Tirdagsopgaven

Opgave 1

```
libname ssvs "/courses/d284cd65ba27fe300/Sommerskole/Data";

proc import datafile='/courses/d284cd65ba27fe300/Sommerskole/Data/ESS9e03_renset.sav'
out=ud replace;
run;

proc format;
value   STRAT
        1='unge m nd '
        2='mid. m nd '
        3=' ldre    m nd '
        4='unge kvinder '
        5='mid. kvinder '
        6=' ldre    kvinder '
;

data DE9;
set ud;
where cntry='DE';
* nedenst ende udarbejdes 6 Strata baseret p   k n og alder **;
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;
* Der skal ikke her erstattes missings v rdier      *;
run;



*Opgave 1;
proc standard data=DE9 replace out=DE9_stand;
var trstep trstprl;
run;
proc means data=DE9;
var trstep trstprl;
run;
proc means data=DE9_stand;
var trstep trstprl;
run;

proc standard data=DE9 replace out=DE9_stand;
var trstep trstprl;
run;
```

Først beregner vi gennemsnittet for tillid i Tyskland til EP og deres eget parlament.

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
trstep	Trust in the European Parliament	2277	4.5621432	2.4017108	0	10.0000000
trstprl	Trust in country s parliament	2327	5.0988397	2.4761343	0	10.0000000

I nedenstående standardisere vi datasættet.(Generer alle de missing data points), men fastholder gennemsnittet.

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
trstep	Trust in the European Parliament	2358	4.5621432	2.3600818	0	10.0000000
trstprl	Trust in country s parliament	2358	5.0988397	2.4597970	0	10.0000000

Det kan udledes, at gennemsnittet forbliver ens, men ved at gendanne de manglende observationer, sænkes variansen. Gennemsnittet ændrer sig ikke, da de gendannede variable sættes til gennemsnittet.

Opgave 2

Vi bruger formlen:

$$V(\hat{R}) = \frac{1}{\bar{X}^2} * \frac{N-n}{N*n} * \tau_z^2 \Leftrightarrow$$

$$(\hat{R}) = \frac{1}{\bar{X}^2} * \frac{N-n}{N*n} * (s_y^2 + \hat{R}^2 s_x^2 - \hat{R} * \hat{\rho}_{xy} * s_y s_x) \Leftrightarrow$$

$$(\hat{R}) = \frac{1}{5,10^2} * \frac{2358-100}{2358*100} * (2,36^2 + 0,89^2 * 2,46^2 - 2*0,89*0,61349*2,36*2,46) \Leftrightarrow$$

$$\tau_z^2 = 4,02 \Leftrightarrow$$

$$(\hat{R}) = \frac{1}{5,10^2} * \frac{2358-100}{2358*100} * 4,02 = 0,0148 = 0,039^2$$

Opgave 3

```
proc surveyselect data=DE9 seed=69 n=100 out=stik1;
run;
proc means data=stik1 n mean std var maxdec=2;
var trstep;
run;
proc surveymeans data=stik1 N=2358;
var trstep;
run;
```

Konstruer første tabel:

Analysis Variable : trstep Trust in the European Parliament			
N	Mean	Std Dev	Variance
97	4.20	2.44	5.95

Statistics						
Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
trstep	Trust in the European Parliament	97	4.195876	0.242378	3.71476057	4.67699201

Det kan udledes, at gennemsnittet er stort set identisk med gennemsnittet for hele gennemsnittet.

Opgave 4

```
proc surveyselect data=DE9 seed=69 n=100 out=stik2;
run;
proc surveymeans data=stik2  mean std;
var trstep trstprl;
ratio trstep/trstprl;
run;
```

Variable	Label	N	Mean	Std Dev	Variance
trstep	Trust in the European Parliament	97	4.20	2.44	5.95
trstprl	Trust in country s parliament	99	4.94	2.33	5.45

Her ses det, at der gennemsnitteligt er større tiltro til landets eget parlament.

Ratio Analysis			
Numerator	Denominator	Ratio	Std Error
trstep	trstprl	0.840909	0.040717

Her er forholdet udregnet.

