projekt

Lumba i Strujići

May 2, 2017

Mjerenje uspješnosti investicijskih fondova

Motivacija i opis problema:

Cilj ovog projektnog zadatka je izmjeriti i usporediti uspješnost odabranih investicijskih fondova u koje hrvatski ulagači mogu ulagati. Performanse investicijskih fondova očituju se kroz promjenu vrijednosti udjela (imovine), te se najčešće izražavaju u dnevnim, tjednim ili mjesečnim povratima, definiranim kao postotna promjena vrijednosti udjela. Dok je rast vrijednosti fonda poželjna stvar, raspršenost povrata (često izražena kao standardna devijacija ili varijanca) se smatra negativnom i predstavlja rizik fonda.

Opis skupa podataka:

Podatci uključuju dnevne prinose nekoliko odabranih mirovinskih i dioničkih fondova dostupnih hrvatskim ulagačima od 2010. do 2017., dnevne prinose referentnog tržišnog indeksa CROBEX, te bezrizične stope izražene preko kamatnih stopa tromjesečnih trezorskih zapisa RH.

Dionički fondovi: ERSTE Adriatic Equity, OTP meridian 20 i ZB Aktiv Mirovinski fondovi: Reiffeisen DMF, ERSTE Plavi EXPERT i ERSTE Plavi PROTECT

1. zadatak: Prije istraživanja potrebno je urediti i poravnati podatke, te napraviti inicijalnu eksplorativnu analizu podataka i usporediti statistička svojstva različitih fondova.

Rješenje problema: Danom podatkovnom skupu podataka u početku smo pristupili osnovnim metodama za pojednostavljenje i prilagodbu poznatom statističkom i matematičkom aparatu. Prvotni problem na koji smo naišli bio je velik broj neradnih dana u kojima se ne trguje pa su pripadni cjenovni indeksi bili prepisani od posljednjeg radnog dana. Shodno tome, pripadni logaritamski dnevni prinosi bili su jednaki nuli što je prouzročilo izraženo odstupanje od normalne razdiobe u nuli. Eliminacijom tog problema logaritamski dnevni prinosi značajno su se približili normalnoj distribuciji što demonstriramo pripadnim Q-Q plotovima. Usprkos tome, zamijećujemo prisutnost izuzetno 'teških repova' na što nemamo utjecaj. Pretpostavka je da zbog relativne robusnosti kasnije provedenih T-testova ovu nesavršenost možemo zanemariti.

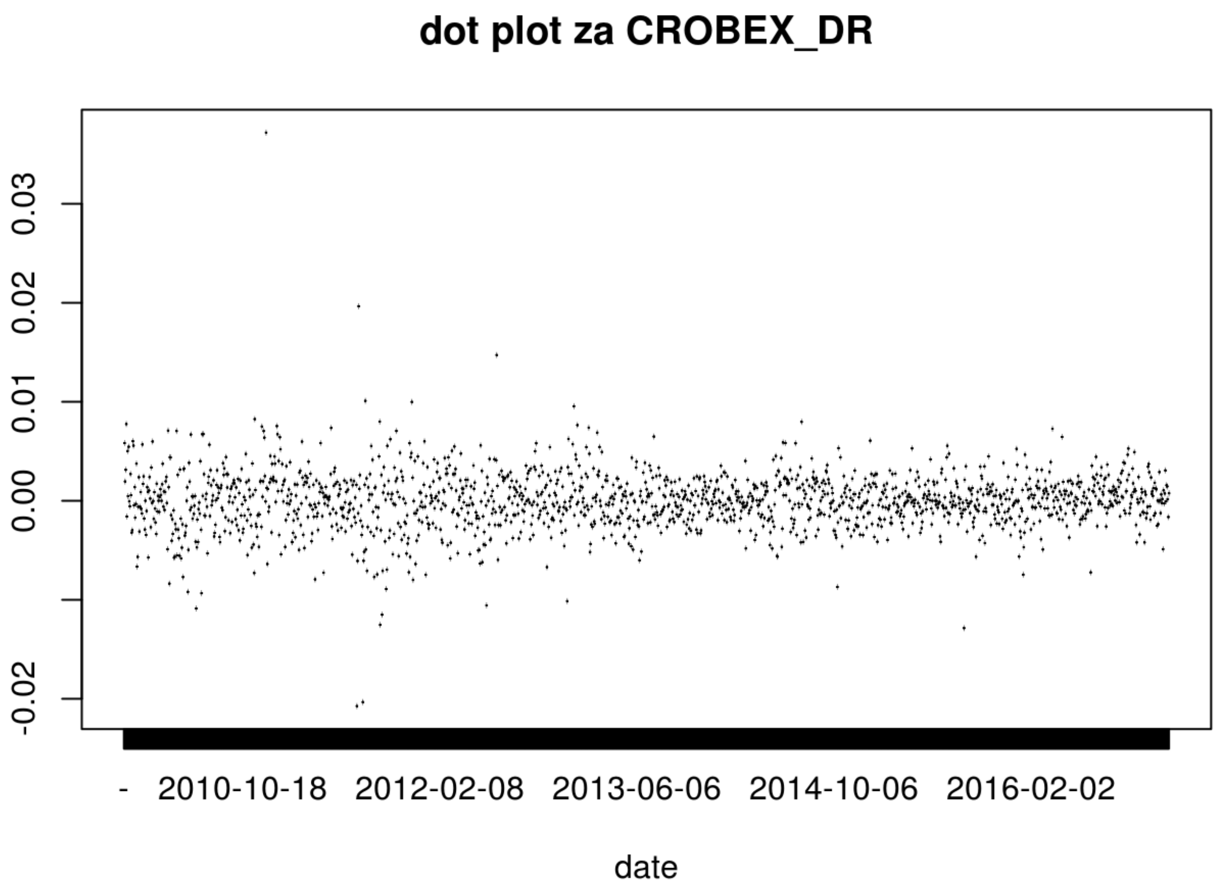
```r
#učitavanje dataseta

library(readxl)
data <- read_excel("investicijski_fondovi_data.xlsx")

#učitavanje nadopunjenog dataseta
data_log <-read_excel("investicijski_fondovi_data_log.xlsx")
```

Uređivanje i poravnavanje podataka i eksplorativna analiza podataka:

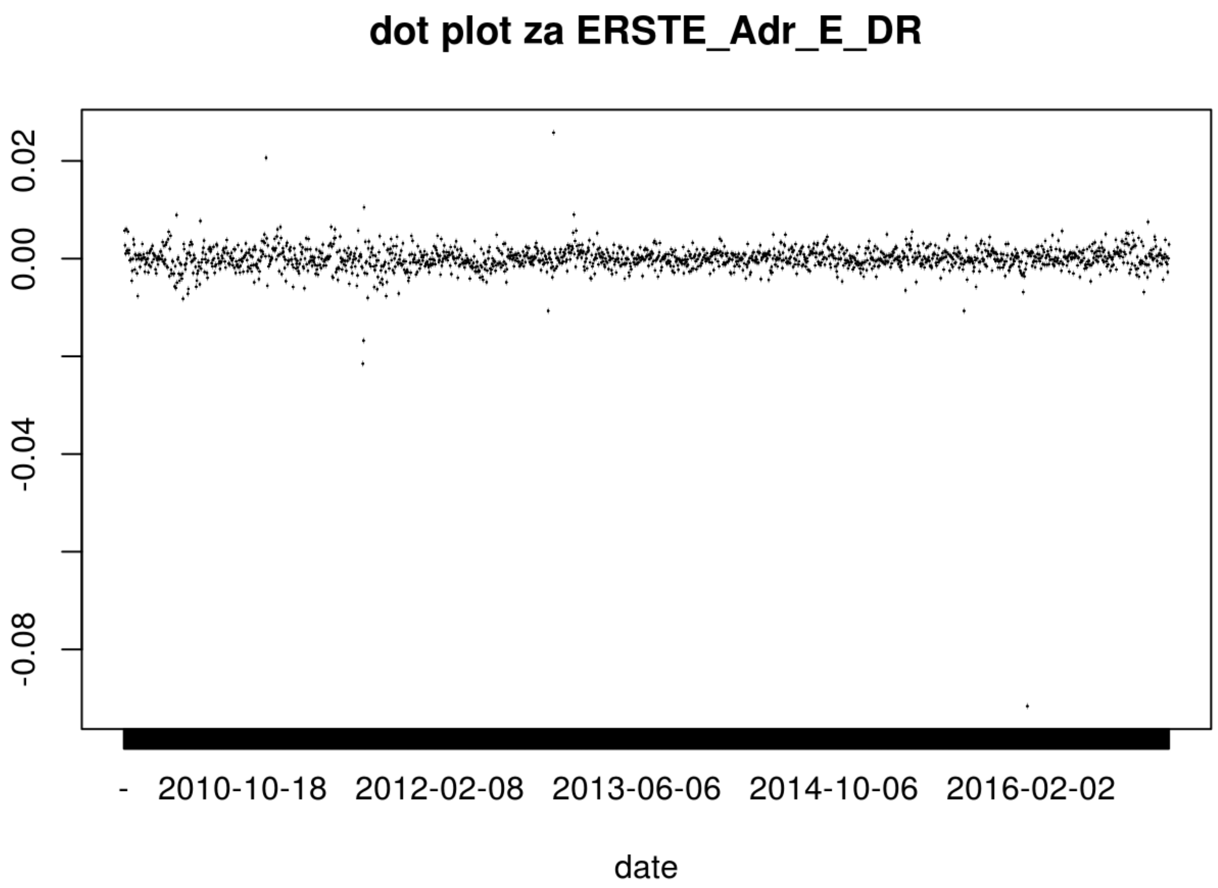
```r
#izbacivanje neradnih dana

library(stringr)

data_log$Date <- as.POSIXlt(data_log$Date, format = "%Y-%m-%d")

cond <- data_log$Date$wday == 0 | data_log$Date$wday == 6

data_log <- data_log[!cond, ]

cond2 <- data_log$CROBEX_DR == 0 & data_log$ERSTE_Adr_E_DR == 0 & data_log$OTP_DR
==0 & data_log$ZB_DR == 0

data_log <- data_log[!cond2, ]


data_log <- cbind(str_split_fixed(data_log$Date, "-", 2), data_log)
data_log$Date <- NULL
names(data_log)[1]<-paste("Y")
names(data_log)[2]<-paste("M")

neradni <- read.table("neradni_dani.txt")

data_log$M <- as.character(data_log$M)
neradni$V1 <- as.character(neradni$V1)

cond3 <- data_log$M %in% neradni$V1

data_log <- data_log[!cond3,]

data_log <- cbind(paste(data_log$Y, data_log$M, sep = "-"), data_log)
names(data_log)[1]<- paste("Date")

data_log$Y <- NULL
data_log$M <- NULL
data_log <- data_log[-1,]
```

Dot-plotovi logaritamskih dnevnih prinosa svih fondova. Načelno, zamijećujemo veću raspršenost dioničkih fondova.

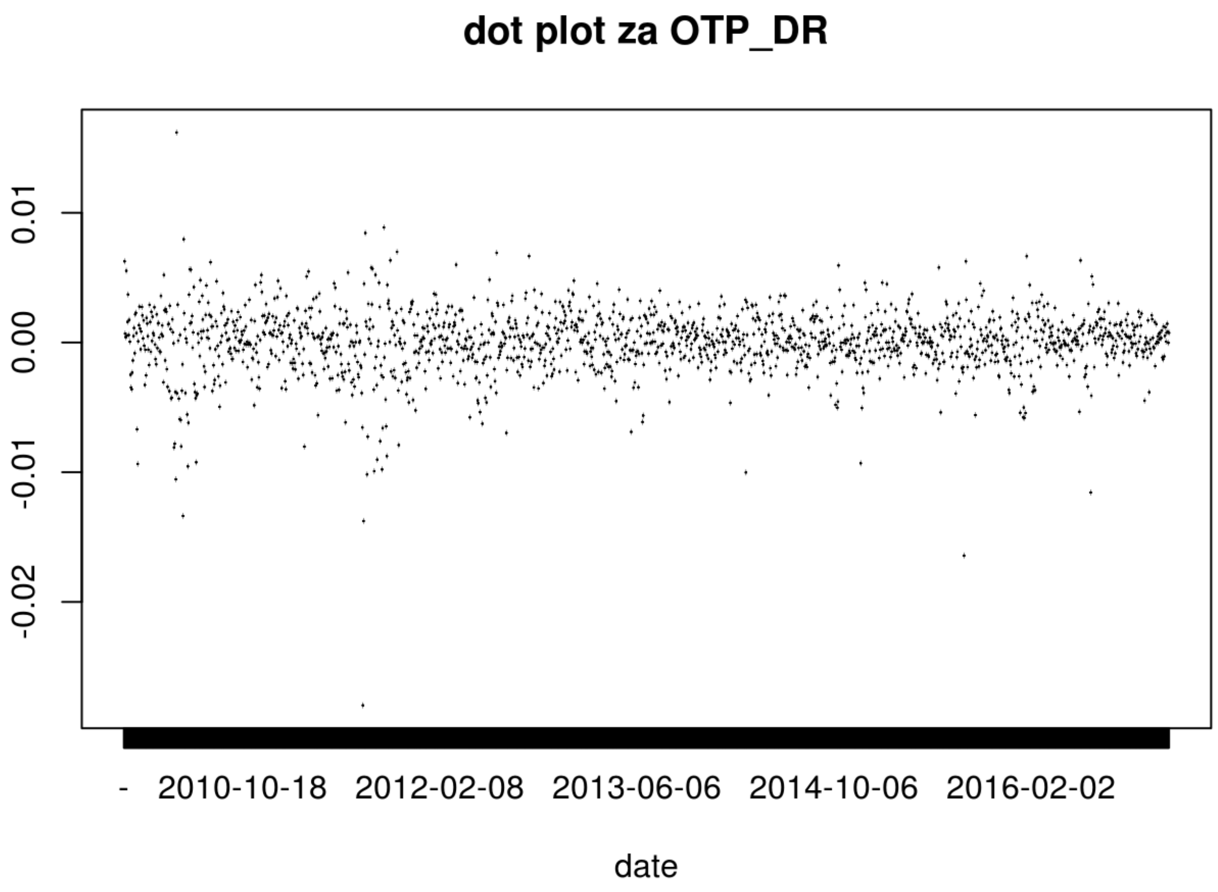
```r
#dotplot za dnevne logaritamske prinose

plot(data_log$Date, data_log$CROBEX_DR, main="dot plot za CROBEX_DR",
   xlab="date" )
```



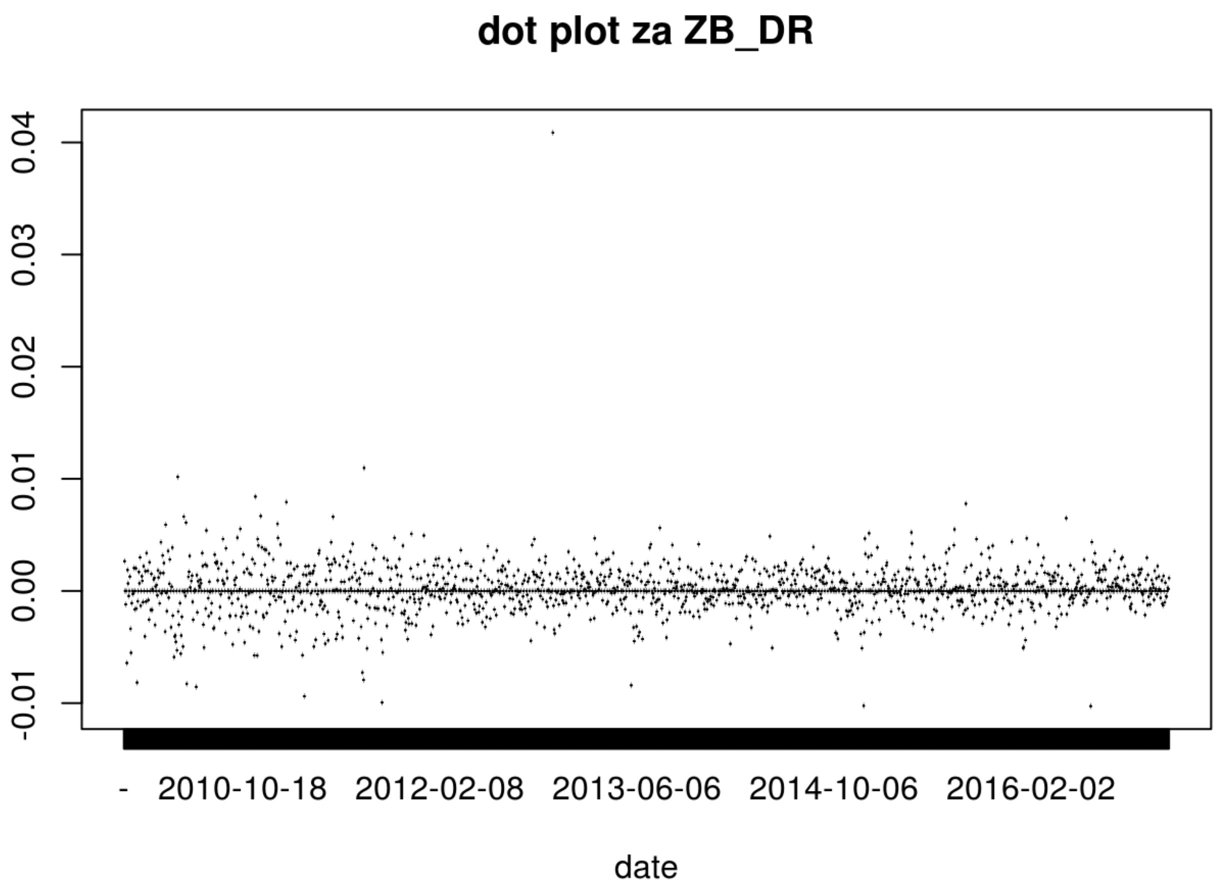

```r
#dionički fondovi

plot(data_log$Date, data_log$ERSTE_Adr_E_DR, main="dot plot za ERSTE_Adr_E_DR",
   xlab="date" )
```



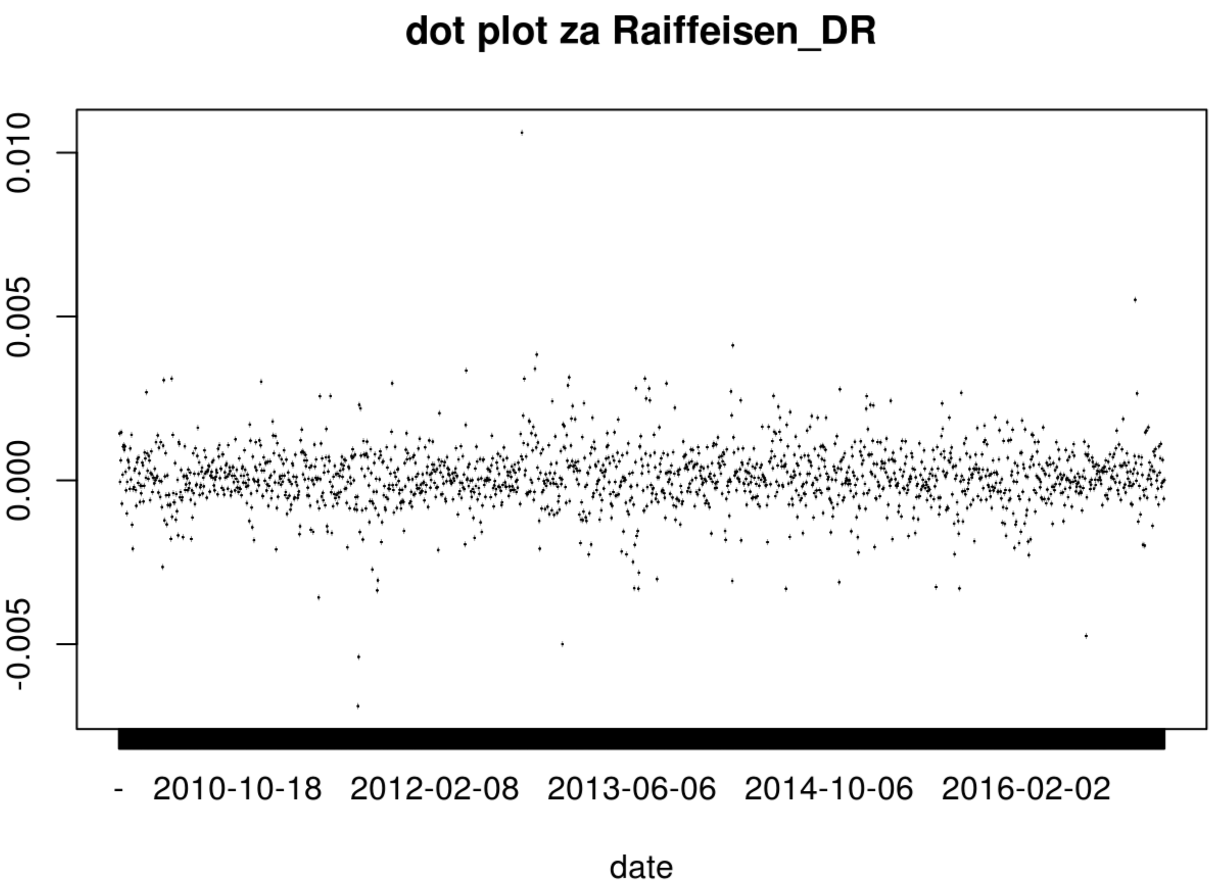

```
plot(data_log$Date, data_log$OTP_DR, main="dot plot za OTP_DR",
   xlab="date" )
```



