

EKSEMPELBESVARELSE MET4 HJEMMEEKSAMEN V18

27. APRIL 2018

Denne besvarelsen ble levert som hjemmeeksamen i faget MET4 vårsemesteret 2018, og er publisert i anonymisert form i samråd med den aktuelle studentgruppen og eksamenskontoret ved NHH. Vi takker for samarbeidet.

Se eget vedlegg for sensors kommentarer til denne oppgaven.



MET4 HJEMMEEKSAMEN

MISLIGHOLD AV KREDITTKORTGJELD FOR KUNDER I TAIWAN



Introduksjon

Over hele verden har det å kunne handle på kreditt muliggjort økt velferd ved å få tilgang på kapital før det er opptjent. Det har resultert i at flere besitter kredittkort, ofte uten å forstå hvilke forpliktelser et slikt lån har (Wang, u.d.). I Taiwan opplevde bankmarkedet en kontant- og kredittkortkrise i 2006, etter at banker utstedte kontanter og kredittkort til en stor andel underkvalifiserte søkere (Yeh & Lien, 2009). I den sammenheng ble det stilt flere spørsmål i markedets banker; Hvilke persongrupper er mest attraktive for å oppnå profittmaksimering, og har dermed lavest misligholdssannsynlighet blant kredittkortkandidater? Hvordan kan de mest kredittverdige kundene identifiseres, og hvordan burde bedriften tilpasse seg for å oppnå en mer kredittverdig kundebase? Formålet med denne oppgaven er å utforske disse problemene, og avgjøre hvilke variabler som kan påvirke en persons kredittverdighet.

Oppgave 1

For å gjennomføre våre analyser bruker vi et datasett fra University of California, Irvine, Machine Learning Repository, som inneholder et utvalg på 30 000 kredittkortkunder i Taiwan fra 2005. Datasettet inneholder ulik deskriptiv data for hver kredittkortkunde. Dataene vi anser som relevante for våre analyser er kjønn, utdanning, sivilstatus, alder og kredittgrense. De tre førstnevnte er nominale data, og de to siste er intervalldata. Da oppgaveteksten oppgir at betalingshistorikk ikke vil være mulig å innhente for nye kunder, anser vi variablene “historikk over tidligere betaling”, “fakturabeløp” og “beløp for forrige betaling” som mindre relevante, og velger å ikke se videre på disse i analysen.

Ettersom enkelte av forklaringene til variablene i datasettet er mangelfulle, gjør vi en selvstendig vurdering i forhold til hvordan vi skal behandle dette. Eksempelvis ser vi at for variablene “Utdanning” og “Sivilstatus”, eksisterer det flere kategorier i datasettet enn hva som er oppgitt i artikkelen (Yeh & Lien, 2009). Vi velger derfor å ekskludere observasjoner som inneholder kategorien “Others”, samt verdier som ikke tilhører en kategori. Disse observasjonene utgjør en liten andel av de 30 000 observasjonene.