Opgave 5

```
proc surveyselect data=DE9 seed=69 n=100 out=out1
reps=10;
run;
proc means data=out1 n mean std var maxdec=2;
class Replicate;
var trstep trstprl;
output out=b mean= var=;
run;
proc means data=b n mean maxdec=2;
var trstep trstprl;
run;
```

Sample Replicate Number	N Obs	Variable	Label	N	Mean	Std Dev	Variance
1	100	trstep	Trust in the European Parliament	97	4.20	2.44	5.95
		trstprl	Trust in country s parliament	99	4.94	2.33	5.45
2	100	trstep	Trust in the European Parliament	98	4.08	2.51	6.28
		trstprl	Trust in country s parliament	98	4.26	2.33	5.43
3	100	trstep	Trust in the European Parliament	94	4.80	2.39	5.71
		trstprl	Trust in country s parliament	98	5.24	2.60	6.78
4	100	trstep	Trust in the European Parliament	94	4.32	2.32	5.36
		trstprl	Trust in country s parliament	98	5.19	2.56	6.55
5	100	trstep	Trust in the European Parliament	97	4.45	2.40	5.75
		trstprl	Trust in country s parliament	98	5.22	2.35	5.54
6	100	trstep	Trust in the European Parliament	96	4.72	2.36	5.57
		trstprl	Trust in country s parliament	100	5.34	2.31	5.32
7	100	trstep	Trust in the European Parliament	98	4.68	2.34	5.46
		trstprl	Trust in country s parliament	98	5.28	2.22	4.92
8	100	trstep	Trust in the European Parliament	93	4.62	2.26	5.11
		trstprl	Trust in country s parliament	97	5.23	2.44	5.95
9	100	trstep	Trust in the European Parliament	95	4.35	2.43	5.91
		trstprl	Trust in country s parliament	99	4.74	2.38	5.64
10	100	trstep	Trust in the European Parliament	96	4.23	2.37	5.63
		trstprl	Trust in country s parliament	99	5.00	2.53	6.41

Der ses en gennemgående tendens til, at man favoritisere hjemlandets politikere end Europa-parlamentets politikere. Dette kan ske, da Brussel og EU-unionen ligger fjernt for de fleste borgere, også selvom Tyskland har mest indflydelse i EU.
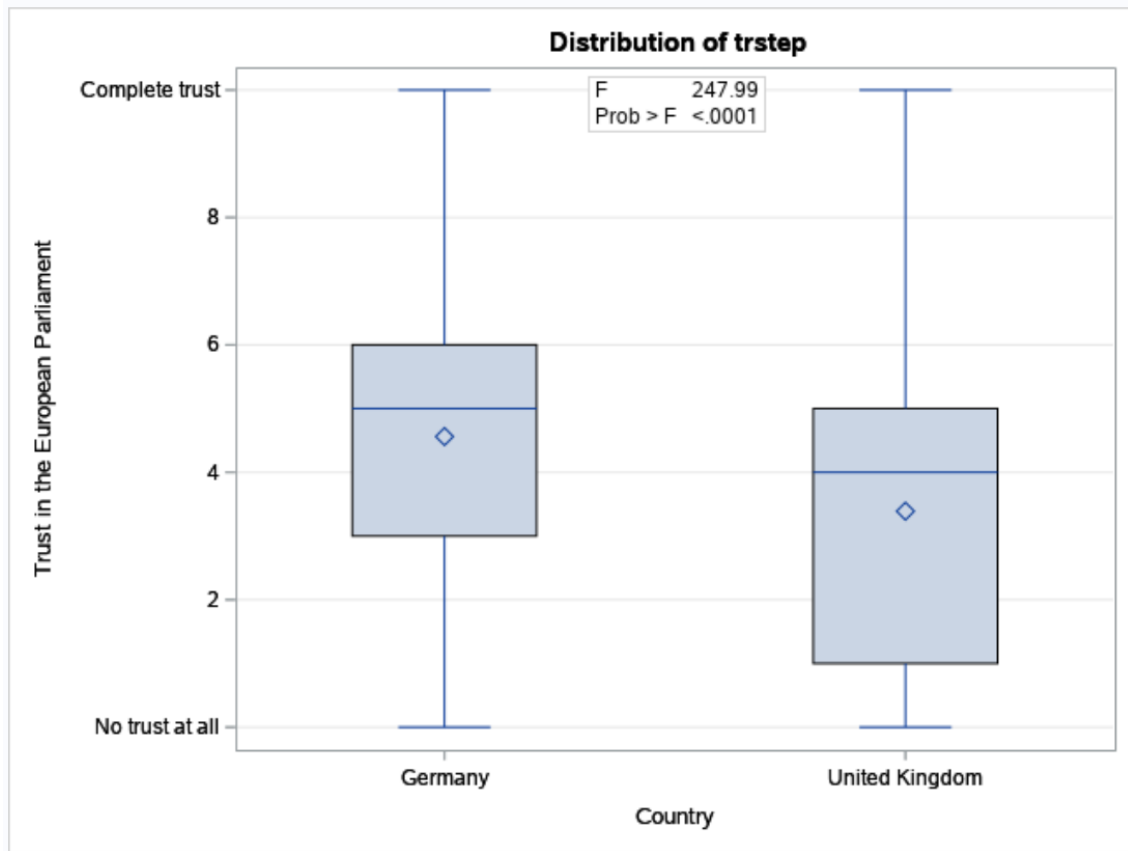
Del 2: Opgave 1

```
proc anova data=ud;
where cntry in ('DE','GB');
class cntry;
model trstep=cntry;
run;
```

