

Del 1

Opgave 2

0 er dansk 1 er anden	N Obs	Variable	Label	N	Mean
Dansk herkomst	5858	score_r score_m score_n ESCS	laese scoren matematik scoren naturfag scoren Index of economic, social and cultural status	5858 5858 5858 5844	504.89 511.91 497.29 0.56
Anden etnisk herkomst	1551	score_r score_m score_n ESCS	laese scoren matematik scoren naturfag scoren Index of economic, social and cultural status	1551 1551 1551 1520	433.92 448.66 422.60 -0.25

Figure 1: Etnicitet

sex	N Obs	Variable	Label	N	Mean
Pige	3816	score_r score_m score_n ESCS	laese scoren matematik scoren naturfag scoren Index of economic, social and cultural status	3816 3816 3816 3724	500.90 494.00 479.11 0.37
Dreng	3841	score_r score_m score_n ESCS	laese scoren matematik scoren naturfag scoren Index of economic, social and cultural status	3841 3841 3841 3707	475.34 500.52 480.09 0.41

Figure 2: Køn

Antal af bøger i hjemmet		antal	Gns. for læsescoren
0-10 books	1	1.274	435,82
11-25 books	2	1.295	463,69
26-100 books	3	2.284	491,13
101-200 books	4	1.252	516,87
201-500 books	5	980	543,87
More than 500 books	6	355	526,63
i alt		7.440	

Bruger ovenstående tabeller til at lave en dekomponering.

Oaxaca-Blinder-dekomponering med én forklarende variabel			
	Parameterestimer	Middelværdi	Middelværdi
Klassifikationsvariabel	Var_1	Var_1	Y_var
0	37.66	0.56	504.89
1	-	-0.25	433.92
Diff.	-	0.81	70.97
Differens	70.97	100%	
Forklaret	30.50	43%	
Uforklaret	40.47	57%	

Dette giver ovenstående resultater.

Opgave 2

Nu tilføjer jeg variablene DISCLIMA (*Disciplinary climate in test language lessons*), JOYREAD (*Joy/Like reading*) og PISADIFF (*Perception of difficulty of the PISA test*) Dette giver en mindre forklaringsgrad end

før. Ikke super gennemtænkte medtaget variable.

Differens	70.97	100%
Forklaret - Var 1	24.97	35%
Forklaret - Var 2	-0.11	0%
Forklaret - Var 3	11.50	16%
Uforklaret	34.61	49%

Opgave 3

Først generes alle observationer uden end ESCS-værdi en værdi vha. SURVEYIMPUTE. Ydermere bruges en proc standard til at fastholde middelværdierne. Her dannes der 226 værdier jf. nedenstående tabel:

Imputation Summary		
Observation Status	Number of Observations	Sum of Weights
Nonmissing	7431	7431
Missing	226	226
Missing, Imputed	226	226
Missing, Not Imputed	0	0

Herefter kører vi en regression på dansk/udenlandsk herkomst og får de nedenstående estimationsresultater.

Parameter Estimates					
Variable	Label	DF	Parameter Estimate	Standard Error	t Value Pr > t
Intercept	Intercept	1	474.82329	1.30470	363.93 <.0001
ESCS	Index of economic, social and cultural status	1	40.06405	1.43810	27.86 <.0001

Figure 3: Dansk Herkomst

Parameter Estimates					
Variable	Label	DF	Parameter Estimate	Standard Error	t Value Pr > t
Intercept	Intercept	1	426.22833	2.12715	200.38 <.0001
ESCS	Index of economic, social and cultural status	1	15.48770	2.14199	7.23 <.0001

Figure 4: Udenlandsk Herkomst

Differens	70.97	100%	Differens	70.97	100%
Forklaret	32.45	46%	Forklaret	12.55	18%
Uforklaret	38.52	54%	Uforklaret	58.42	82%

Figure 5: Dansk Herkomst

Figure 6: Udenlandsk Herkomst

Det ses, at forklaringsgraden for børn af udenlandsk herkomst mindskes, mens der for det danske børn ikke sker det store. En forklaringsgrad på 18% for de udenlandske børn er meget lavt. Tilføjes flere variable forstærkes analysen formentlig ikke.

