

TRIN 1: UDREGN 8. KLASSES NATIONAL TEST SCORE PÅ BAGGRUND AF 2. KLASSES TEST SCORE

Denne del kræver programmet STATA.

Trin 1A kræver adgang til registerdata fra Danmarks Statistik, hvilket er begrænset til forskere på godkendte institutioner (universiteter og tænketanke) og offentlige myndigheder som eksempelvis Finansministeriet og Undervisningsministeriet.

Trin 1B kan dog køres med et hvilket som helst datasæt med testscorer i forskellige år. Fx i en kommune eller på en skole.

TRIN 1A

```
/// INDLÆSER DATA MED NATIONALE TESTSCORER KOBLET TIL BØRNS OPRINDELSES LAND (ietype), FØDSELSÅR (foedselsaar), OG HVIS BOPÆL KAN PLACERES I HELE SKOLETIDEN
```

```
use "data.dta", clear
```

```
forval t = 1/3{
```

```
    bysort testaar fag: egen mean_p`t' = mean(theta_p`t')
```

```
    bysort testaar fag: egen sd_p`t' = sd(theta_p`t')
```

```
    gen zscore_p`t' = (theta_p`t' - mean_p`t') / sd_p`t'
```

```
}
```

```
gen zscore_p123 = (zscore_p1 + zscore_p2 + zscore_p3) / 3
```

```
bysort testaar fag: egen mean_p123 = mean(zscore_p123)
```

```
bysort testaar fag: egen sd_p123 = sd(zscore_p123)
```

```
keep if dan_nattest_gr2 ~= . & dan_nattest_gr8 ~= . /* begrænser til individer med valide svar på både 2. klasse og 8. klasse test score */
```

```
keep if foedselsaar >= 2002 & foedselsaar <= 2004 /* begrænser til individer født i perioden 2002-2004, begge inklusive */
```

```
keep if ietype == 1 /* begrænser til etnisk danske individer, efter Danmarks Statistiks definition */
```

TRIN 1B

```
xtile kvintil_testscore2 = testscore2, n(5) /* deler testscoren i 2. klasse op i fem lige store dele */
```

```
xtile kvintil_testscore8 = testscore8, n(5) /* deler testscoren i 8. klasse op i fem lige store dele */
```

```
tab rank5_dan_nattest_gr2 rank5_dan_nattest_gr8, col nofreq
```

	1	2	3	4	5	Total
1	10,380	5,819	2,935	1,238	412	20,784
	49.94	28	14.12	5.96	1.98	20
2	4,923	5,890	5,182	3,401	1,388	20,784

		23.69	28.34	24.93	16.36	6.68	20
3		2,772	4,421	5,352	5,199	3,039	20,783
		13.34	21.27	25.75	25.01	14.62	20
4		1,691	2,882	4,380	5,966	5,865	20,784
		8.14	13.87	21.07	28.7	28.22	20
5		1,018	1,772	2,934	4,980	10,079	20,783
		4.9	8.53	14.12	23.96	48.5	20
Total		20,784	20,784	20,783	20,784	20,783	103,918
		100	100	100	100	100	100

```
corr rank5_dan_nattest_gr2 rank5_dan_nattest_gr8
```

```
(obs=103,918)
```

```
0.5422
```

TRIN 2: LAV FIGUR (kan laves uden TRIN 1)

Denne del kræver programmet R som kan findes gratis her: <https://cran.r-project.org/bin/windows/>

Herefter skal pakken "networkD3" tilføjes – Det hele kan gøres ved at køre nedenstående kode i R:

```
install.packages("networkD3")
```

```
library(networkD3)
```

```
nodes = data.frame("name" =
```

```
  c("5th quintile", # Node 0
```

```
    "4th quintile", # Node 1
```

```
    "3rd quintile", # Node 2
```

```
    "2nd quintile", # Node 3
```

```
    "1st quintile", # Node 4
```

```
    "5th quintile", # Node 5
```

```
    "4th quintile", # Node 6
```

```
    "3rd quintile", # Node 7
```

```
    "2nd quintile", # Node 8
```

```
    "1st quintile" # Node 9
```

```
  ))
```

```

group <- c("5th quintile", "5th quintile", "5th quintile", "5th quintile", "5th quintile", "4th quintile", "4th quintile", "4th quintile", "4th quintile", "4th quintile", "3rd quintile", "3rd quintile", "3rd quintile", "3rd quintile", "3rd quintile", "3rd quintile", "2nd quintile", "2nd quintile", "2nd quintile", "2nd quintile", "2nd quintile", "1st quintile", "1st quintile", "1st quintile", "1st quintile", "1st quintile")

sankey_color <- 'd3.scaleOrdinal() .domain(["5th quintile", "4th quintile", "3rd quintile", "2nd quintile", "1st quintile"])
.range(["#FF0000", "#00A08A", "#F2AD00", "#F98400", "#5BBCD6"])'

%% INDSÆT TALLENE FRA TRIN 1B

links = as.data.frame(matrix(c(
    0, 5, round(as.numeric(10079)),
    0, 6, round(as.numeric(4980)),
    0, 7, round(as.numeric(2934)),
    0, 8, round(as.numeric(1772)),
    0, 9, round(as.numeric(1018)),
    1, 5, round(as.numeric(5865)),
    1, 6, round(as.numeric(5966)),
    1, 7, round(as.numeric(4380)),
    1, 8, round(as.numeric(2882)),
    1, 9, round(as.numeric(1691)),
    2, 5, round(as.numeric(3039)),
    2, 6, round(as.numeric(5199)),
    2, 7, round(as.numeric(5352)),
    2, 8, round(as.numeric(4421)),
    2, 9, round(as.numeric(2772)),
    3, 5, round(as.numeric(1388)),
    3, 6, round(as.numeric(3401)),
    3, 7, round(as.numeric(5182)),
    3, 8, round(as.numeric(5890)),
    3, 9, round(as.numeric(4923)),
    4, 5, round(as.numeric(412)),
    4, 6, round(as.numeric(1238)),
    4, 7, round(as.numeric(2935)),
    4, 8, round(as.numeric(5819)),
    4, 9, round(as.numeric(10380))),
    byrow = TRUE, ncol = 3))

```

```
links <- cbind(links,group)
names(links) = c("source", "target", "value","group")
sankeyNetwork(Links = links, Nodes = nodes,
              Source = "source", Target = "target",
              Value = "value", NodeID = "name",
              fontSize= 12, nodeWidth = 30,
              LinkGroup = "group", NodeGroup = "name" ,
              colourScale = sankey_color)
```