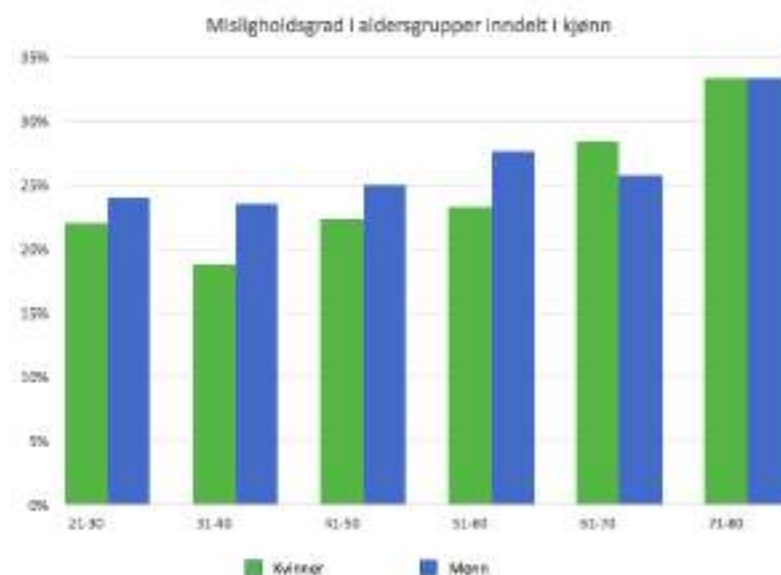
Kategori	Totalt antall	Misligholdsgrad	Statistikker	Kredittgrense (i TWD)	Alder
Kjønn					
Mann	11 575	24,3 %	Min	10 000	21
Kvinne	17 588	21,0 %	Max	1 000 000	79
Utdanning			Median	140 000	34
Graduate	10 531	19,2 %	St.Dev	129 747	9
Universitet	13 862	23,7 %	Gjennomsnitt	167 484	35
Videregående skole	4 770	25,3 %	25% persentil	50 000	28
			75% persentil	240 000	41
Sivilstatus					
Gift	13 425	23,8 %			
Singel	15 738	21,1 %			

Tabell 2 – Deskriptiv statistikk for nominale data

Tabell 1 – Deskriptiv statistikk for intervalldata

Tabell 1 illustrerer antall individer i de ulike kategoriene, samt sannsynlighet for mislighold av kredittkortgjeld gitt hver kategori. Vi ser at en større andel kvinner enn menn i datasettet har kredittkort, samt at de har lavere grad av mislighold enn menn. Videre viser tabellen tre kategorier for utdanningsnivå, hvor vi ser at misligholdsgraden øker med lavere utdanning. Til slutt fremkommer det at grad av mislighold er høyere for personer med sivilstatus gift enn for singel.

Tabell 2, viser informasjon om intervalldataene. Vi påpeker at alder har en svak venstreforskyvning da vi har en lav kvartilbredde, men at vi har høy max-verdi som tilsier at vi har en lang høyrehale. Dette betyr at de fleste kredittkortkundene er relativt unge, men at vi også har noen få eldre kunder. Videre viser tabellen at kredittgrense også har en venstreforskyvning med en lang høyrehale, da median, gjennomsnitt og kvartilbredde er relativt lave i forhold til max-verdien. Igjen tilsier det at de fleste kundene har relativt lav kredittgrense, men vi også har noen få kunder med høy kredittgrense.



Figur 1 - Misligholdsgrad i aldersgrupper inndelt i kjønn

Figur 1 illustrerer misligholdsgrad oppdelt i ulike aldersgrupper basert på total andel av hvert kjønn, og den representative andelen for mislighold. Det fremkommer at forskjellene mellom kjønn i de ulike aldersgruppene viker noe fra hverandre.

Oppgave 2

I denne oppgaven ønsker vi å estimere en regresjonsmodell ved bruk av forklaringsvariablene kjønn, oppnådd “graduate”-utdanning, og sivilstatus som “gift”. Det vil først bli forklart hvilken modell som brukes og hvorfor. Videre vil regresjonsmodellen og koeffisienter bli tolket og forklart. Avslutningsvis vil validiteten til regresjonsmodellen bli diskutert.

Responsvariabelen vår, “Default payment next month”, er en indikatorvariabel og tar kun to verdier, derfor kan vi ikke bruke minste kvadraters metode. I dette tilfellet benyttes logistisk

regresjon for å beskrive en eventuell sammenheng mellom settet av uavhengige variabler og et dikotomt avhengig utfall for neste måned; mislighold (1) eller ikke (0).

For å avgjøre om et individ vil misligholde lånet eller ikke, modellerer vi en lineær skårfunksjon $Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_r r$. Deretter brukes en logit-funksjon som regner om Z-verdiene til sannsynligheter, og gir oss verdier i intervallet (0,1). Vi kan tenke oss at dersom et individ får over en viss skår verdi, Z^* , vil det henge sammen med en høy sannsynlighet for mislighold, og vi får $Y=1$.

