

# BLM3590

## Statistical Data Analysis

### Homework

**COURSE NAME:** İstatistiksel Veri Analizi

**COURSE GROUP:** 1

**INSTRUCTOR NAME:** Prof. Dr. Nizamettin Aydın

**STUDENT ID:** 16011038

**STUDENT NAME AND SURNAME:** Talha Bacak

**E-MAIL:** talhadhh@gmail.com

## KODLAR

### T1 :

```
library(readxl)

SdA_HW <- read_excel("D:/MASA/DERS/3.sınıf-1/istatistiksel veri analizi/proje/SdA-HW.xls")

View(SdA_HW)
```

### T2 :

```
SdA_HW$type=as.factor(SdA_HW$type)

SdA_HW <- SdA_HW[-5]

n <- which(is.na(SdA_HW$rrt))

SdA_HW <- SdA_HW[-n,]

n <- which(is.na(SdA_HW$frt))

SdA_HW <- SdA_HW[-n,]

View(SdA_HW)
```

### T3 :

```
type1 <- subset(SdA_HW,subset = (type ==1) )

type2 <- subset(SdA_HW,subset = (type ==2) )

View(type1)

View(type2)
```

```
fivenum(type1)

summary(type1)

fivenum(type2)

summary(type2)
```

### T4 :

```
boxplot(type1[,2:6], main = "type = 1 için")

boxplot(type2[,2:6], main = "type = 2 için")
```

### **T5 :**

```
hist(type1$tpthrt, main = "type = 1 için")
hist(type1$pkthrt, main = "type = 1 için")
hist(type1$dfdrtr, main = "type = 1 için")
hist(type1$rrt, main = "type = 1 için")
hist(type1$frt, main = "type = 1 için")
```

```
hist(type2$tpthrt, main = "type = 2 için")
hist(type2$pkthrt, main = "type = 2 için")
hist(type2$dfdrtr, main = "type = 2 için")
hist(type2$rrt, main = "type = 2 için")
hist(type2$frt, main = "type = 2 için")
```

### **T6 :**

```
normalize <- function(x){
  return ((x - min(x)) / (max(x) - min(x)))
}
```

```
type1$tpthrt <- normalize(type1$tpthrt)
type1$pkthrt <- normalize(type1$pkthrt)
type1$dfdrtr <- normalize(type1$dfdrtr)
type1$rrt <- normalize(type1$rrt)
type1$frt <- normalize(type1$frt)
```

```
type2$tpthrt <- normalize(type2$tpthrt)
type2$pkthrt <- normalize(type2$pkthrt)
type2$dfdrtr <- normalize(type2$dfdrtr)
type2$rrt <- normalize(type2$rrt)
type2$frt <- normalize(type2$frt)
```

```

plot(type1$tpthrt, type = "o", col = "red", ylab = "tpthrt", main = " red -> type=1   blue -> type=2")
lines(type2$tpthrt, type = "o", col = "blue")
plot(type1$pkthrt, type = "o", col = "red", ylab = "pkthrt", main = " red -> type=1   blue -> type=2")
lines(type2$pkthrt, type = "o", col = "blue")
plot(type1$dfdrft, type = "o", col = "red", ylab = "dfdrft", main = " red -> type=1   blue -> type=2")
lines(type2$dfdrft, type = "o", col = "blue")
plot(type1$rrt, type = "o", col = "red", ylab = "rrt", main = " red -> type=1   blue -> type=2")
lines(type2$rrt, type = "o", col = "blue")
plot(type1$frrt, type = "o", col = "red", ylab = "frrt", main = " red -> type=1   blue -> type=2")
lines(type2$frrt, type = "o", col = "blue")

```

## T7 :

```

cor(type1$tpthrt,type1$pkthrt, method = "kendall")
cor(type2$tpthrt,type2$pkthrt)
cor(type1[1:64,2],type2$tpthrt)

```

## ÇIKTILAR

### T1 Read Excell :

	type	tpthrt	pkthrt	dfdrft	time	rrt	frrt
1	1	14.611816	7.155748	20.501258	NA	8.527265	17.054550
2	1	0.000000	0.000000	0.000000	NA	NA	NA
3	1	10.387063	3.173866	21.038962	NA	4.538626	7.564388
4	1	14.494864	5.891724	28.135395	NA	8.425162	6.017975
5	1	6.899807	1.335272	25.239697	NA	3.182401	4.773602
6	1	0.000000	0.000000	0.000000	NA	NA	NA
7	1	14.929119	5.168594	16.150102	NA	2.639675	9.238854
8	1	19.474133	7.977139	33.421967	NA	3.001923	8.148078
9	1	27.051413	11.370301	12.855528	NA	5.081104	0.967829
10	1	24.253508	10.279815	28.548077	NA	2.722247	3.340940
11	1	6.834928	2.835058	1.557392	NA	10.135340	10.135340
12	1	3.300711	0.332214	12.665563	NA	2.375332	1.187666
13	1	15.991371	2.930654	21.597374	NA	3.492362	0.952463
14	1	0.000000	0.000000	0.000000	NA	NA	NA
15	1	0.000000	0.000000	0.000000	NA	NA	NA
16	1	11.061486	4.428561	20.758657	NA	15.832120	6.332840
17	1	28.863992	19.906871	26.068455	NA	9.488942	10.166720
18	1	0.000000	0.000000	0.000000	NA	NA	NA

Showing 1 to 20 of 200 entries, 7 total columns

## T2 Data process:

deneme1.R × SdA\_HW ×

Filter

	type	tpthrt	pkthrt	dfdrtr	rrt	frt
1	1	14.611816	7.155748	20.501258	8.527265	17.054550
2	1	10.387063	3.173866	21.038962	4.538626	7.564388
3	1	14.494864	5.891724	28.135395	8.425162	6.017975
4	1	6.899807	1.335272	25.239697	3.182401	4.773602
5	1	14.929119	5.168594	16.150102	2.639675	9.238854
6	1	19.474133	7.977139	33.421967	3.001923	8.148078
7	1	27.051413	11.370301	12.855528	5.081104	0.967829
8	1	24.253508	10.279815	28.548077	2.722247	3.340940
9	1	6.834928	2.835058	1.557392	10.135340	10.135340
10	1	3.300711	0.332214	12.665563	2.375332	1.187666
11	1	15.991371	2.930654	21.597374	3.492362	0.952463
12	1	11.061486	4.428561	20.758657	15.832120	6.332840
13	1	28.863992	19.906871	26.068455	9.488942	10.166720
14	1	14.339388	6.110674	14.108886	7.281885	8.738260
15	1	14.979864	6.487209	19.371903	4.216685	15.461200
16	1	27.837549	18.991576	32.403461	8.486661	9.699269
17	1	11.757704	6.918127	9.102406	16.488220	24.732330
18	1	19.641946	8.269784	36.352688	4.223497	4.223497