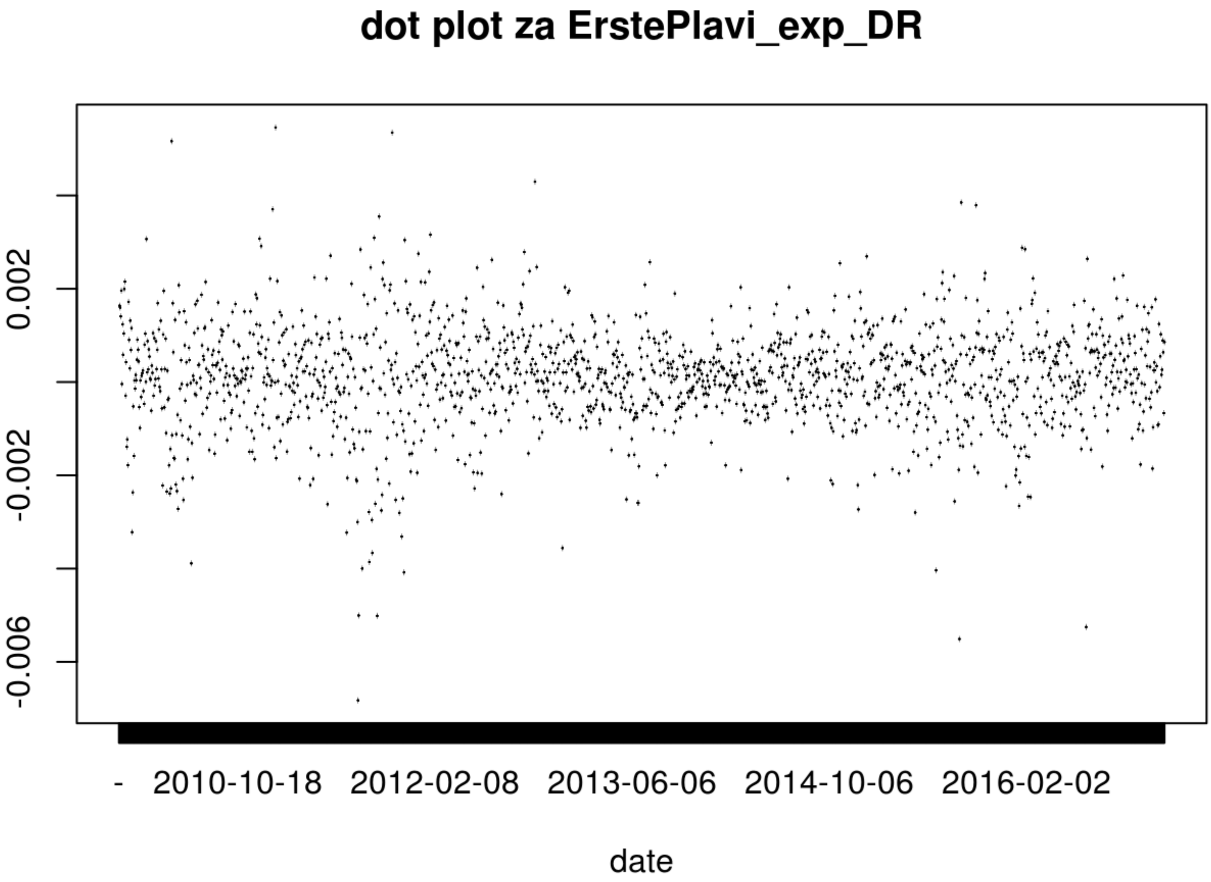

```
plot(data_log$Date, data_log$ZB_DR, main="dot plot za ZB_DR",
   xlab="date" )
```



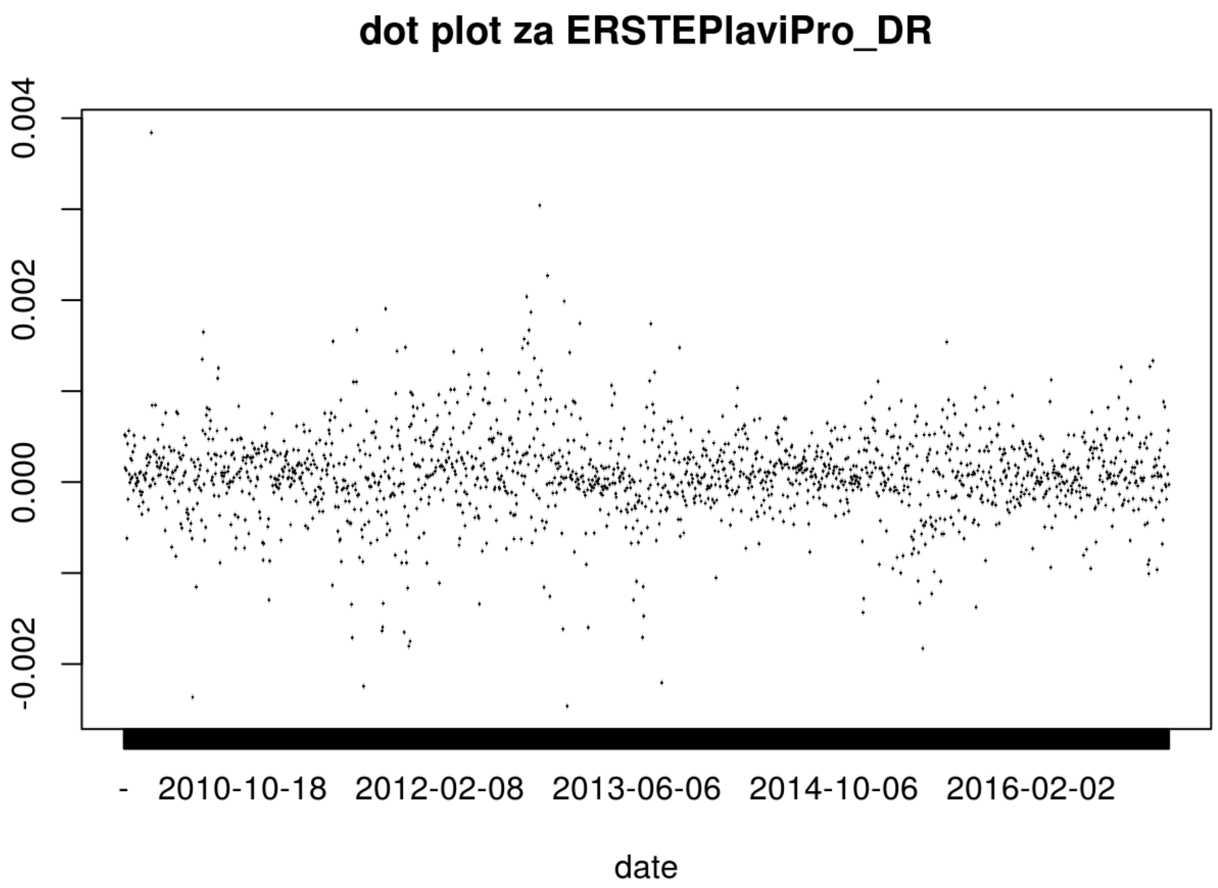

```
#mirovinski fondovi

plot(data_log$Date, data_log$Raiffeisen_DR, main="dot plot za Raiffeisen_DR",
   xlab="date" )
```



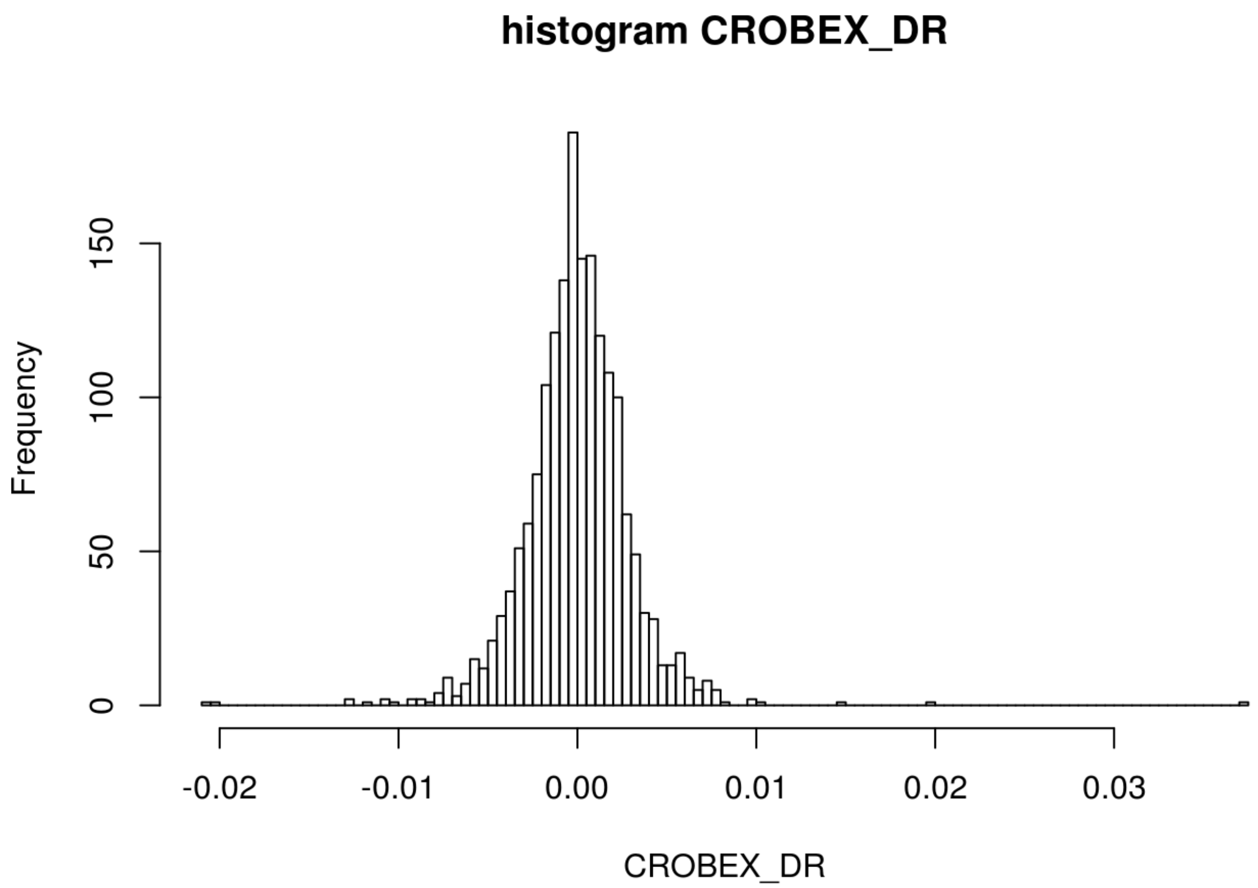

```
plot(data_log$Date, data_log$ErstePlavi_exp_DR, main="dot plot za ErstePlavi_exp_DR
",
   xlab="date" )
```



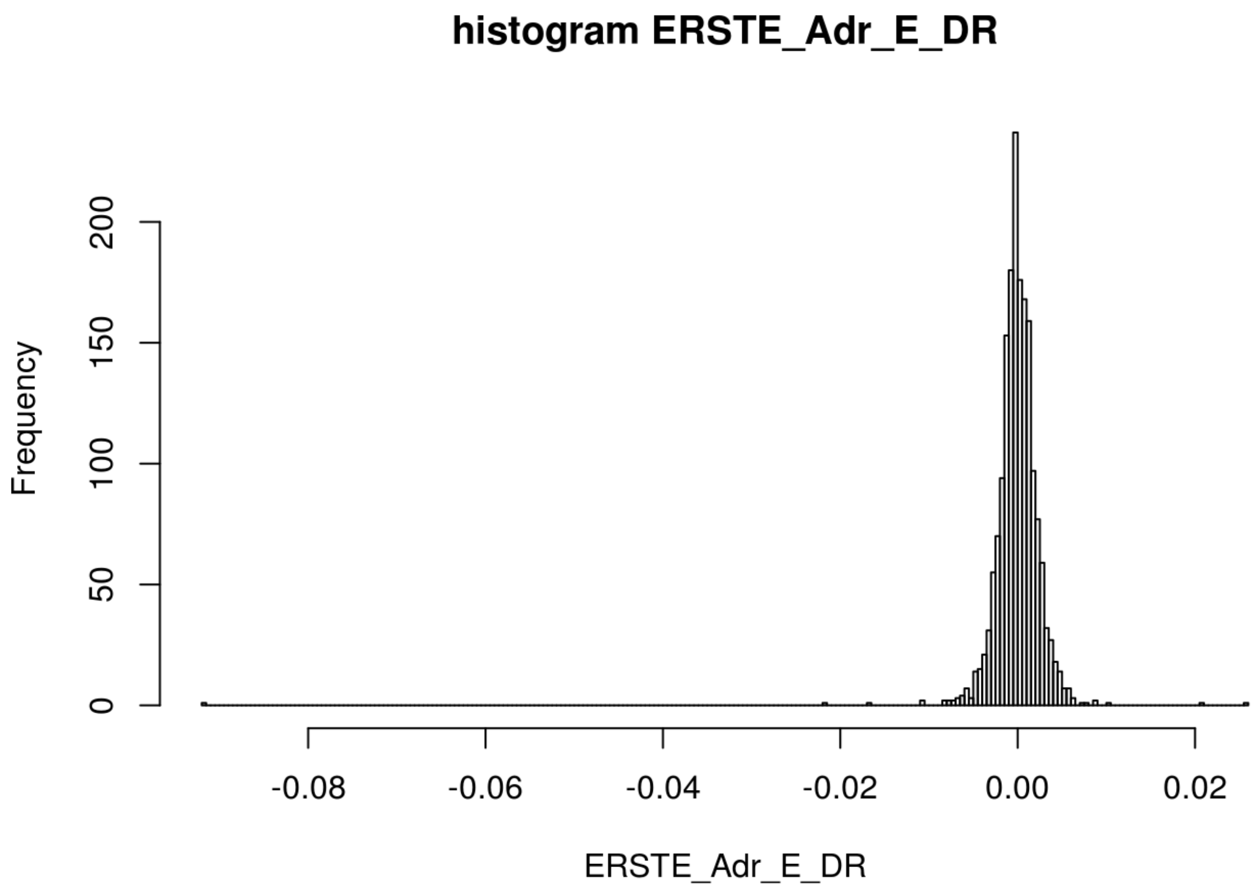

```
plot(data_log$Date, data_log$ERSTEPlaviPro_DR, main="dot plot za ERSTEPlaviPro_DR",
   xlab="date" )
```



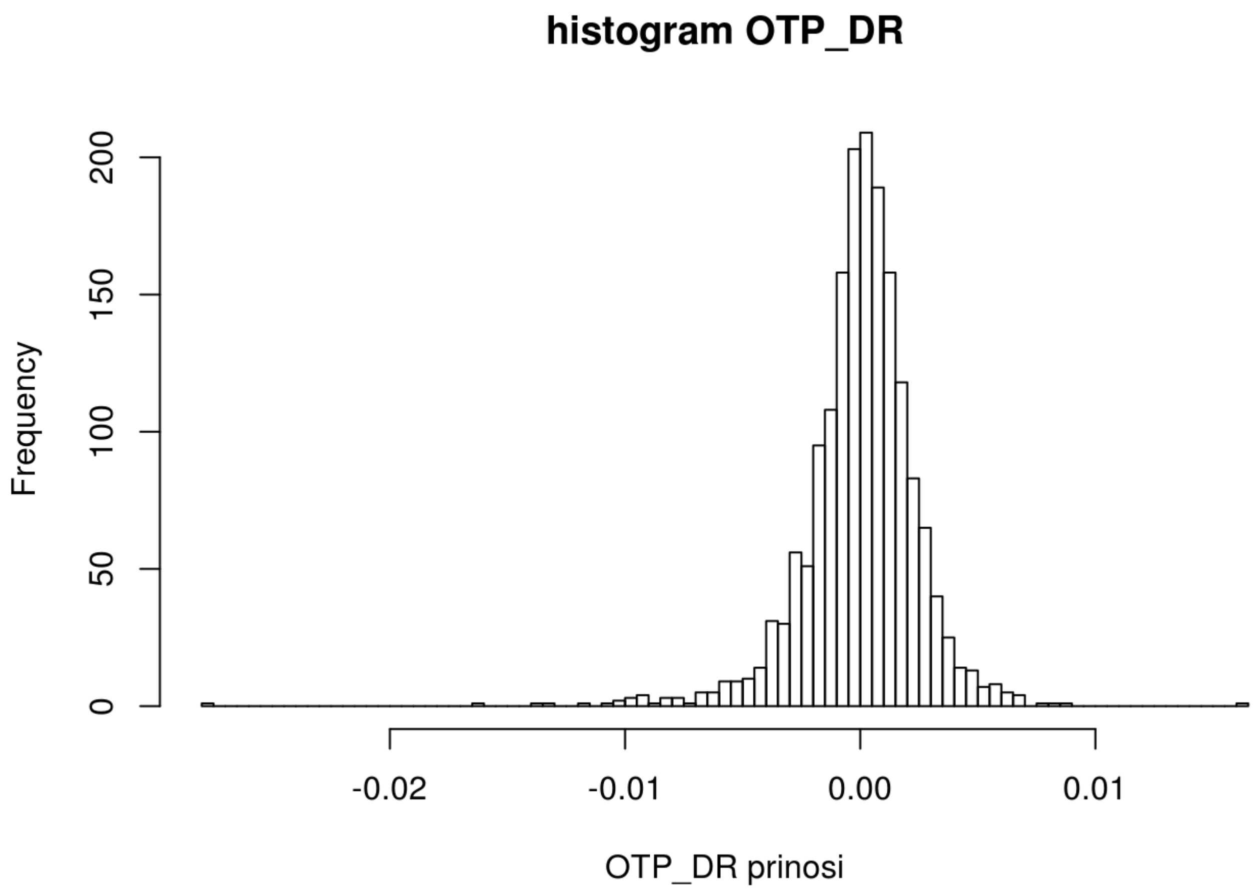

```r
#histogrami dnevnih prinosa i logaritamskih dnevnih prinosa

data_log$Date <- as.POSIXlt(data_log$Date,format="%Y-%m-%d")

hist(data_log$CROBEX_DR, main="histogram CROBEX_DR",breaks = 100,
xlab = "CROBEX_DR", ylab = "Frequency" )
```



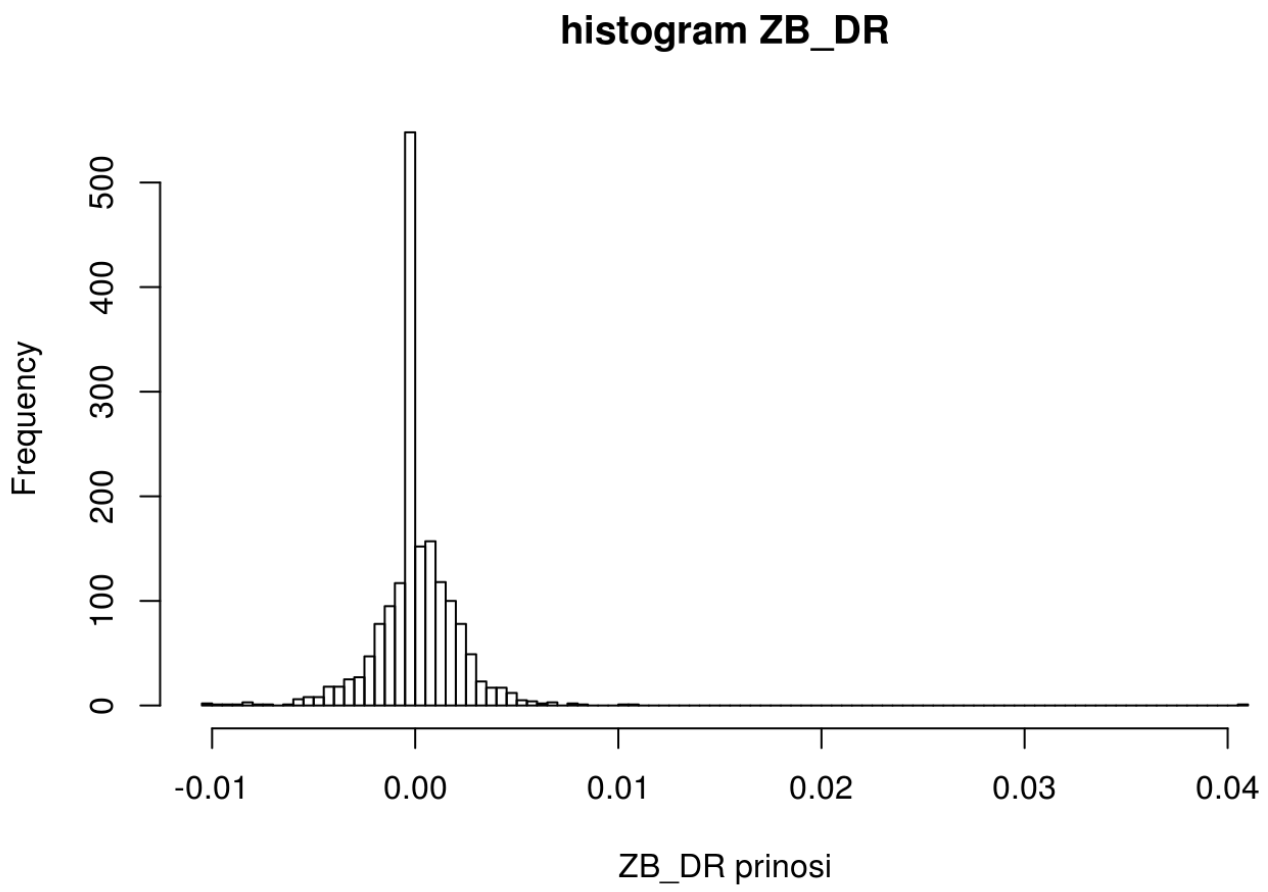

```
hist(data_log$ERSTE_Adr_E_DR, main="histogram ERSTE_Adr_E_DR",breaks = 200,
xlab = "ERSTE_Adr_E_DR", ylab = "Frequency" )
```



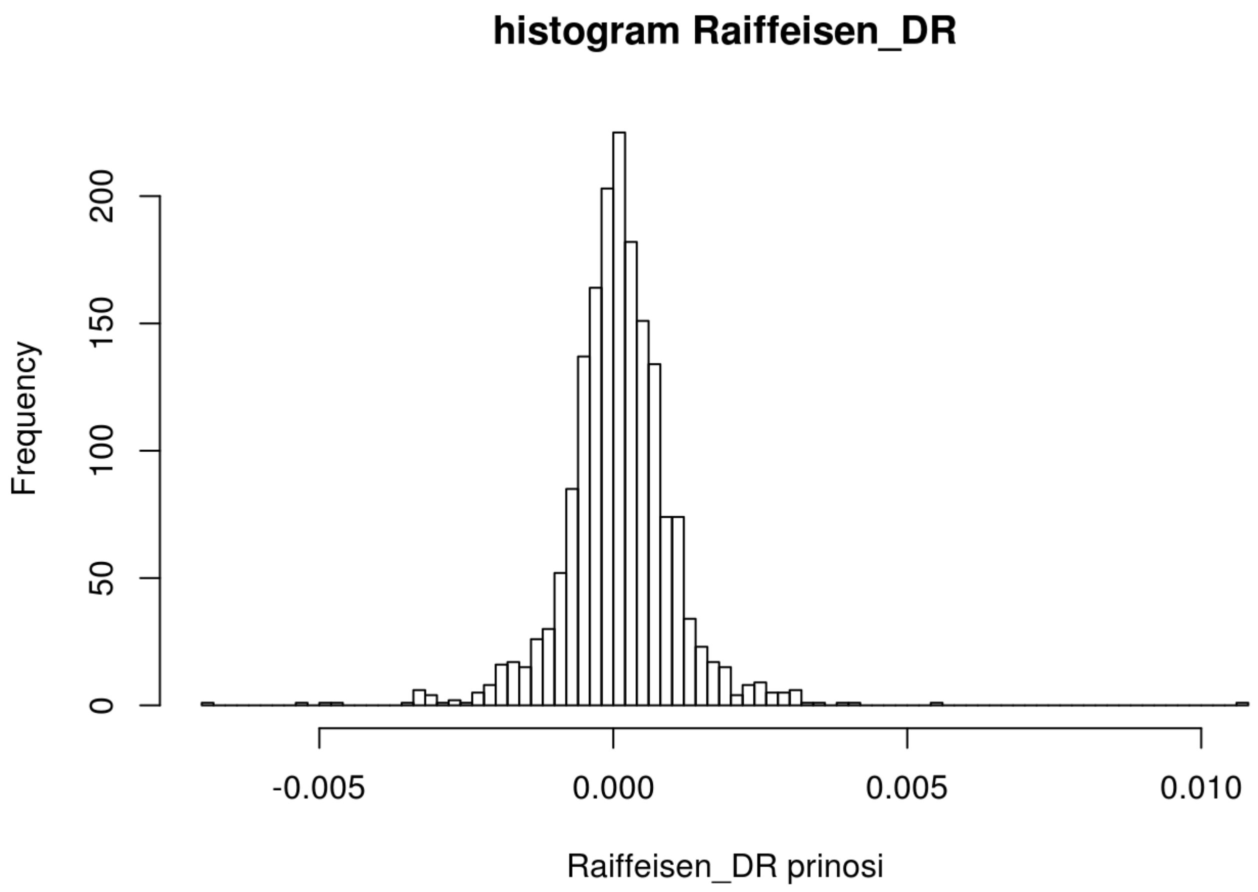

```
hist(data_log$OTP_DR, main="histogram OTP_DR",breaks = 100,
xlab = "OTP_DR prinosi", ylab = "Frequency" )
```



