den måde beregner man, hvilken værdi, respondenterne ville have angivet ved det manglende spørgsmål.

Da jeg har 7.745 respondenter totalt set, må antallet af missing-values ikke overstige 775 respondenter. Opfyldes dette, er der ingen grund til at erstatte de manglende svar med andre værdier, da de resterende respondenter udgør en stor nok stikprøvestørrelse. Hvis det viser sig, at antallet af bortfaldne respondenter er for stort, vil jeg tage stilling til dette, når det forekommer.

Ser man på datasættet, er det spørgsmålene hvor respondenterne skal uddybe eller skrive kommentarer, der bliver negligeret. Disse felter bliver ikke udfyldt, men efterlades tomme, da man ikke har en mulighed for at sige noget om hvad der kunne være skrevet. Andre spørgsmål, der har en forekomst af missing values, er blandt andet de demografiske spørgsmål, men også de underliggende tilfredshedsspørgsmål (nr. 2-11) bliver sprunget over.

Dog er antallet heraf ikke alarmerende stort, og derfor bliver disse manglende spørgsmål ikke udfyldt med andre værdier.

Når man taler om indsamling af spørgeskemaer til KTU-undersøgelsen, er der flere faktorer at tage højde for. Instore-restauranter får udleveret 700 spørgeskemaer, mens DT-restauranter får udleveret 2.000 spørgeskemaer, som skal uddeles til gæsterne og danne grundlag for analysen. Hvordan de enkelte restauranter vælger at uddele og inddrive spørgeskemaerne igen er op til dem selv. Der tilbydes dog en hamburger eller cheeseburger som tak for hjælpen til de deltagende (dette kan også have en stor betydning for sammensætningen af respondenterne, jf. Histogrammet). Nogle restauranter sætter ekstra personale på til at sørge for at skemaerne bliver uddelt og returneret, mens andre overlader dette til det daglige personale. Da datoerne for undersøgelsen er kendt af restauranten et par måneder før afholdelse, er det muligt for restauranten at manipulere med resultat i en vis udstrækning. Der kan skemalægges ekstra personale på selve dagen, hvormed man kan forbedre rammerne for en god kundeoplevelse, men hvorvidt nogle restauranter gør dette, og om det har nogen effekt, må stå hen i det uvisse.

5.5 Spørgeskemaets spørgsmål

Når man laver et spørgeskema er der mange ting, der skal tages højde for. Hvem er målgruppen, spørgeskemakonstruktionen, spørgsmålene – herunder spørgsmålstyper, formulering, farvede spørgsmål, indforståede spørgsmål osv. – men også en test af spørgeskemaet er vigtig før man går i gang med sin undersøgelse. Der er skrevet mange bøger omkring spørgeskemaer og deres faldgrupper, så det vil jeg ikke bruge tid på her, dog vil jeg lige belyse to problemstillinger man skal tage højde for i de skemaer som McDonald's har benyttet ved deres kundetilfredshedsundersøgelse.

For det første findes der ikke en midterkategori. Dette kan være meget godt, således at man ikke får svar tilbage, hvor respondenterne ikke har forholdt sig til spørgsmålet, de skal tage stilling. Men, hvordan ved man om de har taget stilling og ikke bare har sat sit kryds i den kategori, som de tror er midterkategorien? Langt de fleste opfatter tallet 5 som midten, men i virkeligheden ligger midten imellem 5 og 6. Dette giver anledning til endnu et spørgsmål, hvad gør man ved de respondenter, der har læst skalaen korrekt og "har valgt side" ved at vælge enten 5 eller 6?

En anden problemstilling der kan rejses, er brugen af talkategorier. I spørgeskemaet er der vist en glad smiley akkompagneret med teksten "Særdeles tilfreds" over 10-tallet, mens en sur smiley toner frem over 1-tallet med teksten "Overhovedet ikke tilfreds". Der er altså 10 værdier af tilfredshed man kan vælge i denne undersøgelse – hvad der lige skelner om man befinder sig mellem 7 og 8 eller 4 og 5, er op til den enkelte respondent. Pointen her er bare, at jo flere værdier man kan vælge imellem, jo vanskeligere er det at skelne mellem dem. Dette bliver man også nød til at have for øje, når man begynder at analysere datasættet.

6.0 Datasættets opbygning

Det datasæt som jeg bruger i denne opgave, er sammensat af datasæt fra flere forskellige restauranter. Det er således, at man har samlet alle spørgeskemaer sammen hos McDonald's på hovedkontoret, og derfra videresendt dem til et eksternt firma i USA, IRI Consulting, hvor spørgeskemaerne er blevet scannet ind og dataene er blevet bearbejdet til rapporter, der sendes til de enkelte restauranter.

Da jeg ikke har modtaget alle datasættene på samme tid, har jeg selv integreret dem, således at jeg er endt ud med et datasæt bestående af 7.745 observationer fordelt på 14 forskellige restauranter.

De oprindelige datasæt består af 23 variable, som alle er svar på spørgsmålene fra skemaet. Derudover er der så blevet tilføjet en variabel, der hedder restaurant, til at angive, hvilken restaurant, der er tale om, og en variabel, der hedder dato, idet der er to KTU-undersøgelser i løbet af et år. Det er op til de enkelte restauranter, hvorvidt de ønsker at deltage to gange om året, men kravet er, at de skal deltage i mindst én om året.

Da bearbejdningen af spørgeskemaet er foretaget i USA, er der ikke nogen navngivning af variablene, men de er nummereret i stedet for. Jeg har tildelt variablene navne, således at de repræsentere det, som de angiver. De første 15 variable er kvalitets attributter, mens de sidste 7 variable er svar på de demografiske spørgsmål.

Idet datasættet består af brudstykker, har der også været enkelte uoverensstemmelser med, hvordan man angiver et manglende svar. I nogle datasæt har man angivet et manglende svar til at være det næste hele tal udover svarmuligheden, dvs. 11 i de kategorier, hvor man havde mulighed for at svar 1-10. For at det skal give mening i de kommende analyser, har jeg valgt at omdanne sådanne svar (eller rettere manglende) til at være et punktum (.) som svarer til en missing value i SAS.

Jeg vil herunder kort gennemgå de 25 variable, som bliver brugt i denne undersøgelse af kundetilfredsheden.

Date

Denne variabel er medtaget, idet McDonald's foretager gæstetilfredshedsundersøgelser i restauranterne årligt. Hvis man har data tilbage fra sidste år fra samme restaurant, kan man undersøge om gæsternes tilfredshed har ændret sig. Derudover er det også muligt for restauranterne at gennemføre undersøgelsen enten i foråret eller i efteråret. Dermed kan man også undersøge om årstiderne spiller ind på tilfredsheden.

I dette datasæt har jeg "kun" data fra undersøgelserne i 2010, men til gengæld har jeg både fra forår og efterår.

3744 – angiver at det er fra foråret 2010, nærmere bestemt perioden fra d.

3958 – angiver at det er fra efteråret 2010, nærmere bestemt perioden fra d.

Restnr

Denne variabel er med til at identificere, hvilken restaurant, der er tale om. Alle restauranter har et nummer som identificere dem, og nummeret afspejler rækkefølgen de er åbnet i Danmark. Vesterbrogade som var den første McDonald's restaurant har således nummer 1.

10 – Viby

20 – Viborg

38 – Hjørring

45 – Vejle 2

54 – Svendborg

64 – Frederiksværk

77 - Thisted

85 – Brøndby

98 – Esbjerg 3

103 – Hillerød

201 – Køge

202 – Valby

203 – Kastrup

Tilfreds

Angiver den overordnet tilfredshed med dagens besøg. Kan antage værdien 1 – 10, hvor 1 angiver at man overhovedet ikke er tilfreds, mens man med en 10'er er særdeles tilfreds.

Tempof

Angiver tempoet af betjeningen før man bestiller. Kan også antage værdien 1 -10.

Tempoe

Angiver tempoet af betjeningen efter man bestiller.

Betjen

Denne variabel bruges til at vurdere, hvorvidt man fik en høflig og venlig betjening.

Ren

Angiver renligheden i restauranten/McDrive.

Tempmad

Bruges til at vurdere madens temperatur

Smag

Angiver smagen af maden

Frisk

Bruges til at vurdere friskheden af maden.

Orderkor

Angiver hvorvidt ens ordre var korrekt. Her bruges syntaxen: 1 = Nej og 10 = Ja.

Atmosres

Bruges til at vurdere atmosfæren i restauranten. Denne variabel henvender sig kun til de gæster der er indenfor i restauranten.

DTkom

Bruges til at vurdere kommunikationen med medarbejderne. Denne variabel henvender sig til gæsterne i McDrive. Da stort set alle DT-restauranter har en speakerpost (dvs. et højtaleranlæg) til at modtage ordren, er det vigtigt at kommunikationen er i orden.

Vigtig

Her kan gæsten angive, hvilken variabel af de ovenfor nævnte der er vigtigst for dem.

Pris

Her bedes gæsten vurdere, hvorvidt helhedsindtrykket står mål med prisen.

Konkurrent

Angiver hvordan helhedsindtrykket hos McDonalds vurderes i forhold til lignende steder

Dagbesoeg

Her vurderes hvorvidt dagens besøg har været en god oplevelse, hvorved 1 = ja og 2 = nej og 3 = intet svar (N/A).

Kommentar

Her har gæsten mulighed for at komme med en kommentar knyttet til sit svar i Dagbesoeg.

Dette er der ingen respondenter, der har benyttet sig af.

Besoegsfrekvens

Her angiver gæsten antallet af gange han/hun spiser på McDonalds pr. måned.

1 = mere end 4 gange (4+)

2 = 1-3 gange

3 = sjældnere

Placering

Her angives hvordan gæsten har besøgt McDonalds:

1 = Instore (i restauranten)

2 = McDrive (i bilen)

Koen

Angiver gæstens køn, hvor

1 = Mand

2 = Kvinde

Boern

Fortæller hvorvidt gæsten har haft børn under 12 år med på besøget.

1 = Ja

2 = Nej

Alder

Angiver gæstens alder i et forudbestemt interval:

1 = 13 – 19 år

2 = 20 – 29 år

3 = 30 – 39 år

4 = 40 år +

Tidsrum

Angiver tidsrummet man besøgte restauranten i, hvor

1 = åben – 14.00

2 = 14.00 – 17.00

3 = 17.00 – 20.00

4 = 20.00 – luk

Ugedag

Bruges til at identificere hvilken dag gæsten har besøgt restauranten.

1 = Torsdag

2 = Fredag

3 = Lørdag

4 = Søndag

7.0 Datagennemgang

I dette afsnit vil jeg forsøge at afdække, hvordan respondenterne fordeler sig på de 7 forklarende variable, som er: Køn, Alder, Tidsrum, Ugedag, Besøgsfrekvens, Børn og Placering. Først vil jeg se på restauranterne samlet set, men efterfølgende har jeg udvalgt én restaurant, som jeg vil kigge nærmere på. Det kunne også være interessant at kigge på de 13 andre restauranter enkeltvis, men det vil blive alt for omfattende.

Hos McDonald's har man i kundetilfredshedsundersøgelsen valgt at splitte tilfredsheden op i yderligere underkategorier, netop for at man kan identificere de områder, som restauranten performer dårligt på, og som man så efterfølgende kan sætte ind overfor. Jeg vil holde mit fokus på den overordnede tilfredshed, men i tabellen nedenfor, kan man se den gennemsnitlige score på de 10 underkategorier af tilfredshed for alle restauranterne, men også gennemsnittet af den overordnede tilfredshed.

Variabel	Middelværdi	Median	Std. Afv.	Varsians
Tilfreds	8,5711	9	1,6248	2,6399
Tempof	8,5859	9	1,7766	3,1564
Tempoef	8,6538	9	1,7301	2,9931
Betjen	9,0335	10	1,4728	2,1691
Ren	8,1552	9	1,8588	3,4552
Tempmad	8,2297	9	1,9268	3,7124
Smag	8,4918	9	1,6703	2,7900
Frisk	8,2908	9	1,8340	3,3635
Ordekor	9,8120	10	1,4837	2,2014
Atmosres	7,9346	8	1,8320	3,3564
Dtkom	9,6576	10	1,7884	3,1984

Som det fremgår af tabellen er den gennemsnitlige tilfredshed på 8,57, hvilket er en flot score. Dog fortæller variansen som er 2,64 at der er rimelige store udsving i tilfredsheden, så der må være nogle mindre tilfredse kunder iblandt de besøgende, hvis jeg skal sammenholde denne varians med middelværdien.

Topscoren er "Ordekor" som angiver, at man har fået den korrekte mad (det man bestilte). Her ligger gennemsnittet på 9,81, men det er nu heller ikke helt så mærkeligt, idet alle medarbejdere har vidst, at der har været en tilfredshedsundersøgelse i gang, og denne ene parameter er en af dem som de kan påvirke direkte.

Bundplaceringen går til "Atmosres", som angiver atmosfæren i restauranten. Her opnår man kun en score på 7,93 i gennemsnit, og da variansen tilmed er hele 3,36, har der været en hel

del gæster som ikke har været tilfredse med atmosfæren. Spørgsmålet er så bare, hvor meget man kan bruge denne information til? Atmosfære er et vidt begreb, og hvis jeg først skal til at tolke på, hvad gæsterne opfatter ved ordet, er vi allerede på vildspor. Nogle kan jo mene at det har været for højt musik, for meget larm, designet af restauranten har måske ikke lige været indbydende i deres optik og dermed haft en positiv indflydelse på deres vurdering af atmosfæren (det har ikke været det som de betragter som hyggeligt/godt) osv. osv.

Hvis jeg ser på variablen tilfredshed for alle restauranter, kan jeg få følgende skema frem:

Restaurant	Tilfredshed										Total	n
	1	2	3	4	5	6	7	8	9	10		
Viby	0,87	0,54	1,08	1,52	6,18	4,55	11,92	23,4	15,06	34,89	100,0	923
Viborg	0,00	0,00	0,00	0,41	2,07	4,14	6,35	25,97	21,96	39,09	100,0	724
Hjørring	0,36	0,54	0,18	2,36	3,62	6,70	13,04	22,46	16,12	34,6	100,0	552
Vejle 2	0,22	0,45	1,79	0,67	3,13	5,15	10,07	18,12	13,87	46,53	100,0	447
Svendborg	0,55	0,97	0,41	0,97	2,77	4,43	9,96	21,58	15,35	43,02	100,0	723
Frederiksværk	0,00	0,32	0,00	0,00	2,23	4,78	11,78	29,3	12,42	39,17	100,0	314
Thisted	0,00	0,00	0,31	0,00	1,55	2,80	9,94	19,88	15,53	50,00	100,0	322
Brøndby	0,71	0,00	0,89	1,96	4,63	6,95	11,41	17,65	16,22	39,57	100,0	561
Nyborg	0,00	0,19	0,93	1,12	3,55	5,79	8,97	23,55	18,69	37,20	100,0	535
Esbjerg 3	0,00	0,00	0,00	0,21	0,21	0,85	7,48	23,93	16,03	51,28	100,0	468
Hillerød	0,63	0,00	0,16	0,79	3,47	4,57	11,36	20,82	17,67	40,54	100,0	634
Køge	1,47	0,63	1,26	1,47	7,58	5,26	12,42	22,53	15,79	31,58	100,0	475
Valby	0,00	0,26	0,00	0,00	1,82	3,90	11,17	27,53	13,51	41,82	100,0	385
Kastrup	0,00	0,00	0,00	0,25	0,25	1,01	7,30	24,69	15,62	50,88	100,0	397
Total	0,40	0,31	0,54	0,95	3,35	4,49	10,24	22,80	16,30	40,62	100	7460

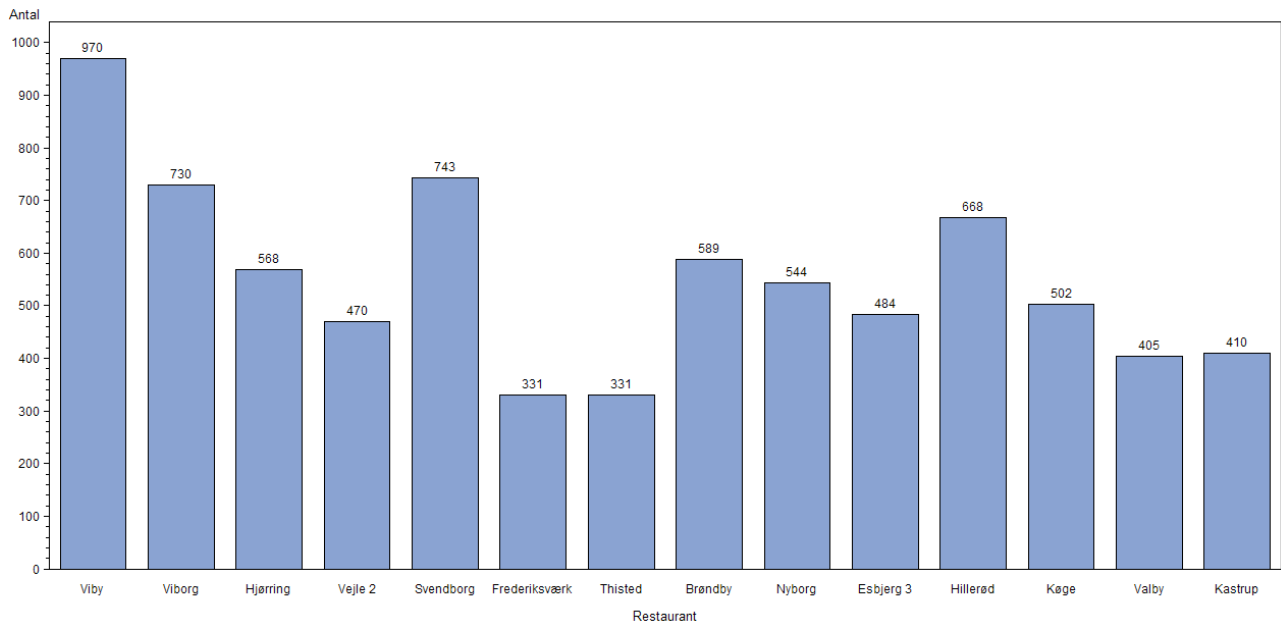
Hermed er det muligt at sammenligne restauranterne, og jeg ser helt tydeligt, at den restaurant, der har den højeste andel af mest tilfredse gæster (andelen af 10-taller) er restauranten i Esbjerg. Modsat er det restauranten i Køge, som har den laveste andel af mest tilfredse gæster. Det er også Køge, som har den højeste procentsats i den dårligste tilfredshed, men det hænger meget godt sammen med at andelen af mest tilfredse gæster er mindre.

Hvis jeg tester for, om restauranterne har den samme tilfredshedsscore i mit statistikprogram (SAS), kommer jeg frem til at dette ikke er tilfældet. Jeg får en p-værdi på $<0,0001$, antal af frihedsgrader på 117 og en værdi af χ^2 på 412,0604. Der er altså ikke noget der tyder på, at tilfredsheden i restauranterne er den samme.

I det følgende afsnit vil jeg se nærmere på baggrundsvariablene, for at se, om jeg kan beskrive de personer, som har været med i kundetilfredshedsundersøgelsen.

7.1 Histogrammer, alle restauranter samlet

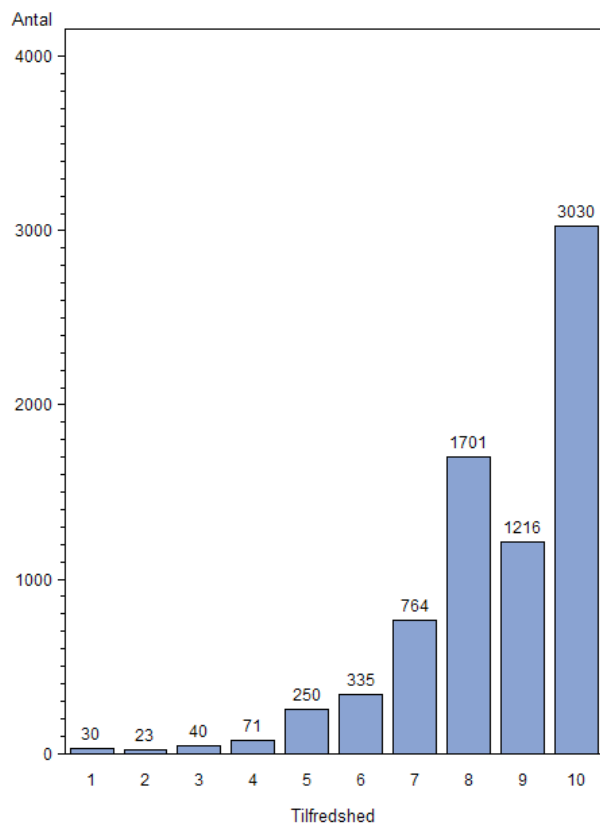
Datasættet indeholder som tidligere skrevet 7.745 respondenter, som fordeler sig på 14 forskellige restauranter:



Af ovenstående histogram fremgår det, at der er et stort udsving i antallet af respondenter mellem restauranterne. Viby ligger i top med hele 970 stk., mens Frederiksværk og Thisted begge ligger i bunden med 331 respondenter i hver restaurant.

Af det datasæt, som jeg har til rådighed, fremgår det ikke, hvor mange besøgende gæster, der har været de pågældende dage, så det er svært at sige noget om, hvorvidt antallet af respondenter er proportionalt med antallet af gæster. Derudover mangler jeg også data på udleverede skemaer, så svarprocenten er heller ikke mulig at identificere for de pågældende restauranter. Dog ved jeg fra McDonald's Danmarks hovedkontor, at hver restaurant har modtaget 2.000 stk. spørgeskemaer til uddeling (for DT-restauranter), mens det tilsvarende tal er 700 stk. for Instore-restauranter. Set i dette lys er svarprocenterne rigtig ringe.

Hvordan fordeler tilfredsheden sig på alle restauranter samlet set?

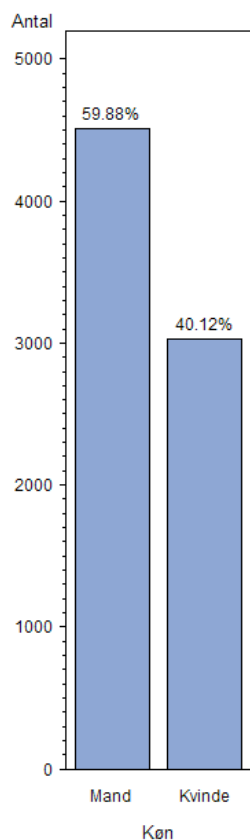


Det er tydeligt, at der er en klar overvægt af tilfredse gæster hos McDonald's. Andelen af dem, der er særdeles tilfredse ligger på hele 40,62%, og hvis jeg regner intervallet fra 8-10 som godt tilfredse, kommer jeg helt op på 79,72%, hvilket også må betragtes som særdeles flot. Hvis man betragter intervallet 1-4 som gæster, der ikke har været tilfredse med deres besøg, udgør det en andel på 2,20%, svarende til 164 gæster. Dette er en forsvindende lille andel, og det kan McDonald's kun være stolte af.