Del 2

Opgave 4

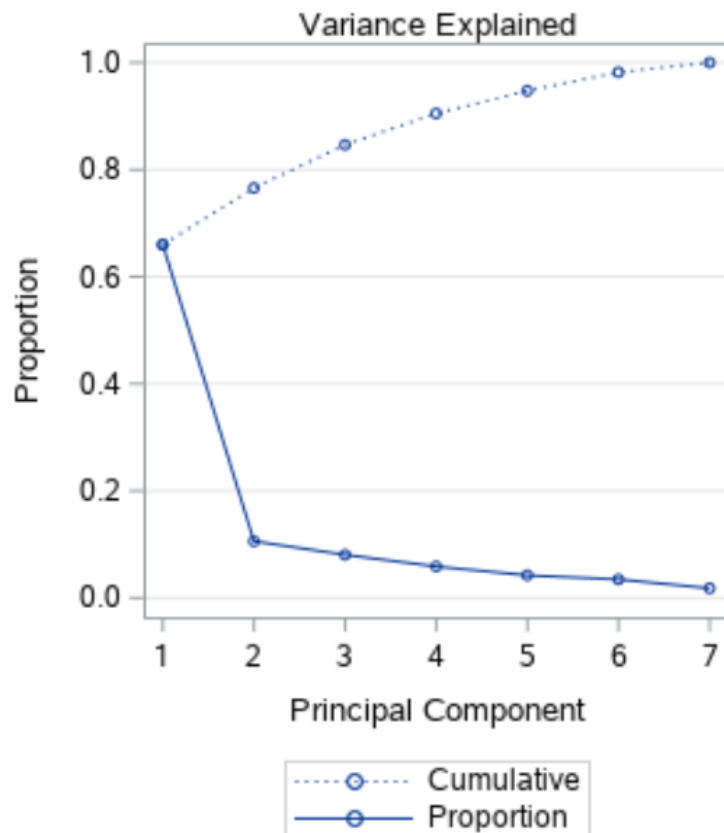
Country=Austria									
7 Variables:	trstlgl trstp1c trstp1t trstp1r trstp1t trstep trstun								
<table> <tr> <th colspan="2">Cronbach Coefficient Alpha</th></tr> <tr> <th>Variables</th><th>Alpha</th></tr> <tr> <td>Raw</td><td>0.888107</td></tr> <tr> <td>Standardized</td><td>0.888924</td></tr> </table>		Cronbach Coefficient Alpha		Variables	Alpha	Raw	0.888107	Standardized	0.888924
Cronbach Coefficient Alpha									
Variables	Alpha								
Raw	0.888107								
Standardized	0.888924								

Ovenfor er et eksempel for landet Østrig indsat. Her ses det, at Cronbach-Alpha-værdien er højere end 0,7. Dette indikerer, at vi kan sammenlægge de syv trust variable. Der gentager sig et mønster for alle lande, hvor Cronbach-Alpha-værdien for alle lande er over 0,7. Derfor kan vi slå dem sammen til en skala, da de ligner/opfører sig på samme måde. Hermed forklarer de samme mængde, som er spredt ud.

Opgave 5

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	4.61970640	3.87832792	0.6600	0.6600
2	0.74137848	0.17770408	0.1059	0.7659
3	0.56367440	0.15298734	0.0805	0.8464
4	0.41068706	0.11489315	0.0587	0.9051
5	0.29579391	0.05337839	0.0423	0.9473
6	0.24241552	0.11607130	0.0346	0.9820
7	0.12634422		0.0180	1.0000