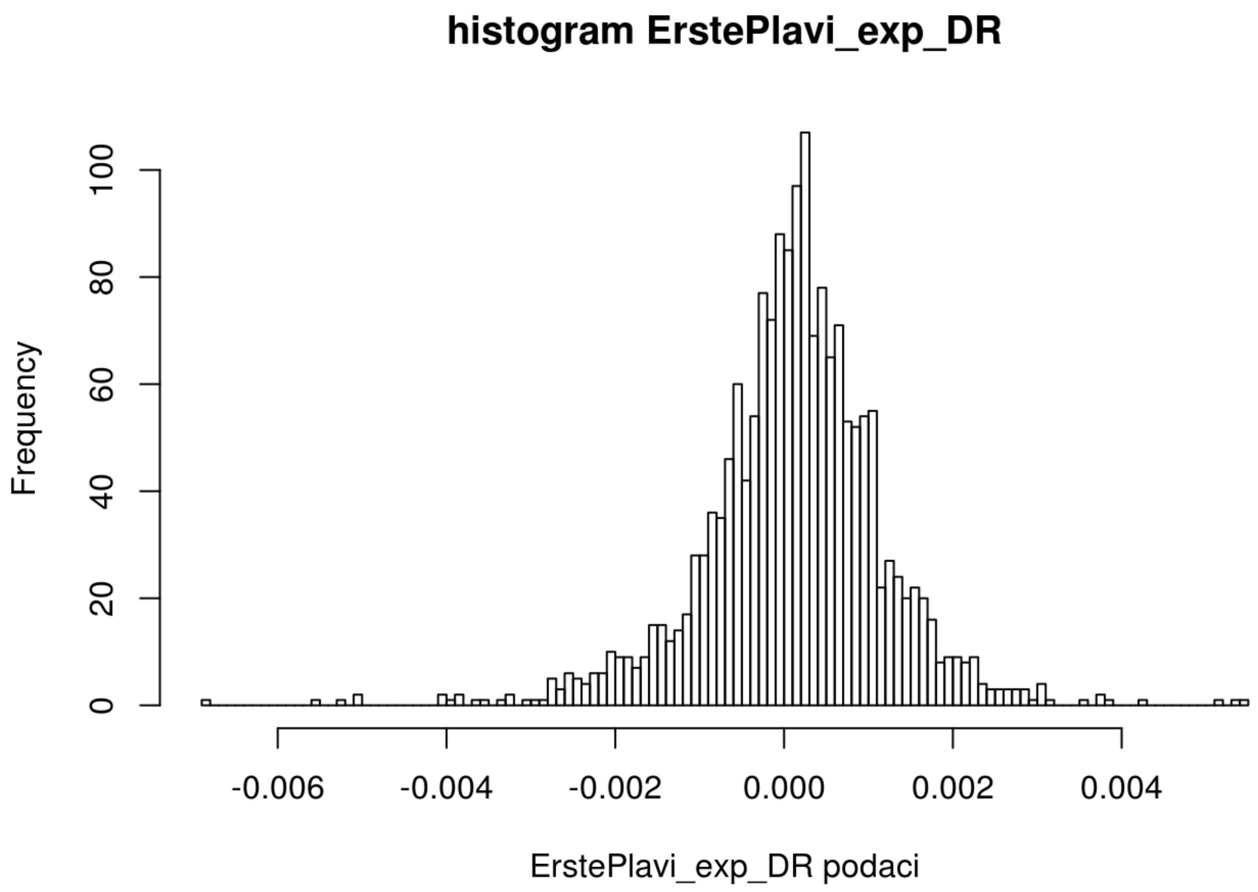

```r
hist(data_log$ZB_DR, main="histogram ZB_DR",breaks = 100,
xlab = "ZB_DR prinosi", ylab = "Frequency" )
```



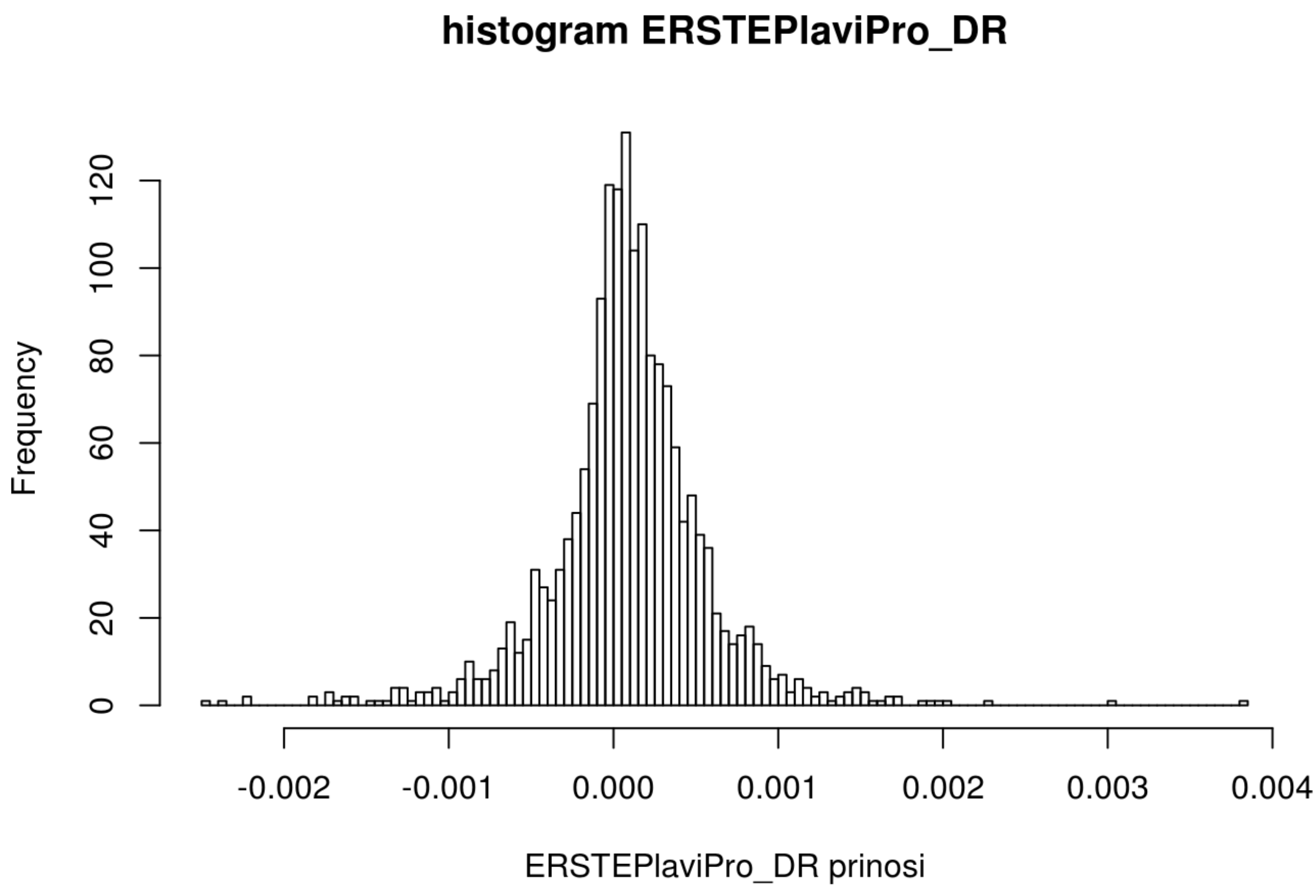

```
hist(data_log$Raiffeisen_DR, main="histogram Raiffeisen_DR",breaks = 100,
xlab = "Raiffeisen_DR prinosi", ylab = "Frequency" )
```




```
hist(data_log$ErstePlavi_exp_DR, main="histogram ErstePlavi_exp_DR",breaks = 100,
xlab = "ErstePlavi_exp_DR podaci", ylab = "Frequency" )
```

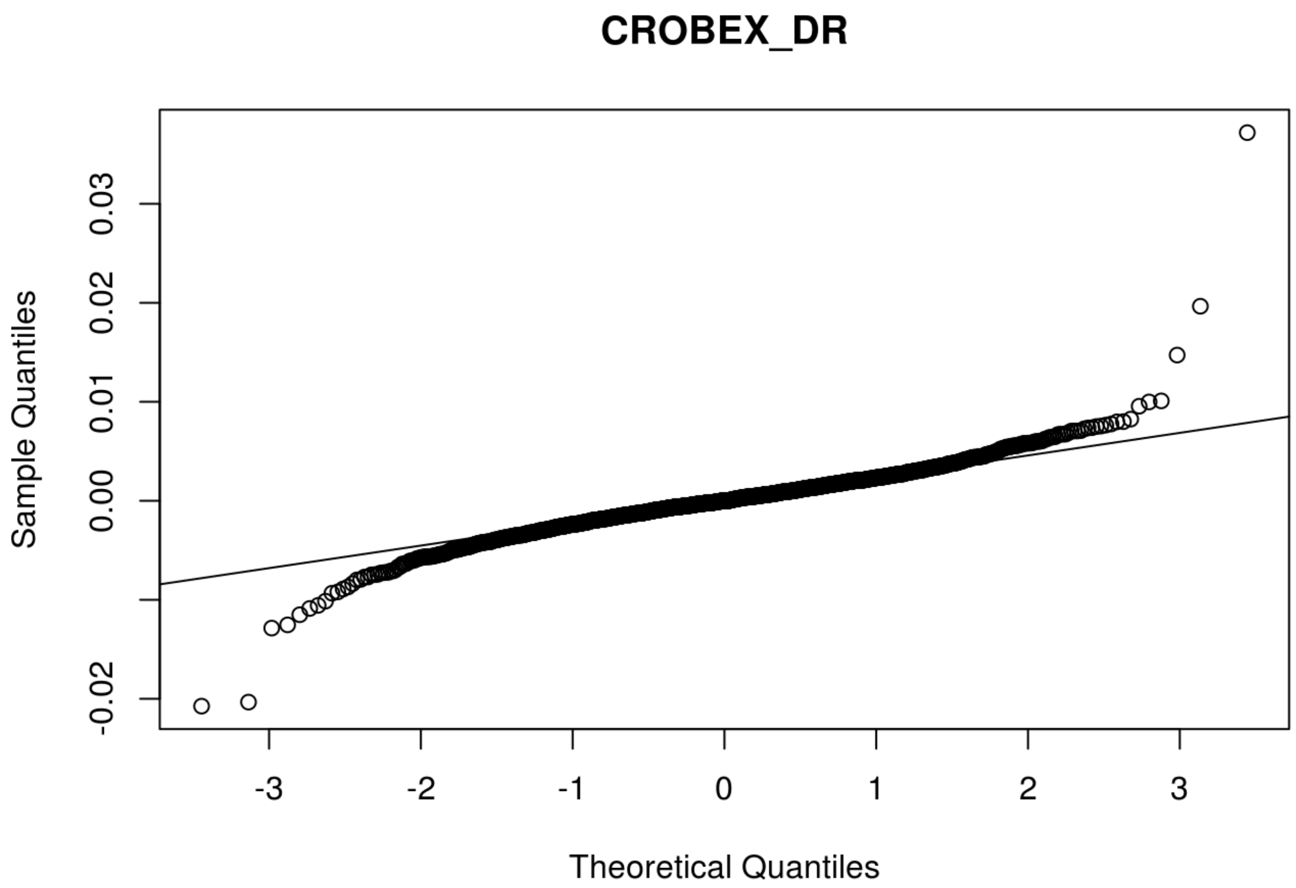



```r
hist(data_log$ERSTEPlaviPro_DR, main="histogram ERSTEPlaviPro_DR",breaks = 100,
xlab = "ERSTEPlaviPro_DR prinosi", ylab = "Frequency" )
```

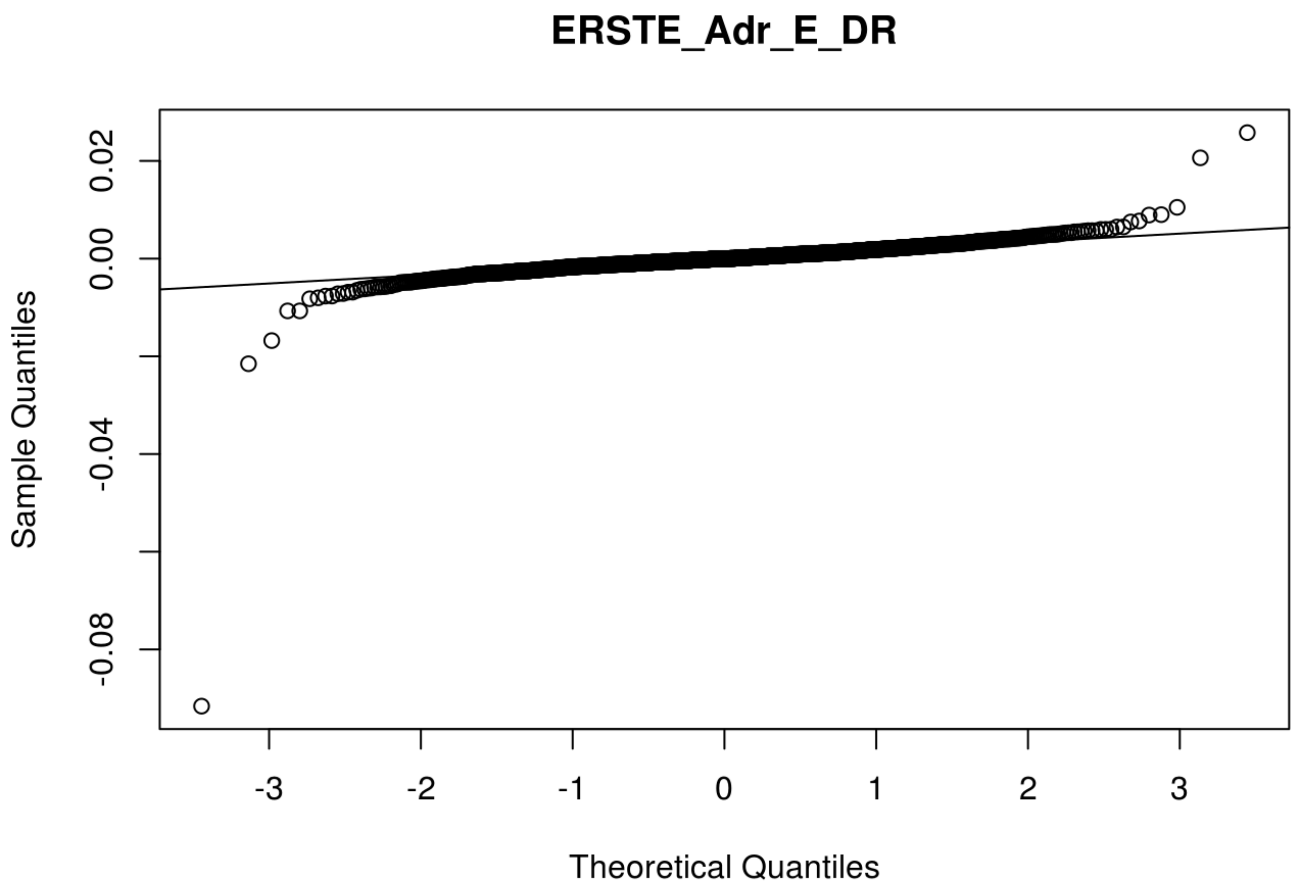


Iscrtavanje Q-Q plota pojedinog fonda u usporedbi s odgovarajućim pravcem normalne razdiobe.

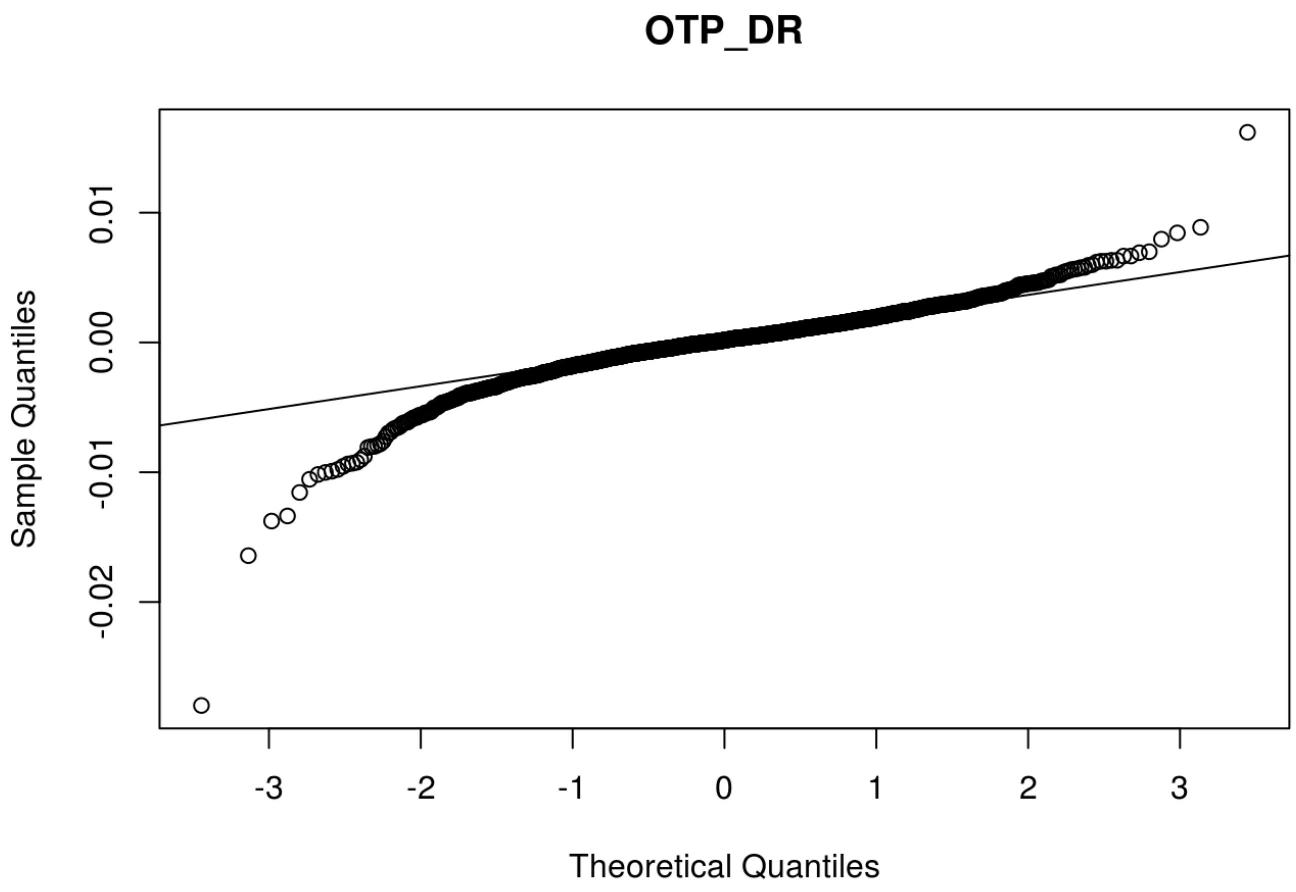
```
#Q-Q plotovi za logaritamske dnevne prinose (usporedba s normalnim qq plotom)

qqnorm(data_log$CROBEX_DR, main = "CROBEX_DR",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles",
       plot.it = TRUE, datax = FALSE)
qqline(data_log$CROBEX_DR)
```