Vi tester om tilliden er ens i de to lande, og får derfor

Source	DF	Anova SS	Mean Square	F Value	Pr > F
cntry	1	1501.800317	1501.800317	247.99	<.0001

Vi tester for tillieden i begge lande er ens. I ovenstående tabel ses det, at $p-værdien < 0,0001$, hvorved vi forkaster nulhypotesen. Ud fra dette kan vi konkluderer vi, at tilliden i landende ikke er ens. Dette

understøttes af nedenstående boxplot. Her ses det, at UK er længere nede end DE.

Del 2: Opgave 2

```
proc surveymeans data=ud N=80000000; where cntry='DE';
var trstep trstprl;
ratio trstep/trstprl;
run;
proc surveymeans data=ud N=65000000; where cntry='GB';
var trstep trstprl;
ratio trstep/trstprl;
run;
```

Ratio Analysis						
Numerator	Denominator	N	Ratio	Std Error	95% CL for Ratio	
trstep	trstprl	2266	0.895255	0.008245	0.87908776	0.91142300

Ratio Analysis						
Numerator	Denominator	N	Ratio	Std Error	95% CL for Ratio	
trstep	trstprl	2092	0.801901	0.011392	0.77955969	0.82424164

Figure 1: Tyskland
Figure 2: United Kingdom

Herved findes forholdet vha. formlen:

$$\frac{\hat{R}_{DE} - \hat{R}_{GB}}{\sqrt{SE_{DE}^2 + SE_{GB}^2}} = \frac{0,8947 - 0,8058}{\sqrt{0,0082^2 + 0,01139^2}} = \frac{0,0889}{\sqrt{0,00006724 + 0,0001297}} \Leftrightarrow$$

$$\frac{0,0889}{0,01403} = 6,34$$

Da dette overskrider et signifikansniveau på 95%, kan det konkluderes, at forholdet mellem E.P og PRL i landene er forskellige fra hinanden.

Del 3: Opgave 1

Programstumpen viser den tidligere model med de forskellige stratifikationer for hhv. alder og køn. Koden er dog uden en vægt. For at tilføje denne, skal man indsætte *weight vgt;*.

Del 3: Opgave 2

```
proc freq data=dk;
table strat/
testp=(0.19 0.228 0.077 0.183 0.228 0.094); weight vgt;
format strat strat;
run;
proc freq data=dk;
table strat/
testp=(0.19 0.228 0.077 0.183 0.228 0.094); format strat strat;
run;
```

STRAT	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
1	290.0339	18.43	19.00	290.0339	18.43
2	368.0899	23.39	22.80	658.1238	41.82
3	117.2872	7.45	7.70	775.411	49.27
4	291.4384	18.52	18.30	1066.849	67.78
5	355.8509	22.61	22.80	1422.7	90.39
6	151.1724	9.61	9.40	1573.873	100.00

STRAT	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
1	279	17.75	19.00	279	17.75
2	420	26.72	22.80	699	44.47
3	147	9.35	7.70	846	53.82
4	230	14.63	18.30	1076	68.45
5	359	22.84	22.80	1435	91.28
6	137	8.72	9.40	1572	100.00

Figure 3: Med vægte   Figure 4: Uden vægte

Det kan udledes, at ved at bruge vægte bliver frekvenserne mere ligefordelte. Ændringen er dog minimal. Altså kan det siges, at vægtene hjælper en smule.

