> head(mydata)

y x1 x2 x3 x4 x5

1 0.3833838 82.19762 128.38572 2.963222 8.148025 66.49398

2 14.8224896 83.84911 117.74896 2.415674 5.945306 59.39393

3 15.5270754 92.79354 98.81826 2.682626 5.933339 69.98642

4 -3.3207493 85.35254 108.51833 2.985579 2.967865 72.07055

5 10.0196414 85.64644 97.02792 3.335348 7.580771 65.93392

6 26.9836869 93.57532 96.28504 2.174727 5.578532 65.00873

> names(newdata)

[1] "y" "x1" "x2" "x3\_f\_factor" "x4" "

x5"

> x3\_categories <- cut(x3, breaks = category\_centers, labels = c("A", "B", "C", "D"))

> newdata <- data.frame(y, x1, x2, x3\_categories, x4, x5)

> head(newdata)

y x1 x2 x3\_categories x4 x5

1 0.3833838 82.19762 128.38572 B 8.148025 66.49398

2 14.8224896 83.84911 117.74896 B 5.945306 59.39393

3 15.5270754 92.79354 98.81826 B 5.933339 69.98642

4 -3.3207493 85.35254 108.51833 B 2.967865 72.07055

5 10.0196414 85.64644 97.02792 C 7.580771 65.93392

6 26.9836869 93.57532 96.28504 B 5.578532 65.00873

> unique\_values <- unique(newdata$x3\_categories)

> print(unique\_values)

[1] B C A D

Levels: A B C D

> print(missing\_columns)

character(0)

> #kategorik değişken için frekans analizi yapıyorum

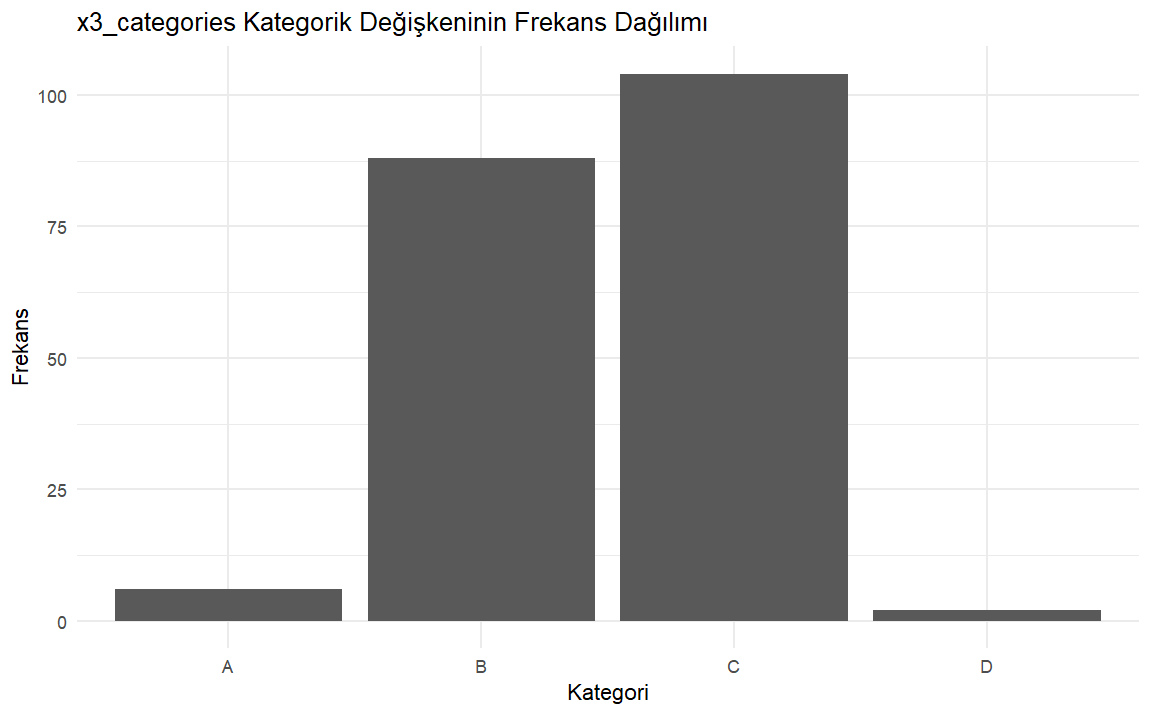
> category\_counts <- as.data.frame(table(newdata$x3\_categories))