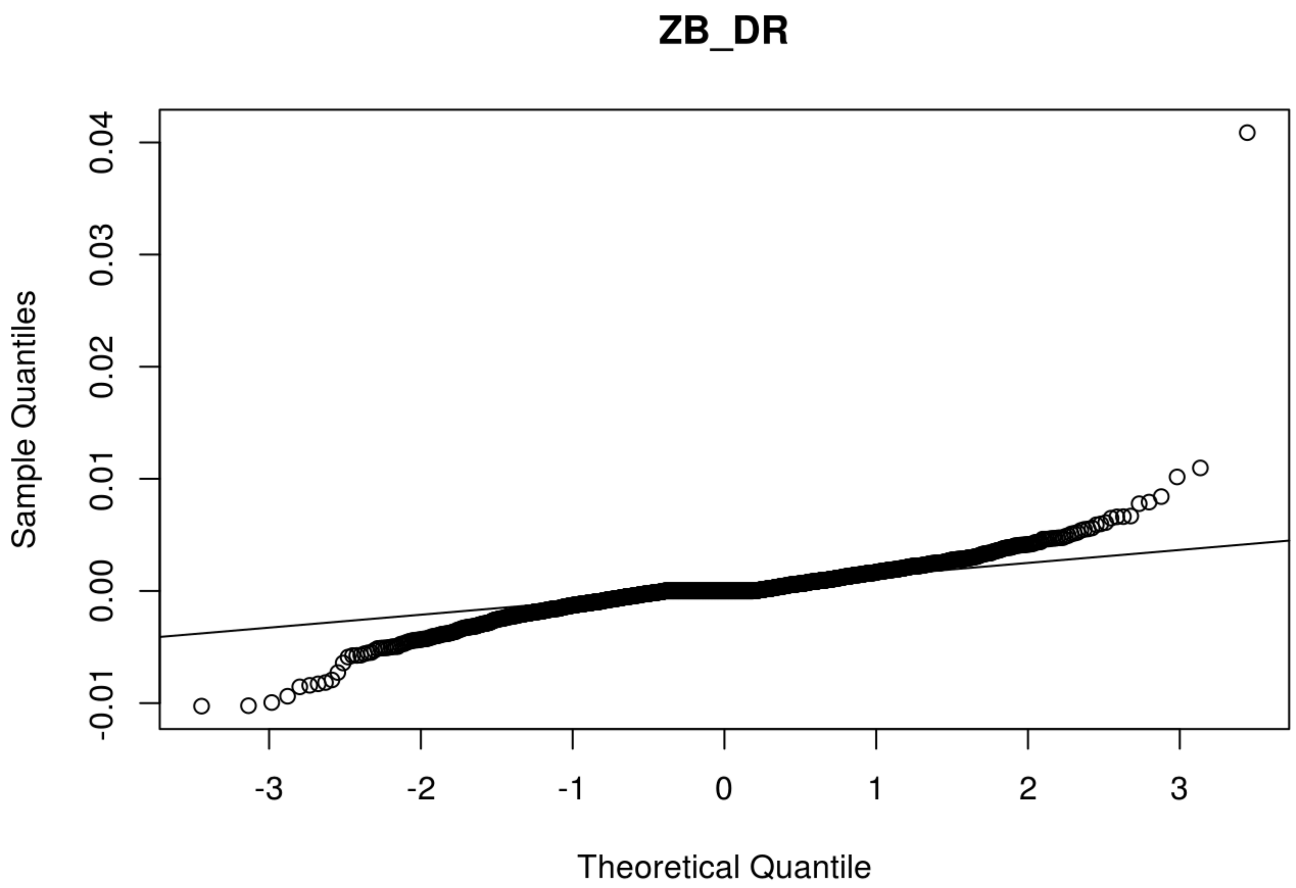

```r
qqnorm(data_log$ERSTE_Adr_E_DR, main = "ERSTE_Adr_E_DR",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles",
       plot.it = TRUE, datax = FALSE)
qqline(data_log$ERSTE_Adr_E_DR)
```



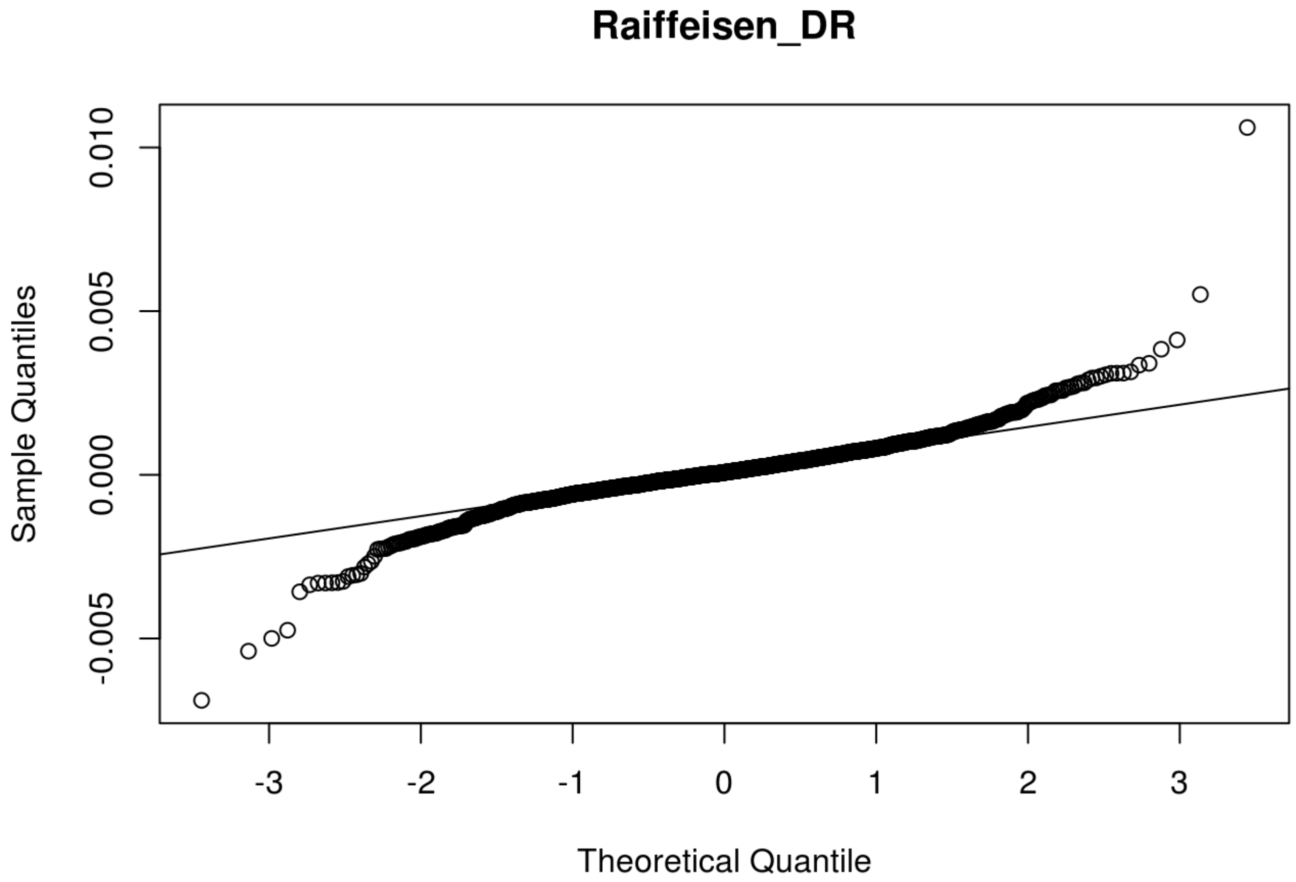

```
qqnorm(data_log$OTP_DR, main = "OTP_DR",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles",
       plot.it = TRUE, datax = FALSE)
qqline(data_log$OTP_DR)
```



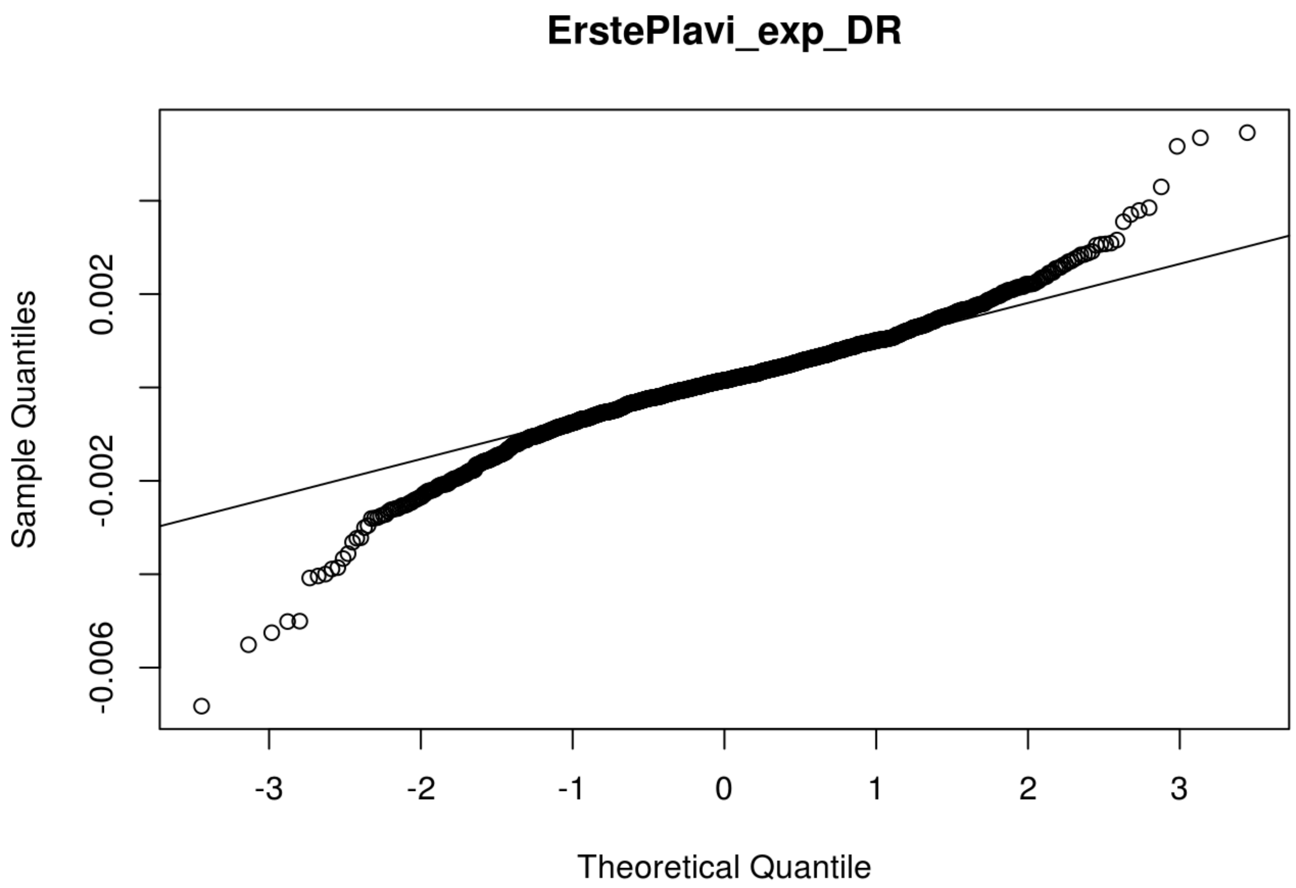

```
qqnorm(data_log$ZB_DR, main = "ZB_DR",
       xlab = "Theoretical Quantile", ylab = "Sample Quantiles",
       plot.it = TRUE, datax = FALSE)
qqline(data_log$ZB_DR)
```



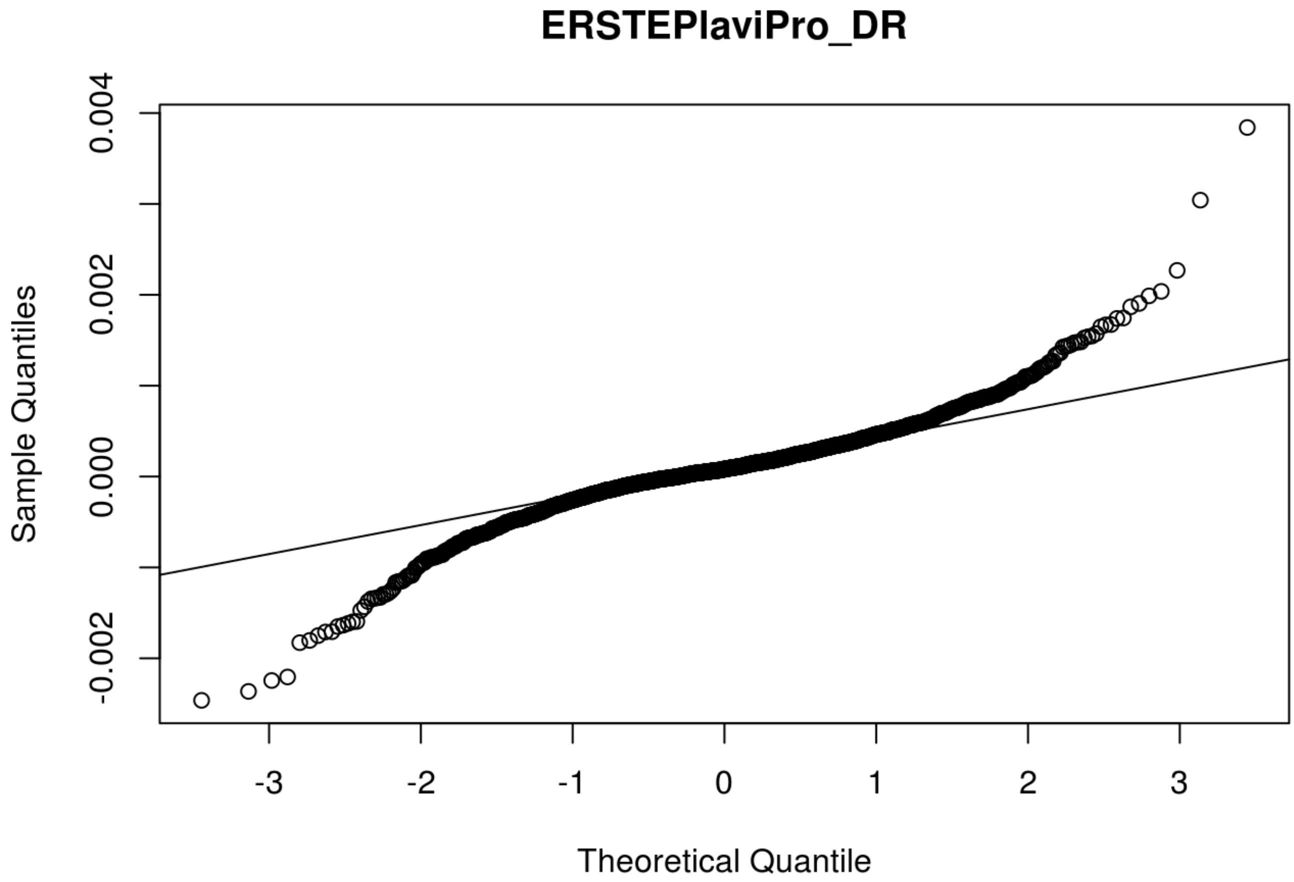

```
qqnorm(data_log$Raiffeisen_DR, main = "Raiffeisen_DR",
       xlab = "Theoretical Quantile", ylab = "Sample Quantiles",
       plot.it = TRUE, datax = FALSE)
qqline(data_log$Raiffeisen_DR)
```




```
qqnorm(data_log$ErstePlavi_exp_DR, main = "ErstePlavi_exp_DR",
       xlab = "Theoretical Quantile", ylab = "Sample Quantiles",
       plot.it = TRUE, datax = FALSE)
qqline(data_log$ErstePlavi_exp_DR)
```

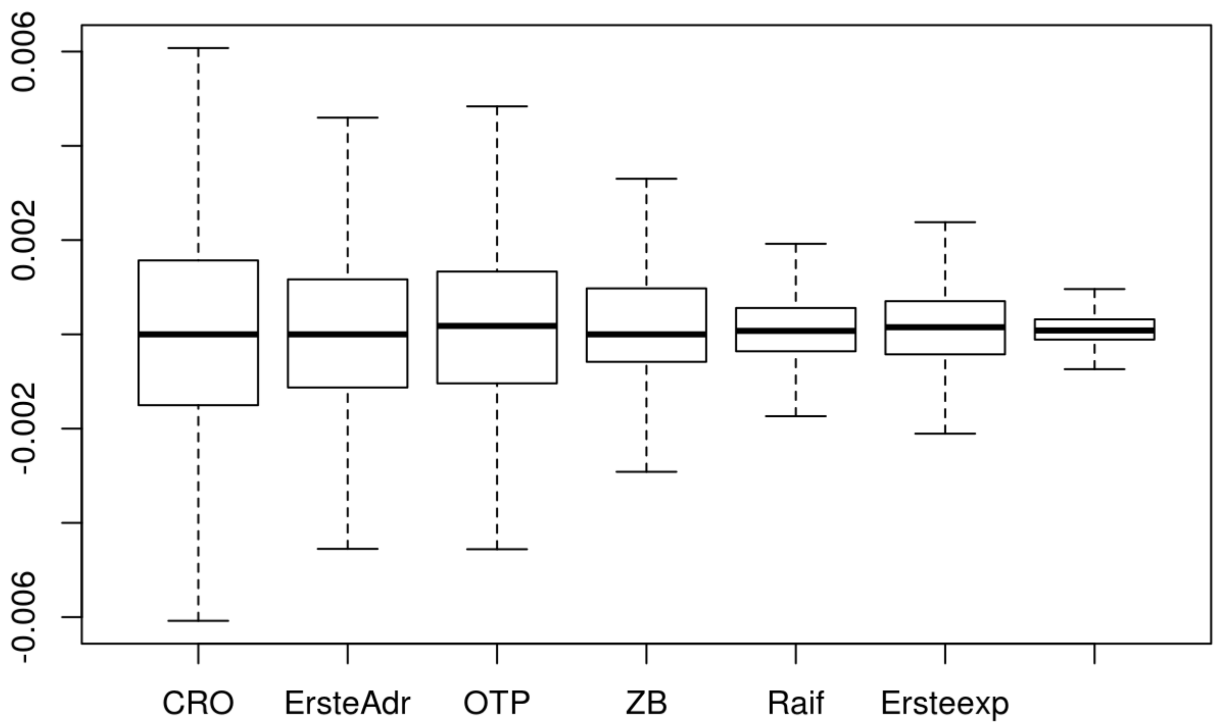



```
qqnorm(data_log$ERSTEPlaviPro_DR, main = "ERSTEPlaviPro_DR",
       xlab = "Theoretical Quantile", ylab = "Sample Quantiles",
       plot.it = TRUE, datax = FALSE)
qqline(data_log$ERSTEPlaviPro_DR)
```



Analiza srednjih vrijednosti i standardne devijacije uzoraka. Vizualizacija Box-plotom bez istaknutih stršećih vrijednosti. Razlog tome je bolja vizualizacija sličnosti i razlika u interkvartilnim rangovima pojedinih fondova. Širinu interkvartilnog ranga, uz standardnu devijaciju, uzimamo kao dobru (i robusnu) mjeru raspršenosti. Ponovno primjećujemo veću raspršenost podataka o dioničkim fondovima.

```r
#aritmetičke sredine i standardne devijacije logaritamskih dnevnih prinosa

CROBEX_mean <- mean(data_log$CROBEX_DR)
CROBEX_sd <- sd(data_log$CROBEX_DR)

Ersteadr_mean <- mean(data_log$ERSTE_Adr_E_DR)
Ersteadr_sd <- sd(data_log$ERSTE_Adr_E_DR)

OTP_mean <- mean(data_log$OTP_DR)
OTP_sd <- sd(data_log$OTP_DR)

ZB_mean <- mean(data_log$ZB_DR)
ZB_sd <- sd(data_log$ZB_DR)

Raiffeisen_mean <- mean(data_log$Raiffeisen_DR)
Raiffeisen_sd <- sd(data_log$Raiffeisen_DR)

Ersteplaviexp_mean <- mean(data_log$ErstePlavi_exp_DR)
Ersteplaviexp_sd <- sd(data_log$ErstePlavi_exp_DR)

Ersteplavipro_mean <- mean(data_log$ERSTEPlaviPro_DR)
Ersteplavipro_sd <- sd(data_log$ERSTEPlaviPro_DR)

boxplot(data_log$CROBEX_DR, data_log$ERSTE_Adr_E_DR, data_log$OTP_DR, data_log$ZB_DR, data_log$Raiffeisen_DR, data_log$ErstePlavi_exp_DR, data_log$ERSTEPlaviPro_DR, outline = FALSE, names = c("CRO", "ErsteAdr", "OTP", "ZB", "Raif", "Ersteexp", "Erstepro"))
```



Motivacija i opis problema:

Uspješnost fondova se stoga može mjeriti u odnosu na bezrizični prinos u Hrvatskoj, čija je vrijednost definirana kamatnim stopama na trezorske zapise RH, budući da investitori mogu birati štednju umjesto investicije u fondove. Budući da investicijski fondovi preuzimaju (veći) rizik, investitori primarno očekuju veće prinose od bezrizične stope. Također, budući da investicijski fondovi naplaćuju usluge aktivnog upravljanja, investitori očekuju bolje performanse (više prinose i/ili niži rizik) od tržišnog (referentnog) portfelja. Model koji objašnjava ove odnose zove se capital asset pricing model (CAPM) i svodi se na linearnu regresiju:

```
Rp − Rf = α + β (Rm − Rf) + e
```

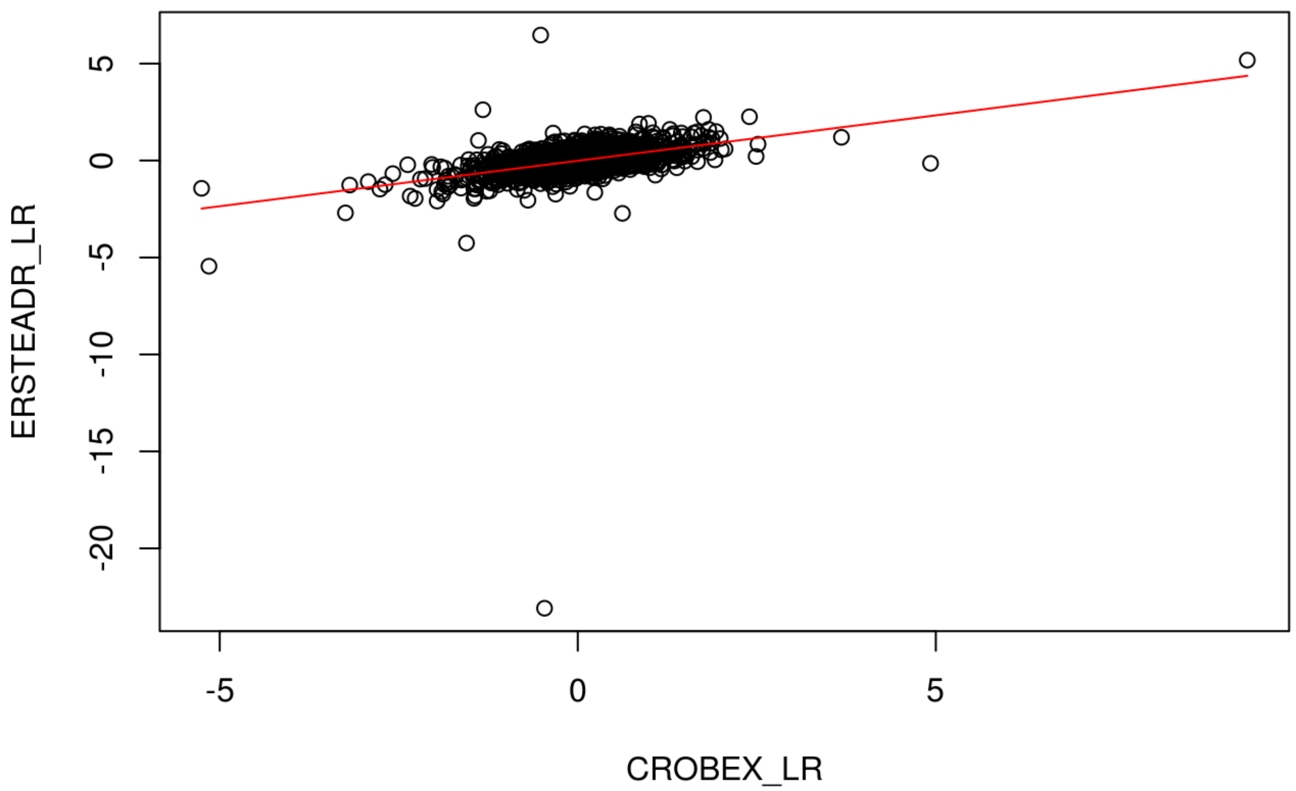
gdje je Rp prinos promatranog fonda (portfelja), Rm prinos tržišnog (referentnog) portfelja, a Rf je bezrizična kamatna stopa. Koeficijent α mjeri koliko je prinos promatranog fonda veći od prinosa tržišnog portfelja, a β mjeri osjetljivost fonda na tržišne prinose i predstavlja rizičnost.