Logistisk regresjon av datasettet gir oss følgende modell: $\ln(\hat{y}) = -1,089 - 0,200X_1 - 0,275X_2 + 0,117X_3$.

<i>Dependent variable:</i>	
<i>`default payment next month`</i>	
Kvinne	-0.200*** (0.029)
Graduate	-0.275*** (0.030)
Gift	0.117*** (0.029)
Constant	-1.089*** (0.028)
Observations	29,163
Log Likelihood	-15,410.430
Akaike Inf. Crit.	30,828.870
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabell 3 - logistisk regresjonsmodell

Ved tolkning av logistisk regresjon vil en ikke legge vekt på størrelsen til koeffisientene, men fokusere på effekten av fortegnene til forklaringsvariablene. Fortegnene forklarer om variablene gir større(+) eller mindre(-) sannsynlighet for mislighold.

Forklaringsvariablene i denne modellen er indikatorvariabler. Variabelen “Kvinne” har negativt fortegn, som forteller at hvis individet er en kvinne vil det gi lavere sannsynlighet for mislighold, alt annet likt. Samtidig indikerer dette at det er større sannsynlighet for at menn misligholder kredittgjelden sin. Dette samsvarer med beskrivende statistikk vist i oppgave 1.

Videre er koeffisienten til forklaringsvariabelen “Graduate” også negativ, hvilket betyr at hvis individet har en “graduate”-utdanning er det mindre sannsynlig for mislighold, alt annet likt. På den andre siden vil lavere utdanningsnivå gi høyere misligholdssannsynlighet. Den siste forklaringsvariabelen, “Gift”, er positiv og forteller at hvis individet er gift, er sannsynligheten større for at dette individet misligholder lånet sitt, alt annet likt. Alle forklaringsvariablene er signifikante i den forståelse at de er klart forskjellig fra null på et 1%-nivå.

Selv med klart signifikante forklaringsvariabler, må det tas hensyn til at datasettet inneholder mange observasjoner (30 000). Med så mange observasjoner, kan det resultere i at selv en liten forskjell kan bli signifikant. Det kan for eksempel være bemerkelsesverdig at de som er gift tilsynelatende har større sannsynlighet for å misligholde, enn de som er single, da dette ikke er rent intuitivt. Det bør og presiseres at det kan være utenforstående variabler som bedre vil kunne forklare sannsynligheten for mislighold.

Det er også viktig å drøfte modellens validitet for å sikre at modellen er god. For logistisk regresjon er det noen forutsetninger som må være oppfylt, blant annet at den avhengige variabelen er binær, og at vi har et stort utvalg observasjoner. Begge forutsetningene er i vårt tilfelle oppfylt. Videre forutsetter vi at observasjonene er uavhengige av hverandre. Det foreligger et krav om at det skal være moderat eller ingen multikollinearitet mellom variablene. For å teste for multikollinearitet i den logistiske regresjonsmodellen, anvendte vi en Variance Inflation Factor (VIF) test. Resultatet ga verdier tilnærmet 1, som tilsier moderat korrelasjon mellom utdanning, kjønn og status. Det er likevel verdt å merke at dette er en relativt grov vurdering av multikollinearitet.

Ut fra denne drøftelsen konkluderer vi med at de tre forklaringsvariablene er signifikante og dermed påvirker sannsynligheten for mislighold. Dersom et individ er en kvinne eller har “graduate”-utdanning vil vi predikere en lavere sannsynlighet for mislighold, men dersom individet er gift, vil det gi en høyere misligholdssannsynlighet. Det bør likevel tas i betraktning at det sannsynligvis finnes andre utenforstående variabler som bedre kan forklare sannsynligheten, og at vi har en moderat grad av korrelasjon. Til tross for dette, anser vi modellen som anvendelig for prediksjon.

