

# Anvendelse af R i Dansk økonomi i Europa

*Simon Fløj Thomsen\**

*september 09, 2022*

## **Abstract**

Formålet med dette dokument er at give en introduktion til anvendelsen af data og R i jeres kursus i Dansk økonomi i europa.

---

\*Aalborg University, [sft@business.aau.dk](mailto:sft@business.aau.dk), MaMTEP

# 1 Hints

## 1.1 Download af data

Formålet med denne øvelse er at gøre den studerende i stand til at finde og downloade data fra Danmarks statistik.

- Find data for BNP for Danmark for perioden 1966-2021
  - Gå ind på statistikbanken <https://www.statbank.dk/statbank5a/default.asp?w=1440>
  - Klik på 'Økonomi' under emner
  - Klik på 'Nationalregnskab'
  - Klik på 'Nøgletal for nationalregnskabet (BNP)'
  - Vælg datasæt 'NAN1'
  - Vælg tidsserien: 'B.1\*g BNP' → prisenhed: '2010-priser' → År: 'markér alle'
  - Klik på 'VIS TABEL'
  - Klik på 'gem som' 'xlsx' og gem filen i *projekt mappen!*

## 1.2 Håndtering af data i excel

- Gør data klar til import i R
  - Åben datasættet i excel
  - Markér området B3:BF4 og kopiér cellerne
  - Højreklik i celle A7 og vælg 'indsæt speciel' → 'transpose' og data er anvist i søljer
  - Slet række 1:6
  - Skriv 'Year' eller 'Aar' i celle A1 og 'BNP' i celle B1
  - Gem excelfilen

## 1.3 Indlæsning af data i R

- Indlæs data i R
  - Sørg for at du øverst i højere hjørne har valgt det rigtige projekt!
  - Åben et R-script (# angiver kommentarer)
  - Klik på 'Files' → Find navnet på det dataset du har gemt → 'Import Dataset' → 'Importer'.
  - Fremadrettet kan I indlæse data via nedenstående kode:

```
library(readxl)
BNP <- read_excel("bnp.xlsx")
View(BNP)
```

- Undersøg datasæt (kan findes øverst i højere hjørne under Environment)
- Klik på pilen ved siden af navnet på datasættet for at undersøge egenskaberne ved data (55 obs., 2 variable (Year og BNP))
- Se bestemt variable (Datasæt\$variabel)

```
BNP$Year
```

```
## [1] "1966" "1967" "1968" "1969" "1970" "1971" "1972" "1973" "1974" "1975"
## [11] "1976" "1977" "1978" "1979" "1980" "1981" "1982" "1983" "1984" "1985"
## [21] "1986" "1987" "1988" "1989" "1990" "1991" "1992" "1993" "1994" "1995"
## [31] "1996" "1997" "1998" "1999" "2000" "2001" "2002" "2003" "2004" "2005"
## [41] "2006" "2007" "2008" "2009" "2010" "2011" "2012" "2013" "2014" "2015"
## [51] "2016" "2017" "2018" "2019" "2020" "2021"
```