Showing 1 to 20 of 143 entries, 6 total columns

## T3 Five-number Data Summary :

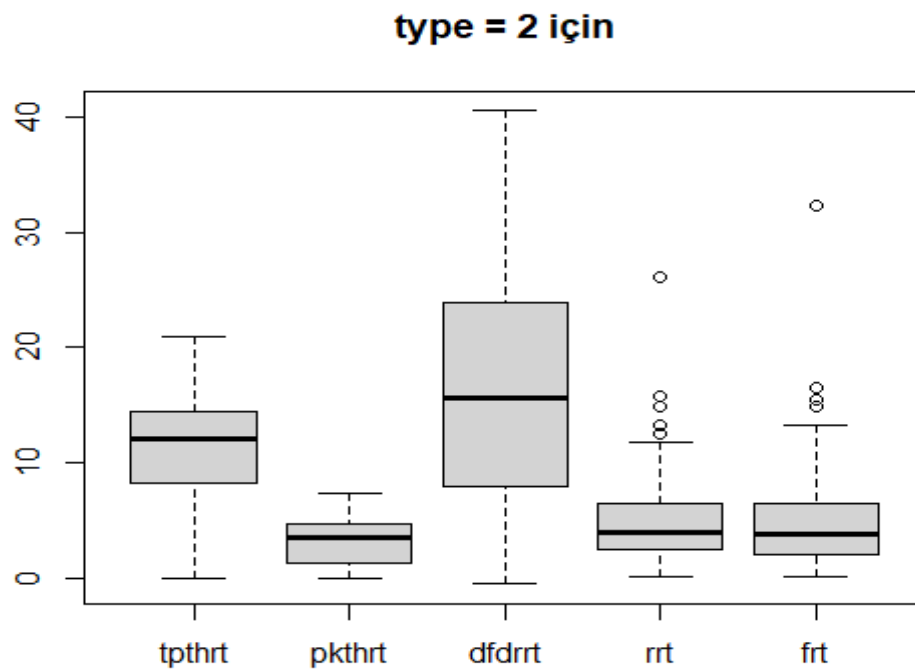
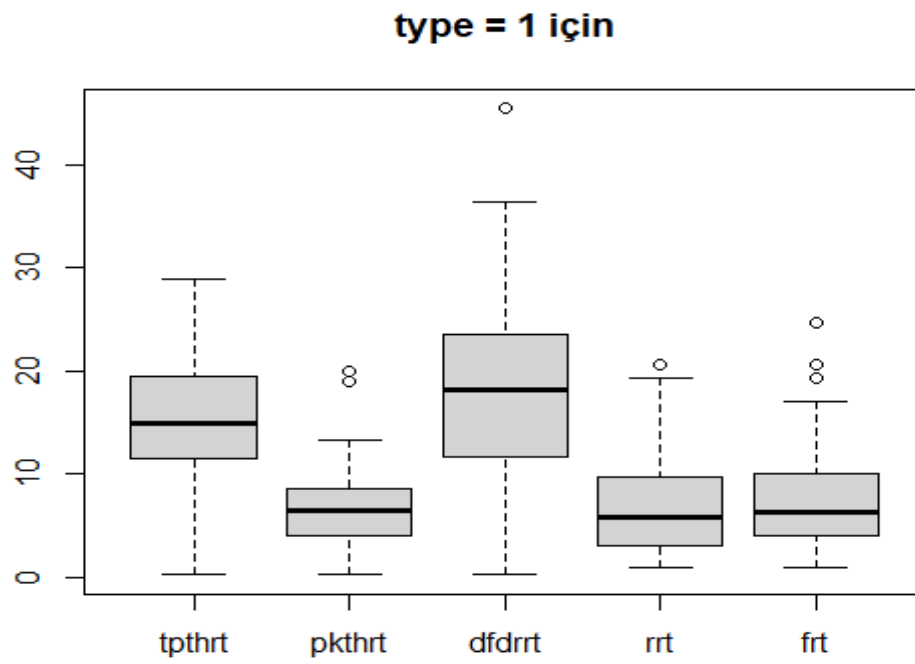
Type=1 :

type	tpthrt	pkthrt	dfdrtr	rrt	frt
1:79	Min. : 0.3364	Min. : 0.3322	Min. : 0.2601	Min. : 1.050	Min. : 0.9525
2: 0	1st Qu.:11.5141	1st Qu.: 3.9943	1st Qu.:11.6244	1st Qu.: 3.119	1st Qu.: 4.0453
	Median :14.9291	Median : 6.4622	Median :18.1353	Median : 5.916	Median : 6.3328
	Mean :15.5227	Mean : 6.5897	Mean :17.9784	Mean : 7.199	Mean : 7.5205
	3rd Qu.:19.5589	3rd Qu.: 8.5579	3rd Qu.:23.5388	3rd Qu.: 9.811	3rd Qu.:10.0627
	Max. :28.8640	Max. :19.9069	Max. :45.4140	Max. :20.610	Max. :24.7323

Type=2 :

type	tpthrt	pkthrt	dfdrtr	rrt	frt
1: 0	Min. : 0.008013	Min. :0.008013	Min. : -0.5604	Min. : 0.05729	Min. : 0.05729
2:64	1st Qu.: 8.253668	1st Qu.:1.357415	1st Qu.: 8.0201	1st Qu.: 2.49535	1st Qu.: 2.00146
	Median :12.101506	Median :3.453576	Median :15.6663	Median : 3.98919	Median : 3.82374
	Mean :11.412612	Mean : 3.377626	Mean :16.4264	Mean : 5.24958	Mean : 5.19507
	3rd Qu.:14.331232	3rd Qu.:4.645590	3rd Qu.:23.9202	3rd Qu.: 6.23552	3rd Qu.: 6.49463
	Max. :20.963047	Max. :7.323946	Max. :40.5450	Max. :26.12703	Max. :32.28238