Oppgave 3

I denne oppgaven tar vi sikte på å undersøke salg av kredittkort til forbipasserende basert på enkle observerbare karakteristika, opp mot personens sannsynlige kredittverdighet. Vi forutsetter at gateselgerne av kredittkort ønsker å vite hvem de skal rette seg mot i første omgang, og deretter vil kunne ha en kort samtale med individet. Derfor presenteres først en modell basert på personkarakteristikkene alder og kjønn, og deretter inkluderer vi variabler som kan innhentes ved en kort samtale. Ettersom hensikten med oppgaven er å prøve å identifisere personer med en misligholdssannsynlighet på under 25%, vil vi trekke ut gruppene som tilfredsstiller kravene.

Basert på den logistiske regresjonsmodellen i oppgave 2, er kjønn statistisk signifikant som forklaringsvariabel for misligholdssannsynligheten. Vi er også interessert i å si noe om alder, men denne variabelen er ikke signifikant på 5% nivå. Dermed kan vi ikke si med sikkerhet om alder påvirker sannsynligheten for mislighold. På den andre siden er alder en lett gjenkjennelig karakteristikk som vi ønsker å ha med. Vi har derfor valgt å gruppere alder i intervaller på 10 år for å tegne et klarere bilde, og undersøke om koeffisientene til de ulike alderskategoriene viser seg å være signifikant forskjellig fra null (Heldal, 2006). Ved å kjøre en regresjon av kjønn og aldersgrupper, får vi at kjønn og enkelte av aldersgruppene er signifikante på 5% nivå.

	Alders kategorier					
	(20-30)	(30-40)	(40-50)	(50-60)	(60-70)	(70-80)
Kvinne	0.215	0.195	0.220	0.236	0.249	0.309
Mann	0.248	0.225	0.254	0.271	0.286	0.350

Tabell 4 – prediksjon for aldersgrupper og kjønn

Av prediksjonene i tabell 4 observerer vi en trend som viser til at kvinner stort sett ligger innenfor grensen for misligholdssannsynlighet på 25%, i motsetning til menn som varierer rundt grensen. Dette samsvarer med funnene i oppgave 2. I tillegg er det visse aldersgrupper som fremkommer mer profittmaksimerende ved å ha lavere misligholdssannsynlighet, hvor kvinner mellom 30-40 er den mest attraktive gruppen. Det kan trekkes frem at kvinner over 60 år og menn over 40 år ligger nær eller over 25% grensen, og derfor bør unngås.

Konklusjonen basert på prediksjonen vår er derfor at gateselgerne bør rette seg mot forbipasserende kvinner under 60 og menn under 40. Dette er det konkrete svaret på spørsmålet i forhold til hvem gateselgerne skal henvende seg til. Imidlertid skal en være noe kritisk til denne grensen, da antall observasjoner er relativt få for det eldre sjiktet. Dette er et faktum som kan forklare hvorfor vi har en betydelig økning i sannsynlighet for mislighold når alderen øker.

Videre ønsker vi å inkludere enda et steg i vår analyse, da vi ser at ulike grupper fra tabell 4 ligger relativt nærme grensen på 25%. Vi anser det derfor som hensiktsmessig å inkludere flere forklaringsvariabler som kan innhentes gjennom raske møter med potensielle nye kunder. Basert på dette ønsker vi å undersøke hvorvidt fokus på individers utdanning og status kan bidra til å redusere sannsynligheten for mislighold. Status og utdanning er av interesse da begge forklaringsvariablene er statistisk signifikante, og er informasjon som enkelt kan innhentes i korte samtaler.