Hvis man kigger på histogrammer for de forklarende variable for alle restauranterne får jeg følgende information:

7.1.1 Køn

Ser man nu på kønsfordelingen, får man følgende billede:

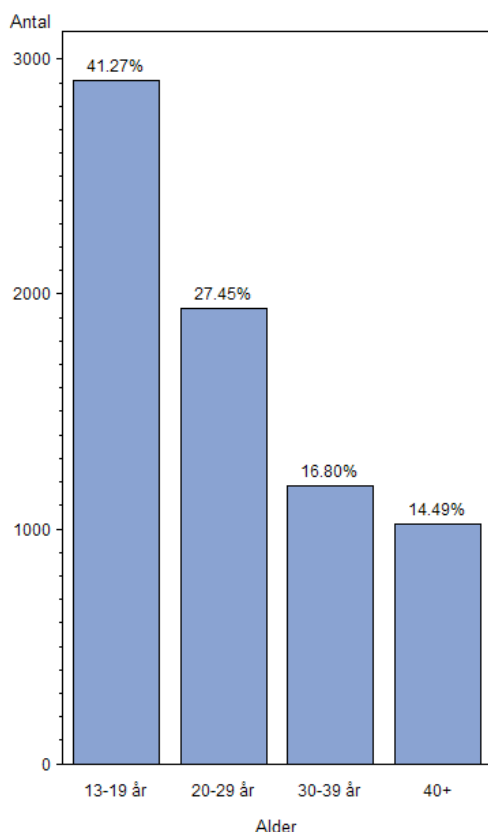


Der er altså en overvægt af mandlige respondenter – hele 59,88% , hvilket svarer til 4.512, mens kvinderne udgør 40,12% , svarende til 3.023 respondenter. De resterende 210 respondenter har ikke angivet deres køn i spørgeskemaet, men selv hvis de skulle vise sig at være kvinder, ville der stadig være en klar overvægt at mandlige respondenter.

Det ville være nærliggende at konkludere, at McDonald's har flest mandlige gæster, men dette er ikke en helt valid konklusion, medmindre man understreger, at dette resultat er fremkommet på baggrund af de data, jeg har haft til rådighed (man ved jo ikke, hvordan kønsfordelingen er blandt dem, der ikke har deltaget i undersøgelsen). Fremover når jeg drager konklusioner, er det på baggrund af datasættet.

7.1.2 Aldersfordeling

Hvordan ser det ud med aldersfordelingen af respondenterne?



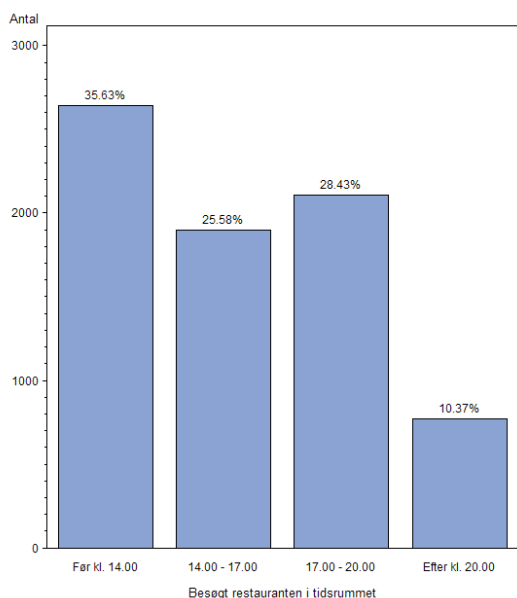
Det fremgår tydeligt af histogrammet til højre, at McDonald's henvender sig til den yngre del af befolkningen. Hele 41,27% - svarende til 2.911 respondenter er i aldersgruppen 13-19 år, mens kun 14,49% (= 1.022 respondenter), befinder sig i aldersgruppen 40+. Antallet af respondenter, der ikke har angivet deres alder, er 691.

En af forklaringerne på at der er en så stor overvægt af unge respondenter, skal selvfølgelig findes i at McDonald's henvender sig til netop denne målgruppe, men man kan heller ikke udelukke, at belønningen med en hamburger/cheeseburger, for

et udfyldt skema, har haft en positiv påvirkning på denne fordeling. For den yngre aldersgruppe vil denne belønning procentvis udgøre et større beløb frem for den ældre aldersgruppe, der har arbejde og en god indkomst, og så gør det jo heller ikke noget, at det er et sted man godt kan lide at spise/opholde sig.

7.1.3 Tidsrum

Hvornår kommer gæsterne på besøg i restauranten?



Lidt mere end entredjedel kommer i restauranten i frokostperioden (her angivet som før kl. 14).

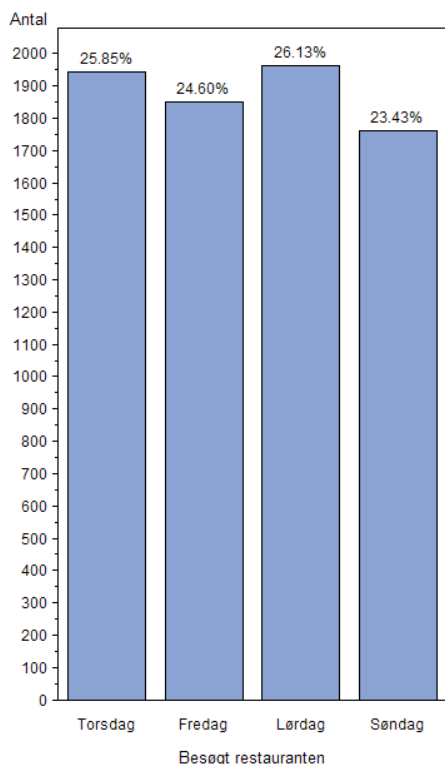
Hvis en af restauranterne havde døgnåbent og serverede morgenmad, ville dette være misvisende, men ingen af de 14 restauranter har åbent før kl. 9.00 om morgenen, og der er ingen af dem, der serverer morgenmad.

Snack-perioden (tidsrummet mellem kl. 14 og 17) har 1.897 respondenter, svarende til 25,58%, mens den traditionelle aftensmadsperiode udgør 28,43% , svarende til 2.108 respondenter.

Andelen af gæster, der kommer efter kl. 20 er på 10,37%. Denne procentsats siger umiddelbart ikke særlig meget, da jeg ikke har noget at holde den op imod (bortset fra de andre 3 perioder). Er der variation i dette forhold, hvis jeg ser på restaurantens beliggenhed (kommer der flere gæster i dette tidsrum, hvis restauranten er placeret i et boligområde kontra ved en landevej) og hvis jeg nu havde lavet min undersøgelse i starten af ugen?

7.1.4 Ugedag

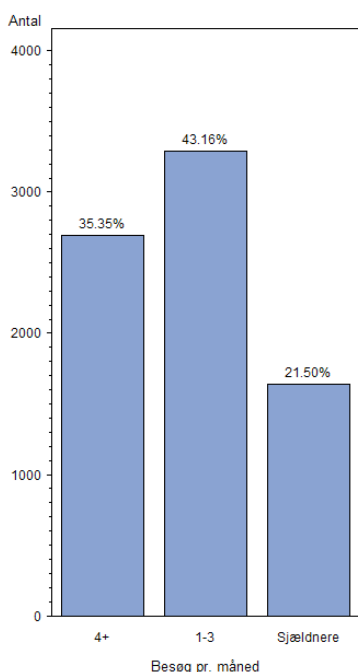
Hvilke dage besøger gæsterne McDonald's?



Da spørgeskemaundersøgelsen kun finder sted på ugens sidste fire dage, er det svært at sige noget præcist herom. Dette ses også tydeligt af histogrammet til venstre, hvor alle dagene mere eller mindre ligger på 25%, altså stort set en lige fordeling af dagene. Da der også er uddelt lige mange skemaer hver dag (i teorien), ville det også være mærkeligt, om der var store udsving heri. Dog kan det undre lidt, at den traditionelt set største omsætningsdag og den dag med flest gæster, søndag, er den dag med færrest tilbagemeldinger – ”kun” 1.761, mod 1.964 om lørdagen. En af forklaringerne skal muligvis findes i, at det netop er den travleste dag på hele ugen, og der kan være blevet glemt at blive uddelt spørgeskemaer i farten.

7.1.5 Besøgsfrekvens

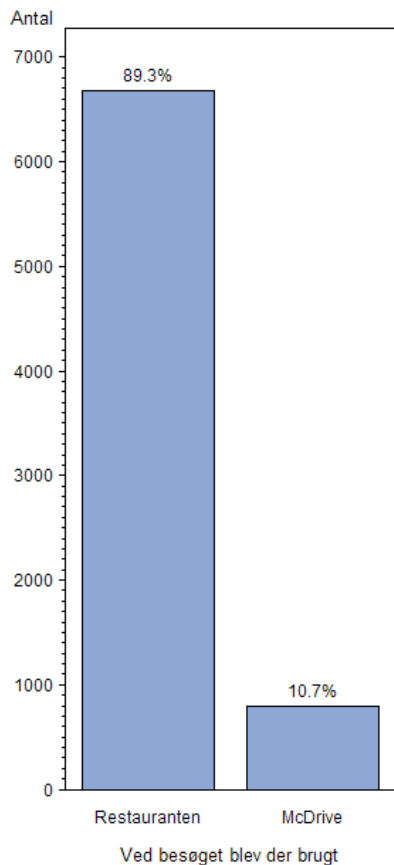
Hvor ofte kommer gæsterne i restauranten?



Som jeg kan se af histogrammet til venstre, kommer lige godt og vel 80% af gæsterne hos McDonald's minimum én gang om måneden, hvilket må siges at være en meget høj procentsats. Af de adspurgte kommer hhv. 2.693 oftere end 4 gange om måneden (hvilket i gennemsnit svarer til 1 gang om ugen), 3.288 kommer mellem 1 til 3 gange om måneden, mens 1.638 kommer hos McDonald's sjældnere end 1 gang om måneden. Man må sige at det er nogle flotte guest-recovery's, der finder sted, og det beviser at McDonalds må gøre noget rigtig for de 80% af gæsterne, som vælger at komme tilbage så ofte.

7.1.6 Besøgssted

Hvordan fordeler gæsterne sig i forhold til deres besøg hos McDonald's?

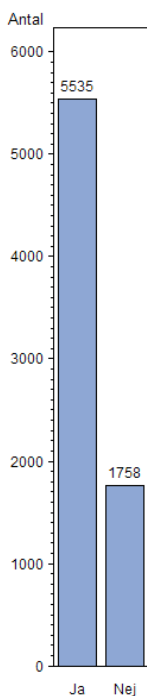


Det er helt tydeligt, at der i denne kundetilfredshedsundersøgelse er en klar overvægt af gæster, der har været inde i restauranten, frem for at benytte McDrive. Hele 89,3% af de adspurgte har været instore, hvilket svarer til 6.685 respondenter, mens kun 801 har benyttet McDrive. Dette betyder at jeg får et misvisende billede af restauranternes samlede tilfreds. Da alle restauranterne er DT-restauranter (fritstående restauranter med mulighed for at benytte bil ved bestilling) og da salget som oftest fordeler sig 50/50, hvilket som regel også betyder, at halvdelen af gæsterne kommer instore, mens den anden halvdel benytter sig af McDrive, kan det ikke være det reele billede af restauranten jeg får tegnet med så få besvarelser. De spørgsmål, der knytter sig til McDrive, har derfor en stor usikkerhed forbundet med sig, pga. den meget lille svarprocent. Ofte er fordelingen af kunderne

således, at der kommer flere gæster i McDrive, og dette gør slet ikke billedet bedre.

7.1.7 Familieforhold

Hvordan ser gæsternes familieforhold ud?



Det er tydeligt at se, at langt de fleste af dem der kommer og besøger McDonald's, har taget deres barn/børn med. Hele 5.535 af de adspurgte, svarende til 75,89%, har medtaget deres barn, hvilket stemmer godt overens med at McDonald's er en familierestaurant. De resterende 24,11%, eller 1.758 respondenter, har således været i restauranten alene. Spørgsmålet er så, om de har været der i formiddagstimerne og om aftenen, hvor man kan forvente at børn er henholdsvis i børnehave/skole eller i seng (i hvert fald på hverdage), eller

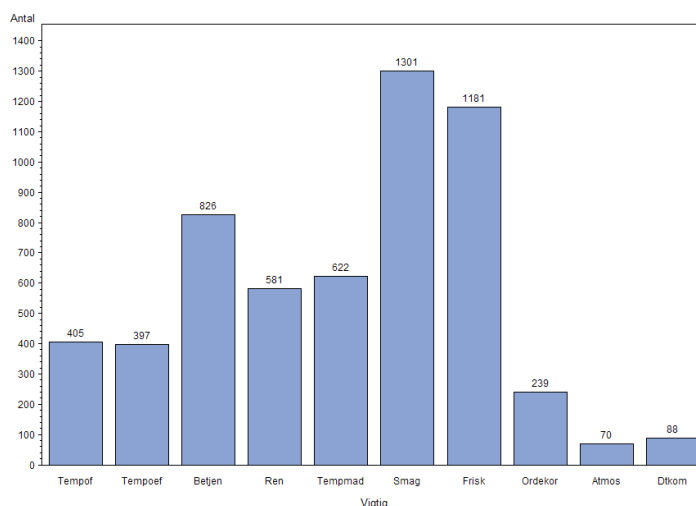
om det er ungegruppen, der er repræsenteret her, da det nok er de færreste i aldersgruppen 13-19 år, der har børn.

7.1.8 Prioritering

Når man laver en kundetilfredshedsundersøgelse, har den selvfølgelig altid til formål, at undersøge tilfredsheden blandt kunderne. I det spørgeskema som McDonald's bruger i deres undersøgelse, er der et mindst ligeså interessant spørgsmål, som er medtaget i undersøgelsen. Dette drejer sig om spørgsmål nr. 12, som lyder "Hvis du ser tilbage på spørgsmål 2 – 11, hvad er så mest vigtigt for dig?" – her får man nemlig en mulighed for at se, hvad gæsten finder vigtigst. Efterfølgende har man så muligheden for at se, om det også er det område, man performer godt på.

Det er lidt trist at være rigtig god til at ekspedere kunderne, når det de finder vigtigst, faktisk er smagen af maden.

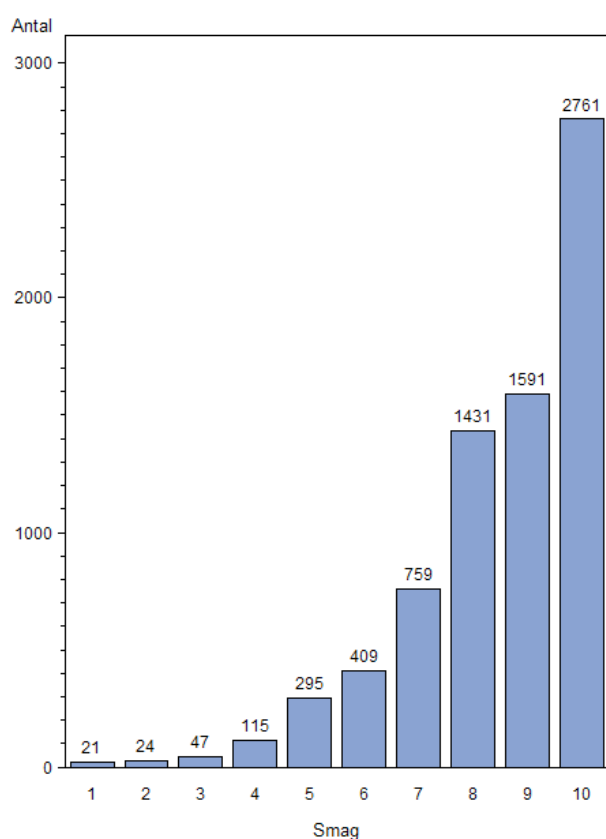
Fordeling af hvad gæsterne prioriterer som det vigtigste for tilfredsheden



Det kunne være interessant at se, hvilken forklarende variabel, kunderne prioriterer som det vigtigste for tilfredsheden. På histogrammet ovenfor, er det tydeligt, at "smag" er det som kunderne prioriterer højest, men efterfulgt af "frisk", altså om maden var frisk da den blev serveret. Begge variable har altså noget at gøre med selve produktet, men servicedelen kommer ind på en tredjeplads. Altså hvorvidt betjeningen var høflig og venlig, så det er ikke helt uvigtigt, hvordan man bliver serviceret. Hele 826 personer har haft dette som deres topprioritet, hvilket ikke er så få. Det fremgår også af histogrammet, at 88 personer har angivet "Dtkom" som deres førsteprioritet. Når man ser dette sammen med alle de andre variable, virker det ikke af mange, men sammenholder man det med, at der kun har været 801

respondenter igennem McDrive (svarende til 10,7%), er det en relativt stor andel. Man må også sige, at kommunikationen er særdeles vigtig i Driven, da man stort set kun kan bestille over et højtaleranlæg, og dermed ikke har mulighed for at have øjenkontakt med personalet. Det er også almindelig viden, at kropssproget er med til at understøtte kommunikationen, men denne er man helt fraskåret i McDrive.

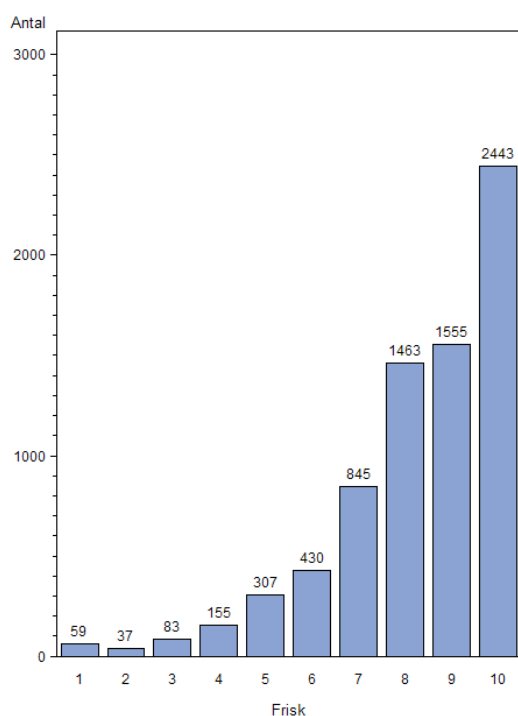
Hvordan er fordelingen på smag, som var den kategori flest gæster fandt vigtigst?



Ligesom fordelingen af den overordnede tilfredshed, ser jeg her, at der er en klar overvægt i de sidste tre kolonner, som man kan betegne som rigtig godt. Dog er andelen af 10-taller næsten 300 stk. mindre set i forhold til tilfredshed, mens 9-tallerne er blevet tilsvarende flere. Procentvis udgør de tre sidste søjler 77,59% af de afgivne karakterer. Igen er det kun meget få, der har været utilfredse med smagen – de første fire søjler svarer til 2,78%. Det er ikke nogen dårlige resultater man her har opnået, men der er plads til forbedringer, når det drejer sig

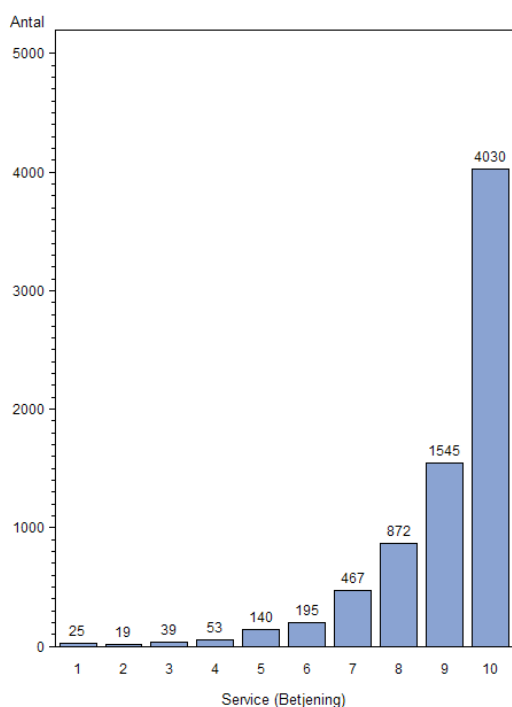
om smagen set i forhold til den overordnede tilfredshed.

Hvordan ser det ud med friskheden af maden?



Igen ser man en klar overvægt på de tre sidste søjler, om end det er noget mindre end både tilfredshed og smag. Her udgør de tilsammen 74,03%, altså lige godt og vel trefjerdedele af alle de afgivne karakterer. Det er dog bekymrende, hvis man ser i den modsatte ende, her er andelen af 1-taller steget lige godt og vel med en faktor 3, og de første 4 søjler udgør her 4,53%, hvilket er godt 2 procentpoint mere end tilfældet var ved smag. Der kan altså godt arbejdes med friskheden af maden.

Hvordan ligger det med servicedelen?



Her fremgår det helt tydeligt, at service overfor kunderne er noget medarbejderne hos McDonald's er særdeles gode til. Lidt over halvdelen, nærmere bestemt 54,57% af de adspurgte giver servicen toppoint, mens 87,30% af respondenternes besvarelse ligger i kategorien 8-10. Ser man på de første 4 søjler udgør de tilsammen 1,84%, hvilket svarer til at 136 gæster har haft en mindre god serviceoplevelse.

Som de tre ovenforstående histogrammer illustrerer, er det en god idé at supplere middelværdi og varians for en "variabel" med et histogram, for at se den faktiske fordeling.

Herved kan man få en forståelse for, hvordan man er kommet frem til værdierne, og den grafiske fremstilling kan også være med til, at man bedre kan forstå dette.

Det er vigtigt her at påpege, at de tre topprioriteter jeg har kigget nærmere på, udelukkende er fundet ud fra respondenternes besvarelse af spørgsmål 12. Her bliver de kun bedt om at give én topprioritet, hvorfor det er muligt, at de tre prioriteter jeg her har kigget nærmere på, ville være nogle andre, hvis respondenterne blev bedt om at angive deres top 3 prioriteter. Derudover er det en meget lille andel, som har svaret på dette spørgsmål, som også kan være med til at betvivle pålideligheden af disse prioriteringer.

Jeg vil i næste afsnit se nærmere på restauranten i Brøndby. Er kunderne i denne restaurant ligeså tilfredse, som resten af restauranterne tilsammen, eller gør de det bedre i denne restaurant? Hvordan er kundesammensætningen i denne restaurant? Disse og mange flere spørgsmål vil blive besvaret i næste afsnit.

7.2 Histogrammer, Brøndby

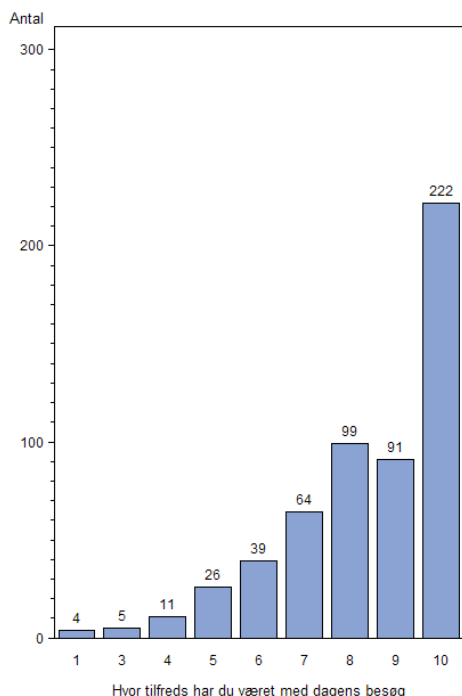
Hvis jeg kigger på restauranten, der ligger i Brøndby, og ser på dens fordeling af de underordnede kategorier af tilfredshed, får man følgende tabel:

Variabel	Middelværdi	Median	Std. Afv.	Varsians
Tilfreds	8,3886	9	1,8152	3,2952
Tempof	8,3573	9	1,9079	3,6401
Tempoef	8,4144	9	1,8493	3,4200
Betjen	8,7996	9	1,6164	2,6126
Ren	8,1825	9	1,8129	3,2866
Tempmad	8,1802	9	1,9161	3,6715
Smag	8,2065	9	1,8415	3,3910
Frisk	8,1538	9	1,9133	3,6607
Ordekor	9,6932	10	1,6347	2,6723
Atmosres	7,8540	8	1,9307	3,7276
Dtkom	8,2464	9	1,8583	3,4531

Det er helt tydeligt, at jeg her har at gøre med en restaurant, der performer dårligere på tilfredshedsparametrene, end gennemsnittet af alle restauranter samlet set. Alle middelværdierne ligger under dem, som jeg fandt frem til for alle restauranter samlet set, og varianserne er også blevet større. Da McDonald's Brøndby også var en del af det datasæt, som blev brugt for alle restauranterne samlet set, tyder det på, at denne restaurant har været med til at hive niveauet ned.