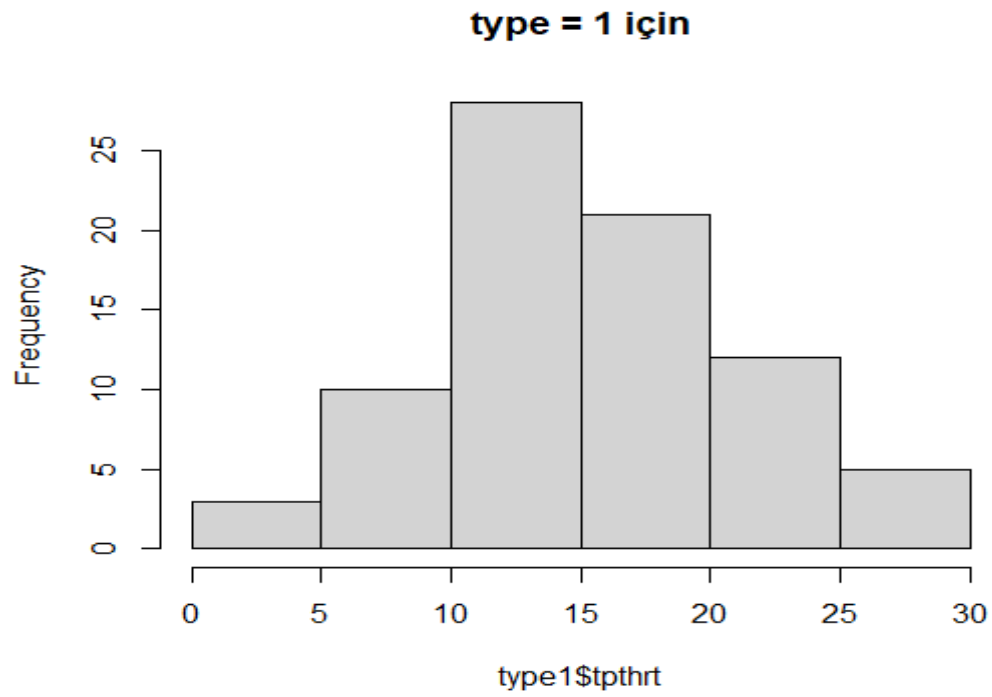
#### T4 Boxplot:



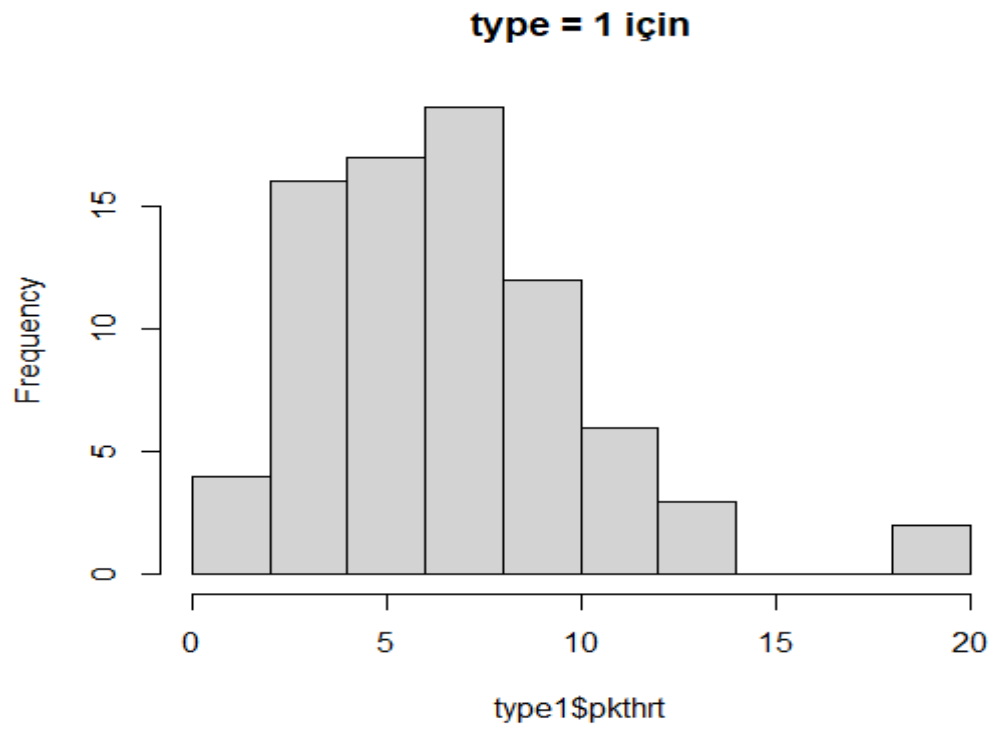
## T5 Histogram :