	Graduate	Universitet	Videregående skole
Kvinne	0.180	0.223	0.238
Mann	0.210	0.259	0.275

Tabell 5 - prediksjon for utdanning og kjønn

	Gift	Singel
Kvinne	0.224	0.198
Mann	0.260	0.231

Tabell 6 - Prediksjon for sivilstatus og kjønn

For å undersøke hvilke grupper i utvalget som har lavest sannsynlighet for mislighold, har vi brukt to ulike regresjonsmodeller for å predikere sannsynlighet. Av tabell 5 ser vi predikerte sannsynligheter for mislighold neste måned med utgangspunkt i kjønn og utdanning. Vi ser her en gjennomgående trend hvor både kvinner og menn med høyere utdanning har lavere sannsynlighet for mislighold, enn hvis de har lavere utdanning. Tabell 6 er basert på en regresjonsmodell som vurderer sannsynlighet for mislighold med hensyn til kjønn og sivilstatus. Med utgangspunkt i tabellen ser vi at single individer har en lavere sannsynlighet for mislighold enn gifte, men hvor kjønn også spiller en sentral rolle. Det er en gjennomgående trend i begge tabellene at kvinner generelt har en lavere sannsynlighet enn menn.

På bakgrunn av våre funn, finner vi det mest hensiktsmessig å lage en logistisk regresjonsmodell for prediksjon med forklaringsvariabler for alder, kjønn, utdanning og sivilstatus, vist i tabell 7. Variablene som er med i regresjonen er indikatorvariabler og derfor må vi utelate en av kategoriene for å unngå perfekt koellinearitet. Den utelatte kategorien fanges opp av konstantleddet slik at vi kan måle de andre kategoriene i forhold til denne. For alderskategoriene ser vi at sannsynligheten for mislighold er større ved eldre alderskategorier, og ellers er fortegnene konsistente med modellen i oppgave 2. Vi kan se fra tabellen at variablene, Gift, Kvinne og Utdanning er signifikante på 1% nivå. Videre er ikke alle alderskategoriene signifikante. Vi velger likevel å inkludere disse variablene fordi det gir gateselgerne mer presise karakteristikk å se etter. Det kan imidlertid påvirke troverdigheten av funnene, noe vi kommer tilbake til i oppgave 4.

<i>Dependent variable:</i>	
	<i>'default payment next month'</i>
Gift	0.162*** (0.033)
Kvinne	-0.207*** (0.029)
Universitet	0.250*** (0.032)
Videregående skole	0.314*** (0.043)
Alderskategori [30,40]	-0.191*** (0.036)
Alderskategori [40,50]	-0.106** (0.044)
Alderskategori [50,60]	-0.033 (0.062)
Alderskategori [60,70]	0.028 (0.147)
Alderskategori [70,80]	0.298 (0.550)
Constant	-1.286*** (0.035)
Observations	29,163
Log Likelihood	-15,392.710
Akaike Inf. Crit.	30,805.420
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabell 7 - Logistisk regresjonsmodell

For å få en indikasjon på hvor presise estimatene våre er, har vi laget prediksjoner for to fiktive individer, og estimert deres sannsynlighet for mislighold og tilhørende konfidensintervallet. Bredden til konfidensintervallet er en funksjon av utvalgets standardavvik, konfidensnivå og utvalgsstørrelse (Keller, 2017). Et bredt intervall indikerer høy usikkerhet knyttet til estimatet, og motsatt for et smalt intervall.

Det første individet er en singel kvinne med “graduate”-utdanning i alderen 30-40 år. Kvinnens sannsynlighet for mislighold vil være mellom [17.15%, 18.8%] med 95% konfidensintervall. Det andre individet er en gift mann med “high school”-utdanning i alderen 50-60 år. Hans sannsynlighet for mislighold vil være mellom [28.7%, 31.6%] med 95% konfidensintervall. Utvalget vårt inneholder svært mange observasjoner, som igjen gir oss mye informasjon, og kan være med på å forklare at begge konfidensintervallene er relativt smale og modellen gir presise prediksjoner.

Konklusjon

Etter å ha analysert og vurdert problemet vi står overfor - å enkelt kunne identifisere hvilke personer som har lavest sannsynlighet for å misligholde kredittgjelden sin - har vi kommet frem til noen kundegrupper som fremkommer som mest attraktive kunder.