Eigenvectors								
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
trstlgl	Trust in the legal system	0.358646	0.438797	0.114938	-.716882	0.383835	0.058369	0.031028
trstplc	Trust in the police	0.306992	0.758867	-.188621	0.521378	-.113032	-.091809	0.035478
trstplt	Trust in politicians	0.404499	-.241277	-.457092	0.068559	0.192778	0.177188	-.704254
trstprl	Trust in country s parliament	0.399770	-.079290	-.027076	-.304168	-.856764	0.075783	0.029283
trstprt	Trust in political parties	0.403051	-.316760	-.391870	0.109766	0.228124	0.149558	0.705120
trstep	Trust in the European Parliament	0.400903	-.233962	0.280751	0.083816	0.095166	-.829157	-.046203
trstun	Trust in the United Nations	0.361275	-.109553	0.713556	0.312974	0.087111	0.491077	-.040225

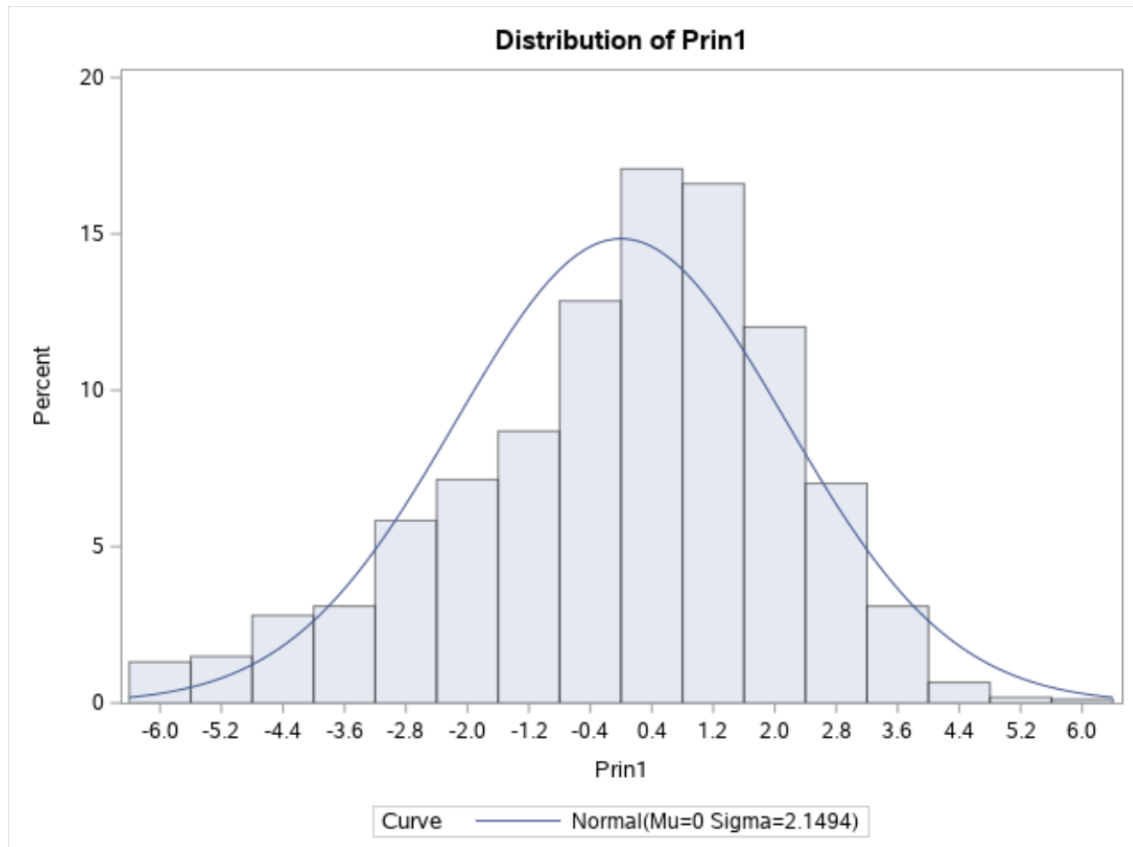


Af figuren "Variance Explained" kan det udledes, at de første tre principale komponenter forklarer mere end 80% af variationen i y-værdierne. Ydermere ses dette af de tre første egenverdier, fordi disse har en høj værdi, hvilket er et mål for deres forklaringskraft. Af det første principale komponent ses det, at den vægter alle cirka lige højt i dannelsen af den nye variabel(Prin1).

