### Exercise 2  
  
remove(list = ls())  
  
library(rJava)

## Warning: package 'rJava' was built under R version 3.4.4

library(iplots)

## Warning: package 'iplots' was built under R version 3.4.4

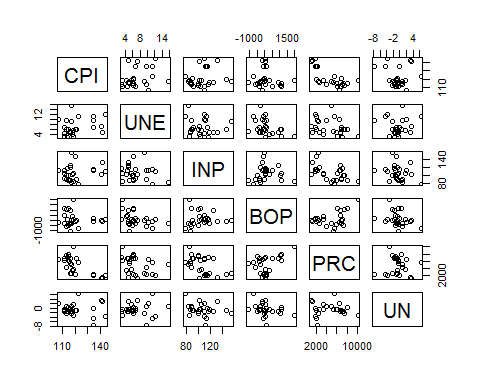
getwd()

## [1] "C:/Users/B-C-Herbert/Documents/Studium/Mannheim/VWL/2019 - 2020 WS/Applied Multivariate Statistics/Assignments/Assignment5/Assignment5"

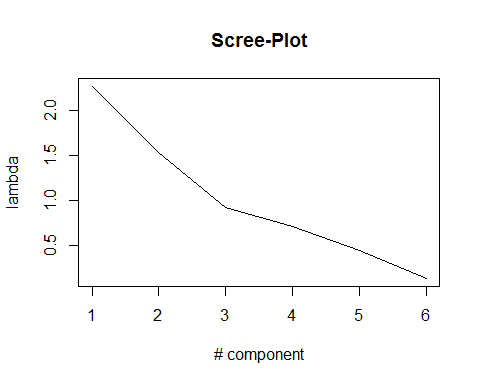
euro = read.table(file = "Europe.txt", header = T, dec =".")  
attach(euro)  
europe = data.frame(euro)  
europe = na.omit(europe)  
ls(europe)

## [1] "BOP" "CPI" "ï..Country" "INP" "PRC"   
## [6] "UN" "UNE"

Country = europe$ï..Country  
CPI = europe$CPI  
UNE = c(europe$UNE)  
INP = c(europe$INP)  
BOP = c(europe$BOP)  
PRC = c(europe$PRC)  
UN = c(europe$UN)  
  
numbers = cbind(CPI,UNE,INP,BOP,PRC,UN)  
S = cov(numbers)  
r = round(cor(numbers), digits = 4)  
  
# a)   
pairs(numbers)



reseigen = eigen(r)  
lambda = round(reseigen$values, digits = 4)  
plot(lambda, type = "l", main = "Scree-Plot", xlab = "# component")



# b)  
round(r, digits = 4)

## CPI UNE INP BOP PRC UN  
## CPI 1.0000 0.2890 0.2117 -0.1089 -0.7064 -0.0779  
## UNE 0.2890 1.0000 -0.1042 -0.4037 -0.3328 0.0255  
## INP 0.2117 -0.1042 1.0000 0.0400 -0.5113 -0.2383  
## BOP -0.1089 -0.4037 0.0400 1.0000 0.4172 -0.3129  
## PRC -0.7064 -0.3328 -0.5113 0.4172 1.0000 -0.0200  
## UN -0.0779 0.0255 -0.2383 -0.3129 -0.0200 1.0000

# c)  
E = round(reseigen$vectors, digits = 4)  
X = scale(numbers, center = TRUE, scale = TRUE)  
Yhat = (X)%\*%E  
y1hat = Yhat[,1]  
y2hat = Yhat[,2]  
ihist(y1hat)

## ID:1 Name: "Histogram (y1hat)"

ihist(y2hat)

## ID:2 Name: "Histogram (y2hat)"

ihist(CPI)

## ID:3 Name: "Histogram (CPI)"

ihist(UNE)

## ID:4 Name: "Histogram (UNE)"

ihist(INP)

## ID:5 Name: "Histogram (INP)"

ihist(BOP)