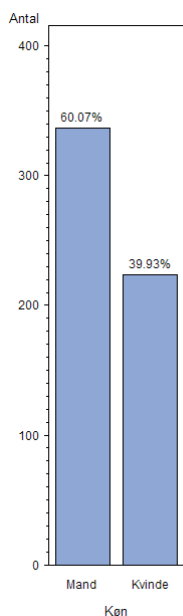
Når det er sagt, er det stadig nogle relativt flotte værdier for tilfredshed, men det er klart, at hvis sammenligningsgrundlaget er alle restauranterne samlet set, er der rum for forbedring.

Ser man på fordelingen af den overordnede tilfredshed, får man følgende fordeling.



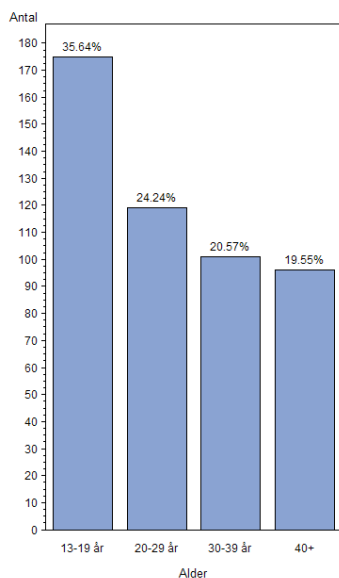
Umiddelbart ser fordelingen flot ud, der er hele 222 kunder, der giver topkarakteren 10, hvilket svarer til 39,57%, men hvis man huske tilbage på tabellen der viser scoren for alle restauranterne, er det tydeligt, at det ikke er helt så imponerende som det nogle af de andre restauranter har præsteret.

7.2.1 Køn



Når man ser på kønsfordelingen hos McDonald's Brøndby, afspejler den stort set det samme mønster, som jeg så hos hele landet. Der er en overvægt af mandlige kunder, som har svaret på spørgeskemaet, helt præcis 337 stk. svarende til 60,07%, mens 224 kvinder har angivet deres køn, svarende til 39,93%. Hele 28 respondenter har ikke angivet køn, hvilket i første omgang kan virke som få. Holder jeg det op imod, at der er 14 restauranter og der kun manglede 210 respondents angivelse i alt, er det dog et procentvis højt tal. Årsagen hertil kan man kun gætte på.

7.2.2 Aldersfordeling

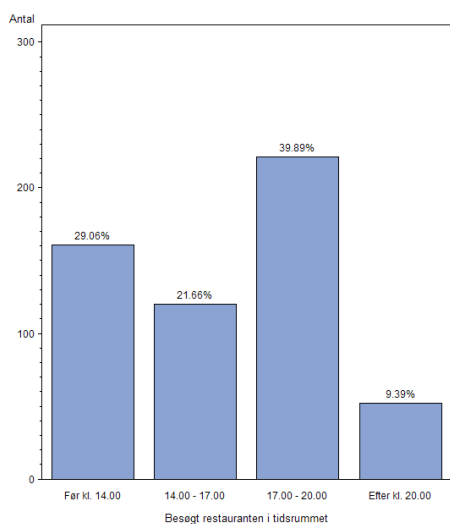


Ved at se på kundernes aldersfordeling hos McDonald's Brøndby, fremgår det, at jeg i denne restaurant har at gøre med et lidt ældre publikum. De to "unge" grupper er hver reduceret med cirka fem procent-point i forhold til landet, mens de to "ældre" grupper tilsvarende hver er forøget med fem procent-point.

Når man ser på aldersfordelingen for restauranten i Brøndby, ser jeg, at den ligger meget op af det man ser for alle restauranterne samlet. Der er en stor overvægt af de helt unge, mens andelen af respondenter, der er fyldt 40 år, er begrænset til kun at udgøre 19,55%

7.2.3 Tidsrum

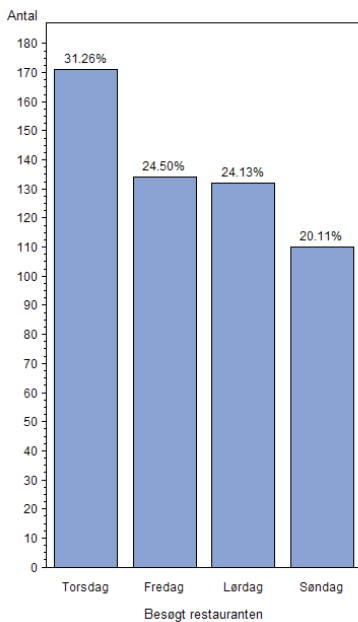
Hvis man ser, hvornår kunderne kommer i Brøndby, er der en klar overvægt i tidsrummet 17-20. Hele 39,89% svarende til 221 respondenter har været i restauranten i dette tidsrum. Det



er mede end 10 %-point flere i dette tidsrum i forhold til landet. Perioden efter kl. 20 ligger på niveau med landet, mens perioderne før kl. 17 ligger et stykke under landet. Det tyder altså på, at restauranten henvender sig til familierne, der spiser ude om aftenen og ikke til de erhvervsaktive, der skal have frokost. Til gengæld må man sige, at det er i frokostpausen, der er potentiale til at øge kundestrømmen, når man sammenligner med landet.

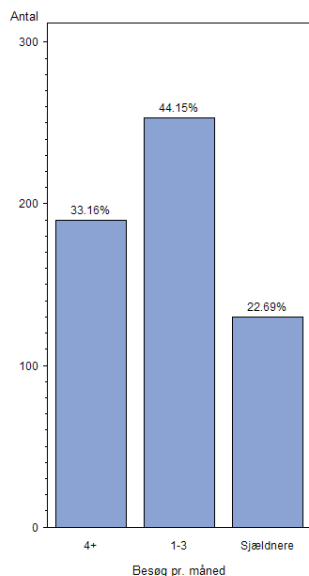
7.2.4 Ugedag

Ser man, hvilke dage kunderne kommer i restauranten, udskiller McDonald's Brøndby sig fra landet. Hvor landet havde stort set lige mange respondenter hver dag, har Brøndby en



overvægt af respondenter om torsdagen – 31,26% svarende til 171, mens søndagen ser noget tyndt besat ud med kun 20,11% svarende til 110 personer. Jeg tror, at dette skal ses som et udtryk for, at der har været en enorm fokus på undersøgelsen den første dag, og så er gassen langsomt sivet ud af ballonen. Hvis man ser på antallet af transaktioner for den uge, hvor undersøgelsen kørte, viser den at søndag har flest transaktioner. Der er altså flest gæster i restauranten denne dag, men hvorfor man så ikke får tilsvarende billede fra spørgeskema-undersøgelsen må ligge hen i det uvisse.

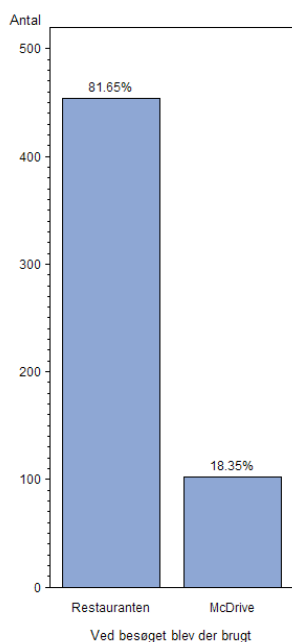
7.2.5 Besøgsfrekvens



Når man ser på besøgsfrekvensen hos McDonald's Brøndby, ser man, at den er stort set identisk med landsplan. 33,18% af kunderne, svarende til 190 respondenter, kommer oftere end 4 gange om ugen, 44,15% kommer 1-3 gange om måneden, mens 22,69% kommer i restauranten sjældnere.

7.2.6 Besøgssted

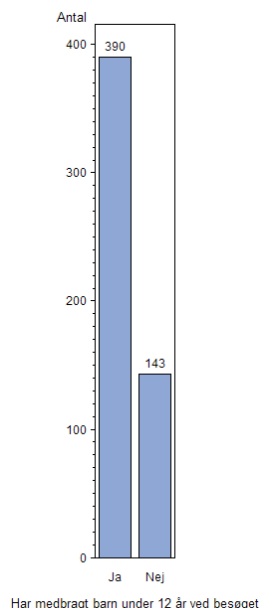
Hvis man kigger på, hvordan gæsterne har besøgt restauranten, fremgår det at hele 18,35% har været igennem McDrive. Dette er det højeste antal respondenter, hvis vi også kigger på de andre restauranter, og det ligger også over gennemsnittet for alle restauranter. Brøndby har



altså været særdeles gode til at få spørgeskemaerne tilbage fra deres kørende gæster. Set i forhold til, hvor mange spørgeskemaer, der er uddelt i McDrive, er det dog ikke et videre positivt resultat. Men, det er også svært at få folk, der har travlt (det er oftest derfor man kommer igennem McDrive), til både at udfylde skemaet, men også at returnere det. Dertil kommer, hvor meget vægt man egentlig kan pålægge besvarelserne fra McDrive? Der er rigtig mange faktorer, der kan spille ind: fokuserer de deres besvarelse kun på dagens besøg? Plejer de at besøge en anden restaurant og sammenligner med denne osv.?

7.2.7 Familieforhold

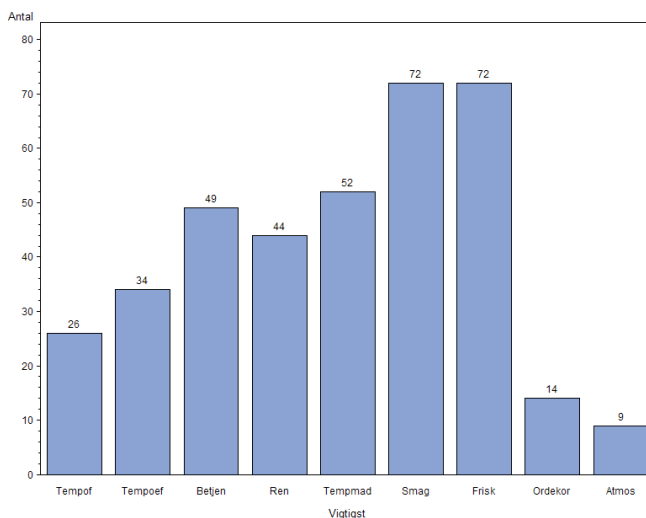
Der er en overvældende stor repræsentation af gæster, der har et barn med på deres besøg



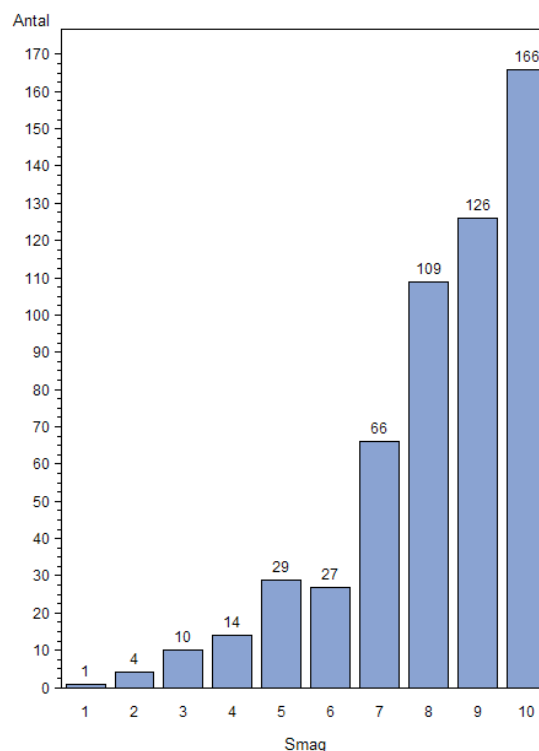
hos McDonald's. Det er også en familierestaurant, og undersøgelsen har kørt over en weekend, så der har været store muligheder for at medbringe børn. Dette vil man kunne se nærmere på i de kommende krydstabeller, der kan angive tidspunkter og dage, for at børn har været med på restauranten. En anden forklaring på den store repræsentation af voksne med børn kan muligvis findes i, at de har haft tid til at udfylde skemaet, mens børnene alligevel legede i børnelandet?

7.2.8 Prioritering

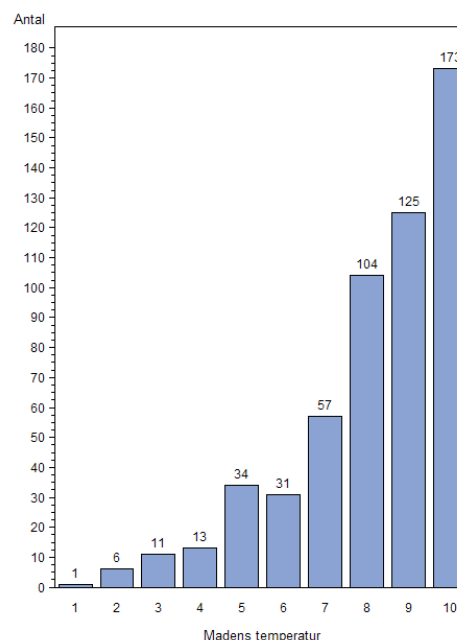
Hvis man kigger på McDonald's Brøndby og ser efter, hvad der prioriteres højest blandt kunderne, når man frem til at smag og frisk er de to vigtigste, mens madens temperatur kommer ind som nummer tre. Dette er en ændring i forhold til alle restauranterne samlet, hvor betjening kom ind på en tredjeplads. Jeg kan dog se, at det kun er tre stemmer, som skulle ændres, hvis man skulle opnå samme resultat som landet, så man kan roligt sige at Brøndby følger landet i prioriteringerne.



Når man ser på histogrammet for fordelingen af smag hos McDonald's Brøndby ser man, at den er næsten identisk med fordelingen på landsplan. Dog er overvægten af 10-taller ikke helt så overvældende, som på landsplan. De tre sidste søjler udgør 72,78% af de afgivne karakterer – holdt op imod 77,59% på landsplan – mens de første fire søjler svarer til 5,26% mod 2,78% på landsplan. Det betyder altså, at kunderne generelt er lidt mindre tilfredse med smagen hos McDonald's Brøndby, og da det er et område, der har topprioritet hos kunderne, er der her noget at arbejde med. Dette stemmer også overens med tabellen for middelværdier og varianser.



Ser man på fordelingen af karaktererne for temperatur, ser billedet ud ligesom for smag. Langt størstedelen er tilfredse med madens temperatur – de tre sidste søjler udgør 72,73%, mens de fire første udgør 5,59%. Det er dog en lille smule foruroligende, at en fjerdedel af alle kunderne ikke er tilfredse med temperaturen, hvis det drejer sig om de varme produkter (og for den sags skyld også de kolde). Restauranten lever af at sælge produkter, så de skal være i orden til hver en kunde, det må være målsætningen og mindstekravet.



7.6 Krydstabeller

En måde at undersøge parvise sammenhænge mellem variable er ved hjælp af krydstabeller. Dette kan være en fordel for variable, der ikke er normalfordelte, f.eks. binære variable og variable med få svarmuligheder/grupper. Hvis der er omtrent lige mange observationer i hver celle, indikerer det, at der ikke er en sammenhæng mellem de to variable. Man kan derefter opstille et statistisk test for uafhængighed, hvor Pearson's Chi-square test (χ^2) er det mest almindelige. Testet udregnes på baggrund af de observerede værdier (O) set i forhold til de forventede værdier (E). Teststørrelsen kan opskrives således:

$$\chi^2 = \sum_r \sum_k \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \chi^2[(r-1)(k-1)]$$

Under Nul-hypotesen (H_0), hvor det antages at de to variable er uafhængige, kan testet antages at være χ^2 -fordelt med antal frihedsgrader svarende til antal rækker (r) minus 1 multipliceret med antal kolonner (k) minus 1.

En lav testsandsynlighed (f.eks. mindre end 0,05) betyder, at de to variable er afhængige. Hvis testsandsynligheden er over 5%, kan man ikke konkludere afhængighed, men man kan heller ikke konkludere uafhængighed.

Hvis antallet af observerede værdier i cellerne er meget små, kan det give et problem ved udregning af Pearson's Chi-square test, da der divideres med de forventede værdier. Som tommelfingerregel bør det laveste forventede antal i en celle være større end 3.

7.7 Krydstabeller for McDonalds

7.7.1 Alder og tilfredshed/vigtig

Når man ser på tabellen for alder*tilfredshed fremgår det, at den celle med den højeste sandsynlighed er en tilfredshed på 10 i aldersgruppen 13-19 år. Helt præcis giver den en sandsynlighed på 19,14%, et resultat, der ikke umiddelbart siger helt så meget. Et resultat jeg finder overraskende er, når man kigger på kolonnesandsynligheder. Her kan man se, at andelen af de helt unge, der giver karakteren 1 for tilfredshed, er på hele 75%. Der er altså en klar overvægt af unge, som er utilfredse set i forhold til aldersfordelingen.

Hvis man kigger på kolonnen med en score på 10 fremgår det, at aldersgruppen 13-19 år giver denne karakter 5%-point oftere end de andre aldersgrupper.

Ligger man scoren for tilfredshed på 8, 9 og 10 sammen (og dermed tager højde for individuelle fortolkninger af tilfredshed) opnår man en tilfredshedsprocent på hhv. 79,53; 77,93; 78,60 og 83,27 for de fire aldersgrupper. Dette viser, at tilfredsheden er stort set identisk i de fire aldersgrupper, men igen et lidt sjovt resultat, at den ældre aldersgruppe er 5%-point mere tilfredse end den yngste aldersgruppe.

Ser man på tabellen over vigtighed fordelt på aldersgrupper, fremgår det, at aldersgruppen 13-19 år synes smagen er det vigtigste, med betjening som en 2. plads, mens de andre aldersgrupper alle har frisk som deres topprioritet med smag som en nummer to. Hvor top 2 handler om maden for aldersgrupperne over 20 år, prioriterer de unge altså betjeningen relativt højt, men det er også bemærkelsesværdigt at smagen scorer hele 27,48% i den unge kategori, der er altså bred enighed om, at det er denne faktor, der er vigtigst. Hvis jeg ser på de andre aldersgrupper, er deres topprioritet flere procentpoint under de helt unges topprioritet. Jeg synes det er overraskende, at tempmad først kommer ind på en fjerdeplads i alle alderskategorierne, men det kan muligvis forklares ved, at ikke alle, der har udfyldt et skema, har købt mad fra køkkenet, der kan være købt kaffe, kage, is eller milkshakes ved besøget.

Frequency Percent Row Pct Col Pct	Table of Alder by Tilfreds											
	Alder	Tilfreds										Total
		1	2	3	4	5	6	7	8	9	10	
13-19 år	21	11	14	26	101	125	277	536	397	1300	2808	
	0.31	0.16	0.21	0.38	1.49	1.84	4.08	7.89	5.85	19.14	41.35	
	0.75	0.39	0.50	0.93	3.60	4.45	9.86	19.09	14.14	46.30		
	75.00	52.38	40.00	41.94	44.89	41.12	38.53	34.58	36.19	47.27		
20-29 år	5	8	15	19	80	89	194	466	264	718	1858	
	0.07	0.12	0.22	0.28	1.18	1.31	2.86	6.86	3.89	10.57	27.36	
	0.27	0.43	0.81	1.02	4.31	4.79	10.44	25.08	14.21	38.64		
	17.86	38.10	42.86	30.65	35.56	29.28	26.98	30.06	24.07	26.11		
30-39 år	1	1	4	11	23	47	158	319	237	344	1145	
	0.01	0.01	0.06	0.16	0.34	0.69	2.33	4.70	3.49	5.07	16.86	
	0.09	0.09	0.35	0.96	2.01	4.10	13.80	27.86	20.70	30.04		
	3.57	4.76	11.43	17.74	10.22	15.46	21.97	20.58	21.60	12.51		
40+	1	1	2	6	21	43	90	229	199	388	980	
	0.01	0.01	0.03	0.09	0.31	0.63	1.33	3.37	2.93	5.71	14.43	
	0.10	0.10	0.20	0.61	2.14	4.39	9.18	23.37	20.31	39.59		
	3.57	4.76	5.71	9.68	9.33	14.14	12.52	14.77	18.14	14.11		
Total	28	21	35	62	225	304	719	1550	1097	2750	6791	
	0.41	0.31	0.52	0.91	3.31	4.48	10.59	22.82	16.15	40.49	100.00	
Frequency Missing = 954												

Alder	Table of Alder by Vigtig										
	Vigtig										Total
	Tempof	Tempoef	Betjen	Ren	Tempmad	Smag	Frisk	Ordekor	Atmos	Dtkom	
13-19 år	157	137	344	237	177	562	292	80	26	33	2045
	3.00	2.61	6.56	4.52	3.38	10.72	5.57	1.53	0.50	0.63	39.01
	7.68	6.70	16.82	11.59	8.66	27.48	14.28	3.91	1.27	1.61	
	41.53	36.93	44.22	44.13	31.05	46.87	28.32	34.78	38.81	40.74	
20-29 år	110	115	207	126	195	290	319	70	10	23	1465
	2.10	2.19	3.95	2.40	3.72	5.53	6.09	1.34	0.19	0.44	27.95
	7.51	7.85	14.13	8.60	13.31	19.80	21.77	4.78	0.68	1.57	
	29.10	31.00	26.61	23.46	34.21	24.19	30.94	30.43	14.93	28.40	
30-39 år	54	71	123	89	108	174	229	54	21	10	933
	1.03	1.35	2.35	1.70	2.06	3.32	4.37	1.03	0.40	0.19	17.80
	5.79	7.61	13.18	9.54	11.58	18.65	24.54	5.79	2.25	1.07	
	14.29	19.14	15.81	16.57	18.95	14.51	22.21	23.48	31.34	12.35	
40+	57	48	104	85	90	173	191	26	10	15	799
	1.09	0.92	1.98	1.62	1.72	3.30	3.64	0.50	0.19	0.29	15.24
	7.13	6.01	13.02	10.64	11.26	21.65	23.90	3.25	1.25	1.88	
	15.08	12.94	13.37	15.83	15.79	14.43	18.53	11.30	14.93	18.52	
Total	378	371	778	537	570	1199	1031	230	67	81	5242
	7.21	7.08	14.84	10.24	10.87	22.87	19.67	4.39	1.28	1.55	100.00
Frequency Missing = 2503											

7.7.2 Køn og tilfredshed/vigtigt

En krydstabel over køn og tilfredshed, samt køn og vigtigt, kan give en identifikation om, hvorvidt, der er forskel på kønnenes tilfredshed, men også om deres præferencer for, hvad der vigtigst.

Table of Koen by Tilfreds											
Koen	Tilfreds										
	1	2	3	4	5	6	7	8	9	10	Total
Mand	25	17	21	46	165	199	476	1090	719	1610	4368 60.12
	0.34	0.23	0.29	0.63	2.27	2.74	6.55	15.00	9.90	22.16	
	0.57	0.39	0.48	1.05	3.78	4.56	10.90	24.95	16.46	36.86	
	89.29	80.95	60.00	64.79	67.35	60.86	63.64	65.27	60.78	54.80	
Kvinde	3	4	14	25	80	128	272	580	464	1328	2898 39.88
	0.04	0.06	0.19	0.34	1.10	1.76	3.74	7.98	6.39	18.28	
	0.10	0.14	0.48	0.86	2.76	4.42	9.39	20.01	16.01	45.82	
	10.71	19.05	40.00	35.21	32.65	39.14	36.36	34.73	39.22	45.20	
Total	28	21	35	71	245	327	748	1670	1183	2938	7266 100.00
	0.39	0.29	0.48	0.98	3.37	4.50	10.29	22.98	16.28	40.43	
Frequency Missing = 479											

Hvis man kigger på tabellen til højre, fremgår det, at andelen af kvinder, der giver tilfredsheden et 10-tal er hele 45,82% mens den tilsvarende procent for mændenes vedkommende kun er på 36,86%. Der er næsten 10%-point i forskel på de to satser, og jeg er meget overrasket over, at det er kvinderne,

der er så positive – normalt vil man kendetegne kvinderne for at være yderst kritiske, så det må være et cadeau til McDonald's at andelen er så høj. Den samme tendens er at finde i den modsatte ende af skalaen, hvor 6,27% af mændene giver en karakter på fem eller derunder, mens det tilsvarende tal for kvinderne er på 4,34%. Man kan altså konkludere, at de kvindelige gæster er mere tilfredse med deres besøg hos McDonald's end deres mandlige gæster.