Første steg er å tilnærme seg kvinner under 60 år og menn under 40 år, og unngå personer over denne alderen. Deretter er det samtalen som gjelder, hvor sivilstatus og utdanningsnivå er kritiske momenter å innhente informasjon om for å treffe målgruppen. I den forbindelse er ideelle kredittkortkunder personer med høyest mulig utdanning, og det er spesielt viktig å vike unna menn med lav utdanning. I forhold til sivilstatus, er single de mest attraktive kundene blant begge kjønn.

Oppsummert er kvinner på generell basis målgruppen vår, hvor vi må være ekstra oppmerksomme på høyt utdanningsnivå og sivilstatus som singel. Blant menn søker vi personer med et utdanningsnivå høyere enn graduate som er single. Det mest vesentlige er å unngå menn med lav utdanning, og spesielt menn over 40 år.

Feilvurdering av hvem som får tilgang på kredittkort vil ha negative konsekvenser for bedriften, og det er derfor vesentlig at dere som selgere tar høyde for både alder, kjønn, utdanning og sivilstatus sett under ett.

Oppgave 4

I denne oppgaven ønsker vi å vurdere validiteten til regresjonsmodellene, og den konklusjonen vi har fått basert på dette. Diskusjonen vil inkludere en vurdering av utgitt datasett, samt drøfte andre forhold vi anser som vesentlige for å gi en vurderingen av potensielle kunders kredittverdighet. Dette er en innledende drøftelse som skal understøtte de poengene vi til slutt fremlegger i et notat til direktøren.

For å vurdere regresjonsmodellenes validitet ønsker vi først og fremst igjen å påpeke at datasettets størrelse kan være problematisk, da selv de minste effekter kan bli signifikante. I forhold til at vi har relativt store forskjeller mellom menn og kvinner, kan også ses i sammenheng med at vi har en andel av kvinner på ca 60 %, og 40 % menn. Videre er det viktig å påpeke at fordelingen av aldersgrupper ikke er ideell, da det er store avvik mellom antall observasjoner i de ulike aldersgruppene. Dette kan resultere i misvisende resultater.

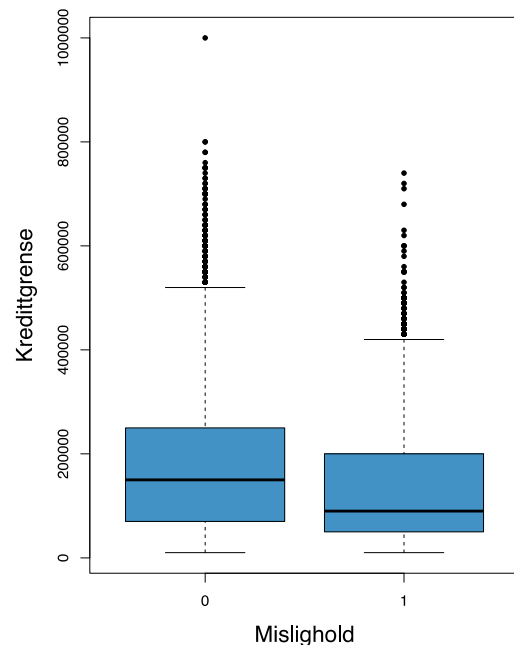
Det kan også være essensielt å drøfte hvorvidt de funnene vi har fått er representative for en bedrift i 2018, sett i lys av tidspunktet informasjonen er innhentet. Datasettet baseres på observasjoner gjort i Taiwan i 2005, en periode preget av at bankene hadde lave krav for utgivelse av kredittkort for å tjene mest mulig penger (Wang, u.d.). Året etter kollapset kontant- og kredittkortmarkedet i Taiwan, og bankene ble tvunget til å innføre strengere krav til kredittkortkunder. Strengere krav til inntekt var en av reguleringene som ble innført. Med det som utgangspunkt er det ikke sikkert at de funnene vi har avdekket i vår analyse er

relevante for en bedrift i dag, da potensielt en stor andel av kundene i datasettet ikke ville blitt godkjent per dags dato.