2. zadatak: Usporediti statistička svojstva različitih fondova, a potom primijeniti CAPM model te usporediti fondove koristeći procijene koeficijenata α i β.

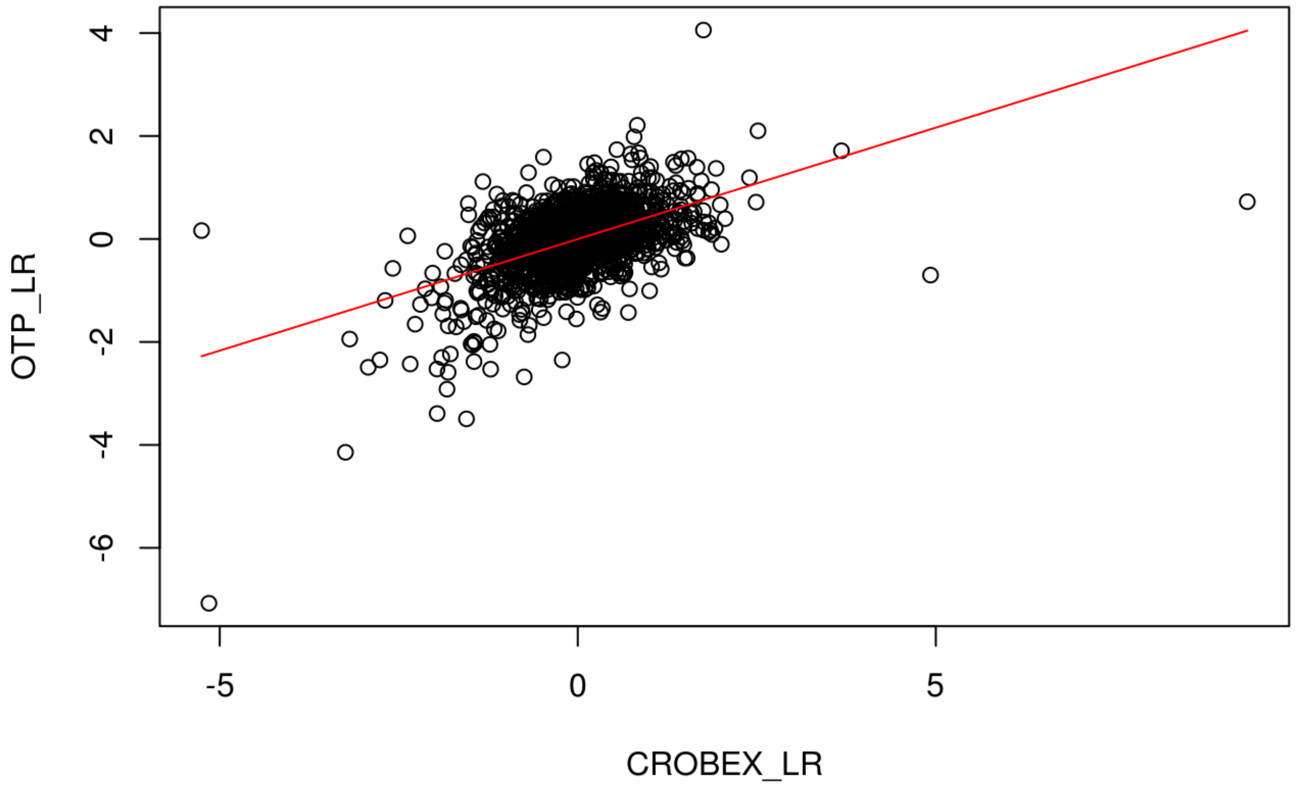
Rješenje problema: CAPM ostvaruje se linearnom regresijom metodom najmanjih kvadrata koja osigurava minimum varijance pojedinih podataka u uzorku u odnosu na procjenjenu linearnu funkcijsku ovisnost očekivanog portfelja o regresoru Rm − Rf. Pretpostavljamo nezavisnost prinosa u različitim danima i normalnu razdiobu reziduala e. Fond smatramo boljim ako je α što veći, a β što manji jer to podrazumijeva sigurne prinose neovisne o fluktuacijama na tržištu.

```
#izračun logaritamskih godišnjih prinosa

CROBEX_AN <- data_log$CROBEX_DR * 252
CROBEX_AN_MEAN <- mean(CROBEX_AN)
CROBEX_LR <- CROBEX_AN - data_log$InterestRate/100
CROBEX_AN_SD <- CROBEX_sd * sqrt(252)

#DIONIČKI

ERSTEADR_AN <- data_log$ERSTE_Adr_E_DR * 252
ERSTEADR_LR <- ERSTEADR_AN - data_log$InterestRate/100
ERSTEADR_AN_MEAN <- mean(ERSTEADR_AN)
ERSTEADR_AN_SD <- Ersteadr_sd * sqrt(252)


OTP_AN <- data_log$OTP_DR * 252
OTP_LR <- OTP_AN - data_log$InterestRate/100
OTP_AN_MEAN <- mean(OTP_AN)
OTP_AN_SD <- OTP_sd * sqrt(252)

ZB_AN <- data_log$ZB_DR * 252
ZB_LR <- ZB_AN - data_log$InterestRate/100
ZB_AN_MEAN <- mean(ZB_AN)
ZB_AN_SD <- ZB_sd * sqrt(252)

#MIROVINSKI

Raiffeisen_AN <- data_log$Raiffeisen_DR * 252
Raiffeisen_LR <- Raiffeisen_AN - data_log$InterestRate/100
Raiffeisen_AN_MEAN <- mean(Raiffeisen_AN)
Raiffeisen_AN_SD <- Raiffeisen_sd * sqrt(252)

Ersteplaviexp_AN <- data_log$ErstePlavi_exp_DR * 252
Ersteplaviexp_LR <- Ersteplaviexp_AN - data_log$InterestRate/100
Ersteplaviexp_AN_MEAN <- mean(Ersteplaviexp_AN)
Ersteplaviexp_AN_SD <- Ersteplaviexp_sd * sqrt(252)

Ersteplavipro_AN <- data_log$ERSTEPlaviPro_DR * 252
Ersteplavipro_LR <- Ersteplavipro_AN - data_log$InterestRate/100
Ersteplavipro_AN_MEAN <- mean(Ersteplavipro_AN)
Ersteplavipro_AN_SD <- Ersteplavipro_sd * sqrt(252)
```

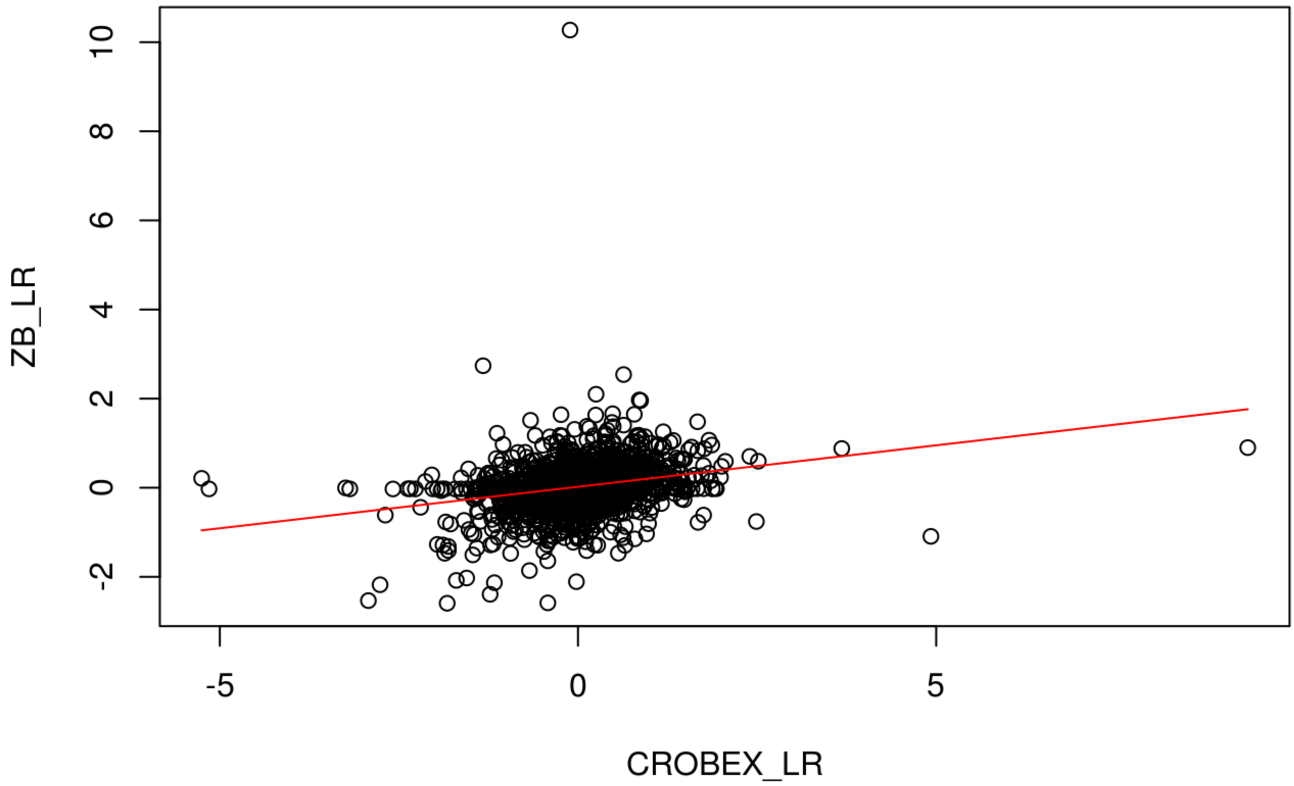
```
#LINEARNA REGRESIJA

#dionički

ERSTEADR_LM <- lm(ERSTEADR_LR~CROBEX_LR)

plot(CROBEX_LR, ERSTEADR_LR) #plot podataka
lines(CROBEX_LR,ERSTEADR_LM$fitted.values,col='red')
```



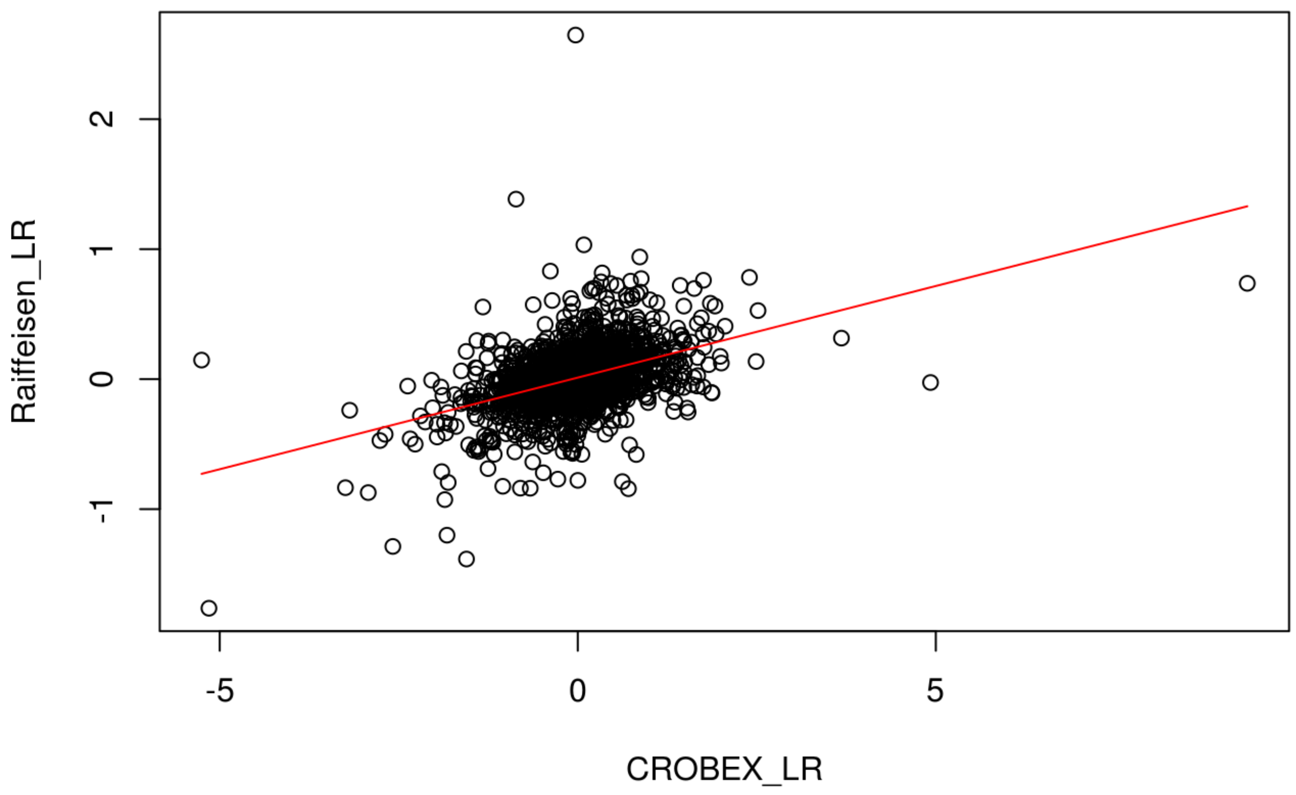

```
OTP_LM <- lm(OTP_LR~CROBEX_LR)

plot(CROBEX_LR, OTP_LR) #plot podataka
lines(CROBEX_LR, OTP_LM$fitted.values,col='red')
```



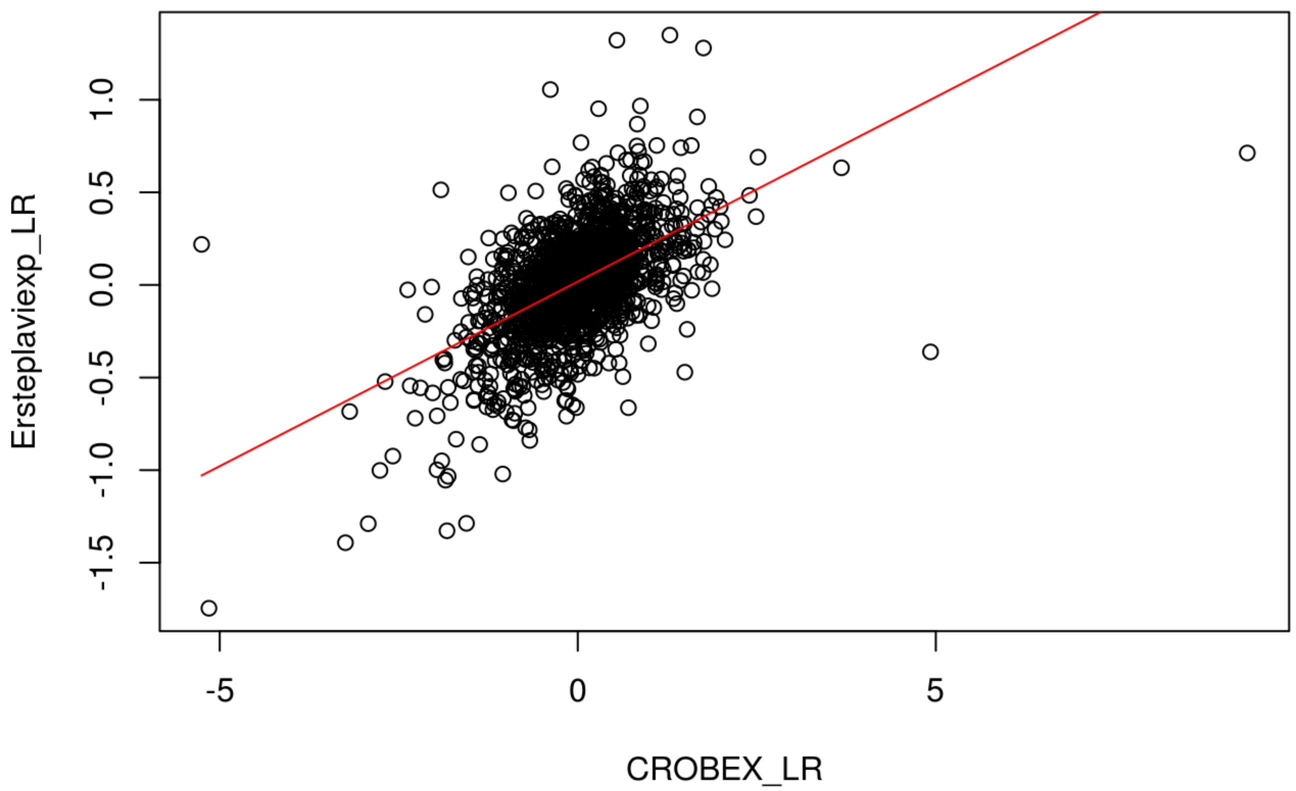

```
ZB_LM <- lm(ZB_LR~CROBEX_LR)

plot(CROBEX_LR, ZB_LR) #plot podataka
lines(CROBEX_LR, ZB_LM$fitted.values,col='red')
```



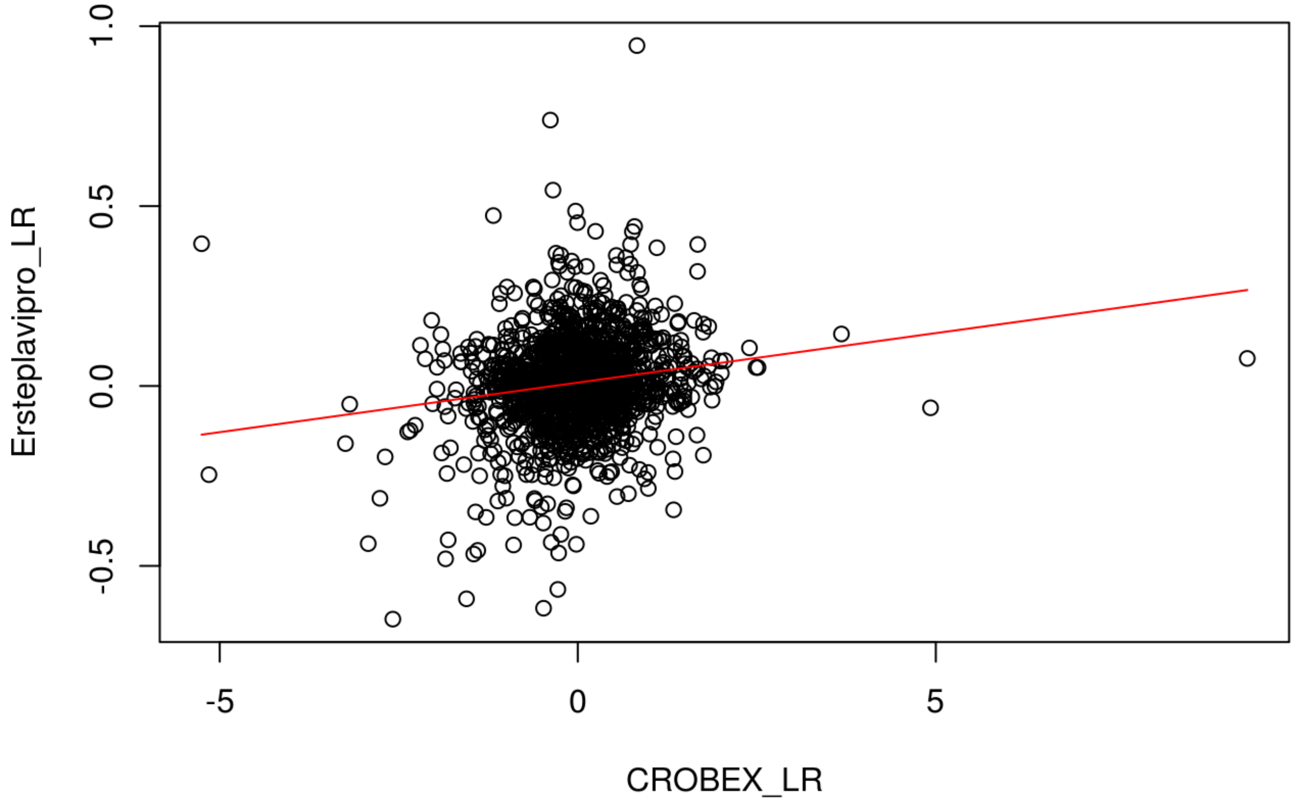

```
#mirovinski

Raiffeisen_LM <- lm(Raiffeisen_LR~CROBEX_LR)

plot(CROBEX_LR, Raiffeisen_LR) #plot podataka
lines(CROBEX_LR, Raiffeisen_LM$fitted.values,col='red')
```




```
Ersteplaviexp_LM <- lm(Ersteplaviexp_LR~CROBEX_LR)

plot(CROBEX_LR, Ersteplaviexp_LR) #plot podataka
lines(CROBEX_LR,Ersteplaviexp_LM$fitted.values,col='red')
```




```
Ersteplavipro_LM <- lm(Ersteplavipro_LR~CROBEX_LR)

plot(CROBEX_LR, Ersteplavipro_LR) #plot podataka
lines(CROBEX_LR,Ersteplavipro_LM$fitted.values,col='red')
```



