Programmet beskriver tilfredsheden med EU i og omkring EU-zonen.

Først importeres datasættes, hvorefter dummy-variable for sex og alder laves. Efterfølgende oprettes variable for Ja eller nej til EU alt efter om individet vil være en del af EU eller ej. Dette opfølges af tilfredshedsvariable for individet. Datasættet kommer fra European Social Survey runde 9.

STRAT betegner en stratifikation, der inddeler respodenter ind i forskellige stratifikationer. I vores tilfælde bliver mænd og kvinder inddelt i køn og aldersgruppe.

Variablen STF betegner tilfredsheden med landets egne institutioner og dens egen økonomi. Her rangeres alle svarmuligheder fra 0-10, hvor nul er den værste og 10 er det bedste. Dette rangerer landets tilfredshed som

Jeg for står ikke spørgsmålet hundrede procent, da det ligner begge variable beregnes på samme måde, hvis altså STF tager de fem variable summeret og dividerer med antallet af paramterer, 5.

EU variablen angiver individernes villighed til at være en del af EU-zonen. Her bliver EU-borgere spurgt, om de fortsat vil være en del af EU. Ikke EU-borgere adspurgt om de vil være en del af EU, og om Storbritanniens borgere gerne vil tilbage i EU-samarbejdet.

Ikke på stående fod.

Belgien har den højeste gennemsnitlighed tilfredshed i datasættet med en værdi på 87%

Aldersgruppen er samlet set ikke godt representeret i det danske datasæt. De udgør dog også den mindste gruppe af deltagerne. Der er flest adspurgte mennesker i aldersgruppen mellem 40 og 70, hvilet også udgør den største del af Danmarks befolkning jf en befolkningspyramide. Ift. køn er spredningen tæt på at være lige, der er dog flere adspurgte mænd i datasættet.

Opgave 2

1

Indlæser datasættet og opretter min mit eget datasæt, som indeholder de ændrede variable. Herefter droppes alle vægte, der ikke skal bruges. Kolon angiver nummeret.

2

Koden angiver elevernes fordelinger ift. region, land og køn.

3

Kode 1 viser afvigelsen i køsnfordelingen i datasættet.

Kode 2 viser afvigelsen i kønsfordelingen, når det antages, at der er 49% kvinder og 51% mænd.

Kode 3 viser den vægtede afvigelse, når det antages, der er 49% kvinder og 51% mænd.

4  
For det ene køn er der 92 missing ud af 3816 og for det andet er der 134 ud af 3841. Derfor fås:

5

Table

Description automatically generated

6

Datasættet indeholder 100 skoler, og der er 7657 respondenter, som alle er elever(antaget). Derfor:

Opgave 3

1Chart, histogram

Description automatically generated

A picture containing text

Description automatically generated

Ved test ser vi, at gennemsnittet i datasættet for årgang 2015 er 0.5 fra landsgennemsnittet.

2

Table

Description automatically generated

Den øverste er fra Sjælland mens den nederste er fra hovedstadsregionen.

Det udledes at begge regioner ligger over lands gennemsnittet. Mens den ligger marginalt over i Region Sjælland, så afviger Hovedstadsområdet med 2,8 procent point fra landsgennemsnittet.

3

Køres programmet:   
*data student1;  
set hurtigt.studenterhue; if aargang\_XXXX=0;  
run;*

Her kan det udledes, at Fanø, Læsø og Samsø har mindst en variabel lig med nul. Dette giver god mening, da der måske ikke er nogen gymnasier i de kommuner pga det lave indbyggertal. Er der ikke nogen gymnasier, så kan der ikke være nogen studenter.

4

Table

Description automatically generated

Her ses det, at antallet af variable er faldet i og med, kommunerne uden nogen studenter er faldet. Det kan udledes, at når aargang\_2013 stiger med 1%, så stiger aargang\_2015 med 0,59. er meget lav, så det er kun en lille del af variationen, der beskrives. P-værdien er også meget lav.

5

Table

Description automatically generated

Den forklarende del stiger, ved at vi indfører en ny variabel. Her ses det, at en ændring i aargang\_2013 giver en mindre ændring i 2015. Beta\_1 her er 0.38, mens det før var 0,59. Beta2 er givet ved 0.51 procent. Standardfejlen for aargang 2013 er dog steget marginalt.

6

Table

Description automatically generated

R-værdien, beta-estimatorne, standardafvigelsen og t-værdierne er alle samme ens ift til sidste spørgsmål. Jeg tænker, jeg har lavet en fejl.