Når en velger variabler til en modell, ønsker en å trekke frem de viktigste trekkene ved virkeligheten (Keller, 2017). Den mest brukte variabelen blant banker for å avgjøre en persons kredittverdighet i tilknytning til utstedelse av kredittkort og kredittgrense er inntekt (Irby, 2018). Det kan derfor diskuteres hvorvidt de regresjonsmodellene utviklet i tidligere oppgaver representerer de viktigste trekkene ved individer når en bedrift ønsker å finne nye kunder.

Ettersom datasettet ikke inkluderer inntekt har vi ingen informasjon om inntekt vil være statistisk signifikant, men det er likevel viktig å påpeke at inntekt er av økonomisk signifikans. Vi forutsetter at kredittgrensen hovedsakelig er basert på inntekten til et individ. Inntekt er også enklere å kvantifisere, vurdere og rettferdiggjøre ved utgivelse av kredittkort, slik at en i høyere grad identifiserer og unngår underkvalifiserte søkere. Ved å utføre en logistisk regresjon med mislighold neste måned som responsvariabel, og kredittgrense som forklaringsvariabel, får vi at kredittgrense er signifikant på 1% nivå. Et spørsmål er hvorvidt kredittgrense samvarierer med en av de andre forklaringsvariablene som tidligere er brukt, eksempelvis utdanning. Ved å legge til høy utdanning (Graduate) i regresjonen, ser vi at utdanning mister sin signifikans når sett sammen med kredittgrense. Dette støtter vår anbefaling - å inkludere inntekt som en variabel.

Dette kan videre bli studert i 'boxplottet' i Figur 2. Denne figuren illustrerer sannsynlighet for mislighold opp mot gitt kredittgrense. Medianen er representert ved den tykke linjen i boksen, og observasjonene over "strekene" viser uteliggere som drar opp gjennomsnittet. Medianen viser at personer som misligholder gjelden sin har en lavere kredittgrense enn personer som ikke misligholder gjelden sin. Basert på våre forutsetninger impliserer dette at personer med høyere kredittgrense, og dermed høyere inntekt, misligholder sjeldnere enn personer med lav inntekt. Personer med høy inntekt, vil kunne gi høyere forventet profitt, og derfor være mer attraktive kunder for bedriften.



Figur 2 - Boxplott

Notat til selskapets direktør:

Analysen fremlagt i oppgave 3 viser til mange interessante funn som selgerne kan innhente informasjon om i korte møter med potensielle kunder på gaten. Samtidig har vi belyst at denne salgsstrategien ikke gir presis nok informasjon i forhold til hvilke kunder som vil bidra til profittmaksimering, slik at muligheten for feilvurdering av profittmaksimerende kunder fortsatt kan være høy. Nevneverdige

momenter som svekker tilliten til konklusjonen er eksempelvis usikkerheten knyttet til forklaringsvariablene, sammen med markedssituasjonen i Taiwan på tidspunktet dataene ble hentet fra. Verdien er basert på tall fra 2005, året før kontant- og kredittkortkrisen, og det har siden blitt innført strengere reguleringskrav. Derfor er tallene vi bruker basert på ekstreme verdier, som trolig er lite representative for dagens markedssituasjon, sett fra et samfunnsøkonomisk perspektiv.

Ettersom presisjonsnivået ikke er tilstrekkelig høyt til å kunne ha tillit til konklusjonene, vil vi anbefale diverse metoder å kunne sikre seg en høyest mulig andel kredittverdige kunder. Først og fremst har vi sett det som vesentlig å inkludere en del andre variabler i vurderingen av personer som skal få tilgang på kredittkort, som inntekt, formue, gjeld og månedlige betalinger. Vi har funnet at høyere kredittgrense, som vi har forutsatt i hovedsak baserer seg på inntekt, gir lavere sannsynlighet for mislighold. Dette bygger på at det bør settes høyere krav til informasjonsinnhenting for salgskorpset ved utstedelse av kredittkort. Videre anbefaler vi å ikke selge inn en spesiell kredittgrense på gaten, men gjøre en mer nøyaktig vurdering basert på kvantitative faktorer, som inntekt, formue, gjeld og månedlige betalinger. I tillegg kan det være mer etisk holdbart å basere vurderingen av individet på nevnte kvantitative momenter, da analysene er basert på eksempelvis kjønn kan anses som diskriminerende.