1.0
0.5
0.0
-0.5
Ersteplavipro_LR
-5
0
5
CROBEX_LR

```r
#tablica koeficijenata linearne regresije i standardne devijacije reziduala
fonds = c("Erste Adriatic", "OTP", "ZABA", "Raiffeisen", "Erste plavi exp", "Erste
plavi pro")

alpha = c(ERSTEADR_LM$coefficients[[1]], OTP_LM$coefficients[[1]], ZB_LM$coefficie
nts[[1]], Raiffeisen_LM$coefficients[[1]], Ersteplaviexp_LM$coefficients[[1]], Ers
teplavipro_LM$coefficients[[1]])

beta = c(ERSTEADR_LM$coefficients[[2]], OTP_LM$coefficients[[2]], ZB_LM$coefficien
ts[[2]], Raiffeisen_LM$coefficients[[2]], Ersteplaviexp_LM$coefficients[[2]], Erst
eplavipro_LM$coefficients[[2]])

r_squared = c(summary(ERSTEADR_LM)$r.squared, summary(OTP_LM)$r.squared, summary(Z
B_LM)$r.squared, summary(Raiffeisen_LM)$r.squared, summary(Ersteplaviexp_LM)$r.squa
red, summary(Ersteplavipro_LM)$r.squared)

alpha_p = c(summary(ERSTEADR_LM)$coefficients[,4][[1]], summary(OTP_LM)$coefficien
ts[,4][[1]], summary(ZB_LM)$coefficients[,4][[1]], summary(Raiffeisen_LM)$coeffici
ents[,4][[1]], summary(Ersteplaviexp_LM)$coefficients[,4][[1]], summary(Ersteplavi
pro_LM)$coefficients[,4][[1]])

beta_p = c(summary(ERSTEADR_LM)$coefficients[,4][[2]], summary(OTP_LM)$coefficient
s[,4][[2]], summary(ZB_LM)$coefficients[,4][[2]], summary(Raiffeisen_LM)$coefficie
nts[,4][[2]], summary(Ersteplaviexp_LM)$coefficients[,4][[2]], summary(Ersteplavip
ro_LM)$coefficients[,4][[2]])

regression_table = data.frame(row.names = fonds, alpha, alpha_p, beta, beta_p, r_s
quared)

save(regression_table, file = "dataframes/regression_table.Rda")

regression_table
```

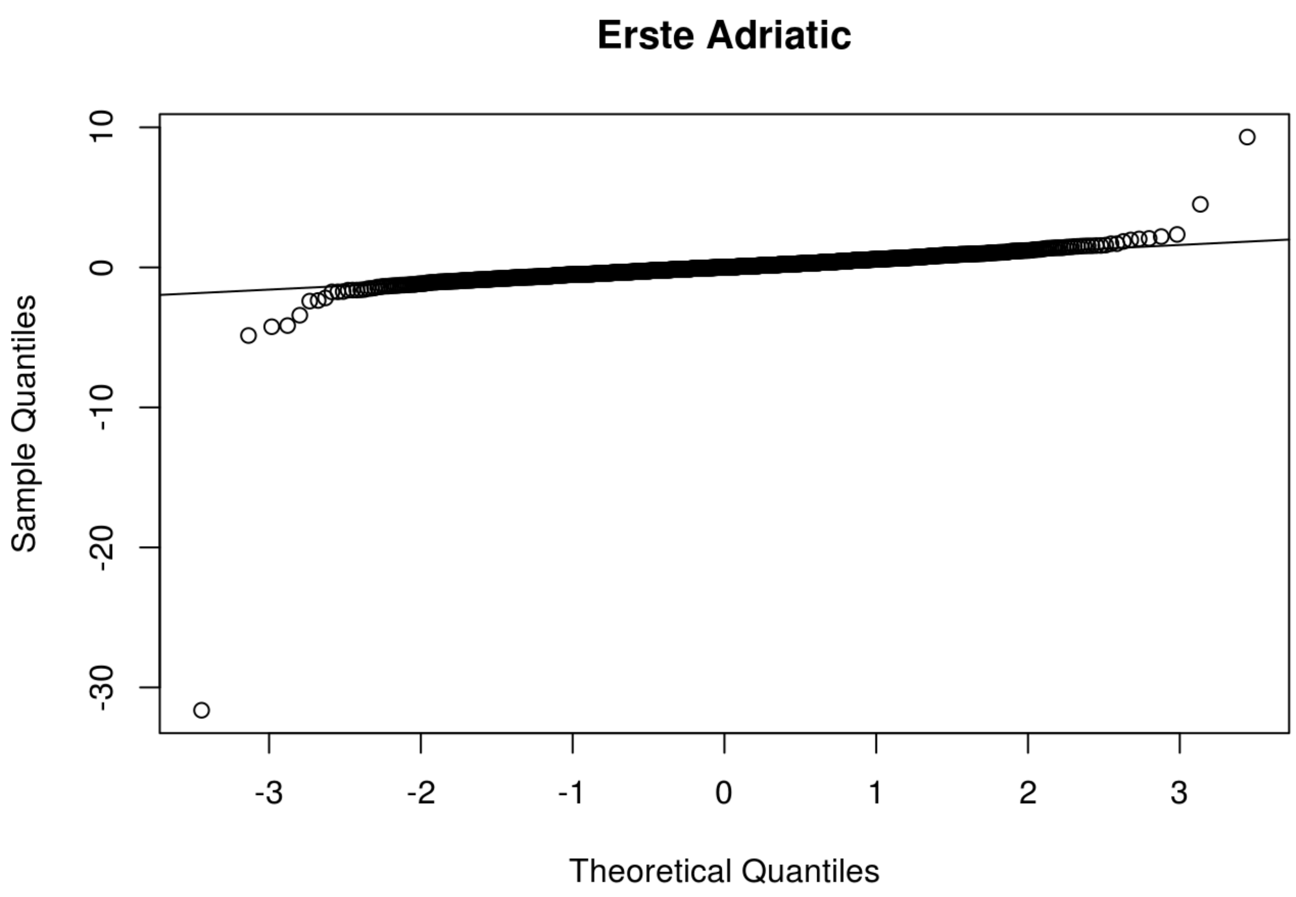
```
##                        alpha      alpha_p       beta        beta_p
## Erste Adriatic  -0.013774980 0.4257393383 0.46868387  3.747153e-82
## OTP             -0.002761662 0.8285228284 0.43308279 2.484227e-121
## ZABA             0.021751975 0.0837553000 0.18621930  1.382838e-27
## Raiffeisen       0.011647813 0.0250737956 0.14095772  2.673960e-82
## Erste plavi exp  0.018118358 0.0009396681 0.19936374 1.390261e-136
## Erste plavi pro  0.009316311 0.0011792341 0.02753062  1.049722e-12
##                  r_squared
## Erste Adriatic  0.19027316
## OTP             0.26958046
## ZABA            0.06557182
## Raiffeisen      0.19058515
## Erste plavi exp 0.29832409
## Erste plavi pro 0.02863935
```

Prethodna dva isječka koda ostvaruju tabličnu i grafičku interpretaciju kvalitete izračunate linearne regresije. Bitno je primijetiti da koeficijenti alpha i beta opisuju pravac linearne regresije, alpha_p i beta_p signifikantnost istih koeficijenata, a r_squared raspršenost reziduala oko linearne funkcije. Očito je da investicijski fondovi generalno pokazuju veću sklonost praćenju stanja tržišta (beta koeficijent je velik), dok mirovinske fondove opisuje manja vrijednost beta koeficijenta. Po pitanju beta koeficijenta i sigurnosti ulaganja, ulaganje u mirovinske fondove je preporučljivije. Što se tiče koeficijenta alpha, značajniju razliku među dvama kategorijama fondova ne primijećujemo pa formiramo zaključak da je u uvjetima hrvatskog tržišta bolje ulaganje u mirovinske fondove. Možda najbolji fond pojedinačno je ipak dionički fond Zagrebačke Banke jer ga obilježava apsulutno najveći koeficijent alpha, a relativno mali (u usporedbi s ostalim dioničkim fondovima) koeficijent beta. Također, bitno je naglasiti da je i raspršenost oko pravca linearne regresije relativno mala što garantira sigurnost ulaganja.

Općenito, signifikatnost koeficijenta beta je izuzetno velika pa ga možemo prihvatiti kao relevantnu procjenu kod svih pojedinih fondova. Koeficijent alpha je relativno nesignifikantan kod Erste i OTP fondova, no to više ne može utjecati na odluke o ulaganju jer smo po ostalim parametrima već odabrali ulaganje u ZBActiv.

U nastavku provodimo analizu normalnosti reziduala linearne regresije korištenjem vizualizacije Q-Q plotom i deterministički - korištenjem Kolmogorov-Smirnovog testa. Normalnost ne može biti potvrđena, ali to značajno ne utječe na rezultate naše procjene.

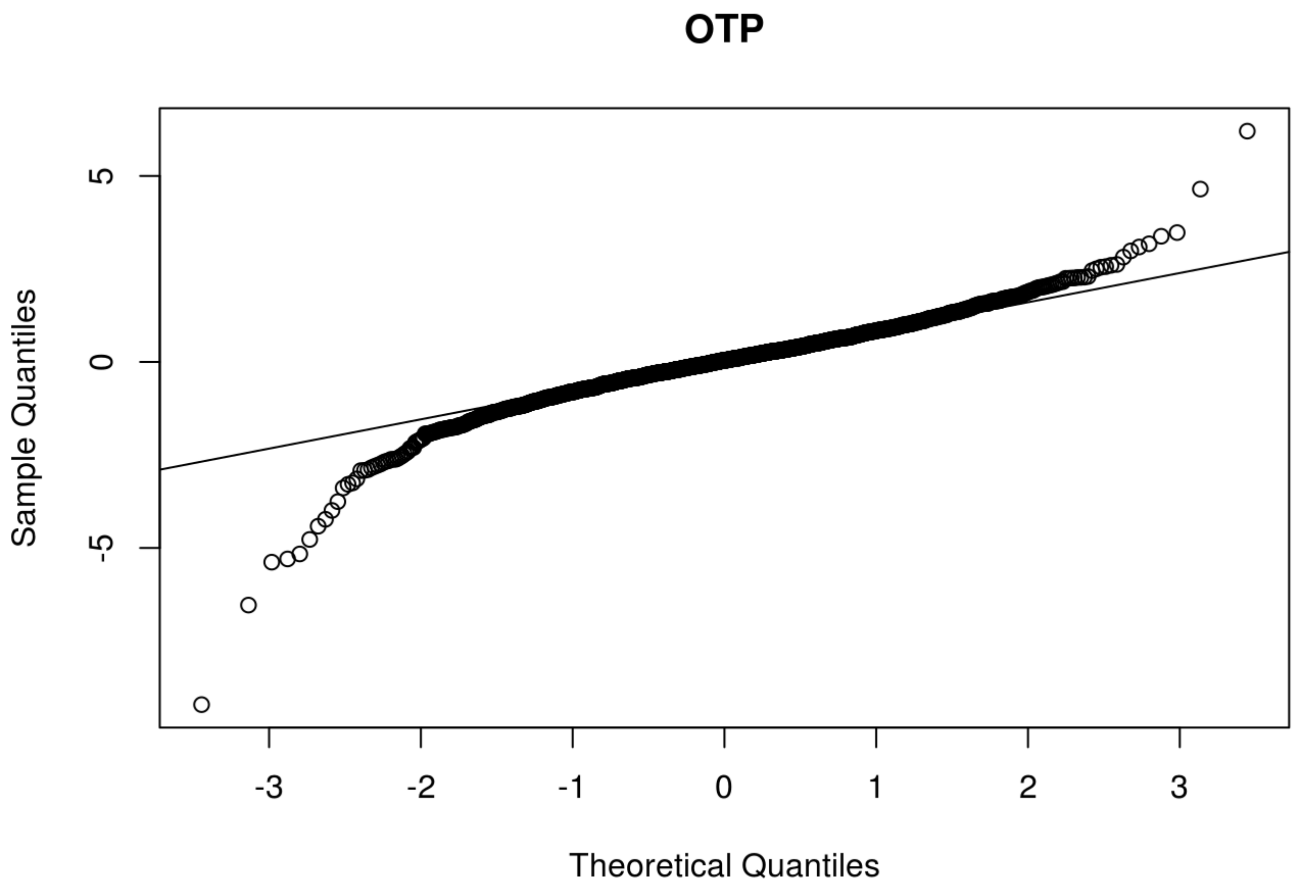
```
#usporedba distribucije reziduala s normalnom razdiobom

qqnorm(rstandard(ERSTEADR_LM), main = "Erste Adriatic")
qqline(rstandard(ERSTEADR_LM))
```




```
ks.test(rstandard(ERSTEADR_LM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(ERSTEADR_LM)
## D = 0.15075, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

```
qqnorm(rstandard(OTP_LM), main = "OTP")
qqline(rstandard(OTP_LM))
```

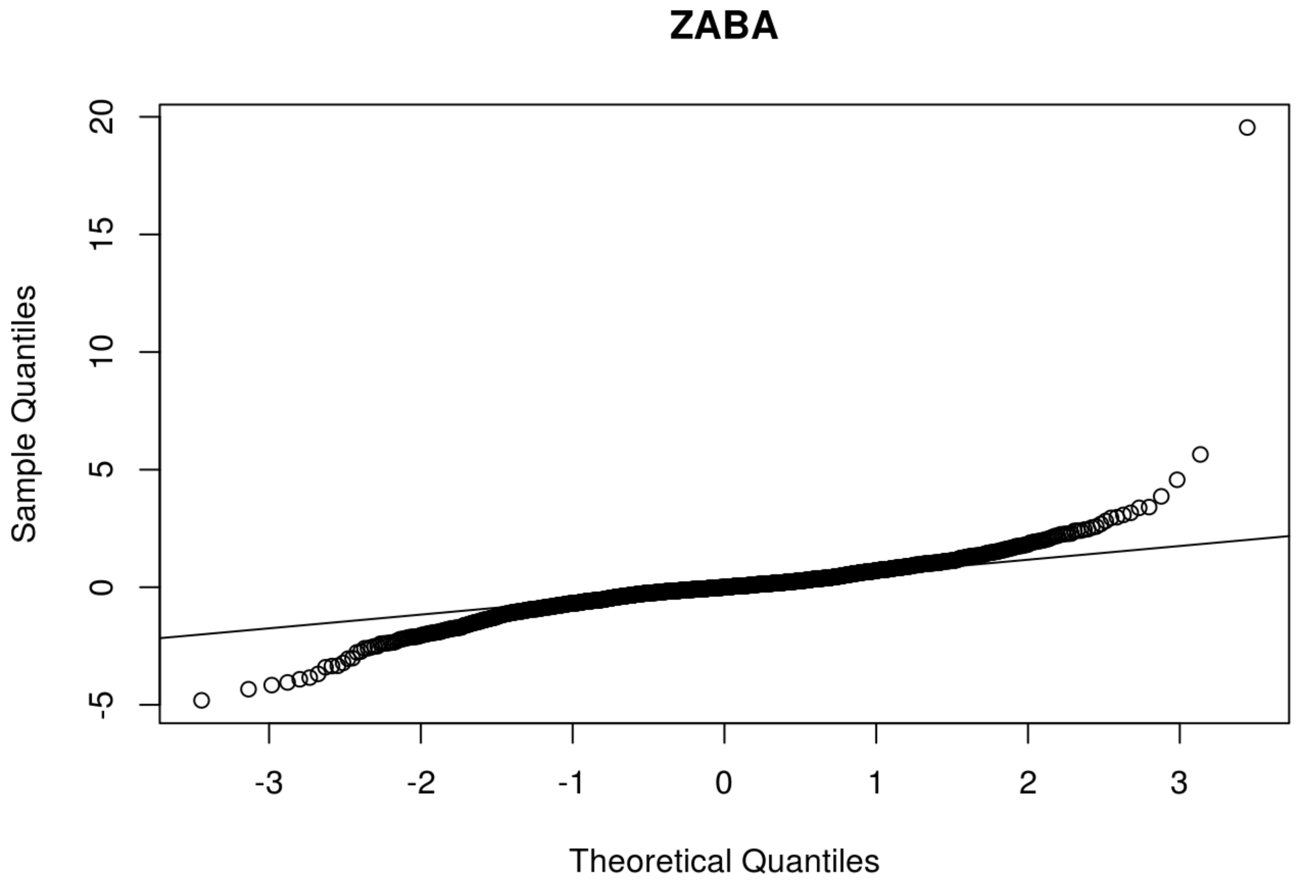



```
ks.test(rstandard(OTP_LM), pnorm)
```

```
## Warning in ks.test(rstandard(OTP_LM), pnorm): ties should not be present
## for the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(OTP_LM)
## D = 0.063397, p-value = 1.568e-06
## alternative hypothesis: two-sided
```

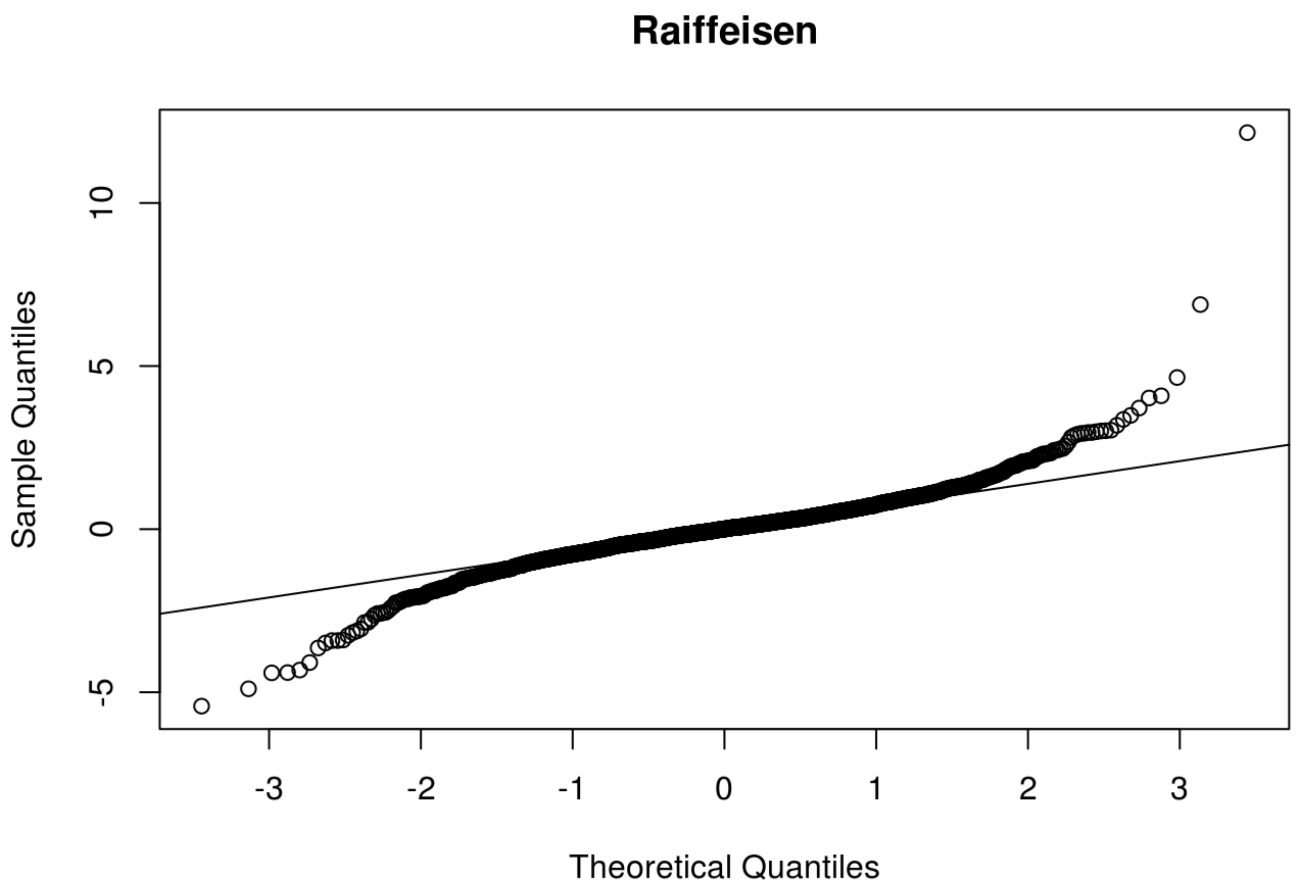
```
qqnorm(rstandard(ZB_LM), main = "ZABA")
qqline(rstandard(ZB_LM))
```




```r
ks.test(rstandard(ZB_LM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(ZB_LM)
## D = 0.10239, p-value = 2.22e-16
## alternative hypothesis: two-sided
```

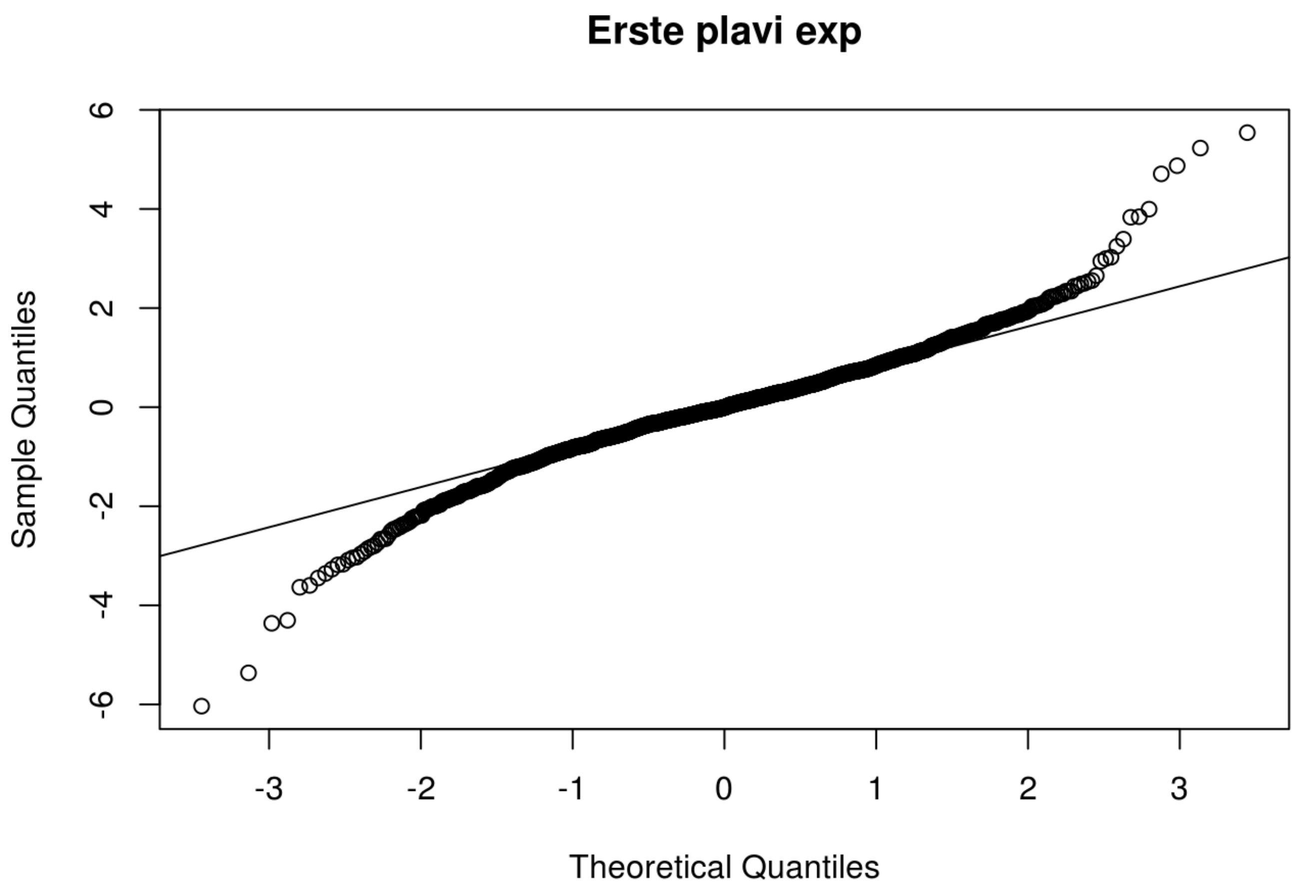
```r
qqnorm(rstandard(Raiffeisen_LM), main = "Raiffeisen")
qqline(rstandard(Raiffeisen_LM))
```