Fra det anden principale komponent (prin2), kan det ses, at denne nye variabel vægter primært tillid til retssystemet middelmådigt, tillid til politiet højt og tillid til FN lavt.

Fra det tredje principale komponent (prin3), kan det udledes, at den nye variabel vægter tillid til retssystemet og det EU-parlamentet lavt, troværdighed til FN meget højt og de resterende negativt.

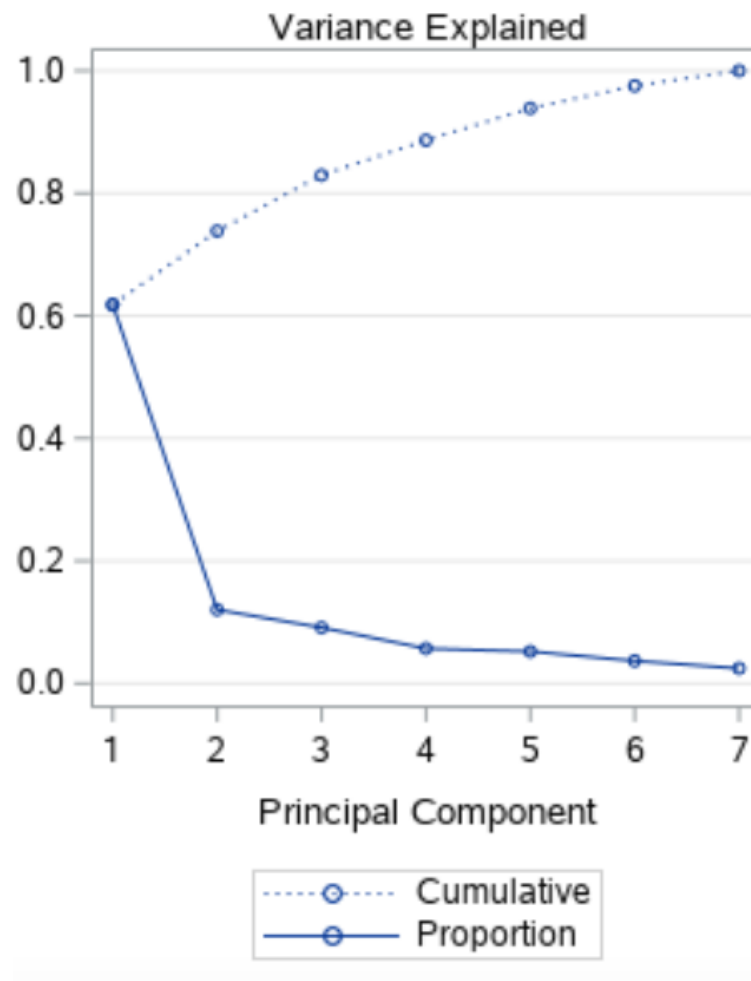
Opgave 6



Pearson Correlation Coefficients, N = 1680 Prob > r under H0: Rho=0			
	Prin1	Prin2	Prin3
Prin1	1.00000 1.0000	0.00000 1.0000	0.00000 1.0000
Prin2	0.00000 1.0000	1.00000 1.0000	0.00000 1.0000
Prin3	0.00000 1.0000	0.00000 1.0000	1.00000 1.0000

Det kan ses, at de tre principale komponenter er uafhængige. Dette ses ved at alle $corr = 0$

Opgave 7



Det kan ses, at over 80% af variationen i y er forklaret af de første tre principaler. Hvorfor de første tre kun undersøges.