Ser man på vigtighed opgjort i køn fremkommer følgende krydstabel:

Table of Koen by Vigtig											
Koen	Vigtig										
	Tempof	Tempoef	Betjen	Ren	Tempmad	Smag	Frisk	Ordekor	Atmos	Dtkom	Total
Mand	289	281	473	296	373	846	549	156	46	41	3350
	5.17	5.03	8.46	5.30	6.68	15.14	9.82	2.79	0.82	0.73	59.95
	8.63	8.39	14.12	8.84	11.13	25.25	16.39	4.66	1.37	1.22	
	72.61	72.42	58.32	51.75	60.95	66.41	47.78	67.53	68.66	47.67	
Kvinde	109	107	338	276	239	428	600	75	21	45	2238
	1.95	1.91	6.05	4.94	4.28	7.66	10.74	1.34	0.38	0.81	40.05
	4.87	4.78	15.10	12.33	10.68	19.12	26.81	3.35	0.94	2.01	
	27.39	27.58	41.68	48.25	39.05	33.59	52.22	32.47	31.34	52.33	
Total	398	388	811	572	612	1274	1149	231	67	86	5588
	7.12	6.94	14.51	10.24	10.95	22.80	20.56	4.13	1.20	1.54	100.00
Frequency Missing = 2157											

Her fremgår det, at der er forskel på de to køns prioriteringer. Mændene finder at smagen er vigtigst (25,25%), mens det er friskheden af maden, der er vigtigst for kvinderne (26,81%). Dog er kønnene enige om, at det er disse to parametre, der er de to vigtigste. Noget andet man

også kan se ud fra skemaet er, at mændene prioriterer tempoet af betjeningen dobbelt så højt som kvinderne, mens kvinderne prioritere renligheden af restauranten højere end mændene.

7.7.3 Besøgsfrekvens og tilfredshed/vigtig

Ser jeg nu på krydstabeller for besøgsfrekvensen, kan jeg se, om der er forskel i tilfredsheden blandt de gæster, der kommer ofte hos McDonald's og dem, der kommer knap så tit, samt se om der er forskel på det de prioriterer som vigtigt.

Table of Besoegsfrekvens by Tilfreds											
Besoegsfrekvens	Tilfreds										Total
	1	2	3	4	5	6	7	8	9	10	
4+	8	7	17	17	93	131	299	611	404	981	2568
	0.11	0.10	0.23	0.23	1.27	1.79	4.08	8.33	5.51	13.37	35.01
	0.31	0.27	0.66	0.66	3.62	5.10	11.64	23.79	15.73	38.20	
	28.57	31.82	45.95	23.94	37.65	39.94	39.71	36.48	33.98	32.86	
1-3	8	6	15	37	94	142	303	783	573	1224	3185
	0.11	0.08	0.20	0.50	1.28	1.94	4.13	10.67	7.81	16.69	43.42
	0.25	0.19	0.47	1.16	2.95	4.46	9.51	24.58	17.99	38.43	
	28.57	27.27	40.54	52.11	38.06	43.29	40.24	46.75	48.19	41.01	
Sjældnere	12	9	5	17	60	55	151	281	212	780	1582
	0.16	0.12	0.07	0.23	0.82	0.75	2.06	3.83	2.89	10.63	21.57
	0.76	0.57	0.32	1.07	3.79	3.48	9.54	17.76	13.40	49.30	
	42.86	40.91	13.51	23.94	24.29	16.77	20.05	16.78	17.83	26.13	
Total	28	22	37	71	247	328	753	1675	1189	2985	7335
	0.38	0.30	0.50	0.97	3.37	4.47	10.27	22.84	16.21	40.70	100.00
Frequency Missing = 410											

Endnu engang forbløffes jeg af de resultater, som jeg kan læse af tabellen til venstre. Her fremgår det, at 49,30% af de gæster, som befinder sig i kategorien "sjældnere" giver tilfredsheden et 10-tal, det er altså næsten halvdelen af de gæster, der kommer sjældent hos McDonald's som er yderst tilfreds med deres besøg. Ser

man på det tilsvarende tal for dem, der besøger restauranten oftere, er det på henholdsvis 38,43% og 38,20%, altså et noget lavere niveau. Jeg havde troet, at disse procenter var højere end for gruppen "sjældnere", da de netop kommer i restauranten oftere, men det er åbenbart ikke repeat business pga. tilfredsheden!

Tabellerne nedenfor angiver vigtigt ift. Besøgsfrekvensen:

Table of Besoegsfrekvens by Vigtig							
Besoegsfrekvens	Vigtig						
	Tempof	Tempoef	Betjen	Ren	Tempmad	Smag	Frisk
4+	132	135	313	219	192	453	457
	2.34	2.39	5.55	3.88	3.40	8.03	8.10
	6.55	6.70	15.53	10.87	9.53	22.48	22.68
	33.08	34.79	38.36	38.02	31.32	35.01	39.33
1-3	168	181	330	253	281	599	524
	2.98	3.21	5.85	4.49	4.98	10.62	9.29
	6.66	7.17	13.08	10.03	11.14	23.74	20.77
	42.11	46.65	40.44	43.92	45.84	46.29	45.09
Sjældnere	99	72	173	104	140	242	181
	1.76	1.28	3.07	1.84	2.48	4.29	3.21
	8.98	6.53	15.70	9.44	12.70	21.96	16.42
	24.81	18.56	21.20	18.06	22.84	18.70	15.58
Total	399	388	816	576	613	1294	1162
	7.07	6.88	14.47	10.21	10.87	22.94	20.60
Frequency Missing = 2105							

Table of Besoegsfrekvens by Vigtig				
Besoegsfrekvens	Vigtig			
	Ordekor	Atmos	Dtkom	Total
4+	72	19	23	2015
	1.28	0.34	0.41	35.73
	3.57	0.94	1.14	
	30.25	27.94	26.74	
1-3	119	30	38	2523
	2.11	0.53	0.67	44.73
	4.72	1.19	1.51	
	50.00	44.12	44.19	
Sjældnere	47	19	25	1102
	0.83	0.34	0.44	19.54
	4.26	1.72	2.27	
	19.75	27.94	29.07	
Total	238	68	86	5640
	4.22	1.21	1.52	100.00
Frequency Missing = 2105				

Her fremgår det, at dem der besøger McDonald's mere end 4 gange om måneden, prioriterer friskheden som det vigtigste, dog efterfulgt af smagen med kun 0,20%-point i forskel, mens dem, der kommer sjældnere end 4 gange om måneden, prioriterer smagen højest.

Der er altså ingen tvivl om, at uanset, hvordan jeg sammensætter krydstabellerne, er det smag og friskhed, der er vigtigst for gæsterne.

7.8 Krydstabeller for McDonald's Brøndby

Det kunne være interessant at se, hvordan kundetilfredsheden fordeler sig i forhold til alder. Hertil kan jeg se på krydstabellen for alder×tilfredshed:

Frequency Percent Row Pct Col Pct	Table of Alder by Tilfreds										
	Alder	Tilfreds									Total
		1	3	4	5	6	7	8	9	10	
	13-19 år	2 0.43 1.17 66.67	0 0.00 0.00 0.00	2 0.43 1.17 22.22	12 2.56 7.02 52.17	10 2.14 5.85 29.41	20 4.27 11.70 36.36	17 3.63 9.94 21.79	25 5.34 14.62 31.65	83 17.74 48.54 45.11	171 36.54
	20-29 år	0 0.00 0.00 0.00	0 0.00 0.00 0.00	4 0.85 3.67 44.44	7 1.50 6.42 30.43	6 1.28 5.50 17.65	7 1.50 6.42 12.73	25 5.34 22.94 32.05	20 4.27 18.35 25.32	40 8.55 36.70 21.74	109 23.29
	30-39 år	1 0.21 1.05 33.33	3 0.64 3.16 100.00	2 0.43 2.11 22.22	1 0.21 1.05 4.35	9 1.92 9.47 26.47	17 3.63 17.89 30.91	13 2.78 13.68 16.67	19 4.06 20.00 24.05	30 6.41 31.58 16.30	95 20.30
	40+	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.21 1.08 11.11	3 0.64 3.23 13.04	9 1.92 9.68 26.47	11 2.35 11.83 20.00	23 4.91 24.73 29.49	15 3.21 16.13 18.99	31 6.62 33.33 16.85	93 19.87
	Total	3 0.64	3 0.64	9 1.92	23 4.91	34 7.26	55 11.75	78 16.67	79 16.88	184 39.32	468 100.00
Frequency Missing = 121											

Af tabellen fremgår det, at hele 39,32% af de adspurgte har været særdeles tilfredse (10) med deres besøg hos McDonald's Brøndby – hvis jeg regner med at en god karakter er fra 8-10, kan denne procent øges til hele 72,87%. Ser jeg derimod på den modsatte ende af skalaen, er der kun 8,11% af gæsterne, der har givet en score på 5 eller derunder.

Det fremgår også af tabellen, at det er den unge aldersgruppe, der trækker godt op i den høje karakter. De udgør hele 45,11% af den samlede topscoreprocent (39,32%) og det fremgår også, at det er netop denne celle, der har den højeste procentsats, med hele 17,74% af alle de afgivne stemmer.

Ser jeg nu på tilfredsheden i forhold til køn, fremkommer følgende tabel:

Table of Koen by Tilfreds										
Koen	Tilfreds									
	1	3	4	5	6	7	8	9	10	Total
Mand	2	4	9	15	20	47	54	56	116	323
	0.37	0.75	1.68	2.80	3.74	8.79	10.09	10.47	21.68	60.37
	0.62	1.24	2.79	4.64	6.19	14.55	16.72	17.34	35.91	
	66.67	100.00	81.82	57.69	52.63	74.60	57.45	63.64	55.77	
Kvinde	1	0	2	11	18	16	40	32	92	212
	0.19	0.00	0.37	2.06	3.36	2.99	7.48	5.98	17.20	39.63
	0.47	0.00	0.94	5.19	8.49	7.55	18.87	15.09	43.40	
	33.33	0.00	18.18	42.31	47.37	25.40	42.55	36.36	44.23	
Total	3	4	11	26	38	63	94	88	208	535
	0.56	0.75	2.06	4.86	7.10	11.78	17.57	16.45	38.88	100.00
Frequency Missing = 54										

Her kan jeg se, at ligesom det var tilfældet for alle restauranterne samlet set, er det kvinderne, der er mest tilfredse med deres besøg hos McDonald's Brøndby. Hvis jeg igen kigger på karakteren 8-10, er andelen af kvinder i dette interval på 77,36%, mens det tilsvarende tal for mændene

er på 69,97%. Det kunne nu være interessant at se på, hvor ofte kvinderne kommer i restauranten ift. mændene. Umiddelbart skulle man tro, at jo mere tilfreds man er, jo oftere kommer man i restauranten. Hertil kan jeg kigge på krydstabellen koen*besoegsfrekvens.

7.8.3 Tilfredshed og besøgsfrekvens

I starten af denne opgave så jeg på en teori, der fortalte, at jo mere tilfredse gæsterne var, jo oftere var sandsynligheden for, at de ville komme tilbage.

Table of Besoegsfrekvens by Tilfreds										
Besoegsfrekvens	Tilfreds									
	1	3	4	5	6	7	8	9	10	Total
4+	1	0	2	11	15	18	33	22	75	177
	0.18	0.00	0.37	2.01	2.75	3.30	6.04	4.03	13.74	32.42
	0.56	0.00	1.13	6.21	8.47	10.17	18.64	12.43	42.37	
	33.33	0.00	18.18	42.31	38.46	28.13	34.02	25.00	35.05	
1-3	0	4	7	9	15	34	47	45	81	242
	0.00	0.73	1.28	1.65	2.75	6.23	8.61	8.24	14.84	44.32
	0.00	1.65	2.89	3.72	6.20	14.05	19.42	18.60	33.47	
	0.00	100.00	63.64	34.62	38.46	53.13	48.45	51.14	37.85	
Sjældnere	2	0	2	6	9	12	17	21	58	127
	0.37	0.00	0.37	1.10	1.65	2.20	3.11	3.85	10.62	23.26
	1.57	0.00	1.57	4.72	7.09	9.45	13.39	16.54	45.67	
	66.67	0.00	18.18	23.08	23.08	18.75	17.53	23.86	27.10	
Total	3	4	11	26	39	64	97	88	214	546
	0.55	0.73	2.01	4.76	7.14	11.72	17.77	16.12	39.19	100.00
Frequency Missing = 43										

Hvis jeg kigger på krydstabellen til venstre, fremgår det, at den gruppe, der opnår den højeste andel af karakteren 10, er dem, som kommer sjældnere i restauranten. Dette kan umiddelbart godt tyde på at teorien ikke holder, men, der er lige nogle mangler i denne tabel, som der ikke er taget højde for. For det første

fortæller teorien ikke *hvor ofte* gæsterne skal komme igen, blot at de kommer tilbage, og for det andet, kan jeg heller ikke aflæse det beløb, som gæsten har købt for. Det er ikke utænkeligt at dem, som kommer 4 gange om måneden, er dem som køber for et meget lille beløb, mens dem som kommer sjældnere, køber for et stort beløb – måske har de gjort det til en familietur. Set i dette lys, er det meget positivt, at man opnår en så høj score i denne kategori. Desværre

opnår kategorien af gæster, som kommer sjældnere, også en dårlig score i den utilfredse del af tabellen.

7.8.4 Tidsrum og tilfredshed:

Hvornår er gæsterne mest tilfredse? Kan dette spørgsmål besvares, kan man identificere, hvad der er sket på det pågældende tidspunkt, og videreføre det til de andre tidspunkter. Det er et godt redskab til restauranten at finde frem til tidspunkterne, således at man kan lave best practice. Det kan også være, at man allerede i restauranten ved, at man har gjort noget anderledes i dette tidsrum (og denne dag) og nu kan man se, hvordan gæsterne reagerer herpå.

Table of Tidsrum by Tilfreds											
Tidsrum	Tilfreds										Total
	1	3	4	5	6	7	8	9	10		
Før kl. 14.00	1	0	3	6	7	14	28	39	53	151	
	0.19	0.00	0.57	1.13	1.32	2.65	5.29	7.37	10.02	28.54	
	0.66	0.00	1.99	3.97	4.64	9.27	18.54	25.83	35.10		
	33.33	0.00	27.27	23.08	18.42	23.33	29.17	44.83	25.98		
14.00 - 17.00	0	1	2	7	6	16	21	21	42	116	
	0.00	0.19	0.38	1.32	1.13	3.02	3.97	3.97	7.94	21.93	
	0.00	0.86	1.72	6.03	5.17	13.79	18.10	18.10	36.21		
	0.00	25.00	18.18	26.92	15.79	26.67	21.88	24.14	20.59		
17.00 - 20.00	1	3	6	8	21	26	43	19	83	210	
	0.19	0.57	1.13	1.51	3.97	4.91	8.13	3.59	15.69	39.70	
	0.48	1.43	2.86	3.81	10.00	12.38	20.48	9.05	39.52		
	33.33	75.00	54.55	30.77	55.26	43.33	44.79	21.84	40.69		
Efter kl. 20.00	1	0	0	5	4	4	4	8	26	52	
	0.19	0.00	0.00	0.95	0.76	0.76	0.76	1.51	4.91	9.83	
	1.92	0.00	0.00	9.62	7.69	7.69	7.69	15.38	50.00		
	33.33	0.00	0.00	19.23	10.53	6.67	4.17	9.20	12.75		
Total	3	4	11	26	38	60	96	87	204	529	
	0.57	0.76	2.08	4.91	7.18	11.34	18.15	16.45	38.56	100.00	
Frequency Missing = 60											

Det er interessant at 40,69% af de afgivne 10-taller er givet i tidsrummet 17-20, mens man kun opnår 21,84% af 9-tallerne i dette tidsrum. Det kunne tyde på, at man i dette tidsrum er meget bevidst omkring differentieringen af tilfredsheden. En anden interessant observation man kan gøre sig er, at 50% af de afgivne karakterer i tidsrummet efter kl. 20 er på 10. Tilsvarende, er der kun 1,92% af de afgivne stemmer, der er under 5.

Enten er restauranten superb i tidsrummet efter kl. 20, ellers er gæsterne i dette tidsrum nemmere at imponere. Det er nok det sidstnævnte, der er tilfældet – de fleste ved godt at når man kommer udenfor "normale" spisetider, kan man ikke forvente at alt spiller. Ens forventninger er sænket, så hvis restauranten outperformer disse forventninger, vil man opnå en god tilfredshed, og det ser jo ud til at være tilfældet her.

7.8.5 Børn og tilfredshed:

McDonald's er en familierestaurant, og den er i særdeleshed kendt som en favorit blandt børnefamilier. Her får man sin mad hurtigt, og der er legeplads til børnene, således at forældrene kan få lidt fred. Men afspejler denne tolkning sig i de tal, jeg kan se i tabellen til

Table of Boern by Tilfreds										
Boern	Tilfreds									
	1	3	4	5	6	7	8	9	10	Total
Ja	2	1	8	21	26	42	67	57	149	373
	0.39	0.20	1.57	4.13	5.11	8.25	13.16	11.20	29.27	73.28
	0.54	0.27	2.14	5.63	6.97	11.26	17.96	15.28	39.95	
	66.67	33.33	72.73	80.77	70.27	71.19	77.91	67.06	74.87	
Nej	1	2	3	5	11	17	19	28	50	136
	0.20	0.39	0.59	0.98	2.16	3.34	3.73	5.50	9.82	26.72
	0.74	1.47	2.21	3.68	8.09	12.50	13.97	20.59	36.76	
	33.33	66.67	27.27	19.23	29.73	28.81	22.09	32.94	25.13	
Total	3	3	11	26	37	59	86	85	199	509
	0.59	0.59	2.16	5.11	7.27	11.59	16.90	16.70	39.10	100.00
Frequency Missing = 80										

venstre? Her fremgår det, at 39,95% af voksne som har haft børn med på deres besøg, er særdeles tilfredse (karakteren 10). Det tilsvarende tal for voksne alene er 36,76%. Ser man på de tre højeste karakterer er andelen 73,19% med børn og 71,32% uden børn. Det er altså ude på marginalerne at man kan sige, at

børnefamilier er mere tilfredse end resten af gæsterne hos McDonald's.

7.8.6 Placering og tilfredshed

Er der forskel på tilfredsheden blandt de gæster, der besøger restauranten i McDrive og dem, der kommer ind i restauranten? Der skulle helst ikke være en forskel i tilfredsheden blandt gæsterne, men ser jeg på tabellen nedenfor, er det tydeligt at se, at der er en forskel.

Table of Placering by Tilfreds										
Placering	Tilfreds									
	1	3	4	5	6	7	8	9	10	Total
Restauranten	3	2	3	16	26	52	73	73	182	430
	0.57	0.38	0.57	3.02	4.91	9.83	13.80	13.80	34.40	81.29
	0.70	0.47	0.70	3.72	6.05	12.09	16.98	16.98	42.33	
	100.00	66.67	33.33	66.67	72.22	82.54	76.84	82.95	87.50	
McDrive	0	1	6	8	10	11	22	15	26	99
	0.00	0.19	1.13	1.51	1.89	2.08	4.16	2.84	4.91	18.71
	0.00	1.01	6.06	8.08	10.10	11.11	22.22	15.15	26.26	
	0.00	33.33	66.67	33.33	27.78	17.46	23.16	17.05	12.50	
Total	3	3	9	24	36	63	95	88	208	529
	0.57	0.57	1.70	4.54	6.81	11.91	17.96	16.64	39.32	100.00
Frequency Missing = 60										

I McDrive opnår man kun karakteren 10 i 26,26% af svarene, mens den tilsvarende procentsats for instore gæster er 42,33%. Der er altså en markant forskel, og den bliver ikke udjævnet ved at tage de 3 top-karakterer under ét, det giver hhv. 63,63% for McDrive og

76,29% for instore gæster. Tilgængæld ligger der rigtig mange "svar" i midterkategoriene, hvilket tyder på, at gæsterne har været nogenlunde tilfredse i McDrive, men der mangler lige den sidste tand for at bringe dem helt til tops. Det ville være ideelt at nedsætte en fokusgruppe, som kunne forklare, hvad der lige netop manglede. Et alternativ hertil er at se på hvad McDrive gæsterne finder vigtigst og så se, om restauranten performer på denne parameter.

Jeg har her set på nogle af de vigtigste krydstabeller, men det er klart at der sagtens kunne produceres mange flere. Det vil dog være alt for omfattende og langtrukket at tage flere af disse tabeller med ind i denne opgave, så jeg har valgt at stoppe her. Dog skal det lige pointeres, at en krydstabel ikke nødvendigvis behøver at være to-dimensional. Man kan

sagtens lave flere dimensioner af tabeller, men det bliver bare vanskeligere at finde rundt i. Nedenfor er angivet en tre-dimensional tabel, der angiver ugedag×tidsrum×tilfredshed, men som man kan se er der nu fire tabeller man skal holde styr på frem for én.

