**Rainfall:**

data=read.csv(file.choose())

summary(data$JAN)

summary(data$FEB)

summary(data$MAR)

summary(data$APR)

summary(data$MAY)

summary(data$JUN)

summary(data$JUL)

summary(data$AUG)

summary(data$SEP)

summary(data$OCT)

summary(data$NOV)

summary(data$DEC)

summary(data$ANNUAL)

plot(data$YEAR,data$ANNUAL,type="l",main="Annual Rainfall in each year",xlab ="Years",ylab ="ANNUAL RAINFALL",col="blue")

**Temperature:**

z=read.csv(file.choose())

summary(z$JAN)

summary(z$FEB)

summary(z$MAR)

summary(z$APR)

summary(z$MAY)

summary(z$JUN)

summary(z$JUL)

summary(z$AUG)

summary(z$SEP)

summary(z$OCT)

summary(z$NOV)

summary(z$DEC)

summary(z$AVERAGE)

plot(z$YEAR,z$AVERAGE,type="l",main="AVERAGE TEMPERATURE VARIATION(1901-2017)",xlab ="Years",ylab ="AVERAGE TEMERATURE",col="red")

**Correlation:**

x=c(24.5766,25.0125,24.7675,24.6125,25.1125,25.189,24.71083,24.6783,24.82,24.7325,24.9125)

y=c(1456.431,1406.439,1514.067,1367.394,1159,1468.194,1392.956,1209.014,1475.172,1244.336,1280.583)

var(temp)

var(rainfall)

r=var(temp,rainfall)

r

cor.test(x,y,method="pearson")

cor.test(x,y,method="spearman")

cor.test(x,y,method="kendall")

analysis=data.frame(temp,rainfall)

analysis

scatter.smooth(x=analysis$rainfall, y=analysis$temp, main="Temperature vs Rainfall",xlab="Rainfall(in mm)",ylab = "Temperature(in degree Celsius)")

**Robust Regression:**

#create data

df <- data.frame(x1=c(24.5766,25.0125,24.7675,24.6125,25.1125,25.189,24.71083,24.6783,24.82,24.7325,24.9125),

x2=c(1456.431,1406.439,1514.067,1367.394,1159,1468.194,1392.956,1209.014,1475.172,1244.336,1280.583),

y=c(2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014,2015))

#view first six rows of data

head(df)

#fit ordinary least squares regression model

ols <- lm(y~x1+x2, data=df)

#create plot of y-values vs. standardized residuals

plot(df$y, rstandard(ols), ylab='Standardized Residuals', xlab='y')

abline(h=0)

library(MASS)

#fit robust regression model

robust <- rlm(y~x1+x2, data=df)

summary(ols)$sigma

summary(robust)$sigma

**Mann Whitney U test:**

MWU <- read.csv(file.choose())

MWU

boxplot(MWU$JAN, MWU$FEB, MWU$MAR, MWU$APR, MWU$MAY, MWU$JUN, MWU$JULY, MWU$AUG, MWU$SEP, MWU$OCT, MWU$NOV, MWU$DEC)

wilcox.test(MWU$JAN, MWU$FEB, MWU$MAR, MWU$APR, MWU$MAY, MWU$JUN, MWU$JULY, MWU$AUG, MWU$SEP, MWU$OCT, MWU$NOV, MWU$DEC, conf.int=T, conf.level=0.95, exact=F)

**OUTPUT**

> z=read.csv(file.choose())

> z=read.csv(file.choose())

> summary(z$JAN)

Min. 1st Qu. Median Mean 3rd Qu. Max.

17.25 18.07 18.35 18.42 18.72 20.92

> summary(z$FEB)

Min. 1st Qu. Median Mean 3rd Qu. Max.

17.79 19.47 19.99 20.14 20.55 23.58

> summary(z$MAR)

Min. 1st Qu. Median Mean 3rd Qu. Max.

21.78 22.84 23.33 23.43 23.93 26.61

> summary(z$APR)

Min. 1st Qu. Median Mean 3rd Qu. Max.

24.84 26.00 26.54 26.51 26.91 29.56

> summary(z$MAY)

Min. 1st Qu. Median Mean 3rd Qu. Max.

26.97 27.95 28.34 28.39 28.82 30.78

> summary(z$JUN)

Min. 1st Qu. Median Mean 3rd Qu. Max.

27.33 28.02 28.25 28.30 28.58 29.88

> summary(z$JUL)

Min. 1st Qu. Median Mean 3rd Qu. Max.

26.48 27.15 27.34 27.37 27.53 28.47

> summary(z$AUG)

Min. 1st Qu. Median Mean 3rd Qu. Max.

26.21 26.73 26.90 26.94 27.15 28.17

> summary(z$SEP)

Min. 1st Qu. Median Mean 3rd Qu. Max.