### A) Type=1

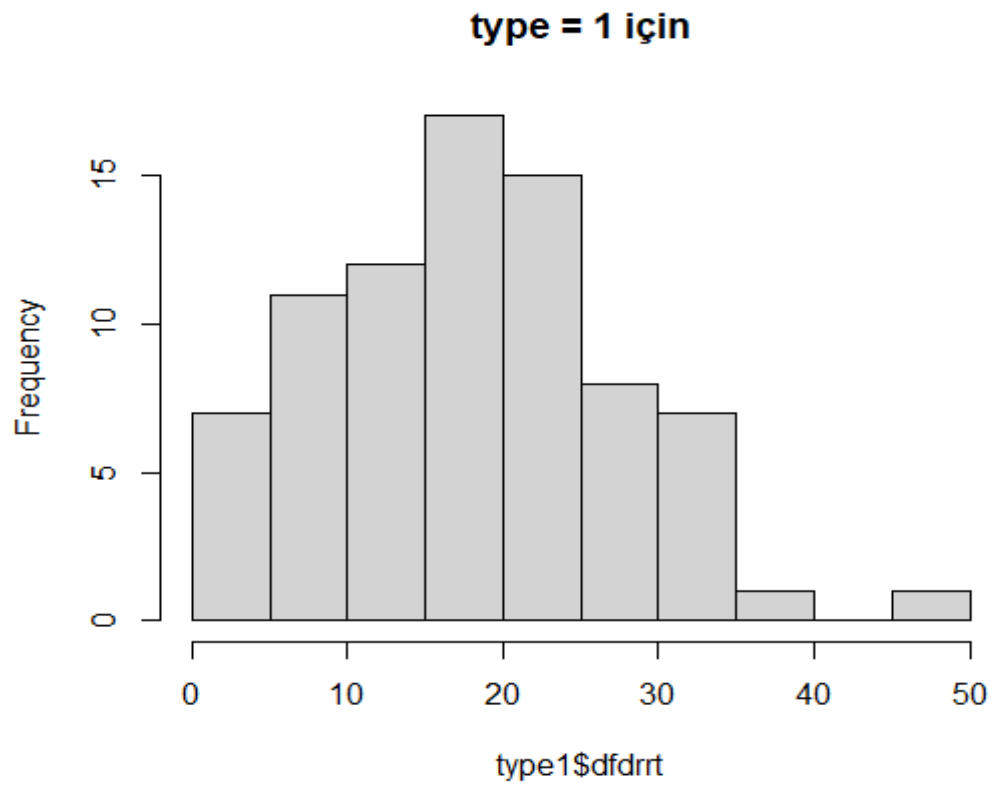
#### 1) Normal Distribution



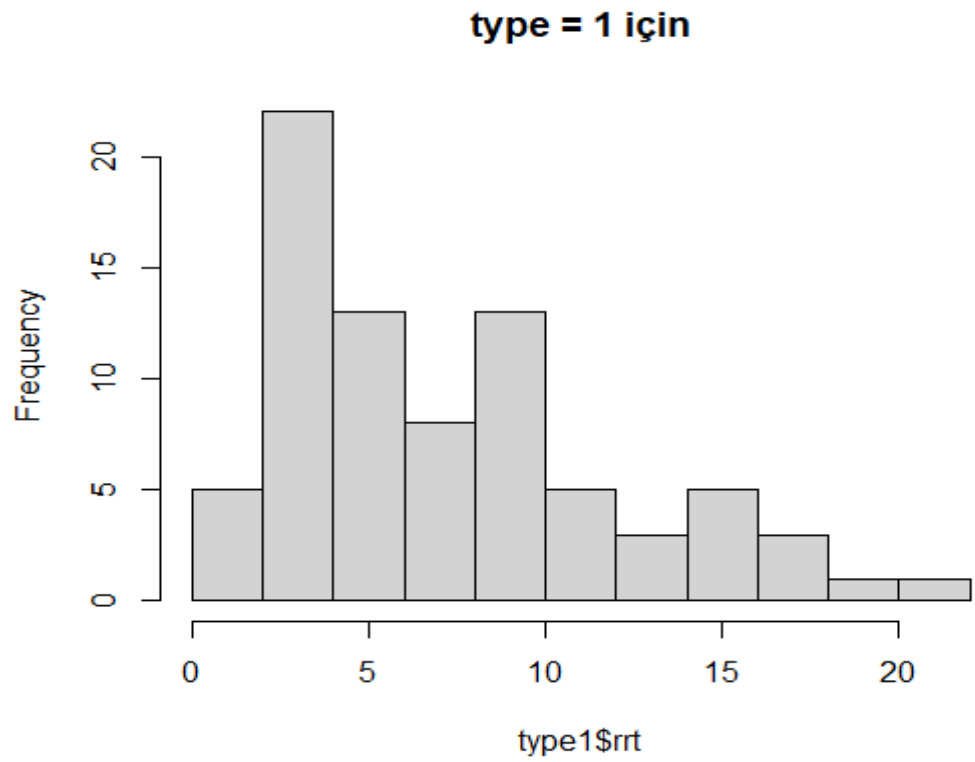
#### 2) Non-normal Distribution



### 3) Non-normal Distribution

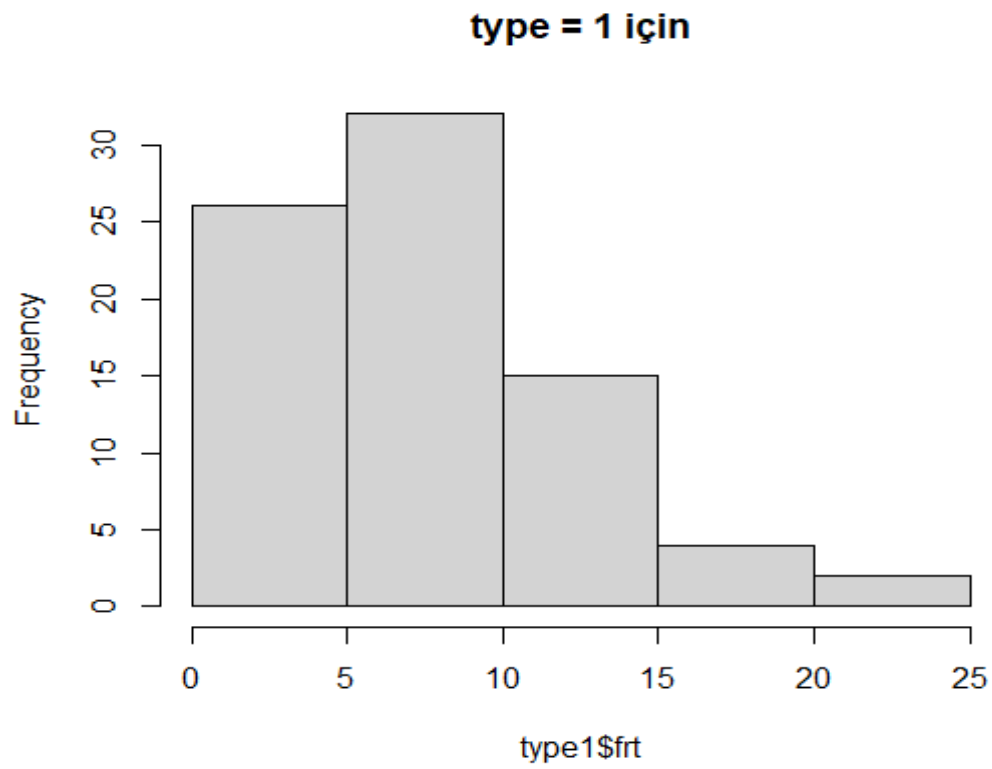