```
ks.test(rstandard(Raiffeisen_LM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(Raiffeisen_LM)
## D = 0.073677, p-value = 1.134e-08
## alternative hypothesis: two-sided
```

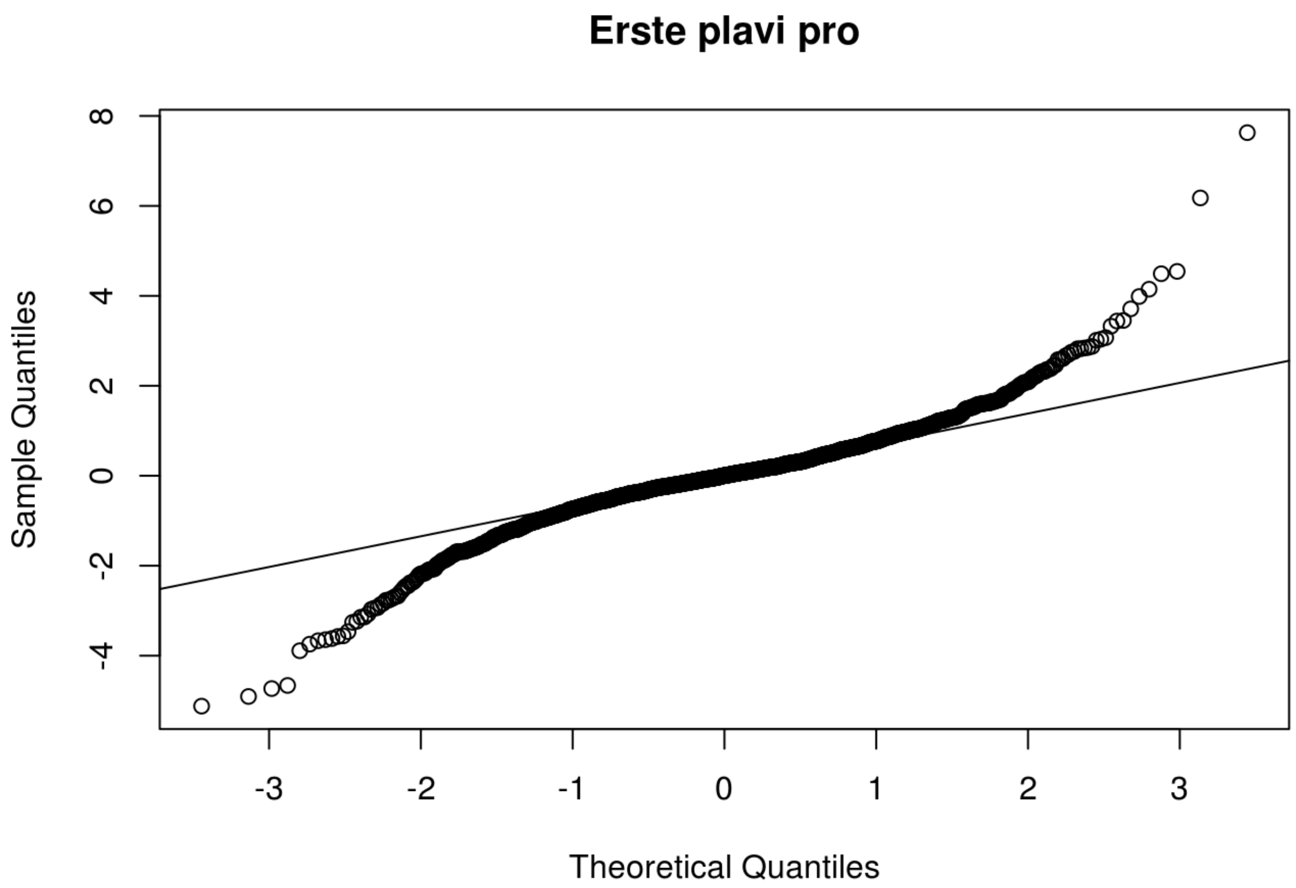
```
qqnorm(rstandard(Ersteplaviexp_LM), main = "Erste plavi exp")
qqline(rstandard(Ersteplaviexp_LM))
```




```
ks.test(rstandard(Ersteplaviexp_LM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(Ersteplaviexp_LM)
## D = 0.056698, p-value = 2.615e-05
## alternative hypothesis: two-sided
```

```
qqnorm(rstandard(Ersteplavipro_LM), main = "Erste plavi pro")
qqline(rstandard(Ersteplavipro_LM))
```




```
ks.test(rstandard(Ersteplavipro_LM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(Ersteplavipro_LM)
## D = 0.081219, p-value = 1.905e-10
## alternative hypothesis: two-sided
```

U nastavku provodimo analizu nelinearnog regresijskog modela uz pretpostavku postojanja koeficijenata gama i delta različitih od nule. Usporedbu s linearnom regresijom provodimo usporedbom R^2 parametra i testom normalnosti razdiobe reziduala. Zaključak je da je nelinearna regresija s kvadratnim i kubnim članom kvalitetnija procjena po pitanju oba kriterija jer daje veću vrijednost R^2, a reziduali pokazuju veću tendenciju praćenja normalne distribucije (veća je p vrijednost Kolmogorov-Smirnovog testa).

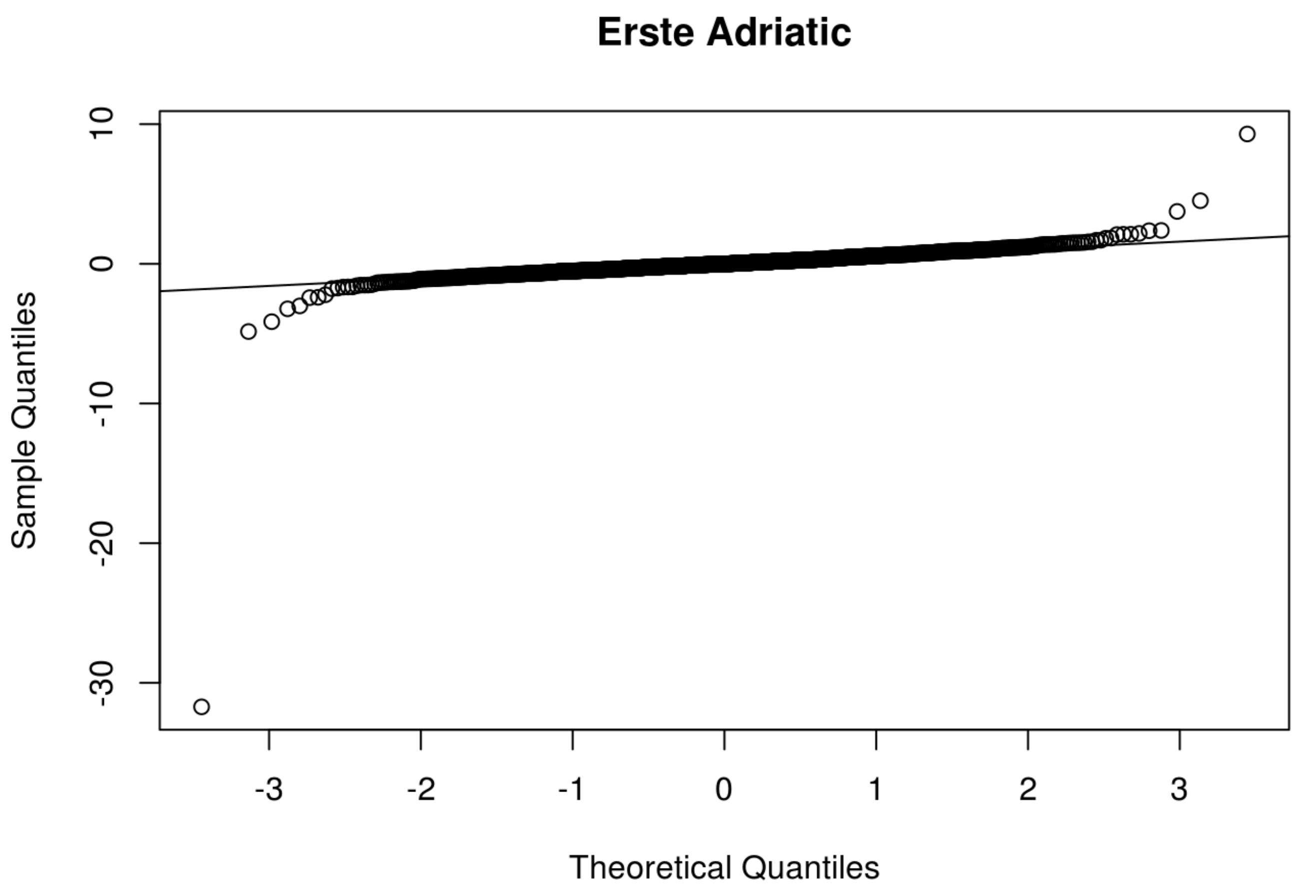
```
#Regresija uz proširenje regresijske funkcije kvadratom i kubom regresora

CROBEX_KR <- CROBEX_LR^2
CROBEX_TR <- CROBEX_LR^3

ERSTEADR_TM <- lm(ERSTEADR_LR~CROBEX_LR +I(CROBEX_KR)+I(CROBEX_TR))
OTP_TM = lm(OTP_LR~CROBEX_LR +I(CROBEX_KR)+I(CROBEX_TR))
ZB_TM = lm(ZB_LR~CROBEX_LR +I(CROBEX_KR)+I(CROBEX_TR))
Raiffeisen_TM = lm(Raiffeisen_LR~CROBEX_LR +I(CROBEX_KR)+I(CROBEX_TR))
Ersteplaviexp_TM = lm(Ersteplaviexp_LR~CROBEX_LR +I(CROBEX_KR)+I(CROBEX_TR))
Ersteplavipro_TM = lm(Ersteplavipro_LR~CROBEX_LR +I(CROBEX_KR)+I(CROBEX_TR))

alpha_n = c(ERSTEADR_TM$coefficients[[1]], OTP_TM$coefficients[[1]], ZB_TM$coeffic
```

```
iients[[1]], Raiffeisen_TM$coefficients[[1]], Ersteplaviexp_TM$coefficients[[1]], E
rsteplavipro_TM$coefficients[[1]])

beta_n = c(ERSTEADR_TM$coefficients[[2]], OTP_TM$coefficients[[2]], ZB_TM$coeffici
ents[[2]], Raiffeisen_TM$coefficients[[2]], Ersteplaviexp_TM$coefficients[[2]], Er
steplavipro_TM$coefficients[[2]])

gama_n = c(ERSTEADR_TM$coefficients[[3]], OTP_TM$coefficients[[3]], ZB_TM$coeffici
ents[[3]], Raiffeisen_TM$coefficients[[3]], Ersteplaviexp_TM$coefficients[[3]], Er
steplavipro_TM$coefficients[[3]])

delta_n = c(ERSTEADR_TM$coefficients[[4]], OTP_TM$coefficients[[4]], ZB_TM$coeffic
ients[[4]], Raiffeisen_TM$coefficients[[4]], Ersteplaviexp_TM$coefficients[[4]], E
rsteplavipro_TM$coefficients[[4]])

# p-vrijednosti
alpha_n_p = c(summary(ERSTEADR_TM)$coefficients[,4][[1]], summary(OTP_TM)$coeffici
ents[,4][[1]], summary(ZB_TM)$coefficients[,4][[1]], summary(Raiffeisen_TM)$coeffi
cients[,4][[1]], summary(Ersteplaviexp_TM)$coefficients[,4][[1]], summary(Erstepla
vipro_TM)$coefficients[,4][[1]])

beta_n_p = c(summary(ERSTEADR_TM)$coefficients[,4][[2]], summary(OTP_TM)$coefficie
nts[,4][[2]], summary(ZB_TM)$coefficients[,4][[2]], summary(Raiffeisen_TM)$coeffic
ients[,4][[2]], summary(Ersteplaviexp_TM)$coefficients[,4][[2]], summary(Ersteplav
ipro_TM)$coefficients[,4][[2]])

gama_n_p = c(summary(ERSTEADR_TM)$coefficients[,4][[3]], summary(OTP_TM)$coefficie
nts[,4][[3]], summary(ZB_TM)$coefficients[,4][[3]], summary(Raiffeisen_TM)$coeffic
ients[,4][[3]], summary(Ersteplaviexp_TM)$coefficients[,4][[3]], summary(Ersteplav
ipro_TM)$coefficients[,4][[3]])

delta_n_p = c(summary(ERSTEADR_TM)$coefficients[,4][[4]], summary(OTP_TM)$coeffici
ents[,4][[4]], summary(ZB_TM)$coefficients[,4][[4]], summary(Raiffeisen_TM)$coeffi
cients[,4][[4]], summary(Ersteplaviexp_TM)$coefficients[,4][[4]], summary(Erstepla
vipro_TM)$coefficients[,4][[4]])

r_squared_n = c(summary(ERSTEADR_TM)$r.squared, summary(OTP_TM)$r.squared, summary
(ZB_TM)$r.squared, summary(Raiffeisen_TM)$r.squared, summary(Ersteplaviexp_TM)$r.sq
uared, summary(Ersteplavipro_TM)$r.squared)

non_linear_regression_table = data.frame(row.names = fonds, alpha_n, alpha_n_p, bet
a_n, beta_n_p, gama_n, gama_n_p, delta_n, delta_n_p, r_squared_n)

save(non_linear_regression_table, file = "dataframes/non_linear_regression_table.Rd
a")

non_linear_regression_table
```

```
##                        alpha_n     alpha_n_p      beta_n       beta_n_p
## Erste Adriatic  0.0001418186 0.9937444942 0.43840594   2.228479e-57
## OTP             0.0316404061 0.0147958563 0.44319226 3.105719e-105
## ZABA            0.0210941511 0.1088575700 0.21159900   2.955757e-27
## Raiffeisen      0.0175518502 0.0012106649 0.14455536   3.099846e-68
## Erste plavi exp 0.0248401042 0.0000115719 0.21459708 5.156521e-126
## Erste plavi pro 0.0099774247 0.0009064769 0.03110089   2.053787e-12
##                         gama_n      gama_n_p        delta_n    delta_n_p
## Erste Adriatic  -0.0277983308 7.495669e-03  0.0038422908 0.006012364
## OTP             -0.0624052897 1.061587e-16  0.0020540436 0.040456893
## ZABA             0.0030931986 6.820616e-01 -0.0022800408 0.024900887
## Raiffeisen      -0.0105715293 6.871262e-04  0.0001893012 0.650857103
## Erste plavi exp -0.0112082477 5.600524e-04 -0.0007603670 0.081500332
## Erste plavi pro -0.0009484446 5.821701e-01 -0.0002563389 0.269062656
##                 r_squared_n
## Erste Adriatic   0.19418214
## OTP              0.31251041
## ZABA             0.07013011
## Raiffeisen       0.20026362
## Erste plavi exp  0.31911345
## Erste plavi pro  0.03163958
```

KS testovi za reziduale nelinearne regresije

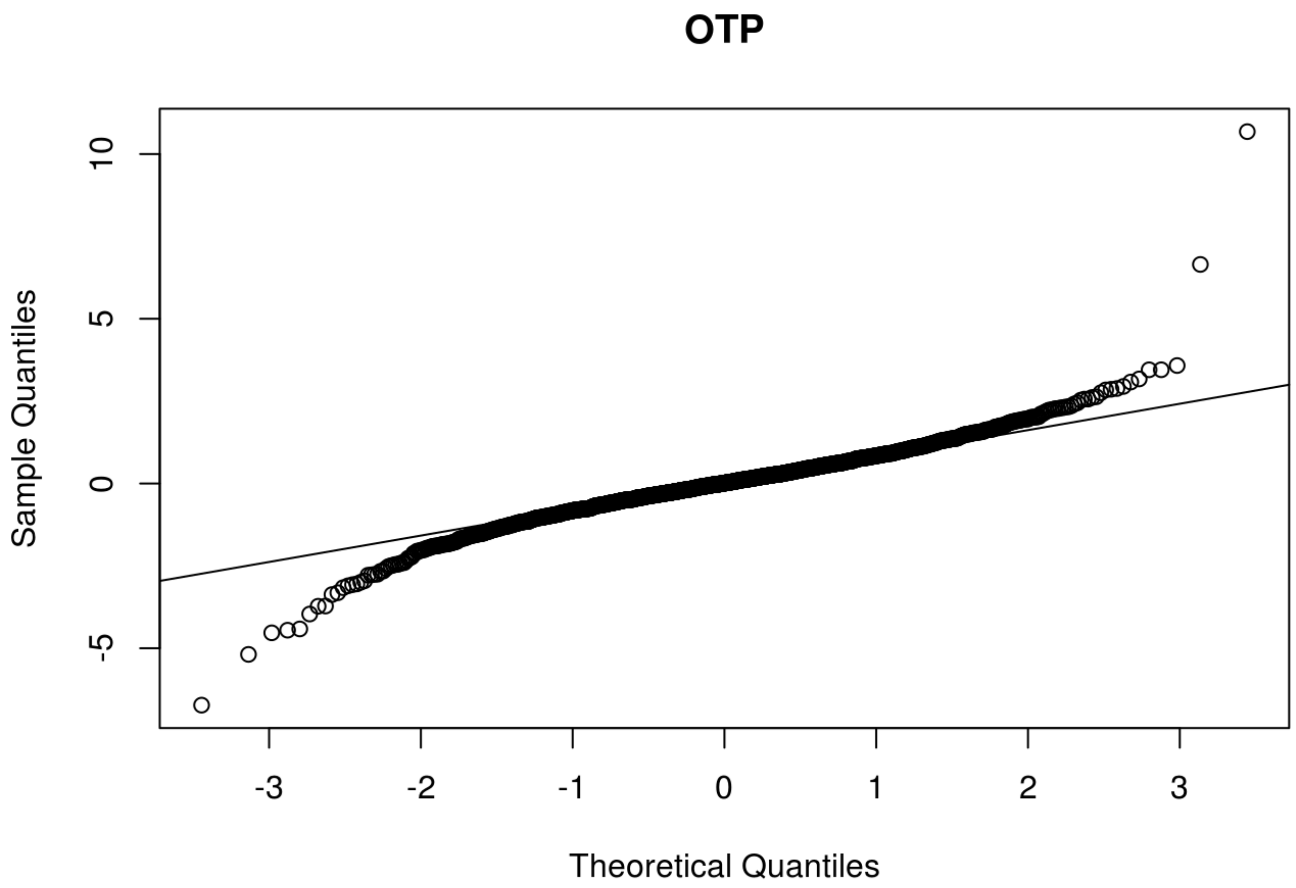
```
qqnorm(rstandard(ERSTEADR_TM), main = "Erste Adriatic")
qqline(rstandard(ERSTEADR_TM))
```



```r
ks.test(rstandard(ERSTEADR_TM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(ERSTEADR_TM)
## D = 0.15179, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

```r
qqnorm(rstandard(OTP_TM), main = "OTP")
qqline(rstandard(OTP_TM))
```

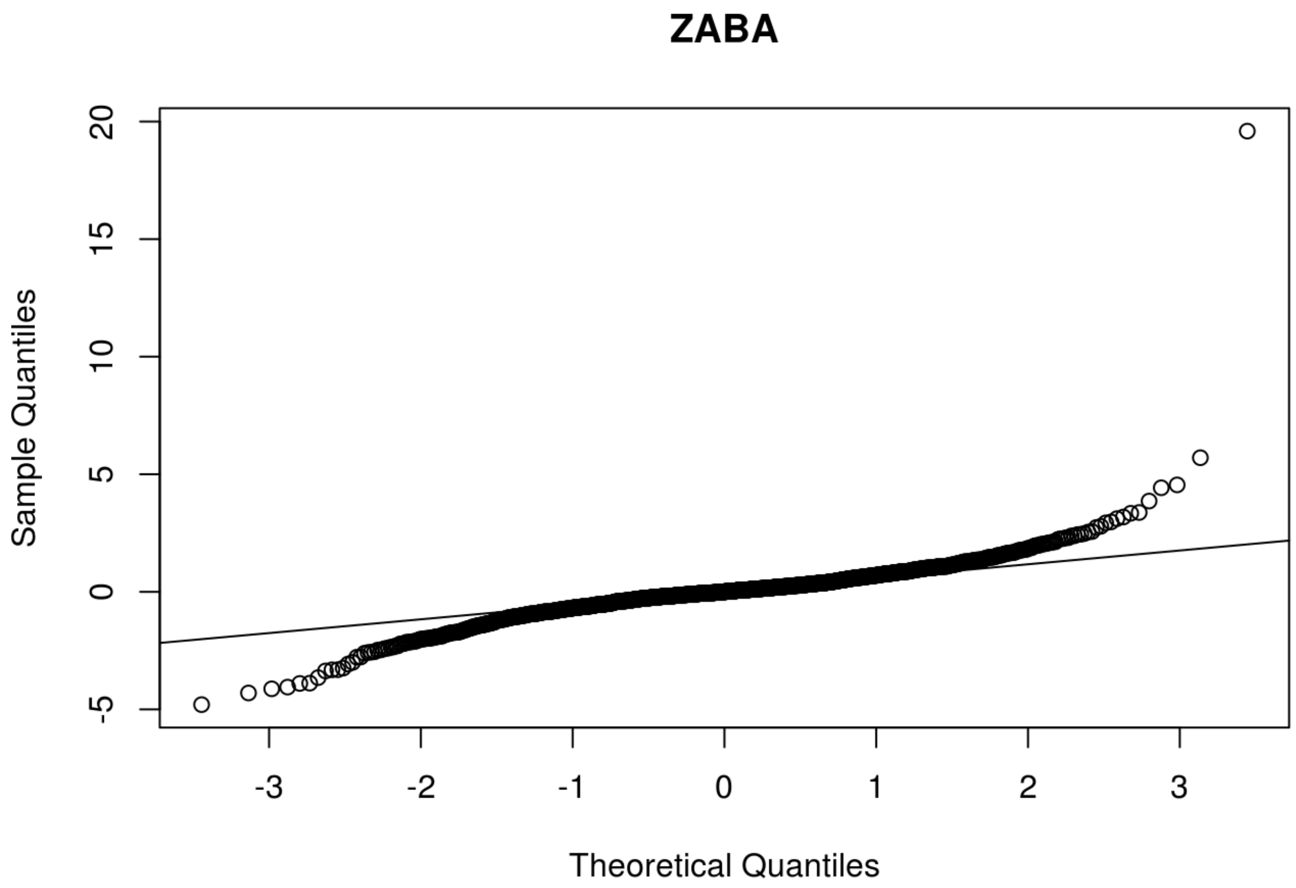



```r
ks.test(rstandard(OTP_TM), pnorm)
```

```
## Warning in ks.test(rstandard(OTP_TM), pnorm): ties should not be present
## for the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(OTP_TM)
## D = 0.052921, p-value = 0.0001113
## alternative hypothesis: two-sided
```