Table 1 of Tidsrum by Tilfreds												
Controlling for Ugedag=Torsdag												
Tidsrum	Tilfreds										Total	
	1	3	4	5	6	7	8	9	10			
Før kl. 14.00	0 0.00 0.00 0.00	0 0.00 0.00 .	3 1.91 5.08 75.00	2 1.27 3.39 33.33	2 1.27 3.39 28.57	6 3.82 10.17 28.57	13 8.28 22.03 46.43	17 10.83 28.81 54.84	16 10.19 27.12 27.12	59 37.58		
14.00 - 17.00	0 0.00 0.00 0.00	0 0.00 0.00 .	0 0.00 0.00 0.00	2 1.27 6.45 33.33	2 1.27 6.45 28.57	3 1.91 9.68 14.29	4 2.55 12.90 14.29	5 3.18 16.13 16.13	15 9.55 48.39 25.42	31 19.75		
17.00 - 20.00	0 0.00 0.00 0.00	0 0.00 0.00 .	1 0.64 2.13 25.00	1 0.64 2.13 16.67	1 0.64 2.13 14.29	10 6.37 21.28 47.62	10 6.37 21.28 35.71	4 2.55 8.51 12.90	20 12.74 42.55 33.90	47 29.94		
Efter kl. 20.00	1 0.64 5.00 100.00	0 0.00 0.00 .	0 0.00 0.00 0.00	1 0.64 5.00 16.67	2 1.27 10.00 28.57	2 1.27 10.00 9.52	1 0.64 5.00 3.57	5 3.18 25.00 16.13	8 5.10 40.00 13.56	20 12.74		
Total	1 0.64	0 0.00	4 2.55	6 3.82	7 4.46	21 13.38	28 17.83	31 19.75	59 37.58	157 100.00		
Frequency Missing = 14												

Table 2 of Tidsrum by Tilfreds												
Controlling for Ugedag=Fredag												
Tidsrum	Tilfreds										Total	
	1	3	4	5	6	7	8	9	10			
Før kl. 14.00	1	0	0	0	3	4	7	12	23	50		
	0.81	0.00	0.00	0.00	2.44	3.25	5.69	9.76	18.70	40.65		
	2.00	0.00	0.00	0.00	6.00	8.00	14.00	24.00	46.00			
	50.00	.	0.00	0.00	75.00	33.33	38.89	50.00	41.07			
14.00 - 17.00	0	0	1	2	0	4	5	9	12	33		
	0.00	0.00	0.81	1.63	0.00	3.25	4.07	7.32	9.76	26.83		
	0.00	0.00	3.03	6.06	0.00	12.12	15.15	27.27	36.36			
	0.00	.	50.00	40.00	0.00	33.33	27.78	37.50	21.43			
17.00 - 20.00	1	0	1	3	1	4	6	3	20	39		
	0.81	0.00	0.81	2.44	0.81	3.25	4.88	2.44	16.26	31.71		
	2.56	0.00	2.56	7.69	2.56	10.26	15.38	7.69	51.28			
	50.00	.	50.00	60.00	25.00	33.33	33.33	12.50	35.71			
Efter kl. 20.00	0	0	0	0	0	0	0	0	1	1		
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.81	0.81		
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00			
	0.00	.	0.00	0.00	0.00	0.00	0.00	0.00	1.79			
Total	2	0	2	5	4	12	18	24	56	123		
	1.63	0.00	1.63	4.07	3.25	9.76	14.63	19.51	45.53	100.00		
Frequency Missing = 11												

Table 3 of Tidsrum by Tilfreds												
Controlling for Ugedag=Lørdag												
Tidsrum	Tilfreds										Total	
	1	3	4	5	6	7	8	9	10			
Før kl. 14.00	0 0.00 0.00 .	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	3 2.44 21.43 17.65	5 4.07 35.71 18.52	1 0.81 7.14 6.67	5 4.07 35.71 11.90	14 11.38		
14.00 - 17.00	0 0.00 0.00 .	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	3 2.44 10.71 25.00	5 4.07 17.86 29.41	7 5.69 25.00 25.93	5 4.07 17.86 33.33	8 6.50 28.57 19.05	28 22.76		
17.00 - 20.00	0 0.00 0.00 .	3 2.44 5.00 100.00	3 2.44 5.00 100.00	2 1.63 3.33 50.00	9 7.32 15.00 75.00	7 5.69 11.67 41.18	13 10.57 21.67 48.15	6 4.88 10.00 40.00	17 13.82 28.33 40.48	60 48.78		
Efter kl. 20.00	0 0.00 0.00 .	0 0.00 0.00 0.00	0 0.00 0.00 0.00	2 1.63 9.52 50.00	0 0.00 0.00 0.00	2 1.63 9.52 11.76	2 1.63 9.52 7.41	3 2.44 14.29 20.00	12 9.76 57.14 28.57	21 17.07		
Total	0 0.00	3 2.44	3 2.44	4 3.25	12 9.76	17 13.82	27 21.95	15 12.20	42 34.15	123 100.00		
Frequency Missing = 9												

Table 4 of Tidsrum by Tilfreds												
Controlling for Ugedag=Søndag												
Tidsrum	Tilfreds										Total	
	1	3	4	5	6	7	8	9	10			
Før kl. 14.00	0 0.00 0.00 .	0 0.00 0.00 .	0 0.00 0.00 .	0 1.96 11.11 28.57	2 0.98 5.56 7.69	1 0.98 5.56 10.00	2 1.96 11.11 9.52	8 7.84 44.44 53.33	4 3.92 22.22 11.76	18 17.65		
14.00 - 17.00	0 0.00 0.00 .	0 0.00 0.00 .	1 0.98 5.00 50.00	2 1.96 10.00 28.57	1 0.98 5.00 7.69	4 3.92 20.00 40.00	5 4.90 25.00 23.81	1 0.98 5.00 6.67	6 5.88 30.00 17.65	20 19.61		
17.00 - 20.00	0 0.00 0.00 .	0 0.00 0.00 .	1 0.98 1.82 50.00	9 8.82 16.36 14.29	1 0.98 1.82 69.23	5 4.90 9.09 50.00	13 12.75 23.64 61.90	6 5.88 10.91 40.00	20 19.61 36.36 58.82	55 53.92		
Efter kl. 20.00	0 0.00 0.00 .	0 0.00 0.00 .	0 0.00 0.00 0.00	2 1.96 22.22 28.57	2 1.96 22.22 15.38	0 0.00 0.00 0.00	1 0.98 11.11 4.76	0 0.00 0.00 0.00	4 3.92 44.44 11.76	9 8.82		
Total	0 0.00	0 0.00	2 1.96	7 6.86	13 12.75	10 9.80	21 20.59	15 14.71	34 33.33	102 100.00		
Frequency Missing = 8												

8.0 Teori

I dette afsnit vil jeg se på den teori, som jeg vil benytte mig af til analysen af spørgeskemaerne. Det drejer sig om den teori, der findes for kategoriserede data.

8.1. Kontingenstabeller

Hvis man har kategoriske data – data, som kan inddeles i kategorier efter forskellige kriterier, kan man opskrive disse data i en tabel. Hvis cellerne i tabellen indeholder frekvenser, kaldes denne for en kontingenstabel. Tabellen viser sammenhænge mellem de kategoriske variable, og er f.eks. af størrelsen $I \times J$, hvor I er antal niveauer af variabel X og J er antal niveauer af variabel Y . π_{ij} er sandsynligheden for at (X, Y) opstår i cellen identificeret ved række i og søjle j .

Sandsynlighedsfordelingen $\{\pi_{ij}\}$ er den *fælles fordeling* af X og Y . Den *marginale fordeling* er givet ved række og søjletotaler af den fælles fordeling, defineret som: $\pi_{i+} = \sum_j \pi_{ij}$ og

$\pi_{+j} = \sum_i \pi_{ij}$, som opfylder:

$$\sum_i \pi_{i+} = \sum_j \pi_{+j} = \sum_i \sum_j \pi_{ij} = 1$$

Hvis den ene variabel er respons, og den anden variabel er en forklarende variabel, som er fast, er der ikke længere tale om en fælles fordeling, men responsen har en sandsynlighedsfordeling. Hvorledes denne sandsynlighedsfordeling ændres, når den forklarende variabel ændres, undersøges ved sandsynlighederne $\{\pi_{1i}, \dots, \pi_{ji}\}$, som danne den *betingede fordeling*. Tabel 1 viser, hvorledes ovenstående kan opstilles:

Variabel X	Variabel Y				
	Niveau 1	Niveau 2	...	J	Total
Niveau I	π_{11}	π_{12}	...	π_{1J}	π_{1+}
	(π_{11})	(π_{21})	...	(π_{J1})	1
Niveau II	π_{21}	π_{22}		π_{2J}	π_{2+}
	(π_{12})	(π_{22})		(π_{J2})	1
.	.	.		.	.
.	.	.		.	.
.	.	.		.	.
I	π_{I1}	π_{I2}	...	π_{IJ}	π_{I+}
	(π_{1I})	(π_{2I})	...	(π_{JI})	1
Total	π_{+1}	π_{+2}	...	π_{+J}	1

8.2 Odds Ratio

Hvis π er sandsynligheden for succes, er odds defineret som: $\Omega = \frac{\pi}{1 - \pi}$

Odds er ikke negative og hvis odds er større end 1, er sandsynligheden for succes større end sandsynligheden for fiasko. For tabeller med celledsandsynligheder π_{ij} udregnes odds ratio ofte.

Dette er for en 2x2 tabel givet ved:
$$\theta^{OR} = \frac{\frac{\pi_{11}}{\pi_{21}}}{\frac{\pi_{12}}{\pi_{22}}} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}$$

Hvis alle celledsandsynlighederne er positive, vil $\theta^{OR} = 1$ svare til, at X og Y er uafhængige. Hvis $1 \leq \theta^{OR} < \infty$ er odds for succes i række 1 større end odds for succes i række 2, og omvendt når $0 \leq \theta^{OR} < 1$. Ofte kan det være en fordel, at betragte log odds ratio (logits), hvor $\log(\theta^{OR}) = 0$ svarer til uafhængighed mellem X og Y. Log odds er symmetrisk omkring 0, dvs. at $\log(5) = 0,69897$ er lig med $\log(1/5)$ blot med modsat fortegn. Odds ratio kan også beregnes vha.

cellefrekvenser x_{ij} , således vil stikprøvens odds ratio kunne beregnes som: $\hat{\theta}^{OR} = \frac{x_{11} \cdot x_{22}}{x_{12} \cdot x_{21}}$.

Dette estimat ændres ikke ved at rækker eller søjler multipliceres med en konstant forskellig fra nul. Odds Ratio kan også defineres generelt for en $I \times J$ tabel som:

$$\theta_{ij}^{OR} = \frac{\pi_{ij} \pi_{kl}}{\pi_{il} \pi_{kj}}, \text{ hvor } i, j = 1, 2, \dots, I-1 \text{ og } k, l = 1, 2, \dots, J.$$

8.3 Test for uafhængighed

Hvis n uafhængige observationer kan krydsklassificeres efter ovenstående faktorer X og Y , kan det observerede antal i celle (i, j) betegnes x_{ij} . Hvis alle enheder i stikprøven er uafhængige, så vil x_{ij} være fra en multinomialfordeling¹² med $I \times J$ kategorier. Disse antagelser er opfyldt, hvis stikprøven er en almindelig SRS med tilbagelægning, og tilnærmelsesvis opfyldt for en SRS uden tilbagelægning, hvis stikprøvestørrelsen n er lille i forhold til populationens størrelse.

Nulhypotesen for uafhængighed er givet ved:

$$H_0 : \pi_{ij} = \pi_{j+} \pi_{+i} \text{ for } i = 1, \dots, I \text{ og } j = 1, \dots, J \quad (2.1)$$

Hvis det forventede antal betegnes med $\mu_{ij} = n\pi_{ij}$, så vil $\mu_{ij} = n\pi_{i+} \pi_{+j}$ og

$$\hat{\mu}_{ij} = n\hat{\pi}_{i+} \hat{\pi}_{+j} = n \frac{x_{i+}}{n} \frac{x_{+j}}{n}$$

$$\text{Hvor } \hat{\pi}_{ij} = \frac{x_{ij}}{n} \text{ og } \hat{\pi}_{i+} = \sum_{j=1}^J \hat{\pi}_{ij}$$

Pearsons χ^2 teststørrelse er givet ved:

$$X^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(x_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}} = n \sum_{i=1}^I \sum_{j=1}^J \frac{(\hat{\pi}_{ij} - \hat{\pi}_{i+} \hat{\pi}_{+j})^2}{\hat{\pi}_{i+} \hat{\pi}_{+j}} \quad (2.2)$$

Likelihood Ratio teststørrelsen er givet ved:

¹² Generalisering af binomialfordelingen, hvor det enkelte forsøg har mere end to mulige udfald.

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J x_{ij} \ln \left(\frac{x_{ij}}{\hat{\mu}_{ij}} \right) = 2n \sum_{i=1}^I \sum_{j=1}^J \hat{\pi}_{ij} \ln \left(\frac{\hat{\pi}_{ij}}{\hat{\pi}_{i+} \hat{\pi}_{+j}} \right) \quad (2.3)$$

Hvis n er tilpas stor, vil X^2 og G^2 under nulhypotesen være approksimativt χ^2 – fordelt med $(J-1)(I-1)$ frihedsgrader (df). Den mindste værdi n kan have, for at være tilpas stor ifølge Fienberg (1979) beregnes som:

1. Det forventede antal i hver celle > 1
2. $n \geq 5 \times$ antallet af celler

H_0 -hypotesen kan også for en $I \times J$ tabel skrives vha. odds ratio, således at:

$$H_0 : \frac{\pi_{ij}\pi_{kl}}{\pi_{il}\pi_{kj}} = 1 \text{ for alle } i,j=1,2,\dots,I-1 \text{ og } k,l=1,2,\dots,J$$

Odds ratio kan estimeres vha. de estimerede proportioner:
$$\frac{\hat{\pi}_{ij}\hat{\pi}_{kl}}{\hat{\pi}_{il}\hat{\pi}_{kj}} \quad (2.4)$$

For et tilpas stort n , vil logaritmen til odds ratio være approksimativt normalfordelt med standard afvigelse givet ved:
$$\sqrt{\frac{1}{x_{ij}} + \frac{1}{x_{kl}} + \frac{1}{x_{il}} + \frac{1}{x_{kj}}}$$

Rao og Scott (1984) har vist at for generelle design, er X^2 og G^2 asymptotisk fordelt som en vægtet sum givet ved: $\delta_1 W_1 + \dots + \delta_{c-r-1} W_{c-r-1}$,

W_i er $c-r-1$ uafhængige χ^2 – fordelte tilfældige variable, hvor c er givet ved antallet af celler i tabellen, og r er antallet af rækker. Vægtene δ_i er egenverdierne af den generaliserede designeffekt matrice Δ ¹³.

8.4 Loglineære modeller

Loglineære modeller benyttes primært, når mindst to variable er responsvariable. Formålet med loglineære modeller er, at analysere association og interaktionsmønstre.

Loglineære modeller for to-vejs-tabeller

¹³ Chambers, R. L. & Skinner, C. J. (2003): Analysis of Survey Data s. 87

Hvis man betragter tilfældet, hvor to kategoriske responser krydsklassificeres i en $I \times J$ tabel er cellesandsynlighederne givet ved $\{\pi_{ij}\}$ og de forventede frekvenser er givet ved $\{\mu_{ij} = n\pi_{ij}\}$, hvor n er antal observationer. Hvis rækkevariablen kaldes X og søjlevariablen for Y , vil den loglineære model for statistisk uafhængighed se således ud:

$$\ln \pi_{ij} = \ln \pi_{i+} + \ln \pi_{+j} = \lambda + \lambda_i^X + \lambda_j^Y$$

hvor λ_i^X er række effekt og λ_j^Y er søjle effekt, og $\sum_{i=1}^I \lambda_i^X = 0$ og $\sum_{j=1}^J \lambda_j^Y = 0$

Modellen for uafhængighed kan også skrives på matrixform, som $\bar{y} = \bar{X}\bar{\beta}$, hvor disse for en 2×2 tabel er givet ved¹⁴:

$$\bar{y} = \begin{bmatrix} \ln \pi_{11} \\ \ln \pi_{12} \\ \ln \pi_{21} \\ \ln \pi_{22} \end{bmatrix}, \quad \bar{X} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \quad \text{og} \quad \bar{\beta} = \begin{bmatrix} \lambda \\ \lambda_1^X \\ \lambda_1^Y \end{bmatrix}$$

Parametrene $\bar{\beta}$ estimeres ved brug af de estimerede cellesandsynligheder $\hat{\pi}_{ij}$ og ud fra de estimerede parametre $\bar{\beta}$ kan de fittede værdier for $\hat{\pi}_{ij}$ findes som:

$$\hat{\pi}_{ij} = \exp(\hat{\lambda} + \hat{\lambda}_i^X + \hat{\lambda}_j^Y)$$

Er de to variable derimod statistiske afhængige, får man den fulde loglineære model:

$$\log \pi_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY},$$

hvor leddet λ_{ij}^{XY} giver de to variables afhængighed, og $\sum_{i=1}^I \lambda_{ij}^{XY} = \sum_{j=1}^J \lambda_{ij}^{XY} = 0$. Modellen er

hierarkisk, dvs. modellen indeholder alle led af mindre orden, som kan dannes af variable indeholdt i led af højere orden¹⁵. Hvis modellen indeholder interaktionsleddet λ_{ij}^{XY} indeholder den også leddene λ_i^X og λ_j^Y . Dette gælder også for loglineære modeller af højere dimensioner. For at teste, hvor godt modellen for uafhængighed fitter data, kan man for en $I \times J$ tabel

¹⁴ Afsnittet er baseret på Lohr, Sharon (1999): Sampling: Design and Analysis s. 337-339.

¹⁵ Agresti, Alan (2002): Categorical Data Analysis, s. 316.

benytte X^2 og G^2 til at teste goodness of fit. Nulhypotesen for uafhængighed kan også testes vha. den fulde model. I dette tilfælde er nulhypotesen givet ved:

$$H_0 : \lambda_{ij}^{xy} = 0 \quad \text{for } i = 1, \dots, I-1; j = 1, \dots, J-1.$$

Almindelige statistiske programpakker, kan for multinomiale stikprøver beregne estimerne for λ_{ij}^{xy} og de asymptotiske standardafvigelser, og ud fra disse kan hypotesen om uafhængighed afvises eller accepteres.

8.5 Logistisk regression

Logistisk regression kan betragtes som en form for statistisk modellering, som ofte er passende, hvis man har kategorisk responsvariabel¹⁶. Modellerne beskriver sammenhænge mellem en kategorisk responsvariabel og et antal forklarende variable. Logistisk regressions modeller kaldes også logit-modeller. Hvis man kun betragter én responsvariabel, er det simpelt og naturligt at benytte logit-modeller. Disse modeller har ofte en binær responsvariabel, og antager derfor binomialfordelt data, men modellerne kan også have en responsvariabel med mere end to kategorier. Den multikategoriske respons kan være enten nominelt eller ordinalt skaleret, mens de forklarende variable kan være kategoriske eller kontinuerte. I det følgende afsnit beskrives logistisk regression med binær responsvariabel.

Logistisk regression med binær respons¹⁷.

Y er en binær responsvariabel og X er forklarende variabel. Lad

$$\pi(x) = P(Y = 1 | X = x) = 1 - P(Y = 0 | X = x)$$

Den logistiske regressionsmodel er givet ved:

$$\pi(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)}$$

tilsvarende fås log odds også kaldet logit, som har følgende lineære sammenhæng:

$$\text{logit}[\pi(x)] = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \alpha + \beta x$$

¹⁶ Stokes, Maura E, Charles S. Davis & Gary G. Koch (2000): Categorical Data Analysis Using The SAS System, s. 183.

¹⁷ Afsnittet er baseret på: Agresti, Alan (2002): Categorical Data Analysis, s. 165-166.

Parametrene i ovenstående ligning skal fortolkes således: Fortegnet på β bestemmer om $\pi(x)$ stiger eller aftager, når x stiger. Graden af stigning eller aftagning øges, når β øges. Dvs. at hvis $\beta \rightarrow 0$ vil hældningen på grafen for linjen flade ud og nærme sig en vandret linje. Når $\beta = 0$ er Y uafhængig af X .

Logistisk regression kan også udvides til at indeholde kvalitative forklarende variable eller faktorer. Dette illustreres i tilfældet med en enkelt faktor X , som har I kategorier. Hvis Y stadig er en binær responsvariabel, så fås følgende logit model med en faktor: $\pi_i = \pi(x_i)$

$$\text{logit}[\pi_i] = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) = \alpha + \beta_i$$

Jo højere β_i jo højere er π_i . Hvis $\beta_1 = \beta_2 = \dots = \beta_I$, vil modellen kun have skæringen α tilbage, hvilket vil sige, at faktoren ingen effekt har. Dette svarer til $\pi_1 = \pi_2 = \dots = \pi_I$, hvilket betyder statistisk uafhængighed mellem responsen Y og faktoren X . Faktoren X har ligeså mange β_i - parametre som kategorier. Når X har I kategorier er den ene kategori overflødig, men der er $I-1$ ikke overflødige parametre. Derfor kan f.eks. β_i sættes lig 0.

8.6 Dummyvariable

Ovenstående model kan også udtrykkes vha. dummyvariable. Sættes $d_i = 1$ i tilfældet, hvor observationerne falder i række i og $d_i = 0$ ellers, hvor $i = 1, \dots, I-1$, vil modellen kunne skrives som:

$$\text{logit}(\pi_i) = \alpha + \beta_1 d_1 + \beta_2 d_2 + \dots + \beta_{I-1} d_{I-1}$$

Modellen tager højde for parameteroverflødhed ved ikke at danne den dummyvariabel for kategori I .

8.7 Multipel logistisk regression

Her udvides modellen til at inkludere multiple forklarende variable. I tilfældet med p forklarende variable givet ved vektoren $\bar{x} = (x_1, x_2, \dots, x_p)$ kan modellen skrives som:

$$\text{logit}[\pi(\bar{x})] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

og tilsvarende

$$\pi(\bar{x}) = \frac{\exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}$$

Hvis de forklarende variable har flere niveauer, kan modellen igen opskrives vha. dummyvariable. I tilfældet hvor alle forklarende variable har to kategorier kan dummyvariablene skrives som $d_1 = d_2 = \dots = d_p = 1$ når observationen falder i kategori 1 og nul ellers. Hvilket giver modellen:

$$\text{logit}[\pi(\bar{x})] = \alpha + \beta_1 x_1 d_1 + \beta_2 x_2 d_2 + \dots + \beta_p x_p d_p$$

eller tilsvarende

$$\pi(\bar{x}) = \frac{\exp(\alpha + \beta_1 x_1 d_1 + \beta_2 x_2 d_2 + \dots + \beta_p x_p d_p)}{1 + \exp(\alpha + \beta_1 x_1 d_1 + \beta_2 x_2 d_2 + \dots + \beta_p x_p d_p)}$$

Hvis responsvariablen har flere end to niveauer, antager den logistiske regressions model en multinomialfordeling i stedet for en binomialfordeling¹⁸. De forklarende variable kan ligeledes være multikategoriske. Når responsvariablen er kategorisk findes der forskellige statistiske modeller som kan anvendes afhængig af om responsen er ordinal eller nominelt skaleret, men disse beskrives ikke nærmere i denne opgave.

8.8 Kummulative logit modeller

En type af lineære regressionsmodeller specielt egnet til ordinale data er de kummulative logit modeller¹⁹. De udnytter, at der er en rangorden mellem svarkategorierne for responsvariablen, som netop er tilfældet med nogle af responsvariablene (der har flere svarkategorier).

De kummulative logit modeller tager logaritmen til de kummulative sandsynligheder,

$$P(Y \leq j | x) = \pi_1(x) + \dots + \pi_j(x), j = 1, \dots, J$$

De kummulative logits defineres som

¹⁸ Agresti, Alan (2002): Categorical Data Analysis, s. 4

¹⁹ Agresti, Alan (2002): Categorical Data Analysis s.???

$$\text{logit}[P(Y \leq j | x)] = \log \frac{P(Y \leq j | x)}{1 - P(Y \leq j | x)} = \log \frac{\pi_1(x) + \dots + \pi_j(x)}{\pi_{j+1}(x) + \dots + \pi_J(x)}, \quad j = 1, \dots, J-1$$

Hvis responsvariablen er binær svarer det til logistisk regression. $\pi_1(x) + \dots + \pi_j(x)$ svarer så til sandsynligheden for succes, mens $\pi_{j+1}(x) + \dots + \pi_J(x)$ svarer til ikke-succes og ovenstående kan derfor skrives som:

$$\log \frac{\pi(x)}{1 - \pi(x)}$$

Den mest anvendte kumulative logit model for ordinale responser er Proportional Odds-modellen.

8.9 Proportional Odds-modellen

Her antages det, at data kommer fra en multinomial fordeling. Modellen modellerer de kumulative sandsynligheder i stedet for at modellere responskategoriernes sandsynlighed enkeltvis²⁰.

Generelt kan responsvariablen Y antage værdierne 1 til J .

For et givet x (vektor), som er de forklarende variable, der hører til en given respondent, kan man udregne den kumulative sandsynlighed for, at denne persons svar falder i en kategori mindre end eller lig j :

$P(Y \leq j | x) = \pi_1(x) + \dots + \pi_j(x)$, $j = 1, \dots, J-1$, hvor $\pi_j(x)$ er sandsynligheden for at falde i den j 'te kategori.

Det er ikke interessant at sætte j lig J , da $P(Y \leq J | x) = \sum_{j=1}^J \pi_j = 1$, dvs. der er 100%

sandsynlighed for at et svar falder i en af de J kategorier.