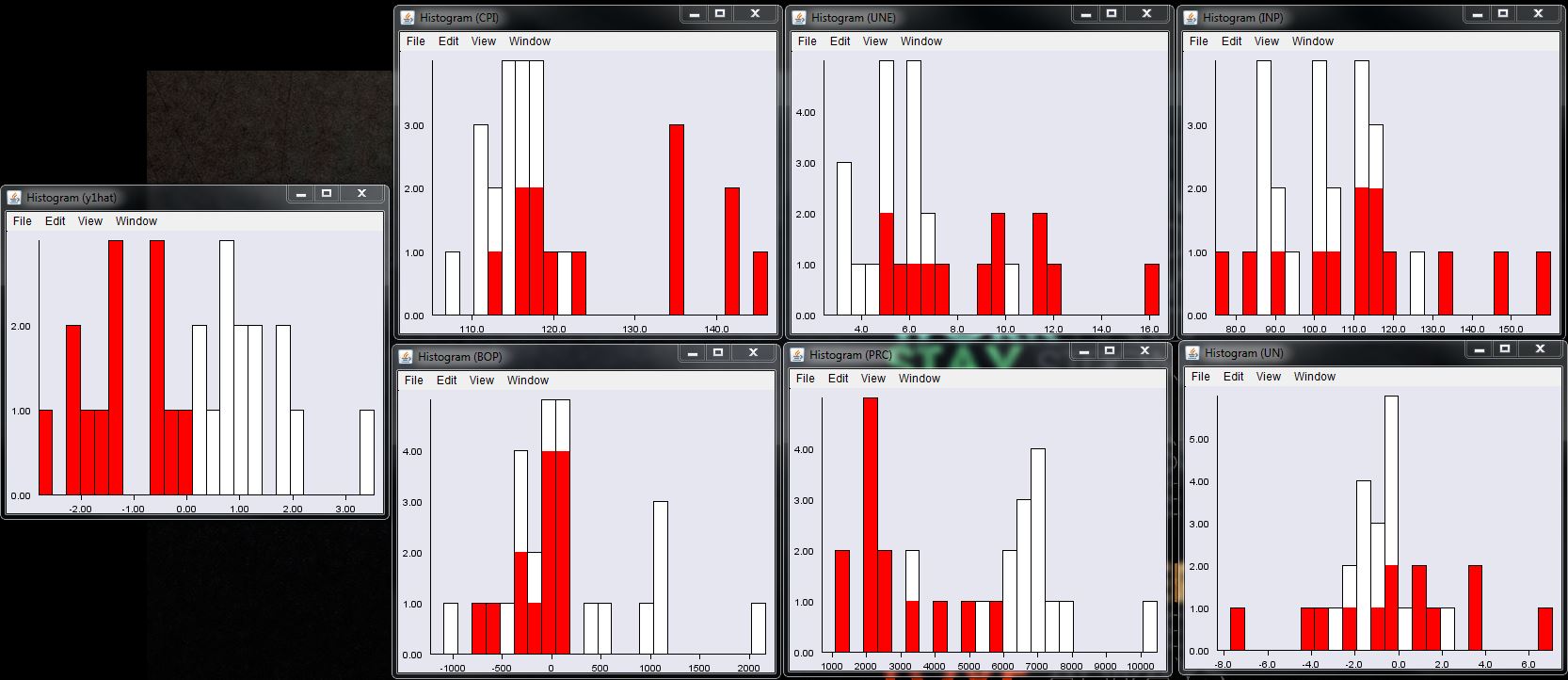
## ID:6 Name: "Histogram (BOP)"