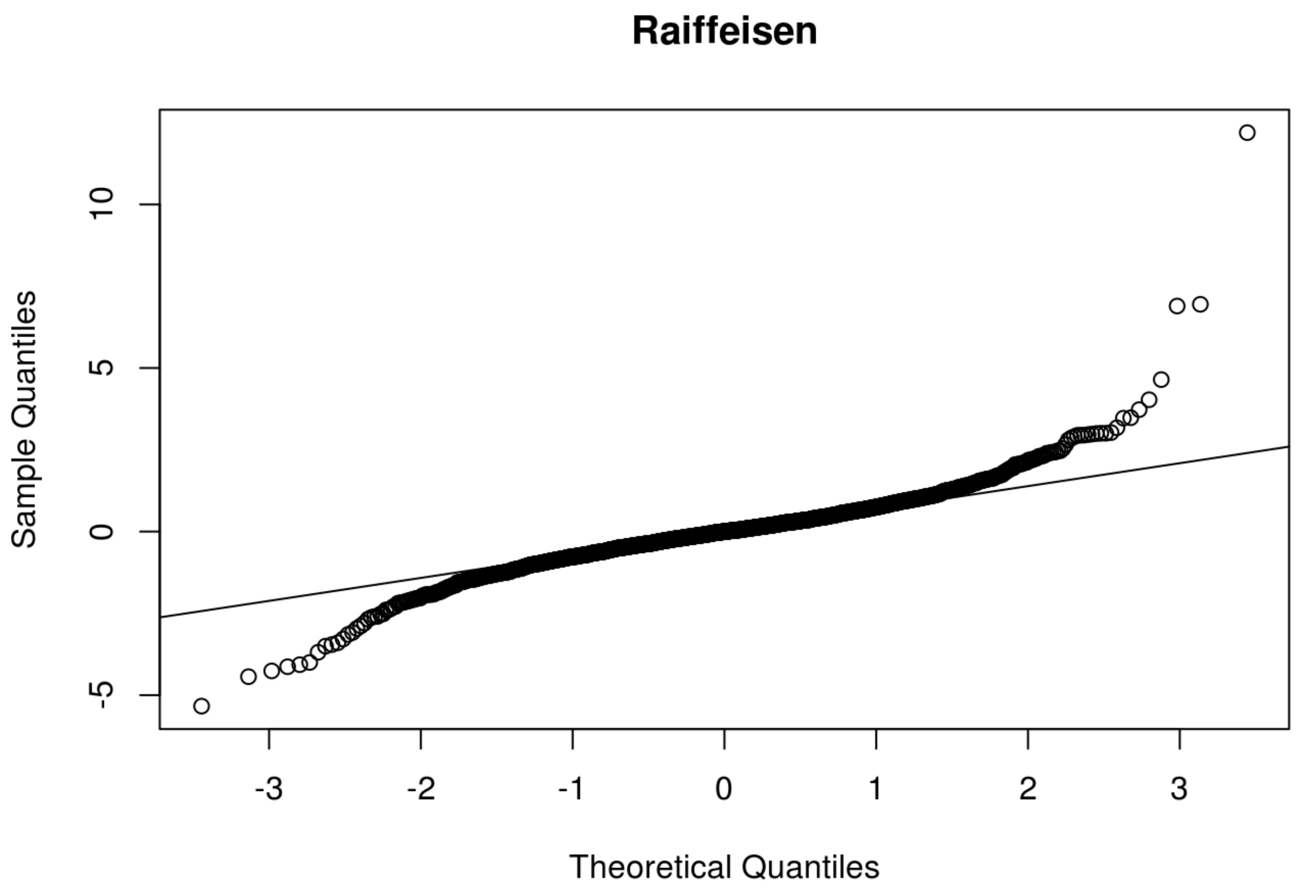
```r
qqnorm(rstandard(ZB_TM), main = "ZABA")
qqline(rstandard(ZB_TM))
```




```r
ks.test(rstandard(ZB_TM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(ZB_TM)
## D = 0.10032, p-value = 9.992e-16
## alternative hypothesis: two-sided
```

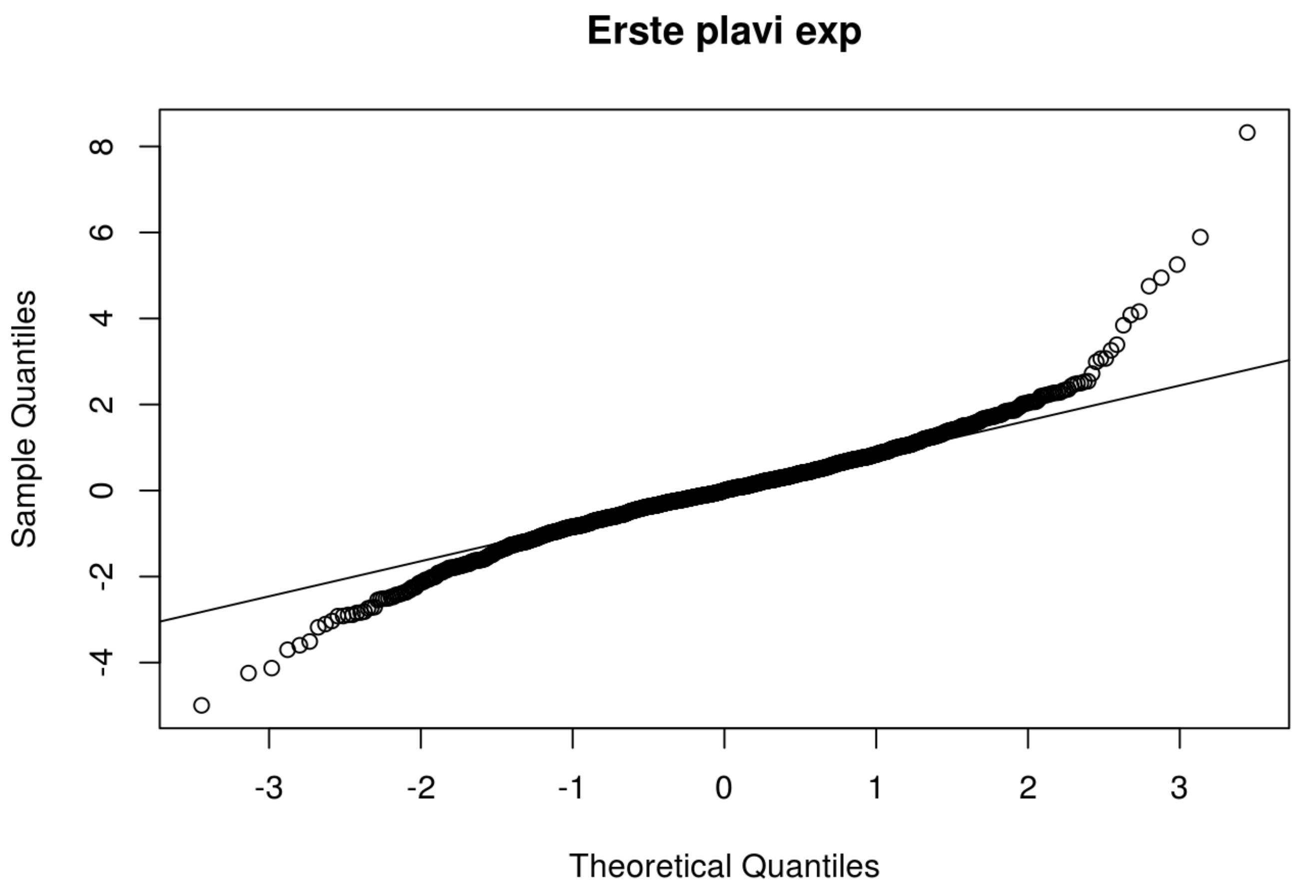
```r
qqnorm(rstandard(Raiffeisen_TM), main = "Raiffeisen")
qqline(rstandard(Raiffeisen_TM))
```




```
ks.test(rstandard(Raiffeisen_TM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(Raiffeisen_TM)
## D = 0.074234, p-value = 8.499e-09
## alternative hypothesis: two-sided
```

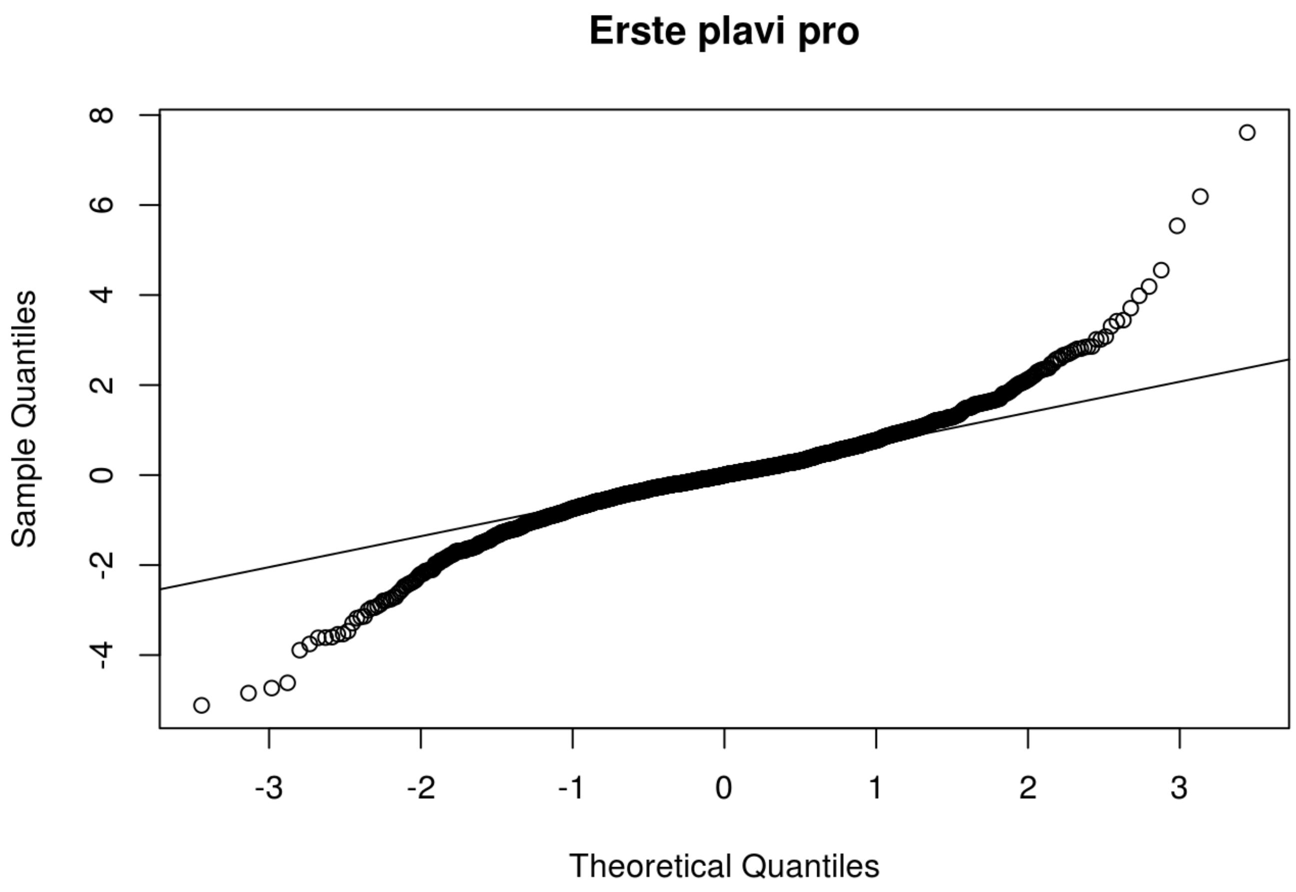
```
qqnorm(rstandard(Ersteplaviexp_TM), main = "Erste plavi exp")
qqline(rstandard(Ersteplaviexp_TM))
```




```
ks.test(rstandard(Ersteplaviexp_TM), "pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(Ersteplaviexp_TM)
## D = 0.048651, p-value = 0.0005074
## alternative hypothesis: two-sided
```

```
qqnorm(rstandard(Ersteplavipro_TM), main = "Erste plavi pro")
qqline(rstandard(Ersteplavipro_TM))
```




```
ks.test(rstandard(Ersteplavipro_TM), "pnorm")
```

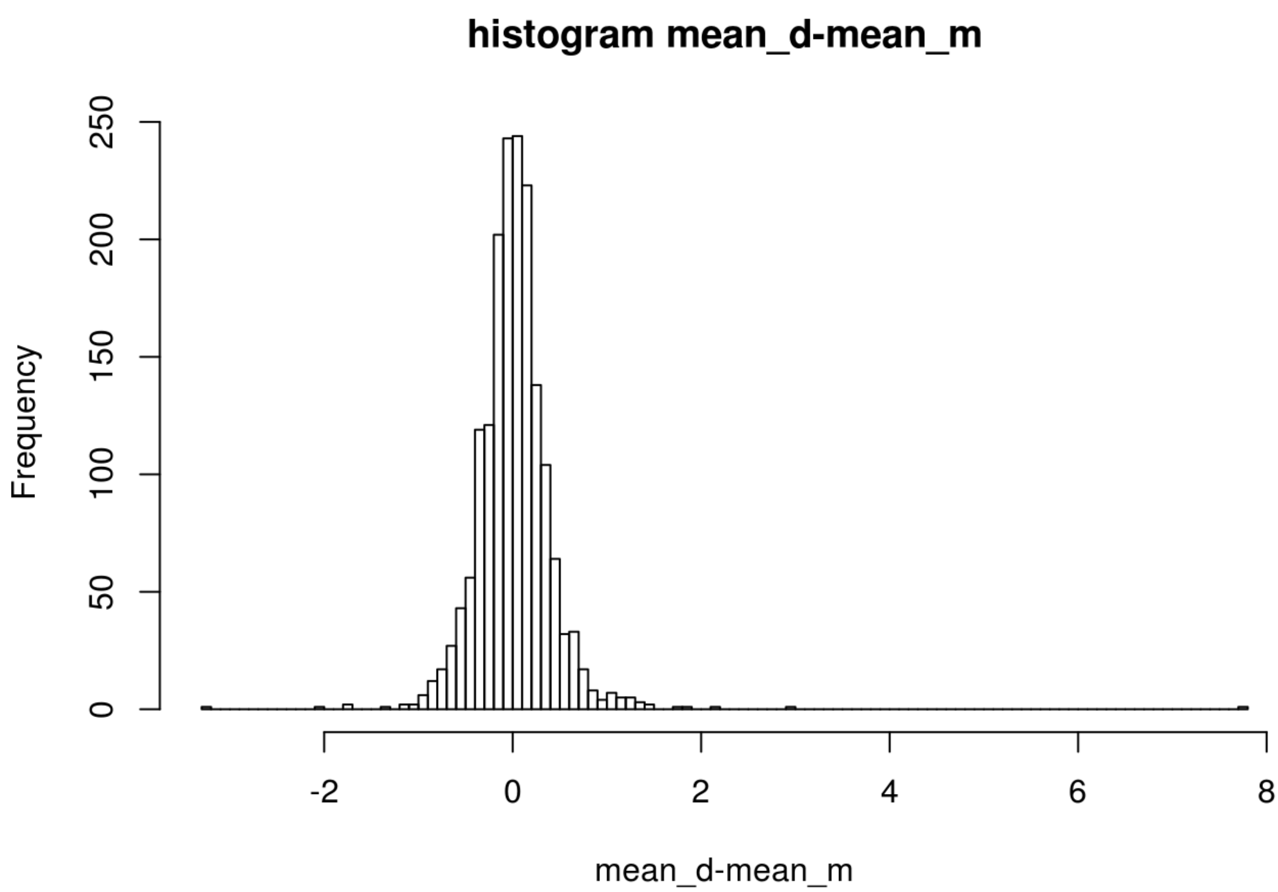
```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(Ersteplavipro_TM)
## D = 0.079551, p-value = 4.867e-10
## alternative hypothesis: two-sided
```

Nakon regresijske analize, provodimo analizu isplativosti ulaganja provođenjem T-testa uparenih podataka (po danima) između dvije velike grupe podataka - zajedno svih mirovinskih i zajedno svih dioničkih fondova. Rezultati potvrđuju ispravnost dosadašnjih odluka. Formalno, zbog velike p-vrijednosti, nismo u mogućnosti odbiti hipotezu H0 da postoji razlika između srednjih prinosa dioničkih i mirovinskih fondova, no neformalno možemo zamijetiti da je povrat mirovinskih fondova najčešće veći od povrata dioničkih. Uz poznatu činjenicu da je rizik ulaganja u mirovinske fondove manji (što pokazuju analizirane mjere raspršenosti), jednoznačna je preporuka da je investicija u mirovinske fondove dugoročno bolja odluka. Iz tih zaključaka uvijek možemo istaknuti specifični izdvojeni slučaj kao što je ZBActiv (dionički fond) koji je istaknut kao preporučljiv za ulaganje.

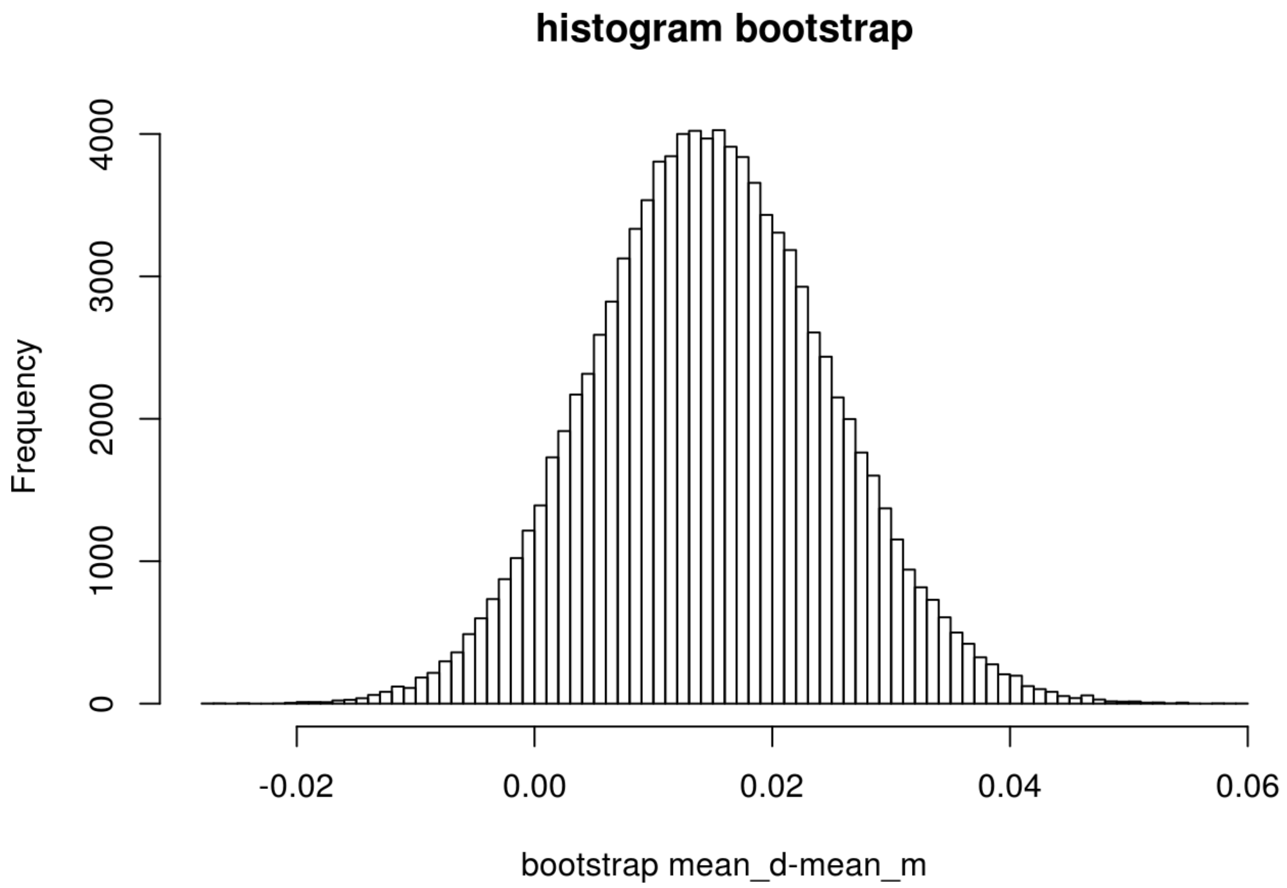
```r
#T-test za usporedbu mirovisnkih i dioničkih fondova kao jedinstvenih grupa podatak
a
library("bootstrap")

mean_d=(ERSTEADR_AN + OTP_AN + ZB_AN)/3
mean_m=(Raiffeisen_AN + Ersteplaviexp_AN + Ersteplavipro_AN)/3
razlika=mean_m-mean_d

#T-test na uparenim podacima
t.test(mean_m, mean_d, mu=0, alt="two.sided", paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  mean_m and mean_d
## t = 1.4921, df = 1748, p-value = 0.1358
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.004696166  0.034565274
## sample estimates:
## mean of the differences
##              0.01493455
```

```r
hist(razlika, main="histogram mean_d-mean_m",breaks = 100,
xlab = "mean_d-mean_m", ylab = "Frequency" )
```



```r
#bootstrap na uparenim podacima

theta=function(razlika){mean(razlika)}
podaci_bootstrap=bootstrap(razlika, 100000, theta)

hist(podaci_bootstrap$thetastar, main="histogram bootstrap",breaks = 100,
xlab = "bootstrap mean_d-mean_m", ylab = "Frequency" )
```




```r
quantile(podaci_bootstrap$thetastar, c(0.025, 0.975))
```

```
##         2.5%        97.5%
## -0.004221517  0.034977347
```

Projekt završavamo bootstrap analizom uparenih podataka dvaju grupa - mirovinskih i dioničkih fondova čime ostvarujemo bolju procjenu distribucije očekivane srednje vrijednosti. Interval povjerenja dobiven bootstrap metodom je vrlo sličan intervalu povjerenja koji je dobiven T-testom što je samo još jedna potvrda robusnosti T-testa i ispravnosti postavljenih zaključaka.