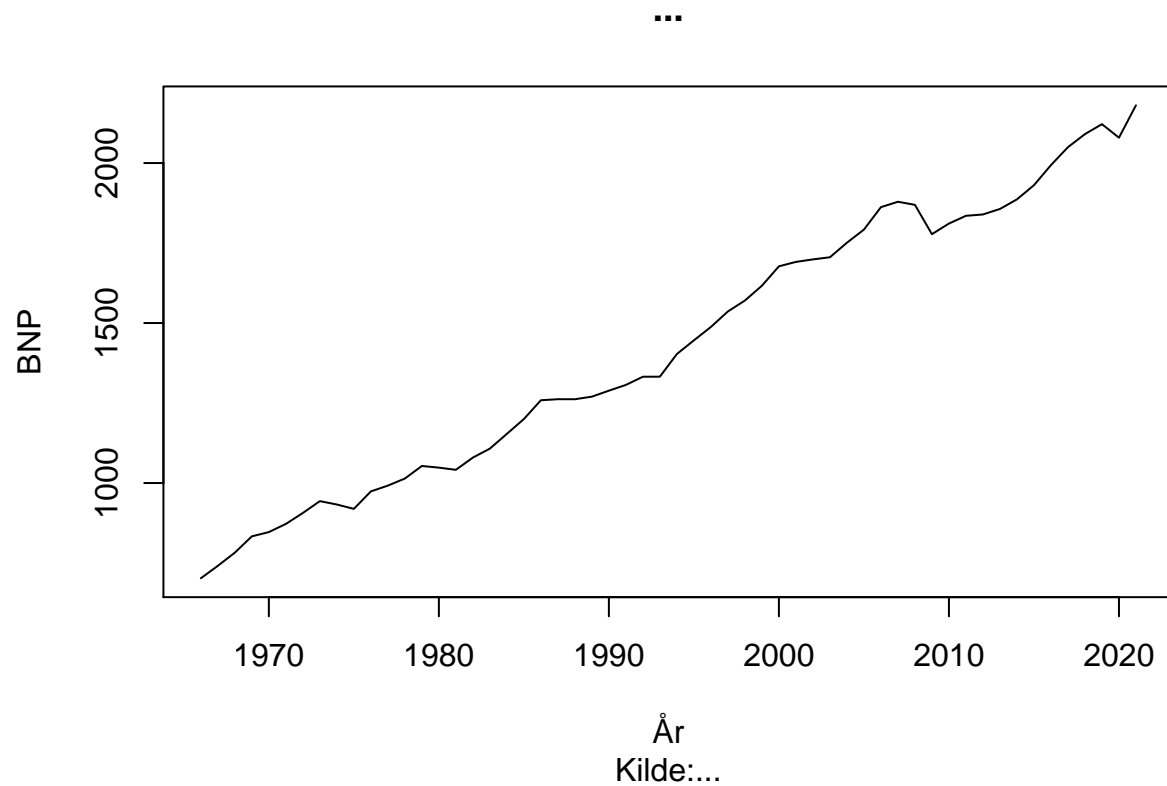
```
BNP$BNP
```

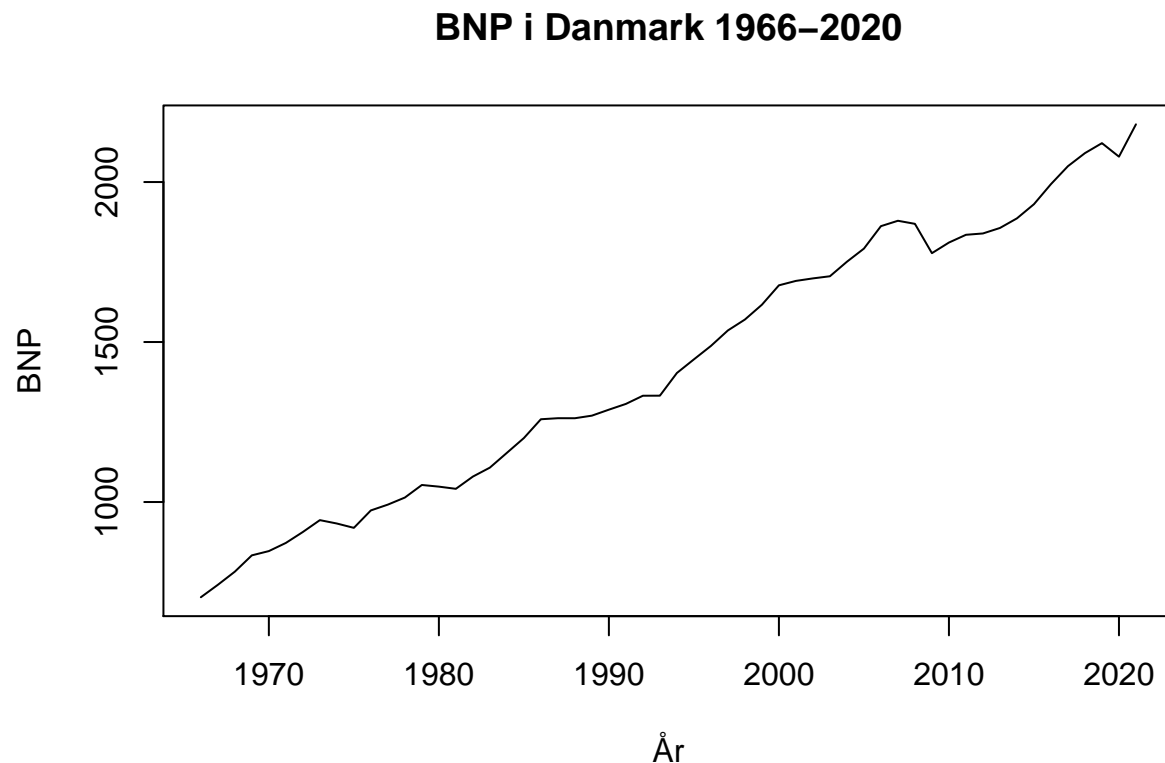
```
## [1] 702.4 741.2 782.4 833.3 846.6 872.1 906.3 943.4 932.8 919.2
## [11] 973.7 991.9 1014.0 1053.2 1048.1 1041.2 1079.5 1107.5 1153.7 1199.9
## [21] 1258.7 1261.9 1261.8 1269.9 1288.6 1306.6 1332.2 1332.3 1403.3 1445.8
## [31] 1487.8 1536.3 1570.3 1616.6 1677.2 1691.0 1698.9 1705.5 1751.0 1792.0
## [41] 1862.1 1879.0 1869.4 1777.7 1810.9 1835.1 1839.3 1856.5 1886.5 1930.7
## [51] 1993.4 2049.6 2090.4 2121.6 2079.3 2180.3
```

## 1.4 Plot figur med 1 linje

- Plot en figur med 'Year' på 1. akse og 'BNP' på 2. akse
- Anvend nedenstående kommando

```
plot(BNP$Year,BNP$BNP,type="l",lty=1,lwd=1,xlab="År",
ylab="BNP",col="black",main="...",sub="Kilde:...")
```





Kilde: Statistikbanken.dk/NAN1

- Forklaring:
  - ‘BNP\$Year’ angiver ‘x-variablen’
  - ‘BNP\$BNP’ angiver ‘y-variablen’
  - ‘type=“l”’ angiver linje-type (l=linje, p=punkter, s= steps, osv.)
  - ‘lty=1’ angiver linje-type (1=linje, 2=stiplet, 3=prikker, osv.)
  - ‘lwd’ angiver tykkelsen på linjen
  - ‘xlab=“år”’ angiver titlen på x-aksen
  - ‘ylab=“BNP”’ angiver titlen på y-aksen
  - ‘col=“black”’ angiver farven (enten “navn” eller tal: 1=sort, 2=rød osv.)
  - ‘main=“x”’ angiver en titel på figuren
  - ‘sub=“y”’ angiver en bemærkning/kilde, som vises under figuren
- Eksporter figur som PDF/PNG
  - Plots → save as image/pdf
  - Indlæs figur i fx Word