### 4) Normal Distribution



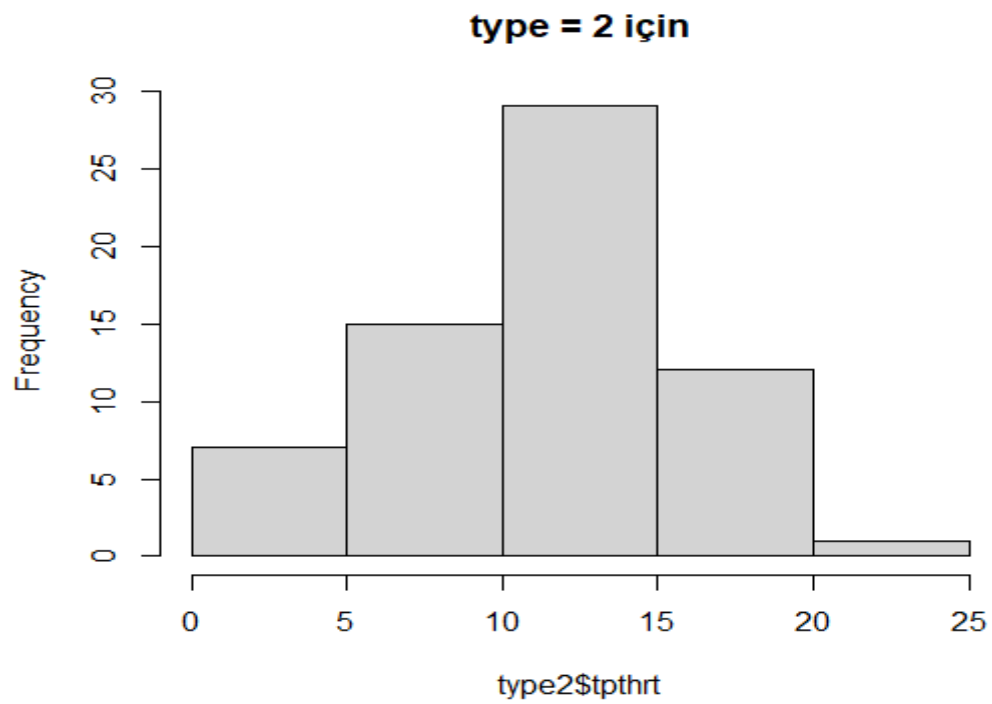


5) Normal Distribution

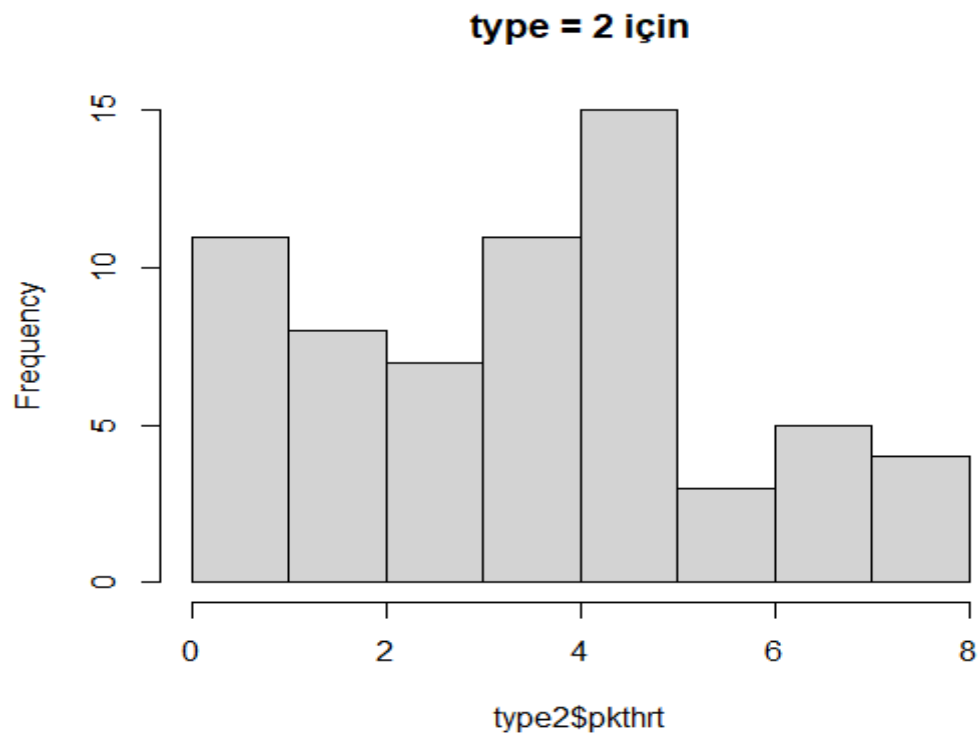


**B) Type=2**

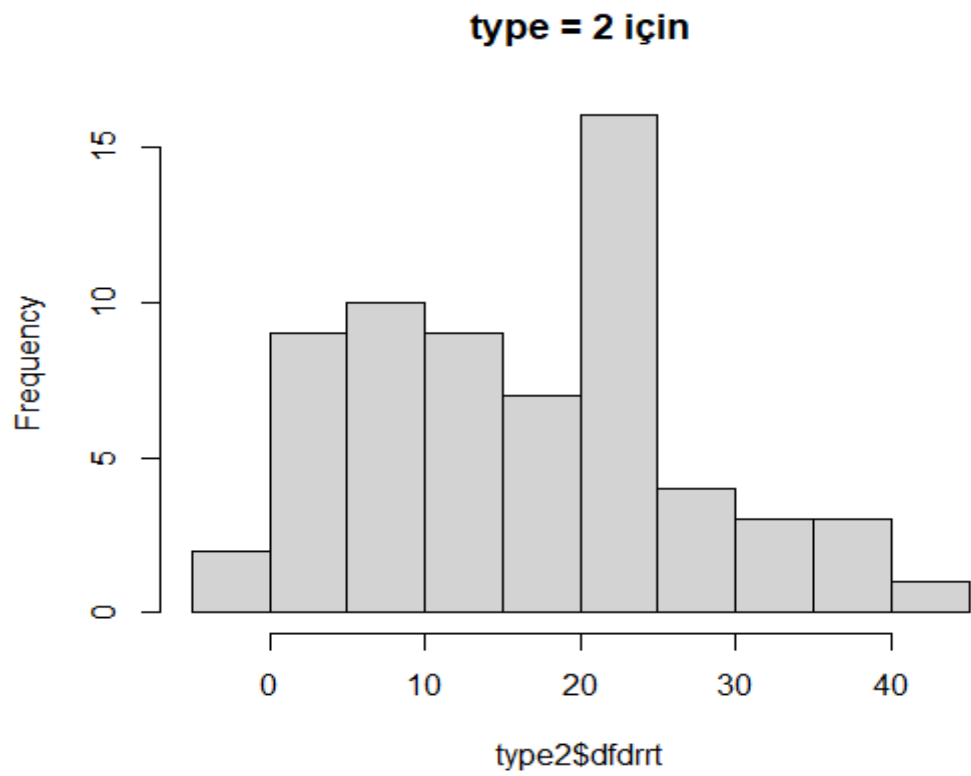
