**Skriftlig Rereeksamen**

|  |  |  |
| --- | --- | --- |
| **Uddannelse:** | **Finansbachelor** | |
|  |  | |
| **Fag:** | **Valgfag Kvantitative metoder i den finansielle sektor** | |
|  |  | |
| **Dato:** | **DEMO 2019** | |
|  |  | |
| **Klokkeslæt:** | **2 timers skriftlig prøve DEMO** | |
| **Hjælpemidler:** | | Alle skriftlige og elektroniske hjælpemidler og hjælpeprogrammer er tilladte – Inkl. Internetadgang, som dog ikke må bruges til kommunikation med andre. |
| Data er vedlagt som Excel bilag i 2 filer. | | |
| **Vægtning af opgaver:** | | |
| Opgavesættet består af 5 delopgaver, der indgår i bedømmelsen af den samlede opgavebesvarelse med følgende omtrentlige vægte: | | |
| Opgave 1 20%  Opgave 2 20%  Opgave 3 20%  Opgave 4 10%  Opgave 5 10%  Opgave 6 20% | | |
| **I alt 100 %** | | |

|  |
| --- |
| **Aflevering af besvarelsen:** |
| Opgavebesvarelsen skal afleveres i henhold til skolens eksamensreglement  *jf. bekendtgørelse nr. 1519 af 16. december 2013.*  Se hjemskolens formalia.  Navn eller eksamensnummer skal fremgå af hver side. |

For at din besvarelse vurderes som komplet, skal du argumentere for dine valg af modeller, vurdere om eventuelle forudsætninger er opfyldte og fortolke dine resultater.

Bemærk data er ikke klargjorte, dvs. man skal selv sortere og opstille data, så man kan benytte analyseværktøjerne.

En bank ønsker på baggrund af historiske data for udviklingen aktiekurs at undersøge følgende:

1. *Benyt tidsrækkeanalyse ARIMA, til at angive den forudsagte aktiekurs for mindst 3 perioder frem. Angiv 95% konfidensinterval, for de forudsagte værdier. Data findes i datasættet* ***ARIMADEMO.xlsx***

*Regn efter i hånden for 3 perioder, om du kan komme til samme resultater.*

*Kode i R:*

*library(forecast)*

*auto.arima(ARIMADEMO)*

*forecast(auto.arima(ARIMADEMO),h=10)*

*plot(forecast(auto.arima(ARIMADEMO)))*

*Output R:*

*Series: ARIMADEMO*

*ARIMA(0,1,0) with drift*

*Coefficients:*

*drift*

*8.3210*

*s.e. 1.1112*

*sigma^2 estimated as 72.88: log likelihood=-206.17*

*AIC=416.34 AICc=416.55 BIC=420.46*

*Point Forecast Lo 80 Hi 80 Lo 95 Hi 95*

*60 589.1710 578.2308 600.1113 572.4394 605.9027*

*61 597.4921 582.0202 612.9639 573.8299 621.1542*

*62 605.8131 586.8640 624.7622 576.8330 634.7932*

*63 614.1341 592.2536 636.0147 580.6708 647.5975*

*64 622.4552 597.9920 646.9183 585.0420 659.8683*

*65 630.7762 603.9782 657.5743 589.7921 671.7603*

*66 639.0972 610.1520 668.0424 594.8294 683.3651*

*67 647.4183 616.4746 678.3620 600.0939 694.7426*

*68 655.7393 622.9185 688.5601 605.5443 705.9344*

*69 664.0603 629.4642 698.6565 611.1501 716.9706*

ARIMA(0,1,0) eller I(1) model. Random walk with a drift på *8.3210*

*Modellen kan opstilles som:*

*Eller hurtig udregning i hånden*

*t=60: 580.85+1\*8.3210*

*t=61: 580.85+2\*8.3210*

*t=62: 580.85+3\*8.3210*

*Et billede, der indeholder kort, tekst

Automatisk genereret beskrivelse*

Samme bank ønsker at undersøge forskellige sammenhænge blandt egne kunder. Man har indsamlet data vedr. følgende variable:

*Forklaring på variablene:*

|  |  |
| --- | --- |
| **Variabel** | **Type** |
| Ansøgt lånebeløb (i 1.000 kr.) | Kvantitativ variabel |
| Boligens salgspris (i 1.000 kr.) | Kvantitativ variabel |
| Salgspris for fraflyttet bolig (i 1.000 kr.) | Kvantitativ variabel |
| Ansøgning imødekommet | Dummy 0: nej, 1: ja |
| Månedlig husstandsindtægt før skat (i 1.000 kr.) | Kvantitativ variabel |
| Samlet nettoformue (i 1.000 kr.) | Kvantitativ variabel |
| Gift | Dummy 0: nej, 1: ja |
| Antal børn | Diskret variabel |
| Selvstændig | Dummy 0: nej, 1: ja |
| Kredithistorik OK | Dummy 0: nej, 1: ja |
| Boligudgift, % af disponibel indkomst (ifølge budget) | Kvantitativ variabel |
| Samlede faste udgifter, % af disponibel indkomst (ifølge budget) | Kvantitativ variabel |
| Arbejdsløshedsprocent i den branche, hvor låneansøger arbejder | Kvantitativ variabel |
| Gymnasial uddannelse eller derover | Dummy 0: nej, 1: ja |
| Ønske om fastforrentet lån? | Dummy 0: nej, 1: ja |
| Ønsket lånetid (måneder) | Kvantitativ variabel |

Resultaterne fra denne stikprøve findes i Excel-filen **Data\_bank.xlsx,** du skal benytte dette datasæt i besvarelsen af Opgave 2 til 6.