25.47 26.11 26.31 26.34 26.50 28.11

> summary(z$OCT)

Min. 1st Qu. Median Mean 3rd Qu. Max.

23.52 24.39 24.67 24.74 24.92 27.24

> summary(z$NOV)

Min. 1st Qu. Median Mean 3rd Qu. Max.

20.59 21.32 21.72 21.77 22.15 23.92

> summary(z$DEC)

Min. 1st Qu. Median Mean 3rd Qu. Max.

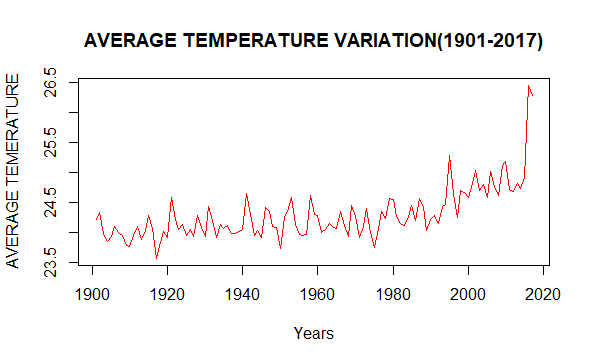
17.98 18.78 19.18 19.17 19.55 21.89

> summary(z$AVERAGE)

Min. 1st Qu. Median Mean 3rd Qu. Max.

23.57 24.01 24.21 24.29 24.46 26.45

> plot(z$YEAR,z$AVERAGE,type="l",main="AVERAGE TEMPERATURE VARIATION(1901-2017)",xlab ="Years",ylab ="AVERAGE TEMERATURE",col="red")



**REGRESSION MODEL**

INPUT

temp=c(24.5766,25.0125,24.7675,24.6125,25.1125,25.189,24.71083,24.6783,24.82,24.7325,24.9125)

rainfall=c(1456.431,1406.439,1514.067,1367.394,1159,1468.194,1392.956,1209.014,1475.172,1244.336,1280.583)

var(temp)

var(rainfall)

r=var(temp,rainfall)

r

cor.test(x,y,method="pearson")

cor.test(x,y,method="spearman")

cor.test(x,y,method="kendall")

analysis=data.frame(temp,rainfall)

analysis

scatter.smooth(x=analysis$rainfall, y=analysis$temp, main="Temperature vs Rainfall",xlab="Rainfall(in mm)",ylab = "Temperature(in degree Celsius)")

OUTPUT

> temp=c(24.5766,25.0125,24.7675,24.6125,25.1125,25.189,24.71083,24.6783,24.82,24.7325,24.9125)

> rainfall=c(1456.431,1406.439,1514.067,1367.394,1159,1468.194,1392.956,1209.014,1475.172,1244.336,1280.583)

> var(temp)

[1] 0.04110419

> var(rainfall)

[1] 14416.97

> r=var(temp,rainfall)

> r

[1] -2.25814

> cor.test(x,y,method="pearson")

Pearson's product-moment correlation

data: x and y

t = -0.27949, df = 9, p-value = 0.7862

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

-0.6561265 0.5369943

sample estimates:

cor

-0.09276215

> cor.test(x,y,method="spearman")

Spearman's rank correlation rho

data: x and y

S = 206, p-value = 0.8601

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

0.06363636

> cor.test(x,y,method="kendall")

Kendall's rank correlation tau

data: x and y

T = 28, p-value = 1

alternative hypothesis: true tau is not equal to 0

sample estimates:

tau

0.01818182

> analysis=data.frame(temp,rainfall)