## 2 Øvelse 1

- I denne øvelse, skal I anvende ovenstående ‘opskrift’ til at illustrere udviklingen i den danske eksport (data er samme sted som BNP)

- Step 1. Importer data

```
Export <- read_excel("Export.xlsx")
View(Export)
```

- Step 2. Observer variable:

```
Export$Year
```

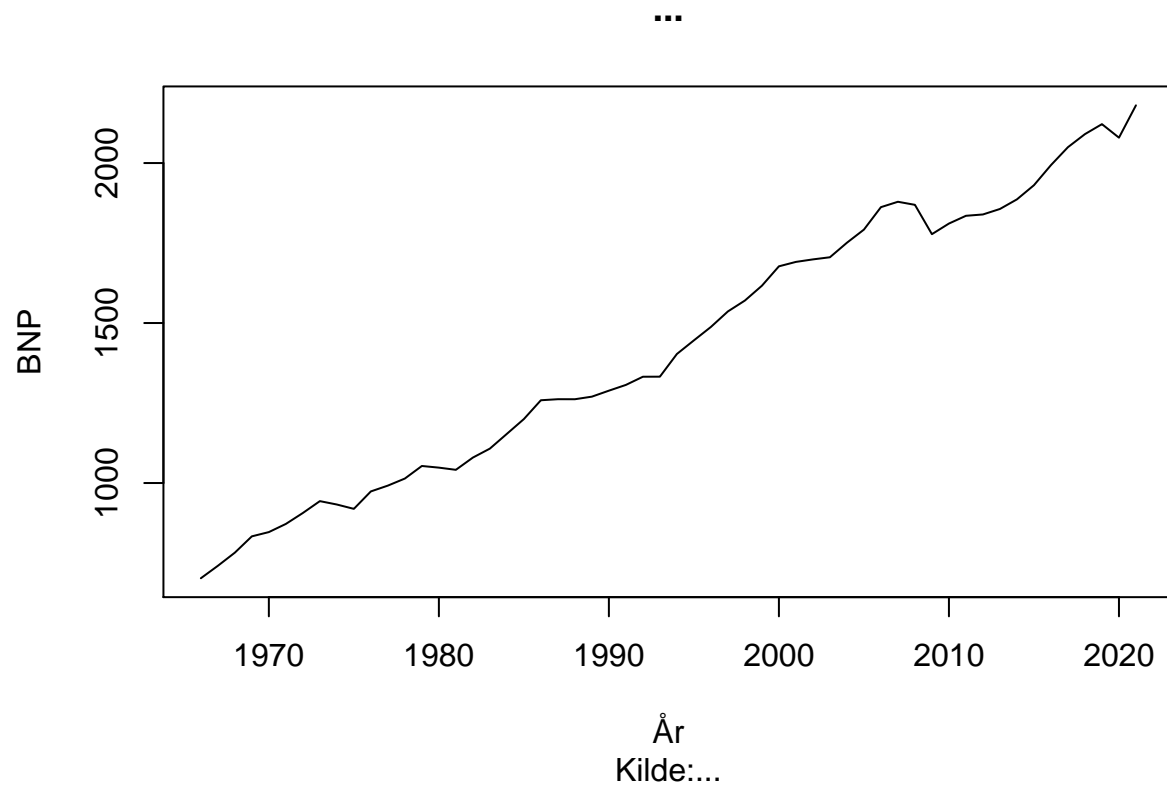
```
## [1] "1966" "1967" "1968" "1969" "1970" "1971" "1972" "1973" "1974" "1975"
## [11] "1976" "1977" "1978" "1979" "1980" "1981" "1982" "1983" "1984" "1985"
## [21] "1986" "1987" "1988" "1989" "1990" "1991" "1992" "1993" "1994" "1995"
## [31] "1996" "1997" "1998" "1999" "2000" "2001" "2002" "2003" "2004" "2005"
## [41] "2006" "2007" "2008" "2009" "2010" "2011" "2012" "2013" "2014" "2015"
## [51] "2016" "2017" "2018" "2019" "2020" "2021"
```

```
Export$X
```