1. *Inddel antal børn i 0, 1, 2 og flere end 2 børn. Er den gennemsnitlige samlede nettoformue ens, når man ser på familiens antal af børn?*

|  |  |
| --- | --- |
| ANOVA/ANAVA | |
| Forudsætninger | Populationerne er normalfordelte dette undersøges vha. normalfraktildiagram.  Populationerne har IKKE samme varians, dette er undersøgt vha. Bartletts test for varianshomogenitet, se nedenfor. |
| Hypotese | *De gennemsnitlige samlede nettoformue er ens.*  *De gennemsnitlige samlede nettoformue er IKKE ens.* |
| Teststørrelse | F teststørrelse 0.0603 |
| P-værdi/signifikanssandsynlighed | 98.06% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien 98.06% er klart større end signifikansniveauet på 5%, kan vi ikke forkaste nulhypotesen. Bemærk forudsætningen om varianshomogenitet er IKKE opfyldt, hvilket kan give problemer med kvaliteten af testet. |
| Ikke-teknisk konklusion | De gennemsnitlige samlede nettoformue er ens uanset familiestørrelse. Bemærk forudsætning om varianshomogenitet er ikke opfyldt hvorfor konklusionen er behæftet med usikkerhed |
| Superløsningen tjekker om populationerne kan antages normalfordelte med normalfraktildiagrammer. | |

1. *Inddel fortsat antal børn i 0, 1, 2 og flere end 2 børn, inddel Ansøgt lånebeløb (i 1.000 kr.) i passende kategorier. Er der sammenhæng mellem børn og ansøgt lånebeløb?*

|  |  |
| --- | --- |
| Chi i anden test | |
| Forudsætninger | De forventede celleværdier skal være større end 5 i næsten alle celler, dog er en enkelt værdi mindre. |
| Hypotese | Der er uafhængighed dvs. ingen sammenhæng mellem antal børn og ansøgt lånebeløb.  Der er afhængighed dvs. sammenhæng mellem antal børn og ansøgt lånebeløb. |
| Teststørrelse | 14.61 |
| P-værdi/signifikanssandsynlighed | 2.35% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien 2.35% er mindre end signifikansniveauet på 5%, kan vi forkaste nulhypotesen. Bemærk konklusionen er niveaufølsom. |
| Ikke-teknisk konklusion | Der er afhængighed dvs. sammenhæng mellem antal børn og ansøgt lånebeløb. 2 cellebidrag med værdi 4.68 og 3.11 er interessante. Folk uden børn ansøger sjældnere om lån på over 2 mio. Folk med flere end 2 børn oftere end forventet. |
|  | |

1. *Afgør ud fra en statistisk analyse, om gifte oftere er selvstændige.*

|  |  |
| --- | --- |
| Chi i anden test | |
| Forudsætninger | De forventede celleværdier er større end 5. |
| Hypotese | Der er uafhængighed dvs. ingen sammenhæng mellem ægtestand om man er selvstændig.  Der er afhængighed dvs. sammenhæng mellem antal mellem ægtestand om man er selvstændig. |
| Teststørrelse | 3.51 |
| P-værdi/signifikanssandsynlighed | 6.11% |
| Signifikansniveau | 5% |
| Teknisk konklusion | Da p-værdien 6.11% er større end signifikansniveauet på 5%, kan vi ikke forkaste nulhypotesen. Bemærk konklusionen er niveaufølsom. |
| Ikke-teknisk konklusion | Der er uafhængighed dvs. ingen sammenhæng mellem mellem ægtestand om man er selvstændig. |
|  | |

1. *Banken ønsker at identificere variable med samme karakteristika, hvilken strukturanalyse, ville banken kunne benytte? (Du skal IKKE gennemføre analysen, men blot argumentere for hvilken strukturanalyse man vil kunne benytte.)*

*Her kan man benytte en faktoranalyse.* *Faktoranalysen kan belyse sammenhænge mellem variable i spørgeskemaet fra banken.*

1. *Banken er interesseret i en dybere analyse af de ugifte kunder. Tegn og indsæt et klyngeplot med 3 klynger, hvilke kunder er i klyngerne*

*VIGTIGT: INDEN du går i gang SKAL du reducere bankdatasættet.*

* *Du skal fjerne alle gifte personer fra datasættet.*
* *Fjern variablen ”Salgspris for fraflyttet bolig (i 1.000 kr.)”*
* *Fjern variablen ”Gift”*

*Datasættet kan reduceres i Excel med fx filter, hvis ikke du reducerer datasættet, kan du ikke gennemføre analysen da klyngeanalysen fordrer kvantitative data. Herunder er benyttet R til at reducere datasættet*

*# R kode*

*#Importer Data\_bank fra Excel til R*

*bank <- Data\_bank*

*bankugift <- subset(bank,bank$`Gift Dummy`==0) #Kun gifte i datasættet*

*bankugift <- bankugift[,c(-3,-7)] #Fjern variablene "Salgspris for fraflyttet bolig (i 1.000 kr.)" og "Gift"*

*pacman::p\_load(factoextra) # Installer og load factoextra pakken*

*km3 <- kmeans(scale(bankugift), 3, nstart = 25)*

*fviz\_cluster(km3, data = bankugift, main = "Klyngeplot bank opdelt i 3 klynger",repel = TRUE)*

**