Da vi ikke kender sandsynlighederne $\pi_j(x)$ omskrives udtrykket til

$$P(Y \leq j | x) = \frac{e^{\alpha_j + \beta x}}{1 + e^{\alpha_j + \beta x}} .$$

Man kan udregne sandsynligheden for at falde i en bestemt kategori som

²⁰ Agresti, Alan (2002): Categorical Data Analysis, s. ???

$$P(Y = j) = P(Y \leq j) - P(Y \leq j-1), j = 2, \dots, J$$

$$P(Y = 1) = P(Y \leq 1)$$

For at finde de kummulative logits tages logaritmen til ovenstående udtryk

$$\text{logit}[P(Y \leq j | x)] = \log\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right) = \log\left(\frac{e^{\alpha_j + \beta x}}{1 - e^{\alpha_j + \beta x}}\right), j = 1, \dots, J-1$$

Den proportionale odds-model kan derfor skrives som

$$\text{logit}[P(Y \leq j | x)] = \log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \alpha_j + \beta x, j = 1, \dots, J-1$$

Hvis J er lig 2, dvs. hvis responsvariablen Y er binær, ville det svare til at lave helt almindelig logistisk regression.

Fortolkning af α og β : Værdien af β siger noget om, i hvor høj grad den tilhørende forklarende variabel x påvirker responsvariablen. Hvis x stiger med en enhed, vil responsvariablen alt andet lige stige eller falde med β . Om responsen stiger eller falder afhænger af β 's fortegn (positivt fortegn fører til en stigning og omvendt). Jo større β , desto større betydning har den tilhørende forklarende variabel for niveauet af responsvariablen. β angiver den partielle effekt på responsen når variabelen øges med 1, og alle andre variable holdes konstant. β -estimerne kan lægges sammen, så de angiver den samlede effekt på Y , når alle variable øges med 1. Hvis variablene har forskellige skalaer, kan de ikke umiddelbart sammenlignes. Ved at fratrække middelværdien og dividere med standardafvigelsen får man de standardiserede beta-værdier. Disse værdier angiver, hvor stor effekt variablene har i forhold til hinanden, men ikke den reelle effekt på responsvariablen.

Hver kummulativ logit har sit eget intercept α_j , der angiver sandsynligheden for kategori j , hvis der ingen påvirkning er fra de forklarende variable. α_j stiger når j stiger, idet sandsynligheden for at falde i den j 'te kategori dermed stiger for samme x , og logit er en voksende funktion af denne sandsynlighed. Effekten for hver forklarende variabel er den samme for hver kumulativ logit, dvs. estimatet β er altså det samme uanset hvilken responskategori respondenter falder i. Denne antagelse kaldes "the Proportional Odds Assumption".

8.9.1 Odds ratio

For den Proportionale Odds-model gælder, at

$$\text{logit}[P(Y \leq j | x_1)] - \text{logit}[P(Y \leq j | x_2)] = \frac{P(Y \leq j | x = x_1)/P(Y > j | x = x_1)}{P(Y \leq j | x = x_2)/P(Y > j | x = x_2)} = \frac{e^{\alpha_j + \beta x_1}}{e^{\alpha_j + \beta x_2}} = e^{\beta(x_1 - x_2)}$$

hvor x_1 og x_2 er to niveauer af en given forklarende variabel²¹.

Odds Ratio = $OR = e^{\beta(x_2 - x_1)}$, dvs. estimatet for den pågældende forklarende variabel multipliceret med forskellen mellem de to variabelniveauer x_1 og x_2 .

Odds ratio for kumulative sandsynligheder kaldes kumulative odds ratio. Oddsene for at lande i en kategori $\leq j$ når $x = x_1$ er $\exp[\beta(x_1 - x_2)]$ gange oddsene når $x = x_2$.

OR er altså proportional med differencen mellem x_1 og x_2 med β som konstant. Idet β også knytter sig til hvert enkelt kumulativ logit, kaldes modellen for den proportionale odds-model.

8.9.2 Korrelationskoefficienter

Korrelationskoefficienter kan bruges til at vurdere parvise relationer mellem de variable, der bruges i modellen. Hvis korrelationskoefficienten er numerisk 1, betyder det at sammenhængen mellem de to variable er perfekt lineær, dvs. at den ene variabel kan skrives som en konstant multipliceret med den anden variabel. Jo nærmere korrelationskoefficienten er på 0, desto mindre er sammenhængen mellem variablene.

Jo nærmere korrelationskoefficienten er på 1 (numerisk), desto større sammenhæng er der mellem variablene. Hvis værdien er positiv betyder det, at den ene variabel vokser når den anden vokser og falder, når den anden falder. Sammenhængen er modsatrettet, hvis fortegnet er negativt.

For at undersøge om sammenhængen mellem to variable er signifikant, testes hypotesen, at korrelationskoefficienten er lig 0. Hvis testsandsynligheden er mindre end 0,05 (alternativt 0,01), er der stor sandsynlighed for, at der virkelig er en relation mellem de to variable. Hvis

²¹ Agresti, Alan (2002): Categorical Data Analysis s. ??

værdien er større end grænseværdien, kan man ikke konkludere, at der er en sammenhæng, men heller ikke, at der ikke er.

9.0 SAS-programmet (databehandling)

Til at undersøge datasættet, har jeg benyttet mig af statistikprogrammet SAS.

GENMOD proceduren i SAS er et statistisk modelleringsværktøj, der fitter generaliserede lineære modeller for kategoriske data. Modellen kan bl.a. være en logistisk regressionsmodel eller proportional odds modellen. Logistisk regression fitter proportional odds modellen, når responsen er ordinal.

Modellerne fittes ved brug af maximum likelihood estimation af parametervektoren β .

Dette gøres ved hjælp af en iterativ proces – enten ved brug af Newton-Raphson eller Fisher-scoring algoritmen.

Standardafvigelser, testsandsynligheder og konfidensintervaller beregnes under antagelse af at maximum likelihood estimaterne er asymptotisk normalfordelte.

Jeg har valgt at benytte SAS i nedenstående analyser af data.

Her følger et uddrag af den kode, som jeg benytter i SAS:

```
proc genmod data=speciale.brondby;  
    class koen alder besøgsfrekvens tidsrum ugedag placering boern;  
    model tilfreds = koen alder besøgsfrekvens tidsrum ugedag placering boern  
        /dist=multinomial link=clogit lrci tyoe1 type3  
run;
```

Først defineres, hvilket datasæt PROC GENMOD proceduren skal køres på. I dette tilfælde er det for restauranten Brøndby.

Herefter defineres, hvilke af de forklarende variable, der skal opfattes som class-variable, hvilket her er køn, alder, besøgsfrekvens, tidsrum, ugedag, placering og børn. Årsagen til dette er, at variablene er kategoriske, og at jeg ønsker at få estimater og testsandsynligheder ud for alle niveauer af disse variable.

Da jeg ikke har givet nogen reference for, hvilket niveau, der skal bruges som baseline, bruger programmet automatisk det sidste niveau i hver variabel. Dette betyder i ovenforstående

eksempel, at søndag vil være baseline for variabelen ugedag, da søndag i datasættet er angivet med et 4-tal. Ønsker man at bruge et andet niveau som baseline, er dette også en mulighed. Her skal man så angive referencen efter variabelen, for eksempel: alder (ref=1), hvorved programmet vil benytte niveauet 1 som baseline-kategori.

Modellen opskrives med angivelse af, hvilken variabel, der skal fungere som responsvariabel og hvilke, der skal være forklarende variable. I dette eksempel er tilfredsheden responsvariabel, mens køn, alder, besøgsfrekvens, tidsrum, ugedag, placering og børn alle er forklarende variable.

Der er desuden angivet, at fordelingsfunktionen skal være multinomialfordelingen og at linkfunktionen skal være den kumulative logit. Kommandoen lrci angiver, at der skal udregnes konfidensintervaller for alle variable, og når der endvidere vælges type3, betyder det, at modellerne ikke afhænger af, i hvilken rækkefølge variable er opskrevet.

Når proceduren køres, kommer der en del output. En stor del af dette output er information til brugeren omkring de relevante variable. Denne information kan ikke bruges til konklusioner og fortolkninger, men er vigtig at holde øje med, da det danner hele baggrunden for resultaterne.

Der er blandt andet defineret det relevante datasæt, link- og fordelingsfunktion, hvilken variabel, der opfattes som respons samt antallet af observationer. Der er desuden angivet antallet af niveauer for hver forklarende variabel, hvilke værdier variabelen kan antage, samt hvor mange observationer, der er i hvert niveau for responsvariablen.

Derudover ses estimater og p-værdier for alle forklarende variable, både for hvert variabelniveau og for variabelen samlet. SAS udregner foruden standardafvigelse og likelihood ratio også 95% konfidensintervaller for estimaterne.

10.0 Analyse af tilfredsheden: modeller og resultater

10.1 Proportional odds modeller for tilfredshed

I dette afsnit vil jeg udvikle modeller til forklaring af tilfredsheden hos McDonald's. Dette gør jeg ved at vurdere modellerne på baggrund af testsandsynlighederne for effekt af de forklarende variable. Hvis der er signifikante effekter af én eller flere af de forklarende variable – restaurant, dato, køn, alder, ugedag, tidsrum, besøgsfrekvens, børn og placering – kan jeg udlede, at respondenter tilhørende denne gruppe, adskiller sig i niveauet for tilfredshed. Insignifikante variable vil blive udeladt af modellerne, da de ikke har noget at bidrage med til forklaringen af tilfredshed.

Jeg vil i denne analyse kun benytte mig af den overordnede tilfredshed, som er den tilfredshedsscore, der er givet i spørgsmål 1 i spørgeskemaet. Man kunne også vælge at bruge et af de andre tilfredshedsspørgsmål som den afhængige variabel (spørgsmål 2-11), men for at begrænse analysen, har jeg valgt kun at se på den overordnede tilfredshed.

Jeg vil starte med at finde en model, der kan fortælle omkring tilfredsheden når man ser på McDonald's som en samlet virksomhed (i dette tilfælde ved at bruge KTU-undersøgelserne fra de 14 restauranter som ét samlet datasæt) for herefter at kigge nærmere på restauranten i Brøndby og se, om jeg kan finde en model, der kan beskrive tilfredsheden der. Vil det være de samme forklarende variable, der indgår i begge modeller, eller er de forskellige, og hvad kan årsagen være herpå/hvor skal jeg finde forklaringen hertil?

10.2 McDonald's Danmark

Inden jeg begynder at opstille modeller for McDonald's Danmark (altså de 14 restauranter samlet set), er det en god idé først at danne sig et overblik over, hvordan variablene påvirker hinanden. Jeg ser derfor på korrelationerne mellem tilfredshed og syv af forklarende variable: køn, placering, besøgsfrekvens, børn, tidsrum, ugedag og alder.

Pearson Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations								
	Tilfreds	Koen	Placering	Besoegsfrekvens	Boern	Tidsrum	Ugedag	Alder
Tilfreds	1.00000 7460	0.08442 <.0001 7266	-0.04770 <.0001 7209	0.04451 0.0001 7335	-0.02937 0.0138 7027	-0.01452 0.2197 7145	-0.03224 0.0061 7243	-0.00366 0.7630 6791
Koen	0.08442 <.0001 7266	1.00000 7535	-0.02504 0.0318 7351	-0.16921 <.0001 7449	0.13416 <.0001 7148	0.00877 0.4549 7265	0.03934 0.0007 7351	0.03673 0.0023 6911
Placering	-0.04770 <.0001 7209	-0.02504 0.0318 7351	1.00000 7486	0.10201 <.0001 7412	-0.01364 0.2499 7121	0.04377 0.0002 7249	-0.00309 0.7919 7308	0.06106 <.0001 6866
Besoegsfrekvens	0.04451 0.0001 7335	-0.16921 <.0001 7449	0.10201 <.0001 7412	1.00000 7619	-0.11812 <.0001 7214	0.01883 0.1069 7334	-0.04607 <.0001 7428	-0.18963 <.0001 6972
Boern	-0.02937 0.0138 7027	0.13416 <.0001 7148	-0.01364 0.2499 7121	-0.11812 <.0001 7214	1.00000 7293	0.04010 0.0008 7037	0.06688 <.0001 7124	0.34127 <.0001 6698
Tidsrum	-0.01452 0.2197 7145	0.00877 0.4549 7265	0.04377 0.0002 7249	0.01883 0.1069 7334	0.04010 0.0008 7037	1.00000 7416	0.08252 <.0001 7253	-0.01791 0.1397 6799
Ugedag	-0.03224 0.0061 7243	0.03934 0.0007 7351	-0.00309 0.7919 7308	-0.04607 <.0001 7428	0.06688 <.0001 7124	0.08252 <.0001 7253	1.00000 7517	-0.00359 0.7662 6881
Alder	-0.00366 0.7630 6791	0.03673 0.0023 6911	0.06106 <.0001 6866	-0.18963 <.0001 6972	0.34127 <.0001 6698	-0.01791 0.1397 6799	-0.00359 0.7662 6881	1.00000 7054

Af tabellen ses, at korrelationskoefficienterne generelt er meget lave, men også at der er en signifikant sammenhæng mellem tilfredshed og både køn, placering, besøgsfrekvens, børn og ugedag. Korrelationskoefficienten mellem tilfredshed og køn er på 0,08442, hvilket betyder at kvinder er mere tilfredse end mænd²². Ser man derimod på sammenhængen mellem tilfredshed og placering, angiver koefficienten -0,04770, at kunderne, der kommer i drive-thru er mindre tilfredse end kunderne, der kommer ind i restauranten. Sammenholder jeg dette resultat med krydstabellen for placering og tilfredshed i kapitel 7, passer dette resultat godt overens. Her fandt jeg også, at der var flere kunder instore, der var tilfredse med deres besøg, end kunder, der kom igennem McDrive.

Tabellen fortæller også, at både tidsrum og alder er insignifikante i sammenhæng med tilfredshed, hvilket betyder, at der ikke er nogen forskel i tilfredsheden, uanset hvor gammel man er, og på hvilket tidspunkt man besøger restauranten. Når man ser på korrelationskoefficienterne, er der altså ikke belæg for at sige, at aldersgruppen 13-19 år er mere tilfredse

²² Da korrelationskvotienten er positiv angiver den, at når den ene variabel (tilfredshed) vokser, vokser den anden også (køn). Da Mand=1 og Kvinde=2 i datasættet, betyder det altså, at kvinder er mere tilfredse end mænd.

end aldersgruppen 20-29 år, og der er heller ikke belæg for at sige, at kunder, der kommer i tidsrummet åben-14 er mere tilfredse end kunder, der kommer i tidsrummet 17-20.

Jeg vil nu undersøge nærmere, om de demografiske variable hver især har en signifikant effekt på tilfredsheden. Dette gøres ved at fitte modellerne:

$$M_1: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \sum_{i=1}^{13} (\beta_i \cdot \text{restaurant}_i), \quad j = 1, 2, 3, \dots, 10$$

$$M_2: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{dato}, \quad j = 1, 2, 3, \dots, 10$$

$$M_3: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{køn}, \quad j = 1, 2, 3, \dots, 10$$

$$M_4: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \sum_{i=1}^3 (\beta_i \cdot \text{alder}_i), \quad j = 1, 2, 3, \dots, 10$$

$$M_5: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \sum_{i=1}^3 (\beta_i \cdot \text{tidsrum}_i), \quad j = 1, 2, 3, \dots, 10$$

$$M_6: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \sum_{i=1}^3 (\beta_i \cdot \text{ugedag}_i), \quad j = 1, 2, 3, \dots, 10$$

$$M_7: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{børn}, \quad j = 1, 2, 3, \dots, 10$$

$$M_8: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{placering}, \quad j = 1, 2, 3, \dots, 10$$

$$M_9: \text{logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \sum_{i=1}^3 (\beta_i \cdot \text{besøgsfrekvens}_i), \quad j = 1, 2, 3, \dots, 10$$

Disse modeller svarer til at teste om $\beta = 0$ for de enkelte variable, altså om man kan sige, at der ikke er nogen forskel i niveauerne for de enkelte variable. I model M_1 svarer det altså til, at jeg tester for om der ikke er nogen forskel i tilfredsheden mellem restauranterne (er tilfredsheden den samme hos kunderne i Viby som tilfredsheden hos kunderne i Køge?), mens model M_9 tester for, at der ikke er nogen forskel i tilfredsheden, uafhængig af om kunden kommer i restauranten 1 gang om måneden eller oftere.

Alle forklarende variable behandles som kategoriske. Dato, køn, børn og placering har hver to kategorier, besøgsfrekvens har tre kategorier, mens alder, ugedag og tidsrum hver har fire kategorier. Restaurant har 14 kategorier, idet datasættet består af besvarelser fra 14 forskellige restauranter. Hver kategori er en dummy, som er 1, hvis respondenter tilhører

kategorien og 0 ellers. Den sidste kategori er sat til at være baseline-kategori, dvs. at de øvrige kategorier sammenlignes med denne. Derfor er estimatet for baseline-kategorien sat til 0 pr. definition.

Jeg benytter proceduren `proc genmod` i SAS-programmet til at udføre denne test – der er læst 7.745 observationer i alt, men der er 1.859 missing values, hvorfor antallet af brugte observationer til at udføre denne test er nede på 5.886.

Skemaet nedenfor opsummere de resultater, jeg er kommet frem til:

	Model 1	Model 2
Variabel	Restaurant	Dato
P-værdi	<0,0001	<0,0001

Det fremgår her helt tydeligt, at tilfredsheden afhænger af både restaurant og dato, dvs. der er altså forskel på tilfredsheden blandt gæsterne i de forskellige

restauranter, og tilfredsheden er også forskellig afhængig af, hvornår kundetilfredshedsundersøgelsen er blevet udført.

Hvis jeg ser på de syv andre variable enkeltvis, får jeg følgende resultater:

	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Variabel	Køn	Alder	Tidsrum	Ugedag	Besøgsfrekvens	Børn	Placering
P-værdi	<0,0001	<0,0001	0,0001	0,0269	<0,0001	0,0002	0,0598

Også her er det helt tydeligt, at flere af variablene har en signifikant indvirkning på tilfredsheden. Det drejer sig om køn, alder og besøgsfrekvens, som alle giver en p-værdi på mindre end 0,0001. Derudover opnår også tidsrum og børn en meget lille testsandsynlighed, og også ugedag opnår en p-værdi, som er mindre end 0,05 som man plejer at sige er grænsen. Da datasættet er så stort som det nu er, vil det også være ok at bruge en mindre p-værdi som grænseværdi. Dette skyldes, at de 0,05 normalt bruges i forbindelse med stikprøver, hvor man ikke har helt så mange observationer, og forholdsmæssigt vil det derfor være fint at bruge en mindre værdi af p til grænseværdi.

Placering når lige op over grænseværdien på de 0,05, men idet den er så tæt på grænsen, vil jeg beholde denne variabel i mine videre undersøgelser. Jeg kan vælge at udelukke placering fra de videre undersøgelser jeg vil foretage, men da p-værdien er så tæt på accept-grænsen, vil jeg benytte mig af denne variabel i de videre tests. Det kan være at den alligevel ryger ud når

man tester den sammen med nogle af de andre variable, men omvendt kan det også hænde at den har en betydning sammenholdt med de øvrige variable.

Jeg har nu set at de forklarende variable, når man ser på dem enkeltvis, alle kan være med til at forklare tilfredsheden, men det er mere interessant at finde en model med flere forklarende variable. Jeg vil altså se, om det er muligt at opskrive en model for tilfredsheden, som afhænger af restaurantnr., dato, køn, alder, tidsrum, ugedag, besøgsfrekvens, børn og placering. Denne model kaldes også en additiv model, idet det indeholder alle de forklarende variable.

Jeg tester derfor den additive model:

$$M_{10}: \text{Logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{Restaurantnr.} + \beta_2 \cdot \text{Dato} + \beta_3 \cdot \text{Køn} + \sum_{i=4}^6 (\beta_i \cdot \text{Alder}_i) + \sum_{i=7}^9 (\beta_i \cdot \text{Tidsrum}_i) + \sum_{i=10}^{12} (\beta_i \cdot \text{ugedag}_i) + \sum_{i=13}^{15} (\beta_i \cdot \text{besøgsfrekvens}_i) + \beta_{16} \cdot \text{børn} + \beta_{17} \cdot \text{placering}, \quad j = 1, 2, 3, \dots, 10$$

Ved at benytte mig af statistik-programmet SAS kommer jeg frem til følgende resultater for denne model:

Model 10									
Variabel	Restnr.	Dato	Køn	Alder	Tidsrum	Ugedag	Besøgsfrekvens	Børn	Placering
P-værdi	<0,0001	.	<0,0001	<0,0001	0,0003	0,0814	<0,0001	0,0321	0,0111
χ^2	99,15	0,00	84,77	27,3700	18,80	6,72	31,96	4,5900	6,45
dF	13	0	1	3	3	3	2	1	1

Af model 10 fremgår det, at variablen ugedag får en p-værdi på 0,0814 – dvs. at ugedag er en insignifikant variabel og kan udelades af modellen. Der er altså ikke belæg for at kundernes tilfredshed afhænger af, hvilken dag de har været i restauranten. Dette må siges at være et godt resultat for McDonald's – uanset om restauranten er blevet besøgt torsdag, fredag, lørdag eller søndag er tilfredsheden blandt kunderne den samme. Dog ligger p-værdien tæt på grænseværdien på 0,05, som afgør, hvorvidt denne variabel skal beholdes eller udelades af modellen.

Jeg vælger at tage den ud af min model og får derved en ny model for tilfredsheden:

$$M_{11}: \text{Logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{Restaurantnr.} + \beta_2 \cdot \text{Dato} + \beta_3 \cdot \text{Køn} + \sum_{i=4}^6 (\beta_i \cdot \text{Alder}_i) +$$

$$\sum_{i=7}^9 (\beta_i \cdot \text{Tidsrum}_i) + \sum_{i=10}^{12} (\beta_i \cdot \text{besøgsfrekvens}_i) + \beta_{13} \cdot \text{børn} + \beta_{14} \cdot \text{placering}, \quad j = 1, 2, 3, \dots, 10$$

Denne model for tilfredshed giver følgende resultater:

Variabel	Model 11							
	Restnr.	Dato	Køn	Alder	Tidsrum	Besøgsfrekvens	Børn	Placering
P-værdi	<0,0001	.	<0,0001	<0,0001	0,0002	<0,0001	0,0165	0,0123
χ^2	96,89	0,00	83,66	28,38	19,39	33,29	5,7500	6,27
dF	13	0	1	3	3	2	1	1

Af tabellen ovenfor ser jeg, at der nu kun er signifikante variable tilbage i modellen, og derfor kan jeg ikke tage flere ud af modellen. Efter at have lavet yderligere tests, hvor jeg har byttet rundt på rækkefølgen af variablene, kommer jeg frem til samme resultat, der er ikke flere variable der kan tages ud.

Jeg vil ikke foretage flere test for tilfredsheden, men det er værd at bemærke, at det er muligt at finde en endnu bedre model til at forklare tilfredsheden, hvis man indtager vekselvirkninger i modellen.

Det bliver alt for omfattende at tage dem med ind i denne opgave, men man skal være klar over, at hvis man ønsker at finde den fulde model, bliver man nødt til også at teste for vekselvirkninger.

En vekselvirkning kunne f.eks. være køn*børn – hvis man tester for denne, vil man måske komme frem til, at der er en sammenhæng mellem køn og børn. Herefter kan man så se, om det er kvinderne, der har børnene med. Det er også muligt, at man finder frem til, at der slet ikke er nogen sammenhæng (en høj p-værdi).

Parameterestimer:

Estimaterne for de kategoriske variable sammenligner hver kategori med den sidste kategori, der er sat til 0. Den sidste kategori er altså sat til at være baseline-kategori for de øvrige.

Da der er ti kategorier af responsvariablen tilfredshed, har modellen ni intercepter. Estimerne for intercepterne er ikke interessante i sig selv, de bruges udelukkende til udregning af de kummulative sandsynligheder.