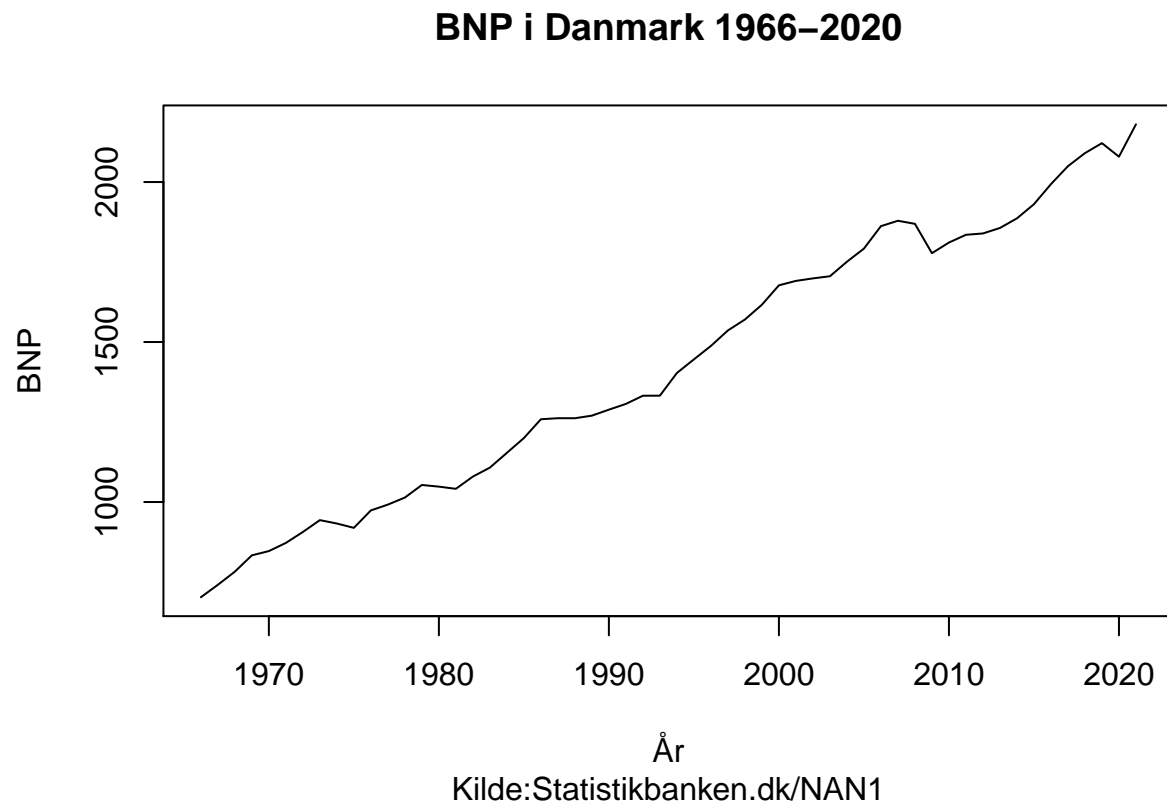
```
## [1] 123.6 128.2 140.7 149.2 154.9 164.7 173.4 187.8 194.2 192.8
## [11] 199.5 206.6 209.5 232.3 245.3 266.5 275.0 287.6 297.1 315.0
## [21] 319.2 334.7 365.3 382.3 407.3 432.4 433.6 438.9 475.1 488.8
## [31] 511.5 534.5 556.4 619.1 696.9 720.3 751.7 742.7 765.1 824.2
## [41] 909.4 942.6 979.1 888.8 914.9 980.8 992.2 1008.1 1039.7 1076.9
## [51] 1121.2 1175.1 1214.5 1269.3 1189.3 1284.3
```

- Step 3. vis udviklingen i eksport

```
plot(BNP$Year, BNP$BNP, type="l", lty=1, lwd=1, xlab="År",
     ylab="BNP", col="black", main="...", sub="Kilde:...")
```



- Step 4: Tilføj beskrivelser



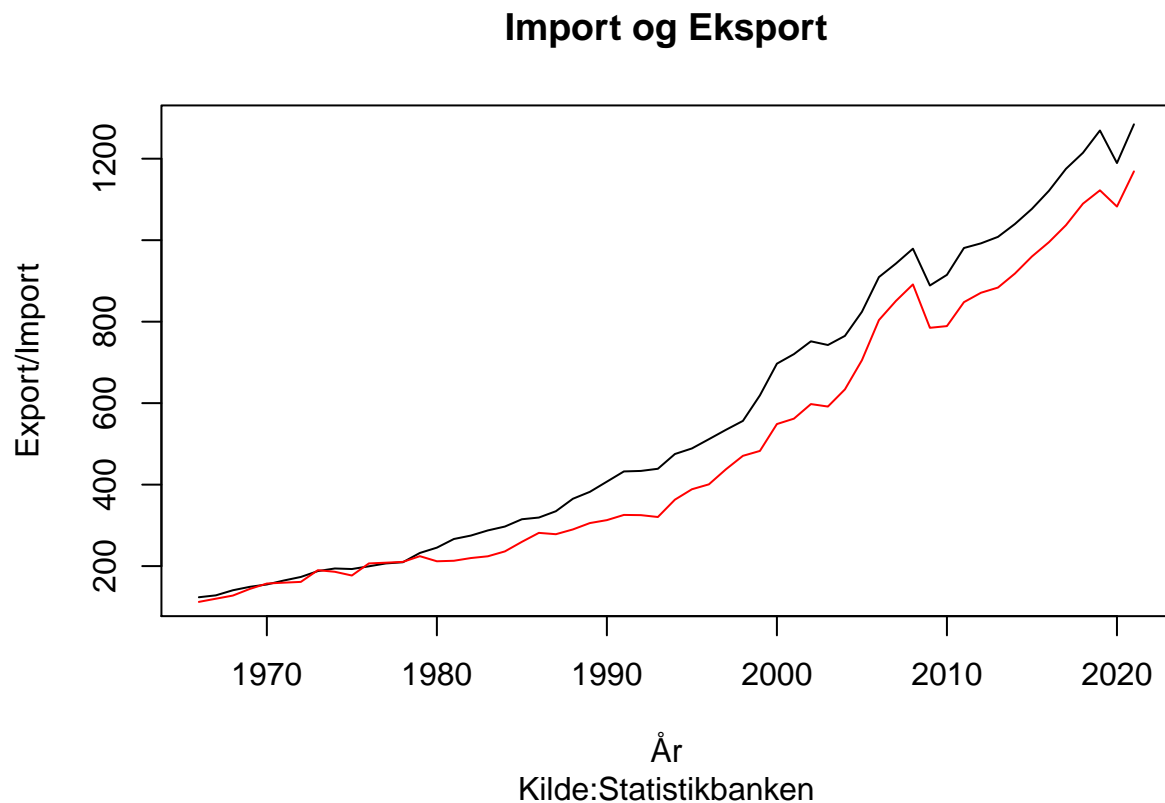
### 3 Øvelse 2

- I skal nu illustrere udviklingen i såvel eksport som import i samme figur
  - Hint: I skal tilføje kommandoen `lines()` efter `plot()`
- Step 1: Load import data:

```
Import <- read_excel("Import.xlsx")
View(Import)
```

```
plot(Export$Year, Export$X, type="l", lty=1, lwd=1, xlab="År",
     ylab="Export/Import", col="black", main="Import og Eksport", sub="Kilde: Statistikbanken")
lines(Import$Year, Import$Import, type="l", lty=1, lwd=1, col="red")
```