Eigenvalues of the Correlation Matrix								
	Eigenvalue	Difference	Proportion	Cumulative				
1	4.32620206	3.48416046	0.6180	0.6180				
2	0.84204160	0.20368222	0.1203	0.7383				
3	0.63835939	0.23933572	0.0912	0.8295				
4	0.39902367	0.03405496	0.0570	0.8865				
5	0.36496871	0.10777288	0.0521	0.9387				
6	0.25719583	0.08498709	0.0367	0.9754				
7	0.17220874		0.0246	1.0000				

Eigenvectors								
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
trstigl	Trust in the legal system	0.373285	0.457919	-.147145	-.095935	-.636451	-.463383	0.017884
trstpic	Trust in the police	0.316289	0.709526	-.094063	0.284474	0.495737	0.246584	0.014210
trstplt	Trust in politicians	0.416310	-.292649	-.282790	-.006806	0.225035	-.200849	-.755013
trstpri	Trust in country s parliament	0.405796	-.157792	-.314668	-.309283	-.278178	0.722030	0.130575
trstprt	Trust in political parties	0.408028	-.341804	-.228762	0.038189	0.331946	-.375504	0.641639
trstep	Trust in the European Parliament	0.366675	-.238727	0.471441	0.700118	-.275235	0.142536	-.008127
trstun	Trust in the United Nations	0.348852	0.050209	0.718279	-.567930	0.187639	-.038546	-.024837

Ud fra prin1 kan det ses, at briterne tillid til alle trust-variable er middelmådig.

For prin2 er tilliden til retssystemet er middelmådig, tilliden til FN er lav og tilliden til deres eget politi er høj. Resten er negativ. Dette stemmer overens med Belgiens prin2.

Ud fra prin3 er tilliden til europa-parlamentet middelmådig og høj tillid til FN. Igen er resterende negativt.

Det kan udledes, at de to lande har mange sammenlignelige punkter. Dette giver god mening, da begge lande er højt industrialiserede vestlige lande, der typisk har de samme værdier.

Opgave 8

Jeg har to gange ramt eksempler, hvori 3 principale komponenter kan beskrive 80% af y-variationen. Derfor mener jeg, at tre principale komponenter vil være nok til at foretage fremtidige analyser.

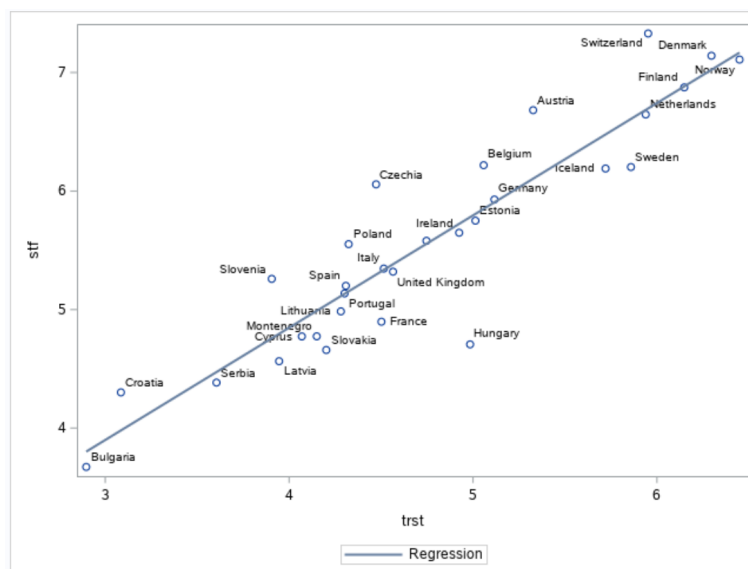
Del 3

Opgave 9

Pearson Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations		
	trst	stf
trst	1.00000 49338	0.67920 <.0001 49322
stf	0.67920 <.0001 49322	1.00000 49489

Det kan udledes, der er en positiv korrelation mellem tillidsvariablen og satisfiedvariablen. Dette ses, da resultaterne er signifikante og de er $0 <$.

Opgave 10+11



Det kan ses, at der er en tydelig sammenhæng mellem alle lande. Kun Ungarn afviger meget fra normen.