Jeg får følgende tabel frem for parameterestimerne for Model 11, der er den model, der bedst beskriver tilfredsheden for restauranterne samlet set uden vekselvirkninger:

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Estimate	Standard Error	Estimat +1,96 SE	Estimat -1,96 SE	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-7,0895	0,3036	-7,7154	-6,5193	545,21	<0,0001
Intercept	2	1	-6,4921	0,2533	-7,0022	-6,0072	656,80	<0,0001
Intercept	3	1	-5,8651	0,2196	-6,3011	-5,4399	713,62	<0,0001
Intercept	4	1	-5,1510	0,1971	-5,5398	-4,7669	682,79	<0,0001
Intercept	5	1	-4,0759	0,1815	-4,4329	-3,7214	504,41	<0,0001
Intercept	6	1	-3,3919	0,1771	-3,7403	-3,0458	366,65	<0,0001
Intercept	7	1	-2,5138	0,1743	-2,8566	-2,1733	208,05	<0,0001
Intercept	8	1	-1,3608	0,1721	-1,6992	-1,0243	62,49	<0,0001
Intercept	9	1	-0,6723	0,1713	-1,0091	-0,3373	15,40	<0,0001
Restaurant	10	1	0,9703	0,1220	0,7322	1,2107	63,22	<0,0001
Restaurant	20	1	0,3962	0,1243	0,1534	0,6408	10,16	0,0014
Restaurant	38	1	0,8904	0,1330	0,6305	1,1520	44,82	<0,0001
Restaurant	45	1	0,5164	0,1439	0,2346	0,7990	12,87	0,0003
Restaurant	54	1	0,4802	0,1271	0,2319	0,7303	14,27	0,0002
Restaurant	64	1	0,6117	0,1565	0,3050	0,9185	15,28	<0,0001
Restaurant	77	1	0,0973	0,1518	-0,2007	0,3947	0,41	0,5216
Restaurant	85	1	0,7050	0,1390	0,4331	0,9782	25,72	<0,0001
Restaurant	87	1	0,6159	0,1340	0,3539	0,8795	21,12	<0,0001
Restaurant	98	1	-0,0141	0,1373	-0,2831	0,2553	0,01	0,9183
Restaurant	103	1	0,5179	0,1301	0,2636	0,7739	15,84	<0,0001
Restaurant	201	1	1,1220	0,1397	0,8489	1,3968	64,46	<0,0001
Restaurant	202	1	0,4778	0,1480	0,1878	0,7682	10,42	0,0012
Dato	1	0	0,0000	0,0000	0,0000	0,0000	0,00	0,0000
Besøgsfrekvens	1	1	0,4036	0,0703	0,2659	0,5416	32,92	<0,0001
Besøgsfrekvens	2	1	0,2373	0,0658	0,1085	0,3664	13,00	0,0003
Placering	1	1	-0,2124	0,0846	-0,3780	-0,0462	6,29	0,0121
Køn	1	1	0,4617	0,0507	0,3624	0,5613	82,82	<0,0001
Børn	1	1	-0,1555	0,0648	-0,2825	-0,0284	5,76	0,0164
Alder	1	1	-0,0291	0,0773	-0,1804	0,1226	0,14	0,7068
Alder	2	1	0,2127	0,0795	0,0570	0,3687	7,15	0,0075
Alder	3	1	0,2751	0,0825	0,1136	0,4369	11,13	0,0008
Tidsrum	1	1	0,1972	0,0870	0,0271	0,3682	5,14	0,0234
Tidsrum	2	1	0,1962	0,0908	0,0185	0,3746	4,67	0,0307
Tidsrum	3	1	0,3729	0,0907	0,1956	0,5511	16,91	<0,0001

Parameterestimerne antages at være asymptotisk normalfordelte med standardafvigelsen (SE) som estimeret spredning. Endepunkterne for 95% konfidensintervaller er $\pm 1,96 \times \text{standardafvigelsen}$.

Parameterestimerne for en variabel fortæller om tilfredsheden, når alle de andre variable holdes konstant. Det betyder at jeg kan rangordne en variabel under forudsætningen at alt andet er lige.

Tager jeg fat i variablen Restaurantnr. kan jeg altså direkte aflæse i tabellen, hvilken restaurant, der er mest tilfreds, når alt andet er lige. Her gælder det ikke om at have det

højeste beta, men det laveste, idet restauranterne ønsker en så høj andel af besvarelsene til at være i den sidste kategori, altså et 10-tal.

Rangerer jeg restauranterne efter deres beta, får jeg følgende tabel:

Restaurant	Beta
Esbjerg	-0,0141
Kastrup	0,0000
Thisted	0,0973
Viborg	0,3962
Valby	0,4778
Svendborg	0,4802
Vejle 2	0,5164
Hillerød	0,5179
Frederiksværk	0,6117
Nyborg	0,6151
Brøndby	0,7050
Hjørring	0,8904
Viby	0,9703
Køge	1,1120

Her ser jeg, at det er restauranten i Esbjerg, der performer bedst, når alt andet er lige, mens det er restauranten i Køge, som performer dårligst, alt andet lige.

Sammenholder jeg dette resultat med den tabel, jeg lavede over tilfredshed for restauranterne i kapitel 7, ser jeg at det passer helt overens. Her fandt jeg frem til, at Esbjerg havde en andel af 10-taller for tilfredsheden, som udgjorde 51,28% af de afgivne karakterer, hvilket var den højeste andel, mens Køge netop lå i bunden med en andel på kun 31,58%.

Ser man derimod på variabelen alder, så er det gruppen af unge, de 13-19 årige, der er mest tilfredse når alt andet er lige, mens det er aldersgruppen 30-39 årige der er mindst tilfredse. Dette giver meget god mening – de unge er ikke helt så kritiske som den næstældste gruppe, de ligger vægt på nogle helt andre parametre i denne aldersgruppe.

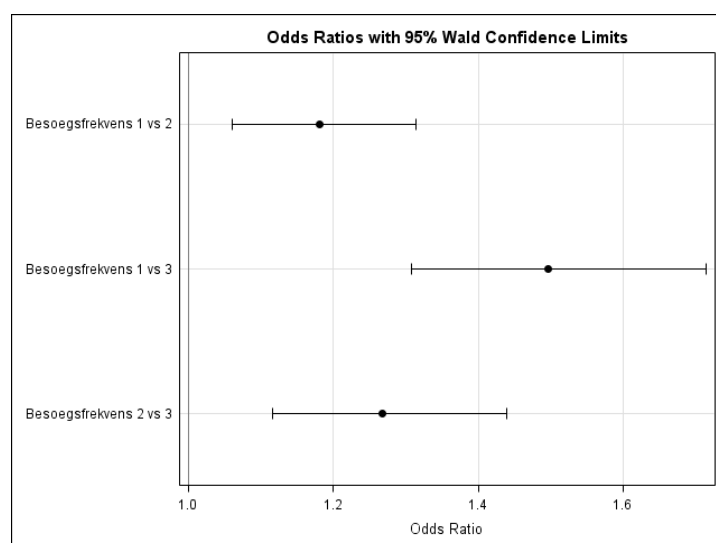
Et lidt overraskende resultat er, at alt andet lige, så er kvinderne de mest tilfredse kunder hos McDonald's (jeg ville have troet at det var mændene), men også at det er den gruppe, der besøger restauranten oftest, som er mindst tilfredse. Her ligger forklaringen måske i, at de er vant til netop at komme i restauranten og sammenligner deres besøg med et tidligere besøg, hvor de oplevede en helt ekseptionel service.

Odds ratio:

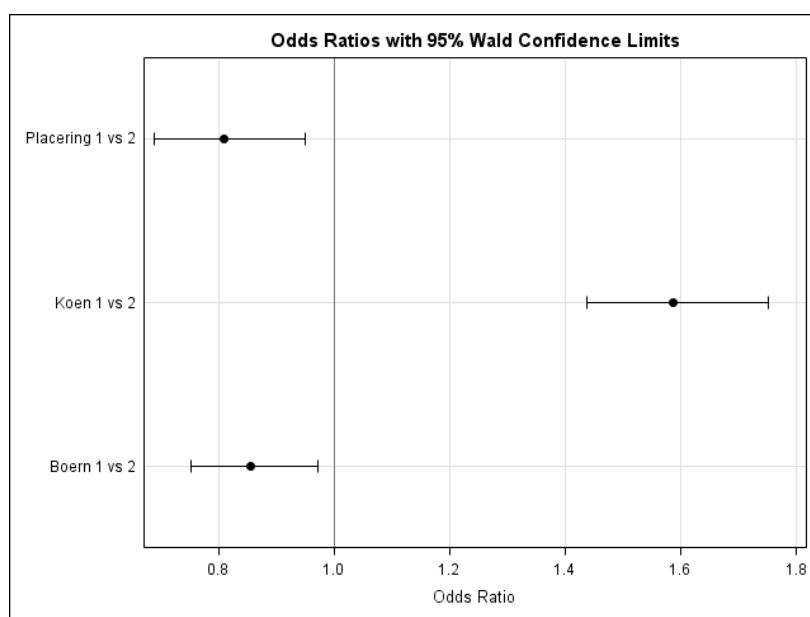
I tabellen nedenfor er vist odds ratio estimeret og det tilhørende konfidensinterval for parametergruppen sammenlignet med baseline-kategorien. Man kunne også have valgt at vise alle odds ratio estimerne, men det vil blive for omfattende. Derimod har jeg valgt, at tage nogle grafer med til at illustrerer odds ratioerne.

Parameter	Gruppe	Odds Ratio	95% konfidensintervaller for Odds Ratio	
Restaurant	10 vs. 203	2,639	2,071	3,362
Restaurant	20 vs. 203	1,486	1,156	1,911
Restaurant	38 vs. 203	2,436	1,871	3,171
Restaurant	45 vs. 203	1,676	1,265	2,220
Restaurant	54 vs. 203	1,616	1,256	2,079
Restaurant	64 vs. 203	1,843	1,351	2,515
Restaurant	77 vs. 203	1,102	0,815	1,491
Restaurant	85 vs. 203	2,024	1,543	2,655
Restaurant	87 vs. 203	1,851	1,417	2,418
Restaurant	98 vs. 203	0,986	0,748	1,300
Restaurant	103 vs. 203	1,678	1,296	2,173
Restaurant	201 vs. 203	3,071	2,330	4,047
Restaurant	202 vs. 203	1,612	1,201	2,164
Besøgsfrekvens	1 vs. 3	1,497	1,307	1,714
Besøgsfrekvens	2 vs. 3	1,268	1,116	1,440
Placering	1 vs. 2	0,809	0,688	0,950
Køn	1 vs. 2	1,587	1,437	1,752
Børn	1 vs. 2	0,856	0,753	0,973
Alder	1 vs. 4	0,971	0,835	1,130
Alder	2 vs. 4	1,237	1,058	1,447
Alder	3 vs. 4	1,317	1,116	1,554
Tidsrum	1 vs. 4	1,218	1,030	1,441
Tidsrum	2 vs. 4	1,217	1,021	1,450
Tidsrum	3 vs. 4	1,452	1,219	1,729

Ser jeg på tabellen ovenfor kan jeg aflæse, at de estimerede odds for at tilfredshed er lavere end en given værdi, når respondenten besøger Restaurant 201 (Køge) er 3,071 gange de estimerede odds, hvis respondenten besøger Restaurant 203 (Kastrup). Respondenterne (kunderne), der besøger McDonald's i Køge, er altså (meget) mindre tilfredse, end dem, der besøger restauranten i Kastrup.

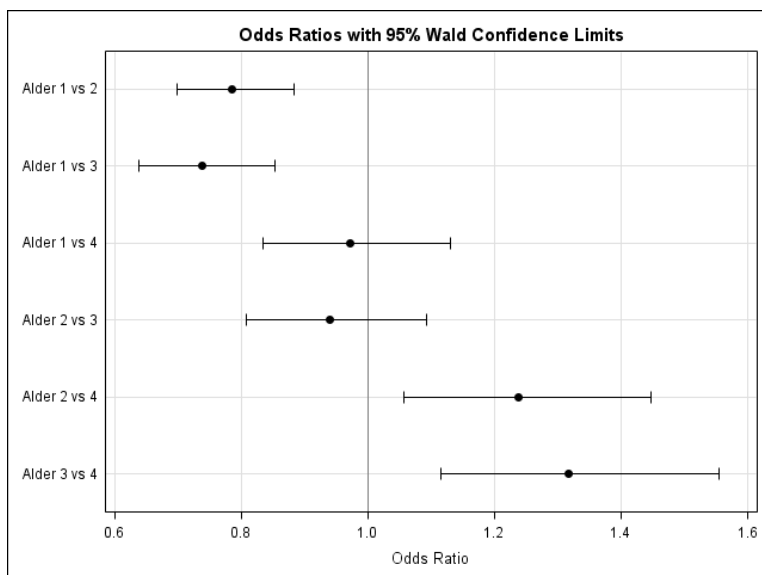


Illustrationen viser, at odds er 1,5 gange så stor for at kunder der kommer ofte (mere end 4 gange om måneden) end dem som kommer sjældent (mindre end 1 gang om måneden), giver en lavere score på tilfredsheden. Dette resultat er lidt overraskende, idet jeg havde forventet, at netop heavy-users var mere tilfredse med deres besøg. De kommer her jo ofte i restauranten, men noget kan tyde på, at det altså ikke er på grund af tilfredsheden at de kommer tilbage. Det kan også være at denne gruppe er mere tilbageholdende med de gode karakterer, hvis de har oplevet en god service ved de andre besøg, men netop på selve dagen er blevet skuffet.

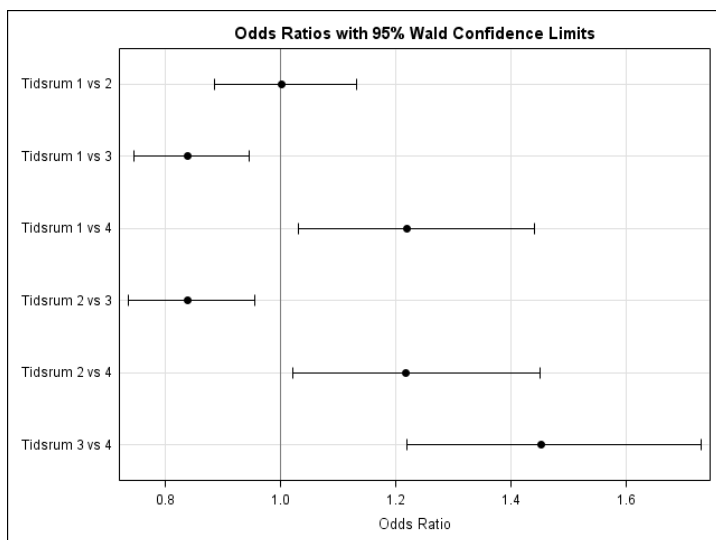


Ser jeg på illustrationen ovenfor, og sammenholder med oddsene givet i tabellen ligeledes ovenfor, fremgår det, at kunder, der besøger restauranten indenfor, er mere tilfredse end kunder, der kommer i driven, mens det er kvinderne, der er mest tilfredse med besøget i restauranten, når jeg ser på køn. Sidst kan jeg se, at dem der har børn med ved deres besøg hos McDonald's er mere tilfredse, end dem, som kommer alene. Noget tyder altså på, at McDonald's, har fat i den lange ende når det kommer til børnene.

Det ses her, at for placering er de estimerede odds for at tilfredshed er lavere end en given værdi, når respondenter er instore (placering er 1) $\exp(-0,1545)=0,857$ gange de estimerede odds, hvis respondenter besøgte restauranten i McDrive. Det vil altså sige, at respondenter, der kommer instore hos McDonald's, er mere tilfredse (idet det jo her er bedre at have en større andel af de høje værdier af j).



Illustrationen ovenfor viser, at aldersgruppe 1 og 4 er mere eller mindre lige tilfredse (odds for alder 1 vs. 4 er lige omkring 1,0), mens aldersgrupperne 2 og 3 er mindre tilfredse end både aldersgruppe 1 og 4. Kunderne i aldersgruppen 20-39 år er altså de mindst tilfredse, når man holder alt andet lige og kun kigger på alder og tilfredshed.



Ser man på odds for tidsrum, er der ingen forskel på, om du kommer i tidsrum 1 eller 2 (i perioden åben til kl. 17), mens det er bedre at kommer til frokost (tidsrum 1) frem for til aftensmad (tidsrum 3) når alt andet er lige og man kun ser på tilfredsheden. Der tegner sig et billede af, at de mindst tilfredse kunder kommer i tidsperioden 3, altså ved middagstid, når man sammenholder denne periode med de øvrige.

Kumulerede sandsynligheder

I stedet for at se på odds ratios kan man se på de kumulerede sandsynligheder, som kan være lettere at forstå. Jeg viser først et eksempel på, hvordan de kumulerede sandsynligheder udregnes og viser derefter alle sandsynlighederne for tre udvalgte restauranter.

Jeg tager udgangspunkt i de to restauranter – Esbjerg (98) og Køge (201) – og beregner sandsynligheden for, at en gæst i den pågældende restaurant giver en karakter for tilfredsheden, som er på 9 eller derunder.

$$\text{Esbjerg: } P(\text{tilfredshed} \leq 9) = \frac{\exp(-0,6723 - 0,0141)}{1 + \exp(-0,6723 - 0,0141)} = 0,3345$$

$$\text{Køge: } P(\text{tilfredshed} \leq 9) = \frac{\exp(-0,6723 + 1,1120)}{1 + \exp(-0,6723 + 1,1120)} = 0,6106$$

Sandsynligheden for at besøget hos McDonald's giver et udfald på tilfredsheden med ni eller derunder, er altså 61,06% hvis man har været i Køge, mens sandsynligheden kun er 33,45% hvis man har været i Esbjerg.

Hvis jeg derimod havde set på en score for tilfredsheden på 5 eller derunder, ville man få følgende beregning:

$$\text{Esbjerg: } P(\text{tilfredshed} \leq 5) = \frac{\exp(-4,0759 - 0,0141)}{1 + \exp(-4,0759 - 0,0141)} = 0,0165$$

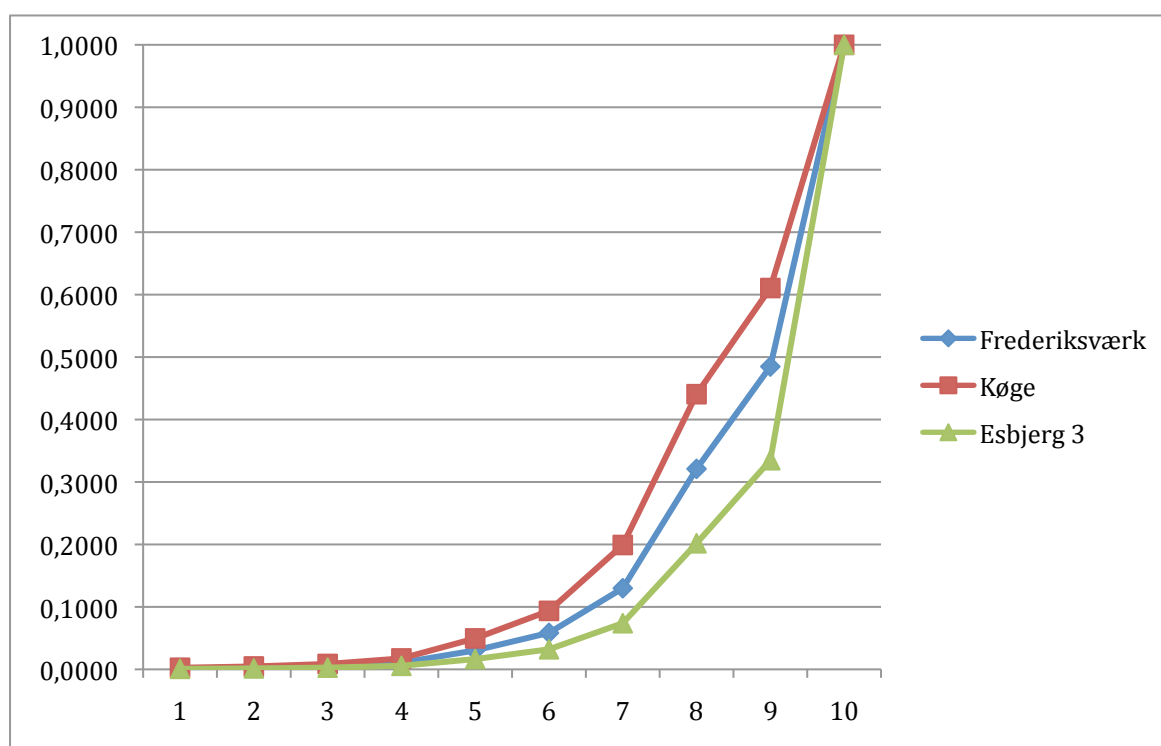
$$\text{Køge: } P(\text{tilfredshed} \leq 5) = \frac{\exp(-4,0759 + 1,1120)}{1 + \exp(-4,0759 + 1,1120)} = 0,0496$$

Sandsynligheden for at besøget hos McDonald's giver et udfald på tilfredsheden med fem eller derunder, er altså 4,96% hvis man har været i Køge, mens sandsynligheden kun er 1,65% hvis man har været i Esbjerg.

Jeg får følgende tabel for de 3 udvalgte restauranter:

Tilfredshed	Frederiksværk	Køge	Esbjerg 3
1	0,0015	0,0026	0,0008
2	0,0028	0,0046	0,0015
3	0,0052	0,0086	0,0028
4	0,0106	0,0175	0,0057
5	0,0303	0,0496	0,0165
6	0,0584	0,0936	0,0321
7	0,1299	0,1991	0,0739
8	0,3210	0,4406	0,2018
9	0,4849	0,6106	0,3348
10	1	1	1

De tre restauranter repræsenterer henholdsvis den bedste og den dårligst performende restaurant, når man ser på tilfredsheden, samt en restaurant (Frederiksværk), som ligger i midterfeltet. Jeg kan også vise dette grafisk, og det er netop hvad illustrationen herunder gør. Restauranterne vil altså befinde sig i området mellem linjen for den bedste restaurant (Esbjerg) og den dårligste restaurant (Køge).



10.3 McDonald's Brøndby

Jeg kigger nu på restauranten i Brøndby. Først vil jeg danne mig et overblik over, hvordan variablene påvirker hinanden. Hertil kigger jeg på korrelationerne mellem tilfredshed og de syv forklarende variable.

Pearson Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations								
	Tilfreds	Koen	Placering	Besoegsfrekvens	Boern	Tidsrum	Ugedag	Alder
Tilfreds	1.00000 561	0.07673 0.0762 535	-0.16404 0.0002 529	0.01430 0.7388 546	-0.01555 0.7264 509	-0.04409 0.3115 529	-0.08504 0.0526 520	-0.05727 0.2162 468
Koen	0.07673 0.0762 535	1.00000 561	-0.06619 0.1248 539	-0.17191 <.0001 555	0.13025 0.0031 515	-0.02553 0.5534 541	0.04902 0.2572 536	-0.02726 0.5530 476
Placering	-0.16404 0.0002 529	-0.06619 0.1248 539	1.00000 556	0.12370 0.0037 549	-0.09816 0.0258 516	0.05083 0.2401 536	-0.07592 0.0811 529	0.15166 0.0009 472
Besoegsfrekvens	0.01430 0.7388 546	-0.17191 <.0001 555	0.12370 0.0037 549	1.00000 573	-0.01842 0.6737 525	-0.06650 0.1200 548	-0.04846 0.2601 542	-0.09985 0.0279 485
Boern	-0.01555 0.7264 509	0.13025 0.0031 515	-0.09816 0.0258 516	-0.01842 0.6737 525	1.00000 533	0.06059 0.1702 514	0.10164 0.0222 506	0.24688 <.0001 452
Tidsrum	-0.04409 0.3115 529	-0.02553 0.5534 541	0.05083 0.2401 536	-0.06650 0.1200 548	0.06059 0.1702 514	1.00000 554	0.19404 <.0001 529	-0.01033 0.8227 473
Ugedag	-0.08504 0.0526 520	0.04902 0.2572 536	-0.07592 0.0811 529	-0.04846 0.2601 542	0.10164 0.0222 506	0.19404 <.0001 529	1.00000 547	-0.04496 0.3339 464
Alder	-0.05727 0.2162 468	-0.02726 0.5530 476	0.15166 0.0009 472	-0.09985 0.0279 485	0.24688 <.0001 452	-0.01033 0.8227 473	-0.04496 0.3339 464	1.00000 491

Af tabellen ses, at korrelationskoefficienterne generelt er meget lave, og at der kun er én korrelation med tilfredshed, som er signifikant, nemlig placering. Derudover er korrelationen mellem tilfredshed og ugedag lige på grænsen til at være signifikant. Da begge korrelationskoefficienter er negative, fortæller det mig, at kunderne, der besøger restauranten indenfor er mere tilfredse end dem, der kommer i bil, og at kunderne, der kommer først på ugen, er mere tilfredse end dem, der kommer sidst på ugen.