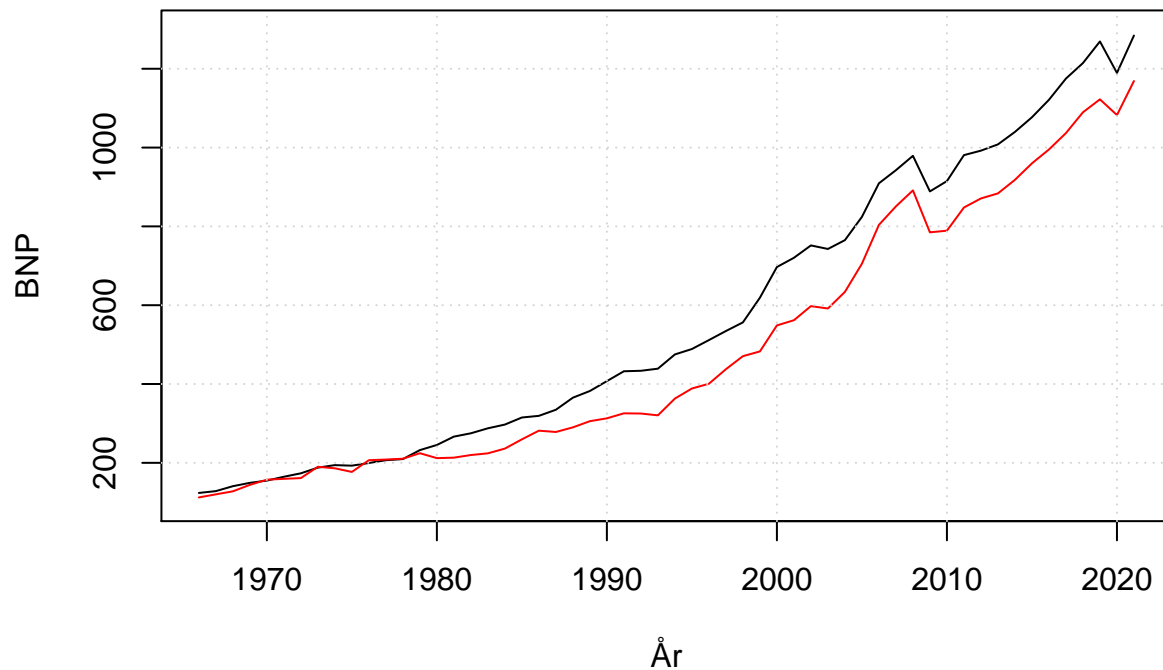


- Forklaring: - x angiver variabel 1 (Year) - z angiver variabel 3 (Import)

- Dimensioner på akserne bestemmes af variabelen angiver i plot(), men kan styres manuelt ved at tilføje kommandoen ylim=c(x1,y1) i plot()

```
plot(Export$Year, Export$X, type="l", lty=1, lwd=1, xlab="År",
     ylab="BNP", col="black", main="Import og Eksport", sub="Kilde: Statistikbanken", ylim=c(100, 1200),
     lines(Import$Year, Import$Import, type="l", lty=1, lwd=1, col="red")); grid()
```