7

Chart, box and whisker chart

Description automatically generated

Det ses, at Hovedstadsområdet har det højeste gennemsnit og regionens fordeling ligger også højest. Region Nordjylland ligger lavest. Her ses det, at Nordjyllands øverste fraktil skærer i Syddanmarks nederste fraktil. Samtidig ligger det Nordjyllandske minimum længst nede og det er kun det Syddanmarks minimu, der kommer i nærheden af det. Afstandsvis, så har Hovedstaden den største spredning fra top til bund.

8

Table

Description automatically generated with low confidence

Det ses, at landsgennemsnittet for kommunerne, der havde afgangselever i 2015, er 16,56. Dette er mindre end landsgennemsnittet fra spørgsmål 1. Minimum og maksimumsværdierne stemmer overens med box-plottet fra forrige spørgsmål.

proc import datafile="/courses/d284cd65ba27fe300/Hurtigt/ESS9e03\_renset.sav" out=ud dbms = sav replace;  
run;

proc contents data=ud varnum; run;

proc format; value STRAT

; run;

1='unge mænd' 2='mid. mænd' 3='ældre mænd' 4='unge kvinder' 5='mid. kvinder' 6='ældre kvinder'

data ess9;  
set ud;  
if agea < 40 then age1=1;  
if 40<= agea< 70 then age1=2;  
if agea =>70 then age1=3;  
if age1=1 and gndr=1 then STRAT =1;   
if age1=2 and gndr=1 then STRAT =2; if age1=3 and gndr=1 then STRAT =3; if age1=1 and gndr=2 then STRAT =4; if age1=2 and gndr=2 then STRAT =5; if age1=3 and gndr=2 then STRAT =6;

EU=0;  
if (vteurmmb=1 or vteumbgb=1 or vteubcmb=1 ) then EU=1; \*\*\*\* tilfredshedsvariable \*\*\*;  
STF=mean(of stfdem stfeco stfedu stfgov stfhlth); STF1=(stfdem +stfeco+ stfedu +stfgov +stfhlth)/5;  
run;

proc freq data=ess9;  
where cntry='DK';  
table strat/  
testp=(0.184 0.234 0.075 0.185 0.226 0.096); format strat strat.;

run;

proc means data=ess9; class cntry;  
var EU;  
output out=b mean=; proc print data=b;

run;

proc import datafile="/courses/d284cd65ba27fe300/Hurtigt/PISA\_DNK\_18\_renset.sav"

out=PISA\_DNK\_18 dbms = sav replace;

run;

proc contents data=PISA\_DNK\_18 varnum;

run;

data a;

set PISA\_DNK\_18;

drop W\_FSTURWT:;

proc contents data=a varnum;

run;

proc freq data=a;

table stratum\*sex/norow nocol nopercent chisq;

run;

proc freq data=a;

table sex/norow nocol nopercent chisq;

run;

proc freq data=a;

table sex/norow nocol nopercent chisq testp=(0.49 0.51);

run;

proc freq data=a;

table sex/norow nocol nopercent chisq testp=(0.49 0.51);

weight W\_FSTUWT;

run;

proc freq data=a;

table ESCS/norow nocol nopercent chisq; where sex=0;

run;

proc freq data=a;

table ESCS/norow nocol nopercent chisq; where sex=1;

run;

proc means data=a;

class sex;

var score\_m score\_r score\_n; run;

Opgave 3

libname hurtigt "/home/u61758272/my\_shared\_file\_links/anders.milhoj/Hurtigt";

data student1;

set hurtigt.studenterhue;

run;

\*\*a\*\*;

proc sgplot data=hurtigt.studenterhue;

histogram aargang\_2015;

run;

proc means data=hurtigt.studenterhue;

var aargang\_2015;

output out=b mean=;

proc print data=b;

run;

\*\*b\*\*;

proc means data=hurtigt.studenterhue;

var aargang\_2015;

where region='SJAELLAND';

run;

proc means data=hurtigt.studenterhue;

var aargang\_2015;

where region='HOVEDSTADEN';

run;

\*\*c\*\*;

data student1;

set hurtigt.studenterhue; if aargang\_2014=0;

run;

\*\*d\*\*;

data student3;

set hurtigt.studenterhue; if aargang\_2013 > 0;

if aargang\_2014 > 0;

if aargang\_2015 > 0;

run;

proc reg data=student3;

model aargang\_2015=aargang\_2013;

run;

\*\*e\*\*;

proc reg data=student3;

model aargang\_2015=aargang\_2013 aargang\_2014/influence partial;

run;

\*\*f\*\*;

proc glm data=student3;

class region;

model aargang\_2015=aargang\_2013 aargang\_2014/solution;

run;

\*\*g\*\*;

proc anova data=student3;

class region;

model aargang\_2015=region;

run;

\*\*h\*\*;

proc means data=student3;

var aargang\_2015;

run;