Del 3: Opgave 3

STRAT	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent
1	279	290.03390	15.75280	18.4280	0.9947
2	420	368.08989	15.38039	23.3875	1.0240
3	147	117.28723	9.21322	7.4521	0.5981
4	230	291.43836	17.76114	18.5173	1.0779
5	359	355.85088	16.50301	22.6099	1.0547
6	137	151.17241	12.34384	9.6051	0.7773
Total	1572	1574	5.54804	100.0000	

Table of STRAT

Erstatningen betyder, at man tager frekvensen af en stikprøve fremfor hele universet. Ved at køre den givne kode, surveyfreq, fås den ovenstående tabel.

Del 3: Opgave 4

prtvtddk	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
Socialdemokratiet - The Social democrats	338	27.48	26.28	338	27.48
Det Radikale Venstre - The Radical Liberal Party	72	5.85	4.58	410	33.33
Det Konservative Folkeparti - Conservative People's Party	52	4.23	3.35	462	37.56
SF Socialistisk Folkeparti - Socialist People's Party	73	5.93	4.19	535	43.50
Dansk Folkeparti - Danish People's Party	172	13.98	21.08	707	57.48
Kristendemokraterne - Christian Democrats	15	1.22	0.83	722	58.70
Venstre, Danmarks Liberale Parti - The Liberal Party	300	24.39	19.47	1022	83.09
Liberal Alliance - Liberal Alliance	48	3.90	7.53	1070	86.99
Enhedslisten - Unity List - The Red-Green Alliance	91	7.40	7.80	1161	94.39
Alternativet - The Alternative	53	4.31	4.80	1214	98.70
Other	16	1.30	0.90	1230	100.00
Frequency Missing = 342					

Party voted for in last national election, Denmark

Ved at køre koden opnåes en tabel for, hvilke partier, der har fået hvilke stemmer. Koden resulterer i ovenstående tabel.

Del 3: Opgave 5

Party voted for in last national election, Denmark					
prtvtddk	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
Socialdemokratiet - The Social democrats	330.7958	27.29	26.28	330.7958	27.29
Det Radikale Venstre - The Radical Liberal Party	70.67596	5.83	4.58	401.4717	33.12
Det Konservative Folkeparti - Conservative People's Party	50.55	4.17	3.35	452.0217	37.29
SF Socialistisk Folkeparti - Socialist People's Party	74.53019	6.15	4.19	526.5519	43.44
Dansk Folkeparti - Danish People's Party	165.7405	13.67	21.08	692.2924	57.11
Kristendemokraterne - Christian Democrats	14.22867	1.17	0.83	706.5211	58.28
Venstre, Danmarks Liberale Parti - The Liberal Party	291.3025	24.03	19.47	997.8236	82.31
Liberal Alliance - Liberal Alliance	49.17131	4.06	7.53	1046.995	86.37
Enhedslisten - Unity List - The Red-Green Alliance	93.33949	7.70	7.80	1140.334	94.07
Alternativet - The Alternative	54.59204	4.50	4.80	1194.926	98.57
Other	17.34772	1.43	0.90	1212.274	100.00
Frequency Missing = 361.59850439					

Figure 5: Gamle vægte

Party voted for in last national election, Denmark					
prtvtddk	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
Socialdemokratiet - The Social democrats	108.623	30.46	26.28	108.623	30.46
Det Radikale Venstre - The Radical Liberal Party	15.24901	4.28	4.58	123.8721	34.74
Det Konservative Folkeparti - Conservative People's Party	13.13692	3.68	3.35	137.009	38.42
SF Socialistisk Folkeparti - Socialist People's Party	18.63583	5.23	4.19	155.6448	43.65
Dansk Folkeparti - Danish People's Party	57.92165	16.24	21.08	213.5665	59.89
Kristendemokraterne - Christian Democrats	4.286579	1.20	0.83	217.853	61.10
Venstre, Danmarks Liberale Parti - The Liberal Party	81.30965	22.80	19.47	299.1627	83.90
Liberal Alliance - Liberal Alliance	14.47545	4.06	7.53	313.6381	87.96
Enhedslisten - Unity List - The Red-Green Alliance	26.85414	7.53	7.80	340.4923	95.49
Alternativet - The Alternative	11.53175	3.23	4.80	352.024	98.72
Other	4.546692	1.28	0.90	356.5707	100.00
Frequency Missing = 125.79812786					

Figure 6: Nye vægte

Vægtningn på tværs af køn og alder virker ikke. Vi kan se, at DF er underrepræsenteret. Dette sker, da vælgerne af DF har tendens til ikke at deltage i disse interviews. Vægtene hjælper derfor en smule på at rette dette.