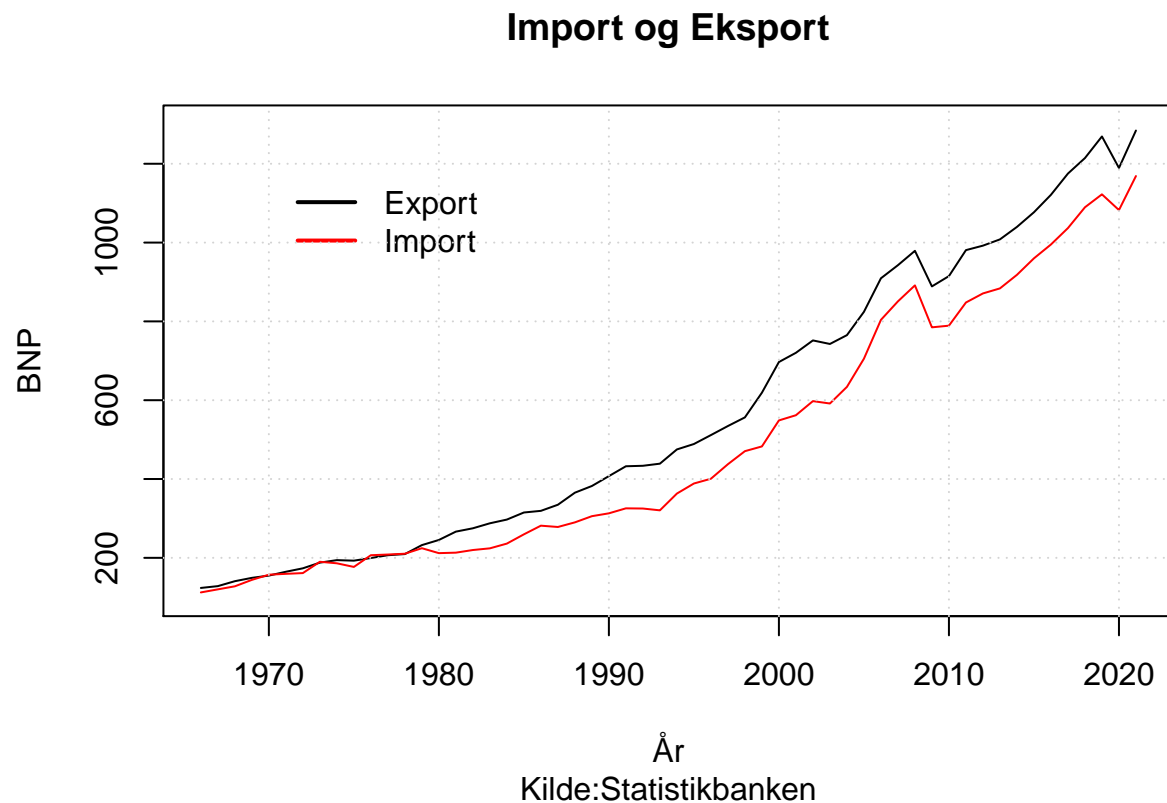
## Import og Eksport



Kilde: Statistikbanken

- x1 angiver her den nedre værdi på akse, mens y1 angiver øvre værdi på akse - Når man har to linjer bør man tilføje en legend, for at øge læsevenligheden af figuren: - Hint: tilføj kommandoen

```
plot(Export$Year, Export$X, type="l", lty=1, lwd=1, xlab="År",
     ylab="BNP", col="black", main="Import og Eksport", sub="Kilde: Statistikbanken", ylim=c(100, 1200),
     lines(Import$Year, Import$Import, type="l", lty=1, lwd=1, col="red"));
legend(x=1970, y= 1200, legend=c("Export", "Import"), lty=1, lwd=2,
      col=c("Black", "Red"), bty="n"); grid()
```



- Forklaring: - x = placering på x-aksen - y = placering på y-aksen - var1 = navn på variabel 1 - var2 = navn på variable 2 - lty = linje-type - lwd = linje-tykkelse - cvar1 = farve på linje for variable 1 - cvar2 = farve på linje for variable 2 - bty="n" - ønskes ramme omkring legend, slettes dette led.

## 4 Øvelse

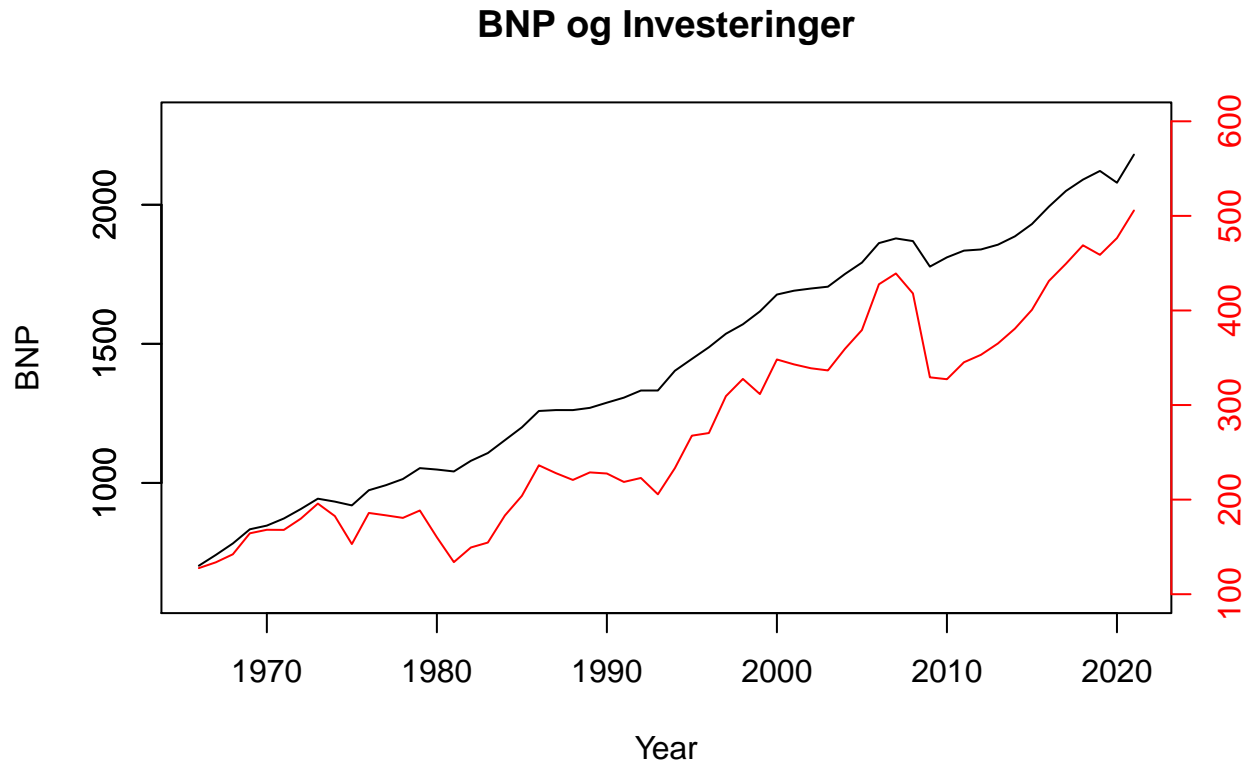
- I skal nu illustrere udviklingen i såvel investeringer som BNP i samme figur
  - Hint: Størrelsesforholdet mellem de to variable taler for at plotte et diagram med to y-akser
- Step 1: Importer data

```
Inv <- read_excel("Inv.xlsx")
```

```
plot(BNP$Year, BNP$BNP, ylim=c(600, 2300), xlab="Year", ylab="BNP",
     type="l", col="black", main="BNP og Investeringer"); axis(2, ylim=c(600, 2300), col="black")

par(new=TRUE)
```

```
plot(Inv$Year,Inv$I, ylab= "", xlab = "", ylim=c(100,600),
     axes=FALSE, type="l", col="red");axis(4, ylim=c(100,600), col="red",col.axis="red")
```