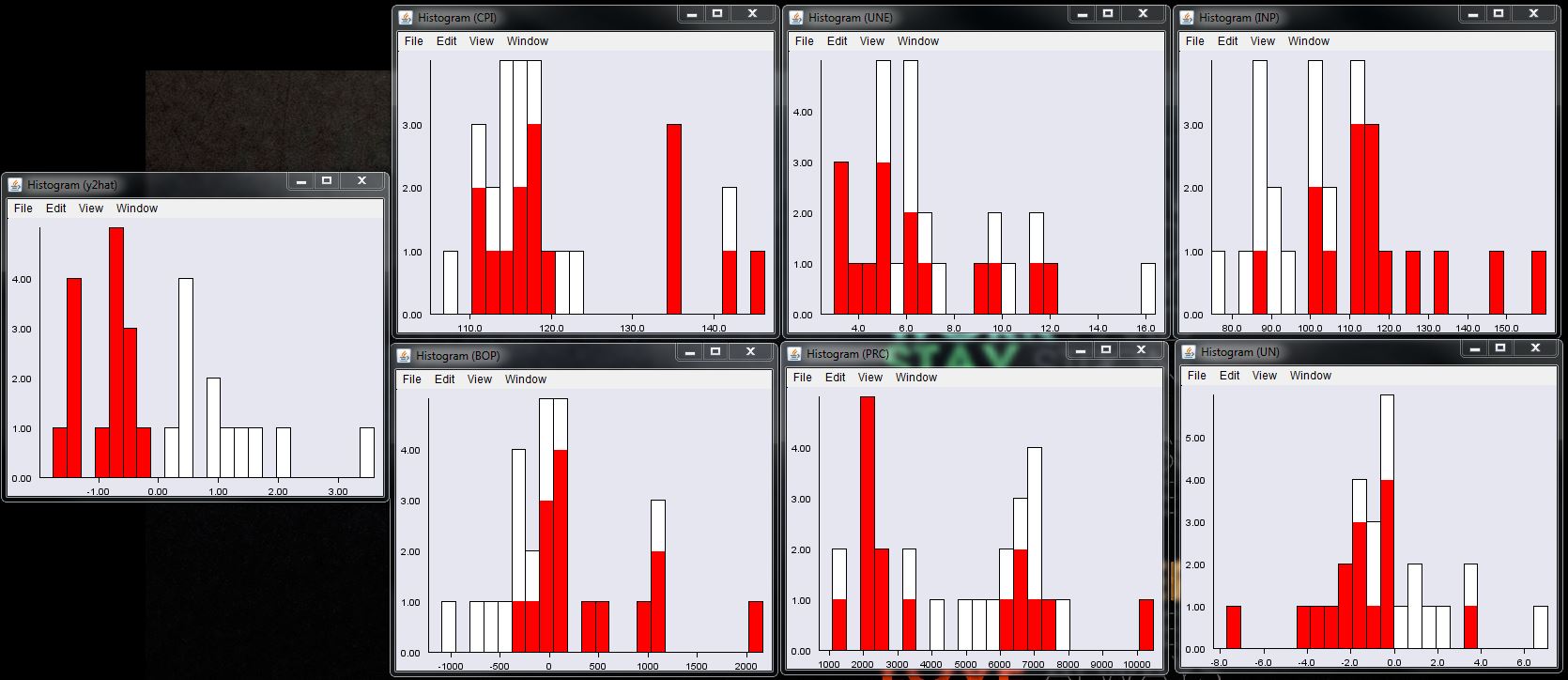
ihist(PRC)

## ID:7 Name: "Histogram (PRC)"

ihist(UN)

## ID:8 Name: "Histogram (UN)"





# d)  
E = round(reseigen$vectors, digits = 4)  
E

## [,1] [,2] [,3] [,4] [,5] [,6]  
## [1,] -0.5100 -0.1699 -0.1673 0.6444 0.1959 -0.4790  
## [2,] -0.3723 0.3358 -0.5979 -0.2128 -0.5867 -0.0402  
## [3,] -0.2900 -0.5340 0.4318 -0.4052 -0.3903 -0.3574  
## [4,] 0.3634 -0.4930 -0.1679 0.4827 -0.5358 0.2768  
## [5,] 0.6202 0.1202 -0.1755 -0.0575 -0.0737 -0.7492  
## [6,] -0.0212 0.5618 0.6075 0.3728 -0.4154 -0.0574

lambda

## [1] 2.2648 1.5356 0.9192 0.7077 0.4417 0.1309

round(cumsum(lambda)/sum(diag(r)),4)

## [1] 0.3775 0.6334 0.7866 0.9046 0.9782 1.0000

e1 = E[,1]  
e2 = E[,2]  
e3 = E[,3]  
e4 = E[,4]  
e5 = E[,5]  
e6 = E[,6]