Derudover fortæller tabellen, at der er en sammenhæng mellem køn og besøgsfrekvens. Her er koefficienten negativ, hvormed den fortæller, at mænd kommer oftere i restauranten. Der er også en sammenhæng mellem køn og børn, her er det kvinderne, der tager børnene med på McDonald's. Umiddelbart fandt jeg dette resultat overraskende, da det ofte er fædre, der har deres børn med på McDonald's uden mor, men forklaringen måske skal findes i spørgeskemaets udformning. Her tager man ikke højde for et besøg af en hel familie, så det kan være moderen, der har udfyldt spørgeskemaet for hele familien, hvormed hun har angivet at være en kvinde og haft et barn med under 12 år. Man skal ikke angive, hvor mange man har

været ved besøget, så medmindre faderen har fået et spørgeskema selv (hvilket han kun vil få hvis han bestiller maden for sig selv), vil han ikke indgå i undersøgelsen.

Hvis jeg ser på andre korrelationer med børn, som er signifikante, er der både placering og ugedag. Her fortæller koefficienterne, at børn kommer inde i restauranten og at der kommer børn først på ugen. Dette resultat kan også virke lidt underligt, idet McDonald's er en familierestaurant og at børnefamilierne har mere tid i weekenden. Ser jeg på restauranten i Brøndby kan forklaringen måske findes i, at der ikke er noget decideret børneland i denne restaurant. Til gengæld har restauranten i Hvidovre (som desværre ikke er med i denne undersøgelse) et kæmpe legeland og ligger kun 10 minutter i bil fra restauranten i Brøndby. Man kunne forestille sig, at hvis familien har besluttet sig til McDonald's i weekenden, så kører de til Hvidovre, hvor børnene kan lege, mens forældrene får et frikvarter.

Jeg vil nu undersøge nærmere, om de demografiske variable har en signifikant effekt på tilfredsheden. Da jeg så på alle restauranterne samlet set, testede jeg først variablene enkeltvis, men denne gang starter jeg med at se på den additive model.

Alle forklarende variable behandles som kategoriske, præcis som det var tilfældet, da jeg så på alle restauranterne samlet set.

Det datasæt, som jeg benytter til at modellere for McDonald's Brøndby, består af 589 respondenter, men da hele 205 observationer har en missing value, bliver antallet af observationer, der benyttes, kun 384.

Jeg starter med at se på den additive model og ser om der er nogle af variablene som kan testes ud af modellen.

$$M_1: \text{Logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{køn} + \sum_{i=2}^4 (\beta_i \cdot \text{alder}_i) + \sum_{i=5}^7 (\beta_i \cdot \text{tidsrum}_i) + \sum_{i=8}^{10} (\beta_i \cdot \text{ugedag}_i) + \sum_{i=11}^{13} (\beta_i \cdot \text{besøgsfrekvens}_i) + \beta_{14} \cdot \text{børn} + \beta_{15} \cdot \text{placering}, \quad j = 1, 2, 3, \dots, 10$$

Ved at benytte SAS-programmet, kommer jeg frem til følgende resultater:

Variabel	Køn	Alder	Tidsrum	Model 1 Ugedag	Besøgsfrekvens	Børn	Placering
P-værdi	0,3882	0,5962	0,6717	0,0441	0,2912	0,1182	<.0001

Jeg ser at variablen Alder har en p-værdi på 0,5962 i denne model, hvilket tyder på at alder ikke har nogen indflydelse på tilfredsheden. Efter at have foretaget flere tests, hvor jeg har byttet om på rækkefølgen af variablen alder, står det klart at denne godt kan tages ud af modellen. Dette sker fordi jeg i alle modellerne opnår en p-værdi, der ligger langt over de 0,05, som jeg bruger som min grænseværdi. Ved denne test kan jeg altså konkludere, at min model for tilfredshed i McDonald's Brøndby, ikke skal indeholde alder. Der er ikke noget statistisk bevis for, at der er forskel på tilfredsheden blandt de fire aldersgrupper – man kan altså hverken sige at de unge er mere eller mindre tilfredse end de ældre.

Jeg tager nu fat i den reducerede model og ser om der er yderligere variable der kan udelades:

$$M_2: \text{Logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{køn} + \sum_{i=2}^4 (\beta_i \cdot \text{tidsrum}_i) + \sum_{i=5}^7 (\beta_i \cdot \text{ugedag}_i) + \sum_{i=8}^{10} (\beta_i \cdot \text{besøgsfrekvens}_i) + \beta_{11} \cdot \text{børn} + \beta_{12} \cdot \text{placering}, \quad j = 1, 2, 3, \dots, 10$$

Jeg kommer frem til følgende resultater:

Variabel P-værdi	Model 2					
	Køn	Tidsrum	Ugedag	Besøgsfrekvens	Børn	Placering
	0,3882	0,7595	0,0373	0,2912	0,1182	<.0001

Her ser jeg at variablen Tidsrum har en relativ høj p-værdi (0,7595), og sammenholder jeg dette resultat med resultatet fra Model 1, ser jeg at p-værdien er forøget når alder er taget ud af modellen. Det vil altså sige, at Alder har haft en påvirkning på Tidsrum, men da værdien langt overstiger de 0,05 kan denne variabel også siges at være insignifikant. Jeg har også ved denne model haft byttet om på rækkefølgen af de forklarende variable i testen, men igen uden at det har haft nogen indflydelse på resultatet.

Derfor kan jeg godt tage variablen Tidsrum ud af min model, der er ikke belæg for at sige, at der er forskel på kundernes tilfredshed, om de kommer i restauranten i tidsrumene 1, 2, 3 eller 4.

Jeg kan nu reducere min model yderligere, således at den nu kun består af fem variable:

$$M_3: \text{Logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{besøgsfrekvens} + \beta_2 \cdot \text{placering} + \beta_3 \cdot \text{børn} +$$

$$\sum_{i=4}^6 (\beta_i \cdot \text{ugedag}_i) + \beta_7 \cdot \text{køn}, \quad j = 1, 2, 3, \dots, 10$$

Jeg kommer frem til følgende resultater med den nye model:

Model 3					
Variable	Besøgsfrekvens	Placering	Børn	Ugedag	Køn
P-værdi	0,2912	<.0001	0,1580	0,0422	0,2260

Her ser jeg, at Køn har en p-værdi på 0,2260, hvilket indikerer at denne variabel også er insignifikant. Ved at foretage yderligere tests, hvor rækkefølgen på de forklarende variable byttes rundt, kommer jeg ikke frem til resultater, der indikerer at variabelen Køn er insignifikant. Det kan dog bemærkes at p-værdien er faldet efter både Alder og Tidsrum er taget ud af modellen, men ikke nok til, at det beretter Køn til at indgå i en model, der beskriver tilfredsheden.

Der er altså ikke belæg for at sige, at mænd, der besøger McDonald's Brøndby, er mere tilfredse med deres besøg end kvinder.

Tager jeg variabelen Køn ud af min model, kommer jeg frem til følgende model til beskrivelse af tilfredsheden:

$$M_4: \text{Logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{besøgsfrekvens} + \beta_2 \cdot \text{placering} + \beta_3 \cdot \text{børn} +$$

$$\sum_{i=4}^6 (\beta_i \cdot \text{ugedag}_i), \quad j = 1, 2, 3, \dots, 10$$

Denne model (M₄) giver følgende resultater:

Model 4				
Variable	Besøgsfrekvens	Placering	Børn	Ugedag
P-værdi	0,6927	<.0001	0,1696	0,0422

Her ses det tydeligt at variablen Besøgsfrekvens har en høj p-værdi (0,6927), hvorfor denne variabel er insignifikant og derfor kan udelades af modellen. Hvor tilfældet med variablen Køn var at p-værdien reducerede sin værdi ved et mindre antal forklarende variable, er tilfældet modsat med Besøgsfrekvens. Det ses tydeligt, at der er sket en stigning i p-værdien, hvilket indikerer, at Besøgsfrekvens kan være tæt forbundet med nogle af de andre forklarende variable.

Denne test fortæller mig altså, at der ikke er forskel på tilfredsheden blandt kunderne uanset om de besøger restauranten ofte (mere end 4 gange om måneden), regelmæssigt (1-3 gange om måneden) eller sjældent (mindre end 1 gang om måneden).

Jeg kan nu reducere min model yderligere, og herved får jeg følgende model:

$$M_5: \text{Logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{placering} + \beta_2 \cdot \text{børn} + \sum_{i=3}^5 (\beta_i \cdot \text{ugedag}_i), \quad j = 1, 2, 3, \dots, 10$$

Dette giver mig følgende resultater når jeg foretager en test-kørsel:

Variable P-værdi	Model 5		
	Placering	Børn	Ugedag
	<.0001	0,1469	0,0376

Jeg ser at variablen Børn opnår en p-værdi, der ligger over min acceptgrænse, hvorfor denne variabel også kan udelades af modellen. Ved at bytte om på rækkefølgen af de forklarende variable i yderligere tests, opnår jeg kun resultater, der støtter op om denne konklusion.

Der er altså ikke noget statistisk bevis for at kunder, der har deres børn med ved besøget, er mere tilfredse end kunder, der kommer alene eller sammen med andre voksne. Selvom McDonald's er en familierestaurant og gør mange tiltag for børnefamilier, er det altså ikke noget, der giver udslag i den overordnede tilfredshed.

Jeg når nu frem til en reduceret model, der kun består af placering og ugedag:

$$M_6: \text{Logit}[P(\text{tilfredshed} \leq j)] = \alpha_j + \beta_1 \cdot \text{placering} + \sum_{i=2}^4 (\beta_i \cdot \text{ugedag}_i), \quad j = 1, 2, 3, \dots, 10$$

Denne model giver mig følgende resultater:

Model 6		
Variable	Placering	Ugedag
P-værdi	<.0001	0,0376

Her ses det at begge variable giver en lav p-værdi, hvorfor det ikke er muligt at tage disse variable ud af modellen. Jeg ender altså op med en model til beskrivelse af tilfredsheden hos McDonald's Brøndby, som består af placering og ugedag.

Parameterestimer

Estimerne for de kategoriske variable sammenligner hver kategori med den sidste kategori, der er sat til 0. Den sidste kategori er altså sat til at være baseline-kategori for de øvrige.

Da der er 9 kategorier af responsvariablen tilfredshed (ingen har givet et 2-tal), har modellen otte intercepter. Estimerne for intercepterne er ikke interessante i sig selv, de bruges udelukkende til udregning af de kumulative sandsynligheder.

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Estimate	Standard Error	Estimat +1,96 SE	Estimat -1,96 SE	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-3,5789	0,6257	-5,0358	-2,5012	32,72	<0,0001
Intercept	3	1	-3,2887	0,5551	-4,5379	-2,3098	35,10	<0,0001
Intercept	4	1	-2,2438	0,3880	-3,0472	-1,5165	33,44	<0,0001
Intercept	5	1	-1,1215	0,3091	-1,7385	-0,5245	13,17	0,0003
Intercept	6	1	-0,3373	0,2915	-0,9139	0,2310	1,34	0,2474
Intercept	7	1	0,4968	0,2902	-0,0741	1,0655	2,93	0,0869
Intercept	8	1	1,2061	0,2951	0,6276	1,7862	16,70	<0,0001
Intercept	9	1	1,9209	0,3036	1,3271	2,5189	40,02	<0,0001
Placering	1	1	-1,1804	0,2372	-1,6485	-0,7172	24,76	<0,0001
Ugedag	1	1	-0,5377	0,2639	-1,0559	-0,0203	4,15	0,0416
Ugedag	2	1	-0,8029	0,2806	-1,3553	-0,2539	8,19	0,0042
Ugedag	3	1	-0,4239	0,2860	-0,9860	0,1364	2,20	0,1383

Parameterestimerne antages at være asymptotisk normalfordelte med standardafvigelsen (SE) som estimeret spredning. Endepunkterne for 95% konfidensintervallerne er $\pm 1,96 \times \text{standardafvigelsen}$.

Parameterestimerne for en variabel fortæller om tilfredsheden, når alle de andre variable holdes konstant. Det betyder at jeg kan rangordne en variabel under forudsætningen at alt andet er lige.

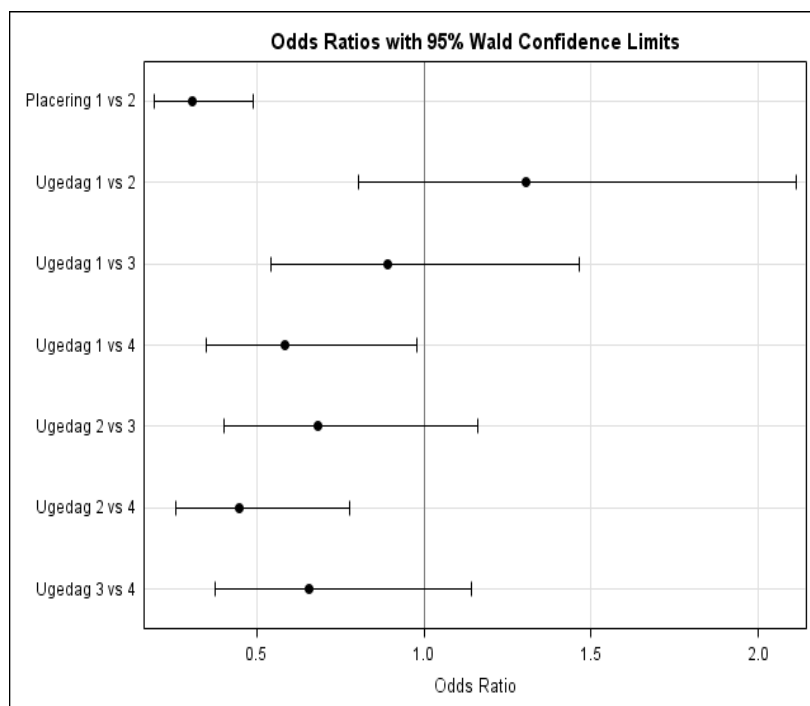
Hvis jeg ser på placering, vil det altså sige, at kunderne er mest tilfredse når de besøger restauranten instore, mens den ugedag der opnår højest tilfredshed, alt andet lige, er fredagen.

Odds Ratio estimer

I tabellen er vist odds ratio estimer og tilhørende konfidensintervaller for alle paramtergrupper sammenlignet med baseline-kategorien.

Parameter	Gruppe	Odds Ratio	95% konfidensintervaller for Odds Ratio	
Placering	1 vs. 2	0,307	0,193	0,489
Ugedag	1 vs. 4	0,584	0,348	0,980
Ugedag	2 vs. 4	0,448	0,259	0,776
Ugedag	3 vs. 4	0,655	0,375	1,144

På grafen nedenfor er endvidere illustreret odds ratio estimer for alle parametergrupper, og her er det helt tydeligt at se, at der er en odds ratio, der skiller sig ud. Det drejer sig om oddsene for ugedag 2 (fredag), den er bedre sammenlignet med de øvrige ugedag, præcis som jeg kunne aflæse under parameterestimerne.



Jeg har nu set på restauranten i Brøndby, og det kunne også være spændende at se, hvilke modeller man ville komme frem til for tilfredshed for de andre restauranter. Det bliver for omfattende at lave en slavisk gennemgang af alle disse, men i tabellen nedenfor har jeg opsummeret, hvilke variable der indgår i modellen for de enkelte restauranter. Igen skal det tilføjes, at jeg ikke har haft kigget på vekselvirkninger.

Jeg er kommet frem til følgende modeller for de enkelte restauranter:

Restaurant	Variable						
	Køn	Alder	Tidsrum	Ugedag	Besøgsfrekvens	Børn	Placering
Viby	x						
Viborg	x	x	x				
Hjørring		x				x	
Vejle 2	x			x			
Svendborg	x	x			x		
Frederiksværk	x	x			x		
Thisted	x				x	x	
Brøndby				x			x
Nyborg	x	x			x		
Esbjerg 3	x	x	x		x		
Hillerød	x		x		x		
Køge				x			
Valby	x	x			x		
Kastrup	x	x			x		

Som man kan se af tabellen, er der mange forskellige modeller af tilfredshed. Der er dog en udgave, som går igen i flere restauranter, og det er den, som indeholder køn, alder og besøgsfrekvens. Ellers ser jeg, at der er stor variation i alle modellerne, både med hensyn til *antallet* af forklarende variable, men også *hvilke* variable, der er indeholdt i modellerne.

Det er også værd at ligge mærke til, at der ikke er nogen modeller, der indeholder mere end 4 variable. Det optimale set fra McDonald's side, ville være, at der slet ikke var forskel i tilfredsheden, og deraf at der slet ikke skulle være nogen model herfor. Derfor er det også flot, at nogle af restauranterne kun har én eller to variable til at forklare tilfredsheden. Dette gør også, at det bliver lettere at sætte ind overfor forbedringer. I stedet for at man skal forbedre sig på alle syv variable, kan man nu "nøjes" med at koncentrere sig om nogle få.

Der er to restauranter, der kun har én forklarende variabel. Det er henholdsvis Viby, hvor tilfredsheden afhænger af kønnet på kunden, og Køge, hvor tilfredsheden afhænger af, hvilken dag på ugen kunden besøger restauranten. Dette er et lidt sjovt resultat, både fordi, det er den restaurant, som performer dårligst, når alle andre parametre holdes lige, men også fordi, at da

jeg så på den model, der havde alle restauranterne inde samlet, var det netop ugedag, der blev testet ud af modellen.

Det er også lidt underligt, at placering kun indgår i modellen hos Brøndby, men at den også indgår i den samlede model for tilfredshed.

Årsagen hertil skal med stor sandsynlighed findes i det bagvedliggende datasæt. Her er hvert spørgeskema angivet som en streng med mange informationer (køn, alder, tidsrum etc.), og hvis man så kun vælger en del af besvarelsene ud, kan det være, at der viser sig en sammenhæng, som man ikke ville få, hvis man så på hele datasættet samlet.

11.0 Konklusion

Jeg har i denne opgave haft fokus på kundetilfredshed hos McDonald's.

Først har jeg haft kigget lidt på restaurationsbranchen og det marked som McDonald's befinder sig i. Her så jeg, at der er stor konkurrence om kunderne, og derfor er det nødvendigt at give en god kundeservice, hvis man vil overleve som virksomhed. Får kunden ikke det som han forventer, finder han bare et andet sted næste gang sulten melder sig.

Jeg så, at ordet tilfredshed kan defineres forskelligt, og deraf have flere betydninger. Derfor er det vigtigt, at når man taler om tilfredshed, at man har samme referenceramme, ellers kan man i princippet tale om to forskellige ting. Dette er også vigtigt at have in mente når man begynder at analysere sine data. Har respondenterne svaret på det samme – er besvarelsene valide og reliable.

"The Service Profit Chain"-teorien viste, at fokus på kundetilfredshed er en god investering, da den har en direkte påvirkning på virksomhedens profit.

Herefter foretog jeg en analyse af kundetilfredsheden hos McDonald's. Jeg fik klarlagt, at tilfredsheden ikke er den samme hos de involverede restauranter. Jeg så at den restaurant, der havde den største andel af mest tilfredse kunder, var restauranten i Esbjerg, mens den restaurant, hvis andel af mest tilfredse kunder var lavest, var restauranten i Køge.

Ved at opskrive en model for tilfredsheden for alle restauranter samlet set, kom jeg frem til at der skulle benyttes hele 6 af de syv forklarende variable, nemlig besøgsfrekvens, placering, køn, børn, alder og tidsrum. Hvilken ugedag kunden kom og besøgte restauranten på, havde ingen indflydelse, når man så på alle restauranterne samlet set.

Så man derimod på restauranten i Brøndby, kunne variationen i tilfredshed tilskrives to variable, nemlig placering og ugedag. Der var altså forskel på kundernes tilfredshed alt efter om de besøgte restauranten instore eller kom i McDrive, men også om de besøgte restauranten torsdag, fredag, lørdag eller søndag.

Tilsvarende så jeg, at variationen i tilfredshed hos de andre restauranter, kunne beskrives ved alt lige fra én til fire variable, og at det var forskellige variable alt afhængig af, hvilken restaurant det drejede sig omkring.

12.0 Perspektivering

Jeg havde begrænset min opgave til ikke at se på vekselvirkninger i de modeller jeg stillede op for tilfredsheden. Det kunne være interessant at se, om der kunne findes en model for tilfredshed indeholdende netop vekselvirkninger.

Derudover kunne det også være interessant at finde modeller for de andre tilfredshedsparametre (spørgsmål 2-11). Vil man komme frem til de samme modeller, som dem jeg lige har fundet for den overordnede tilfredshed, eller vil der komme nogle helt nye modeller på bordet, og hvad ville de bidrage til?

Det kunne også være interessant at lave en kvalitativ analyse af tilfredsheden i restauranterne. Denne opgave har klarlagt, at der findes forskellige niveauer af tilfredshed for diverse parametre, men hvad er årsagen hertil? Hertil kunne man tage ud i restauranterne og få snakket med kunderne omkring deres oplevelse af restauranten.

Det kunne være meget interessant at se nærmere på tidligere kundetilfredshedsundersøgelser for restauranterne. Er der sket en positiv udvikling i tilfredsheden, er der status quo eller er det gået tilbage? Hvor ligger udviklingen henne af og hvordan er det kommet dertil? Disse spørgsmål kræver til dels historiske data, som jeg ikke har haft til rådighed, men det vil også kræve, at jeg fik en samtale i gang med restauranterne og hørte deres versioner af, hvilke tiltag de har haft gjort i restauranterne.

En anden interessant vinkel man kunne bygge videre på i opgaven er at se på tilfredsheden hos medarbejderne. Er der en sammenhæng i tilfredsheden hos medarbejderne og kunderne? Har de restauranter, der opnår den bedste score i kundetilfredshedsundersøgelsen, også den bedste score i medarbejdertilfredshedsundersøgelsen? Er der sket en udvikling i medarbejdernes tilfredshed og kan man finde årsagen til dette? Der er flere teorier der postulerer, at en høj medarbejdertilfredshed giver en høj kundetilfredshed og vice versa, men er det også tilfældet hos McDonald's?

Dette kunne også linkes op på restaurantens overskudsgrad. Er den højere i de restauranter, der har en høj medarbejdertilfredshed eller har det ikke noget at sige?

Der er altså rigtig mange spændende og interessante vinkler man kunne bygge videre på, og jeg håber, at jeg kan få lov at udforske dem en gang i fremtiden.

13.0 Litteraturliste

Nigel Hill, Greg Roche og Rachel Allen, "Customer Satisfaction – the customer experience through the customer's eyes", Cogent Publishing, 2007

Nigel Hill, John Brierley og Rob MacDougall, "How to measure customer satisfaction", Gower Publishing Limited, 1999.

Hayes, Bob E., "Measuring Customer Satisfaction – survey design, use, and statistical analysis methods", ASQ Quality Press, 2. Edition, 1998

Dobson, Annette J., "An Introduction To Generalized Linear Models", Chapman & Hall, 1990

Agresti, Alan, "An Introduction To Categorical Data Analysis", Wiley, 2. Edition, 2007

Agresti, Alan, "Categorical Data Analysis", Wiley, 2. Edition, 2002

Stokes, Maura, Charles Davis and Gary Koch, "Categorical Data Analysis using The SAS System", 2. Edition, 2001

14.0 Bilag