1) Normal Distribution



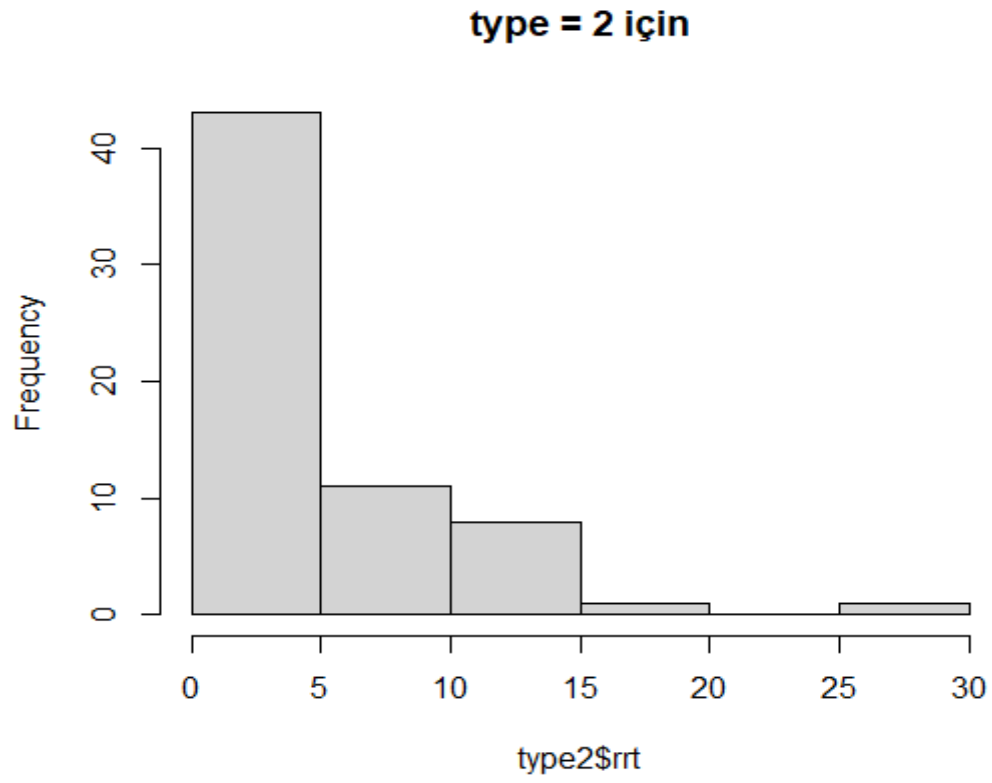
2) Normal Distribution



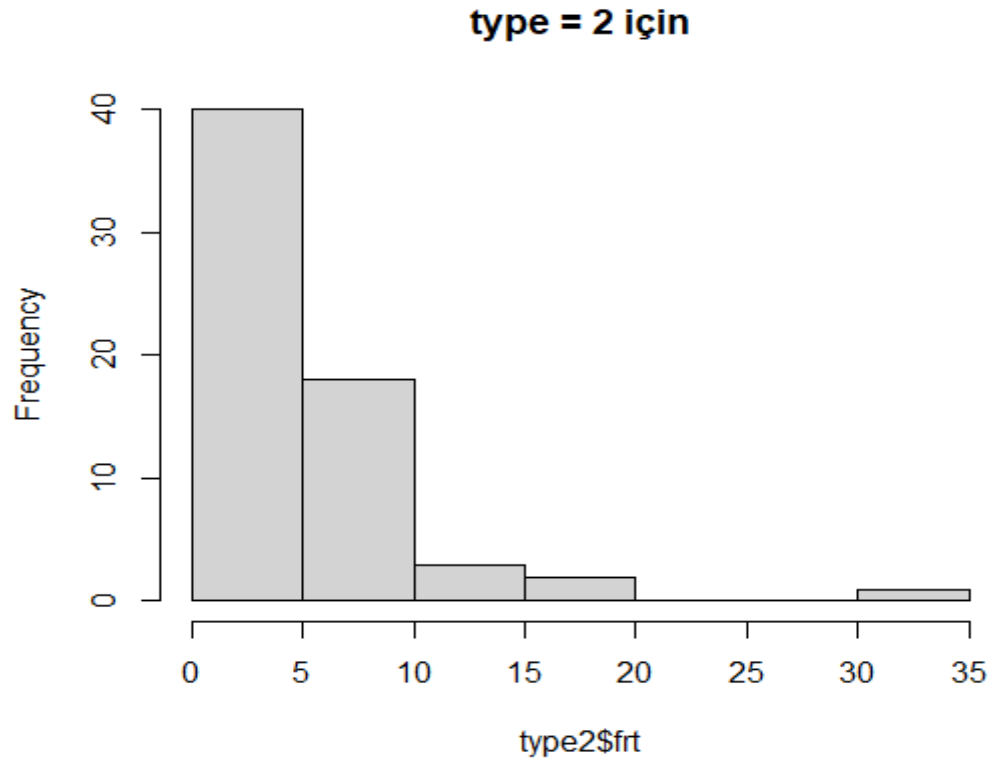
3) Normal Distribution



4) Non-normal Distribution

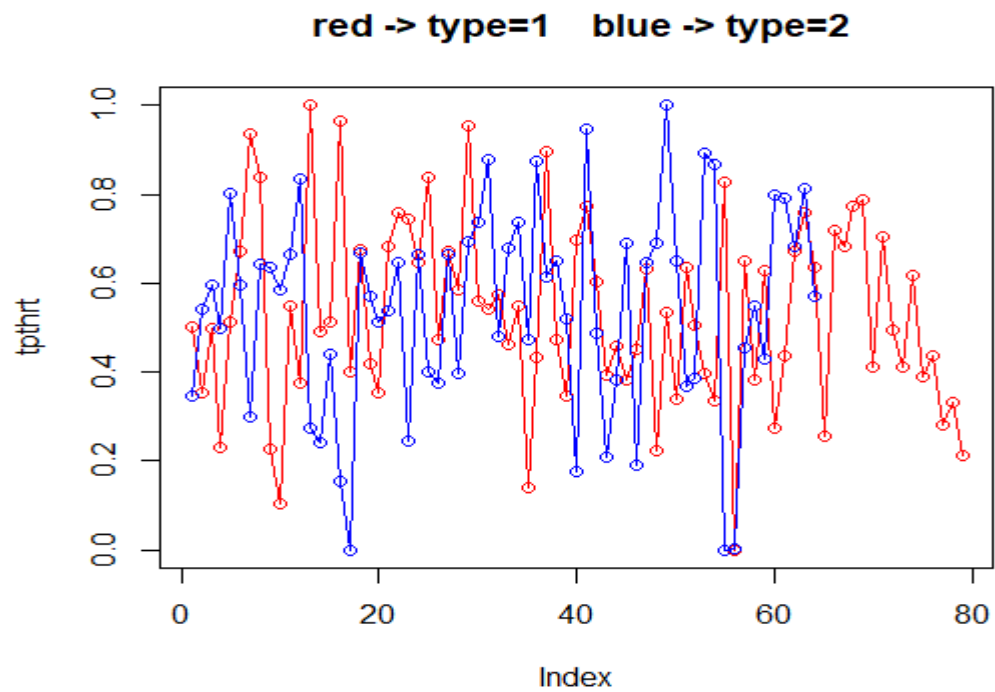