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Var** | **PC1** | **PC2** | **PC3** | **PC4** | **PC5** | **PC6** |
| CPI | -0,5100 | -0,1699 | -0,1673 | 0,6444 | 0,1959 | -0,4790 |
| UNE | -0,3723 | 0,3358 | -0,5979 | -0,2128 | -0,5867 | -0,0402 |
| INP | -0,2900 | -0,5340 | 0,4318 | -0,4052 | -0,3903 | -0,3574 |
| BOP | 0,3634 | -0,4930 | -0,1679 | 0,4827 | -0,5358 | 0,2768 |
| PRC | 0,6202 | 0,1202 | -0,1755 | -0,0575 | -0,0737 | -0,7492 |
| UN | -0,0212 | 0,5618 | 0,6075 | 0,3728 | -0,4154 | -0,0574 |
|  | 2,2648 | 1,5356 | 0,9192 | 0,7077 | 0,4417 | 0,1309 |
| Cum. Prop. | 0,3775 | 0,6334 | 0,7866 | 0,9046 | 0,9782 | 1,0000 |

### e)  
lambda1 = lambda[1]  
lambda2 = lambda[2]  
rec1 = round(lambda1/sum(lambda), digits = 4)  
paste("The first principal component explains", rec1\*100, "% of the total variation. It is mostly driven by CPI and PRC. It is bigger for countriese with a low Consumer Price Index and high final consumption expenditure")

## [1] "The first principal component explains 37.75 % of the total variation. It is mostly driven by CPI and PRC. It is bigger for countriese with a low Consumer Price Index and high final consumption expenditure"

rec2 = round(lambda2/sum(lambda), digits = 4)  
paste("The second principal component explains", rec2\*100, "% of the total variation. It is mostly driven by INP, BOP and UN. It is bigger for Countries with a low industiral Production, low Balances of Payments and high changes in Unemployment")

## [1] "The second principal component explains 25.59 % of the total variation. It is mostly driven by INP, BOP and UN. It is bigger for Countries with a low industiral Production, low Balances of Payments and high changes in Unemployment"

### f)  
y1hatr = round(Yhat[,1], digits = 2)  
y2hatr = round(Yhat[,2], digits = 2)  
y12hat = data.frame(Country,y1hatr ,y2hatr, CPI,UNE,INP,BOP,PRC,UN)  
y12hat

## Country y1hatr y2hatr CPI UNE INP BOP PRC UN  
## 1 Belgium 1.11 -1.42 116.03 4.77 125.59 908.6 6716.5 -1.6  
## 2 Bulgaria -1.99 0.54 141.20 7.31 102.39 27.8 1094.7 3.5  
## 3 CzechRep -0.50 -0.35 116.20 4.88 119.01 -277.9 2616.4 -0.6  
## 4 Denmark 2.06 0.12 114.20 6.03 88.20 1156.4 7992.4 0.5  
## 5 Germany 1.35 -0.60 111.60 4.63 111.30 499.4 6774.6 -1.3  
## 6 Estonia -1.68 -1.76 135.08 9.71 111.50 153.4 2194.1 -7.7  
## 7 Ireland 0.48 1.25 106.80 10.20 111.20 -166.5 6525.1 2.0  
## 8 Greece -0.51 3.36 122.83 11.30 78.22 -764.1 5620.1 6.4  
## 9 Spain -0.70 2.20 116.97 15.79 83.44 -280.8 4955.8 0.7  
## 10 France 0.94 0.86 111.55 6.77 92.60 -337.1 6828.5 -0.9  
## 11 Italy 0.82 0.82 115.00 5.05 87.80 -366.2 5996.6 -0.5  
## 12 Cyprus 0.19 1.36 116.44 5.14 86.91 -1090.6 5310.3 -0.4  
## 13 Latvia -2.55 -0.69 144.47 12.11 110.39 42.3 1968.3 -3.6  
## 14 Lithuania -2.10 -0.77 135.08 11.47 114.50 -77.4 2130.6 -4.3  
## 15 Luxembourg 3.27 -1.44 118.19 3.14 85.51 2016.5 10051.6 -3.0  
## 16 Hungary -1.49 -0.58 134.66 6.77 115.10 156.2 1954.8 -0.1  
## 17 Malta 0.32 -0.39 117.65 4.15 101.65 359.4 3378.3 -0.6  
## 18 Netherlands 1.82 -0.86 111.17 3.23 103.80 1156.6 6046.0 -0.4  
## 19 Austria 1.19 -0.69 114.10 2.99 116.80 87.8 7045.5 -1.5  
## 20 Poland -1.28 -1.30 119.90 6.28 146.70 -74.8 2124.2 -1.0  
## 21 Portugal -0.28 1.67 113.06 9.68 89.30 -613.4 4073.6 0.8  
## 22 Romania -2.23 -0.52 142.34 4.76 131.80 -128.7 1302.2 3.2  
## 23 Slovenia -0.09 0.39 118.33 5.56 105.40 39.4 3528.3 1.8  
## 24 Slovakia -1.48 -1.49 117.17 9.19 156.30 16.0 2515.3 -2.1  
## 25 Finland 0.78 0.54 114.60 5.92 101.00 -503.7 7198.8 -1.3  
## 26 Sweden 1.79 -0.75 112.71 6.10 100.50 1079.1 7476.7 -2.3  
## 27 UnitedKingdom 0.78 0.50 120.90 6.11 90.36 -24.3 6843.9 -0.8