> analysis

temp rainfall

1 24.57660 1456.431

2 25.01250 1406.439

3 24.76750 1514.067

4 24.61250 1367.394

5 25.11250 1159.000

6 25.18900 1468.194

7 24.71083 1392.956

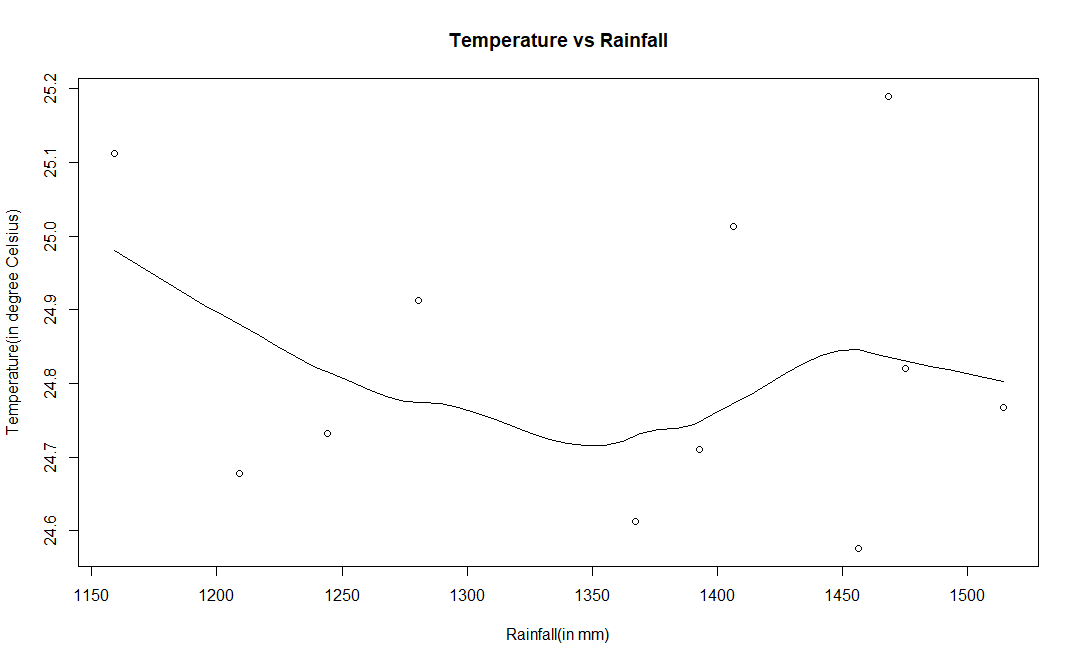
8 24.67830 1209.014

9 24.82000 1475.172

10 24.73250 1244.336

11 24.91250 1280.583

> scatter.smooth(x=analysis$rainfall, y=analysis$temp, main="Temperature vs Rainfall",xlab="Rainfall(in mm)",ylab = "Temperature(in degree Celsius)")



ROBUST REGRESSION

#create data

df <- data.frame(x1=c(24.5766,25.0125,24.7675,24.6125,25.1125,25.189,24.71083,24.6783,24.82,24.7325,24.9125),

x2=c(1456.431,1406.439,1514.067,1367.394,1159,1468.194,1392.956,1209.014,1475.172,1244.336,1280.583),

y=c(2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014,2015))

#view first six rows of data

head(df)

#fit ordinary least squares regression model

ols <- lm(y~x1+x2, data=df)

#create plot of y-values vs. standardized residuals

plot(df$y, rstandard(ols), ylab='Standardized Residuals', xlab='y')

abline(h=0)

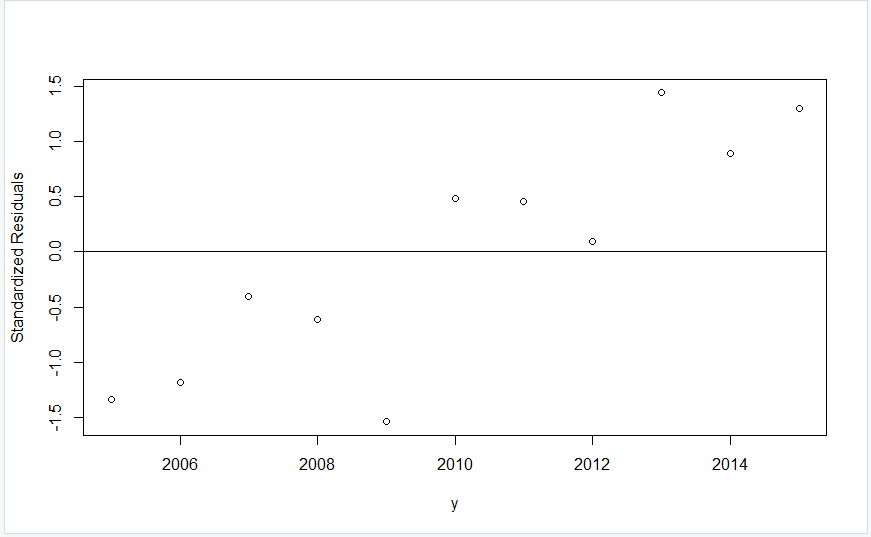
library(MASS)

#fit robust regression model

robust <- rlm(y~x1+x2, data=df)

summary(ols)$sigma

summary(robust)$sigma



Mann Whitney U Test

MWU <- read.csv(file.choose())

MWU

boxplot(MWU$JAN, MWU$FEB, MWU$MAR, MWU$APR, MWU$MAY, MWU$JUN, MWU$JULY, MWU$AUG, MWU$SEP, MWU$OCT, MWU$NOV, MWU$DEC)

wilcox.test(MWU$JAN, MWU$FEB, MWU$MAR, MWU$APR, MWU$MAY, MWU$JUN, MWU$JULY, MWU$AUG, MWU$SEP, MWU$OCT, MWU$NOV, MWU$DEC, conf.int=T, conf.level=0.95, exact=F)