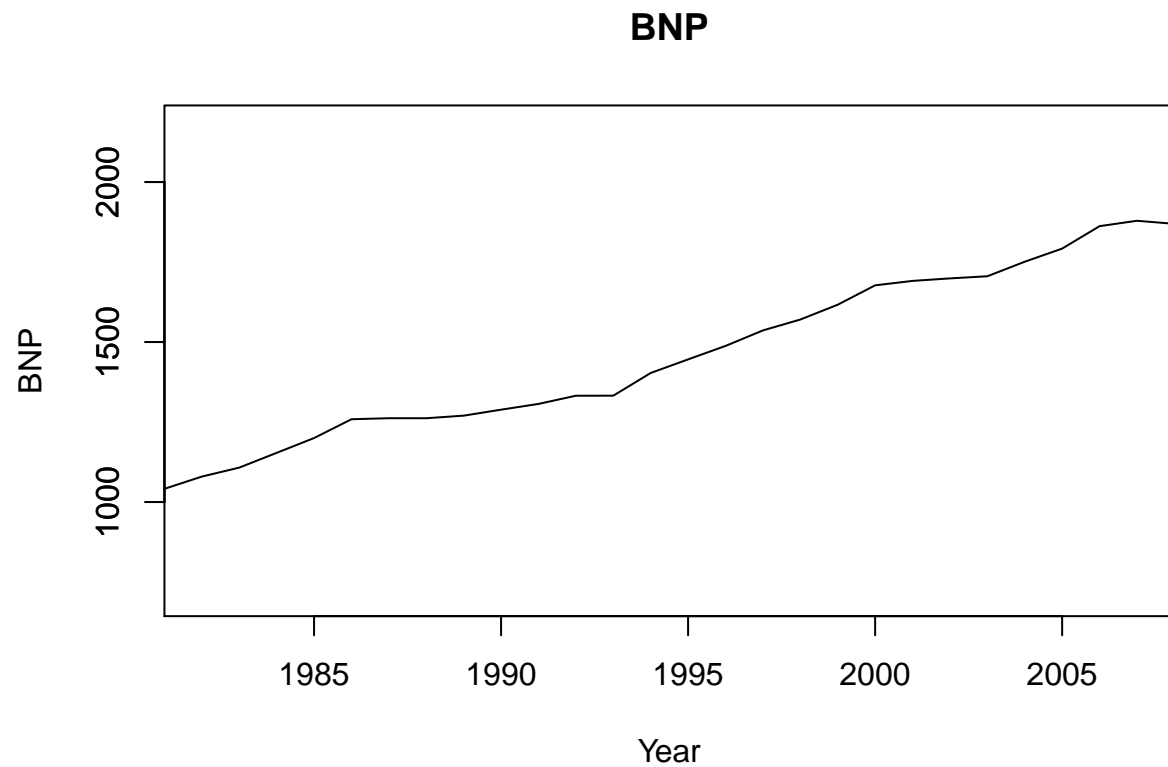
Forklaring: - x = variabel 1 - y = variabel 2 - x1 = nedre værdi y-akse 1 - y1 = øvre værdi y-akse 1 - x = variabel 1 - z = variabel 3 - x2 = nedre værdi y-akse 2 - y2 = øvre værdi y-akse 2

## 5 Øvelse

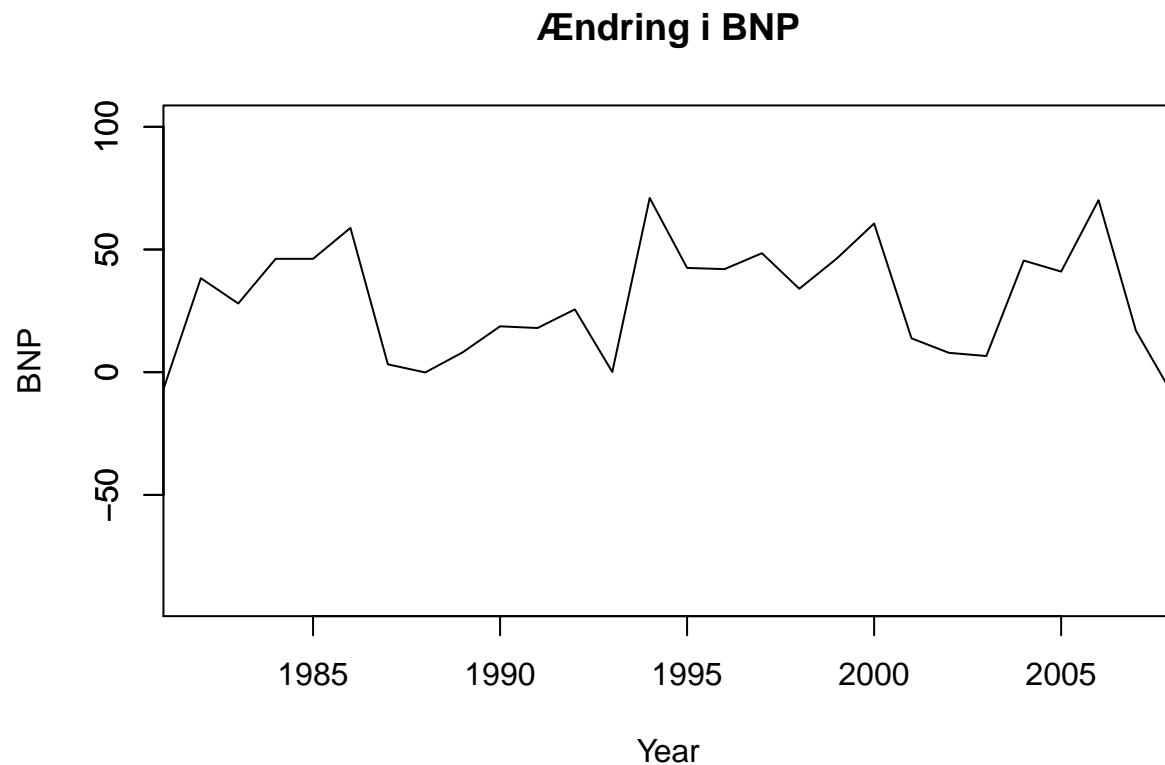
Undersøg grafisk nedenstående to påstande om den danske økonomi:

1. Realvæksten i den årlige BNP var ikke negativ i perioden 1982-2007

```
plot(BNP$Year,BNP$BNP, xlim = c(1982,2007), xlab="Year", ylab="BNP",
     type="l",col="black", main="BNP");axis(2, ylim=c(600,2300),col="black")
```



```
plot(BNP$Year[-1],diff(BNP$BNP), xlim = c(1982,2007), xlab="Year", ylab="BNP",  
type="l",col="black", main="Ændring i BNP");axis(2, ylim=c(600,2300),col="black")
```



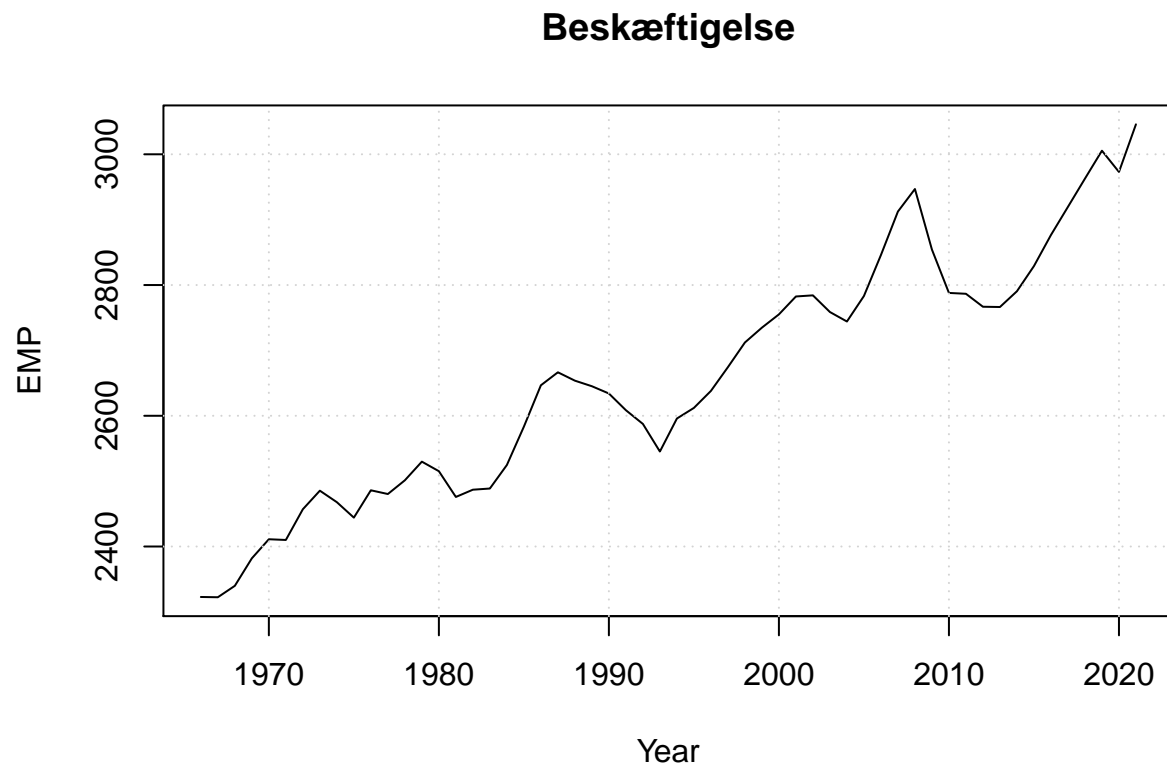
2. Trods effekterne fra Corona, var beskæftigelsen højere i 2020 end i 2018

Indlæser først data for employment:

```
Employment <- read_excel("Employment.xlsx")
```

Vi kan måske se det med et plot?

```
plot(Employment$Year, Employment$Emp, xlab="Year", ylab="EMP",
     type="l", col="black", main="Beskæftigelse"); grid()
```



Hvis det er svært at se kan vi også bare kigge direkte i variabelen

```
View(Employment)
```

Vi kan se 2018 har observation 53 og 2020 observation 55 Dermed kan vi sammenligne disse

```
Emp_2018 = Employment$Emp[53]
Emp_2020 = Employment$Emp[55]
```

Vi kan faktisk direkte spørge R hvilken der er størst

```
Emp_2018 > Emp_2020
```

```
## [1] FALSE
```

```
Emp_2020 > Emp_2018
```

```
## [1] TRUE
```

Vi kan se at employment i 2020 var større end i 2018