### i)   
paste("Denmark, Ireland, France, Italy, Cyprus, Finland, UK")

## [1] "Denmark, Ireland, France, Italy, Cyprus, Finland, UK"

paste("Diese Länder haben die Eigenschaften, dass sie eine geringe Inflation haben und einen hohen Konsum. Außerdem haben sie ein geringe industrielle Produktion und eine geringe Balance of Payments. Weiter eint sie eine starke Änderungen in der Arbeitslosigkeit.")

## [1] "Diese Länder haben die Eigenschaften, dass sie eine geringe Inflation haben und einen hohen Konsum. Außerdem haben sie ein geringe industrielle Produktion und eine geringe Balance of Payments. Weiter eint sie eine starke Änderungen in der Arbeitslosigkeit."

### ii)  
paste("Bulgaria, Greece, Spain, Portugal, Slovenia, Portugal")

## [1] "Bulgaria, Greece, Spain, Portugal, Slovenia, Portugal"

paste("Diese Länder haben eine hohe Inflation und geringen privaten Konsum. Die industrielle Produktion ist gering, wie auch die Balance of Payments. Außerdem ist die Arbeitslosenquote stark schwankend. Dies ist die Gruppe der Südeuropäischen Länder, welche die Finanzkrise stark getroffen hat.")

## [1] "Diese Länder haben eine hohe Inflation und geringen privaten Konsum. Die industrielle Produktion ist gering, wie auch die Balance of Payments. Außerdem ist die Arbeitslosenquote stark schwankend. Dies ist die Gruppe der Südeuropäischen Länder, welche die Finanzkrise stark getroffen hat."

### iii)  
paste("Czech Republik, Estonia, Latvia, Lithuania, Hungary, Poland, Romania, Slovakia")

## [1] "Czech Republik, Estonia, Latvia, Lithuania, Hungary, Poland, Romania, Slovakia"

paste("Hohe Inflation und niedriger privater Konsum. Allerdings sind industrielle Produktion und die Balance of Payments hoch. Die Schwankungen in der Arbeitslosigkeit sind gering. Dies ist die Gruppe der osteuropäischen Länder.")

## [1] "Hohe Inflation und niedriger privater Konsum. Allerdings sind industrielle Produktion und die Balance of Payments hoch. Die Schwankungen in der Arbeitslosigkeit sind gering. Dies ist die Gruppe der osteuropäischen Länder."

### iv)  
paste("Belgium, Germany, Luxembourg, Malta, Netherlands, Austria, Malta, Sweden")

## [1] "Belgium, Germany, Luxembourg, Malta, Netherlands, Austria, Malta, Sweden"

paste("Die Länder eint eine geringe Inflation und hoher privater Konsum, eine ausgeprägte industrielle Produktion, eine hohe Balance of Payments und geringe Schwankungen in der Arbeitslosigkeit. Es ist die Gruppe der nord- und mitteleuropäischen Länder die Volkswirtschaftlich sehr gut dastehen.")

## [1] "Die Länder eint eine geringe Inflation und hoher privater Konsum, eine ausgeprägte industrielle Produktion, eine hohe Balance of Payments und geringe Schwankungen in der Arbeitslosigkeit. Es ist die Gruppe der nord- und mitteleuropäischen Länder die Volkswirtschaftlich sehr gut dastehen."