I salgsbransjen er det vanlig praksis å basere insentivordninger på provisjon basert på kvantitative metoder, som antall salg som utføres, fremfor det kvalitative. Vi anbefaler dere derimot å basere provisjonsordninger på det kvalitative; å oppnå gode kunder med lav misligholdssannsynlighet. I tilknytning til dette er det viktig å ha strenge retningslinjer rundt hvilke kunder man aksepterer for å minimere sannsynligheten for mislighet av kredittgjeld.

Ved å ta hensyn til funnene i rapporten og ta i bruk anbefalingene presentert, vil selskapet kunne ha en bedre mulighet til å oppnå sitt mål om profittmaksimering.

Bibliografi

- Australian National University. (u.d.). *Global Business Immersion study tour – Taiwan*. Hentet April 27, 2018 fra Australian National University: <http://www.anu.edu.au/students/careers-opportunities/global-programs/global-business-immersion-study-tour-%E2%80%93-taiwan>
- Heldal, J. (2006). *Logistisk regresjon - kurskompendium i byråskolens kurs SM507*. Statistisk Sentralbyrå, Metoder og standarder. SSB.
- Irby, L. (2018, Februar 24). *12 Possible Reasons Your Credit Card Application Was Denied*. Hentet April 27, 2018 fra The balance: <https://www.thebalance.com/denied-credit-card-application-960247>
- Keller, G. (2017). *Statistics for Management and Economics*. Boston: Cengage Learning.
- Wang, E. (u.d.). *Taiwan's Credit Card Crisis*. Hentet April 27, 2018 fra Seven Pillars Institute: <https://sevenpillarsinstitute.org/case-studies/taiwans-credit-card-crisis/>
- Yeh, C., & Lien, C.-h. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Science Direct*.

Bedømmelse etter vurderingsskjema for hjemmeeksamen:

Presentasjon av tabeller og figurer:	2
Valg av metode:	3
Anvendelse av metoder:	3
Diskusjon:	3
Etikk:	0
Formalfeil:	0
 Totalscore:	 2.75/3

Intern sensor gjorde følgende notater under første gjennomlesning:

Flott innledning, lover godt. Flott om deskriptiv, tar med momentene om ikke-definerte verdier. To tabeller: en for nominale og en for kontinuerlige data. Gjør logistisk, og ser ut til å ha gjort et utplukk av observasjoner i stedet for å rekategorisere, fint det. Prøver seg ikke på direkte oversettelse av koeffisienter til sannsynligheter. Kommenterer fint at såpass mange observasjoner i seg selv vil gi små p-verdier. Drøfter kort potensiell multikollearitet. Oppgave 3: kjempefin variant, gjør to analyser, en som tar for seg observerbare ting (for å plukke ut) og en som tar for seg ting som kan komme frem gjennom en kort samtale. Fin diskusjon av alder gjennom kategorier etter å ha observert at alder i seg selv ikke er signifikant. Mangler figurer her som kunne vært nyttig å se på, og predikerer bare sannsynligheter for to individer. Konklusjonen ser ut til å være spot on. I oppgave 4 diskuteres potensielle problemer, påpeker at kredittmarkedet på Taiwan har forandret seg (det er jo det som er utgangspunktet for artikkelen). Litt vag konklusjon her, men viser likevel ganske god forståelse. Alt i alt: en meget sterk besvarelse.

Intern sensor foreslo følgende karakter for denne besvarelsen:

A

Merk at den foreslåtte karakteren ikke nødvendigvis er identisk med besvarelsens endelige karakter. Den kan ha blitt justert etter eventuell vurdering av ekstern sensor.