5) Non-normal Distribution

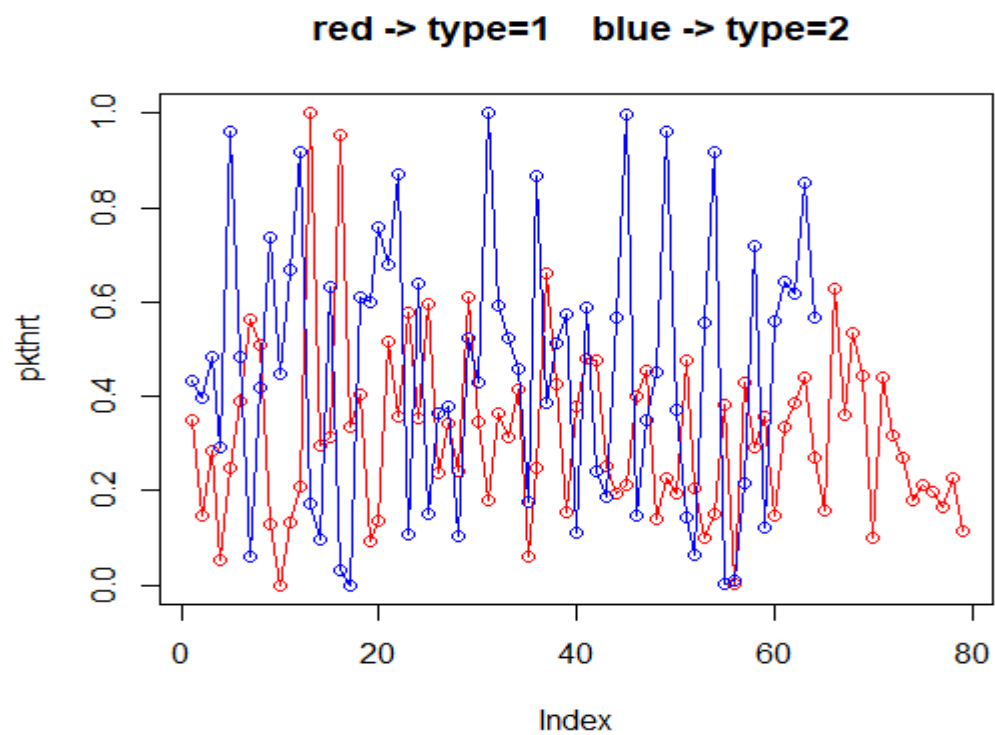


### T6 Line-Plot :

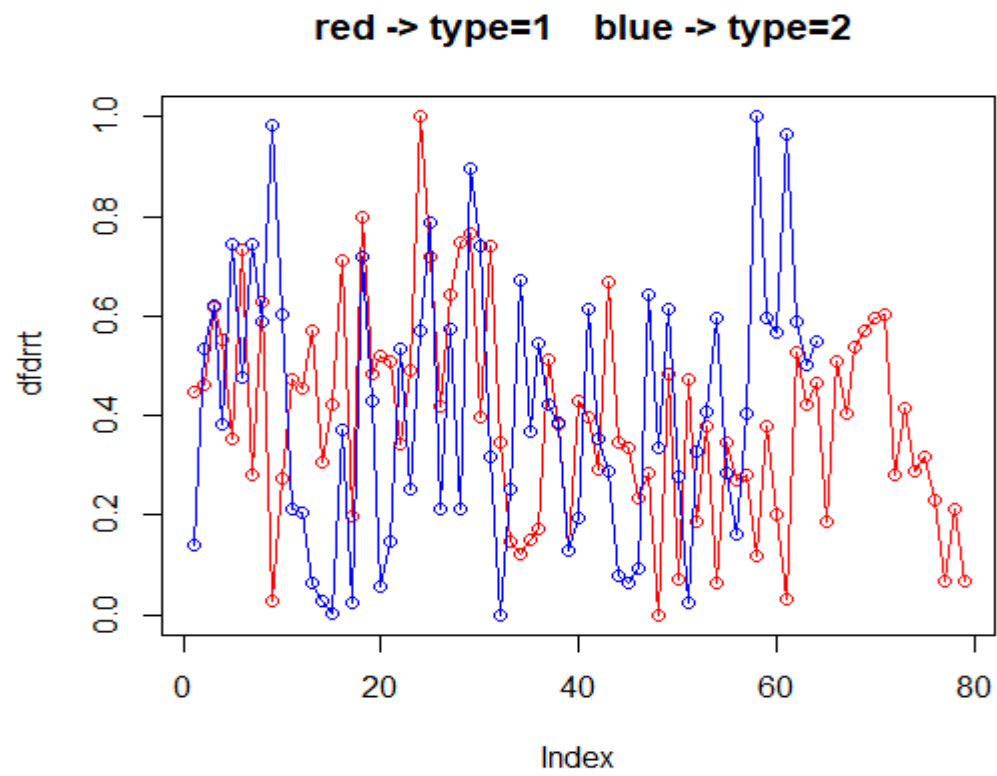
tpthrt :



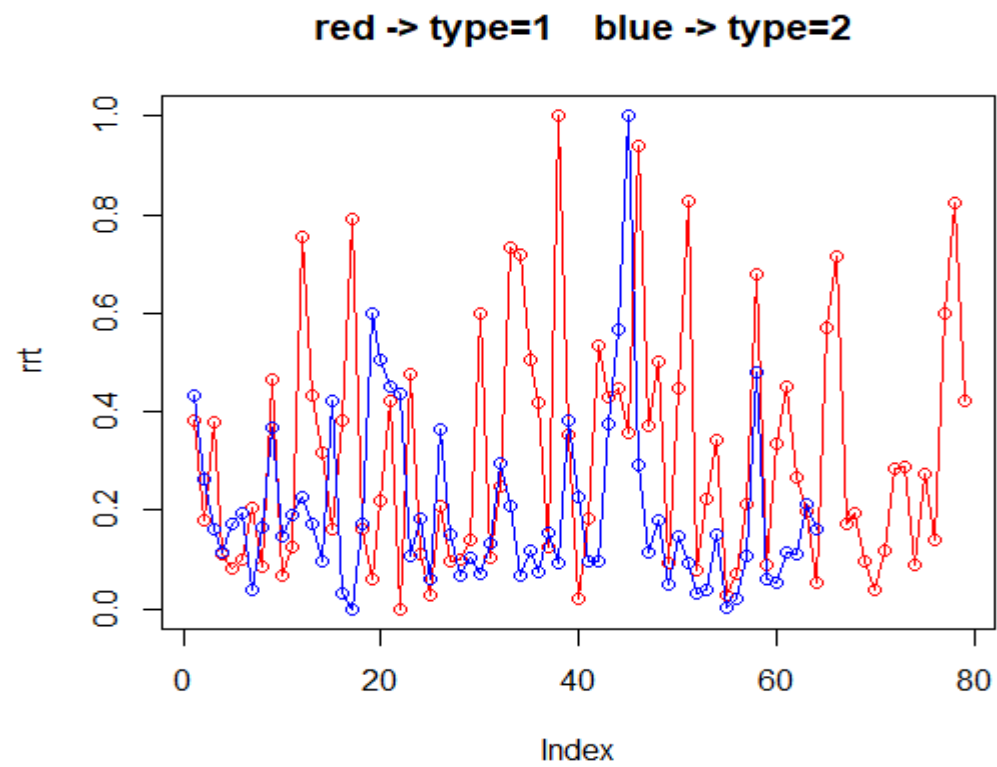
pkthrt :



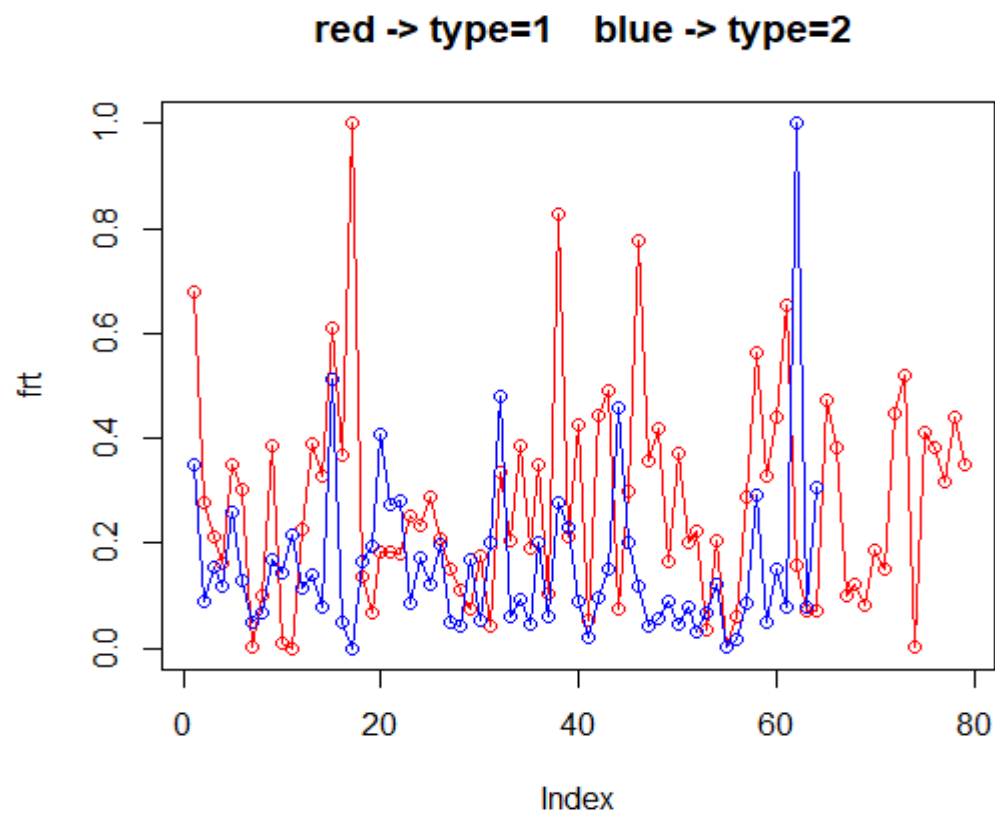
dfdrtrt :



rrt :



frt :



**T7 Correlation :**

```
> cor(type1$tpthrt,type1$pkthrt, method = "kendall")
[1] 0.684518
> cor(type2$tpthrt,type2$pkthrt)
[1] 0.8035611
> cor(type1[1:64,2],type2$tpthrt)
      [,1]
tpthrt -0.08827266
```

## Kodun Kendisi

```
1  ##1
2  library(readxl)
3  SdA_HW <- read_excel("D:/MASA/DERS/3.sınıf-1/istatistiksel veri analizi/proje/SdA-HW.xls")
4  View(SdA_HW)
5
6
7  ##2
8  SdA_HW$type=as.factor(SdA_HW$type)
9  SdA_HW <- SdA_HW[-5]
10
11  n <- which(is.na(SdA_HW$rrt))
12  SdA_HW <- SdA_HW[-n,]
13  n <- which(is.na(SdA_HW$frt))
14  SdA_HW <- SdA_HW[-n,]
15  View(SdA_HW)
16
17
18  ##3
19  type1 <- subset(SdA_HW, subset = (type ==1) )
20  type2 <- subset(SdA_HW, subset = (type ==2) )
21  View(type1)
22  View(type2)
23
24  fivenum(type1)
25  summary(type1)
26  fivenum(type2)
27  summary(type2)
28
29
30  ##4
31  boxplot(type1[,2:6], main = "type = 1 için")
32  boxplot(type2[,2:6], main = "type = 2 için")
33
34
35  ##5
36  hist(type1$tpthrt, main = "type = 1 için")
37  hist(type1$pkthrt, main = "type = 1 için")
38  hist(type1$dfdrtrt, main = "type = 1 için")
39  hist(type1$rrt, main = "type = 1 için")
40  hist(type1$frt, main = "type = 1 için")
41
42  hist(type2$tpthrt, main = "type = 2 için")
43  hist(type2$pkthrt, main = "type = 2 için")
44  hist(type2$dfdrtrt, main = "type = 2 için")
45
46
47
48
49  ##6
50  normalize <- function(x){
51    return ((x - min(x)) / (max(x) - min(x)))
52  }
53
54  type1$tpthrt <- normalize(type1$tpthrt)
55  type1$pkthrt <- normalize(type1$pkthrt)
56  type1$dfdrtrt <- normalize(type1$dfdrtrt)
57  type1$rrt <- normalize(type1$rrt)
58  type1$frt <- normalize(type1$frt)
59
60  type2$tpthrt <- normalize(type2$tpthrt)
61  type2$pkthrt <- normalize(type2$pkthrt)
62  type2$dfdrtrt <- normalize(type2$dfdrtrt)
63  type2$rrt <- normalize(type2$rrt)
64  type2$frt <- normalize(type2$frt)
65
66  plot(type1$tpthrt, type = "o", col = "red", ylab = "tpthrt", main = " red -> type=1   blue -> type=2")
67  lines(type2$tpthrt, type = "o", col = "blue")
68  plot(type1$pkthrt, type = "o", col = "red", ylab = "pkthrt", main = " red -> type=1   blue -> type=2")
69  lines(type2$pkthrt, type = "o", col = "blue")
70  plot(type1$dfdrtrt, type = "o", col = "red", ylab = "dfdrtrt", main = " red -> type=1   blue -> type=2")
71  lines(type2$dfdrtrt, type = "o", col = "blue")
72  plot(type1$rrt, type = "o", col = "red", ylab = "rrt", main = " red -> type=1   blue -> type=2")
73  lines(type2$rrt, type = "o", col = "blue")
74  plot(type1$frt, type = "o", col = "red", ylab = "frt", main = " red -> type=1   blue -> type=2")
75  lines(type2$frt, type = "o", col = "blue")
76
77
78  ##7
79  cor(type1$tpthrt,type1$pkthrt, method = "kendall")
80  cor(type2$tpthrt,type2$pkthrt)
81  cor(type1[1:64,2],type2$tpthrt)
82
```