> ggplot(category\_counts, aes(x = Var1, y = Freq)) +

+ geom\_bar(stat = "identity") +

+ labs(x = "Kategori", y = "Frekans", title = "x3\_categories Kategorik Değişkeninin Frekans Dağılımı") +

+ theme\_minimal()



summary(category\_counts)

Var1 Freq

A:1 Min. : 2

B:1 1st Qu.: 5

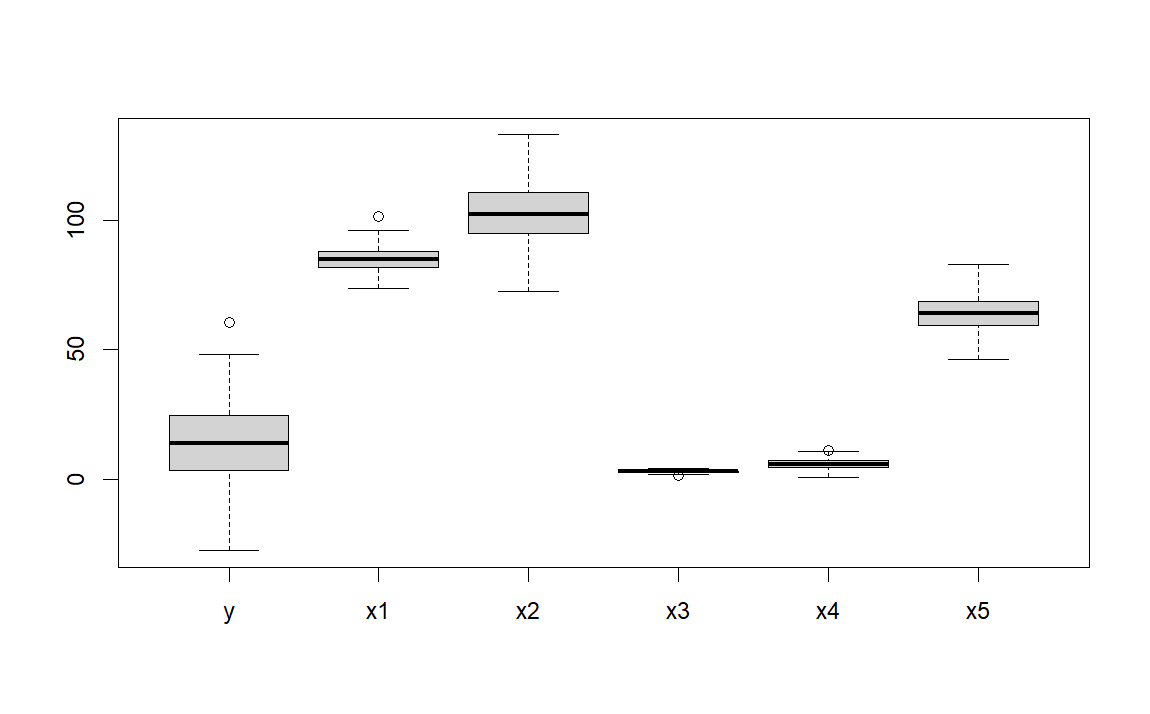
C:1 Median : 47

D:1 Mean : 50

3rd Qu.: 92

Max. :104

> boxplot(mydata)



> summary(outliers)

Mode FALSE TRUE

logical 199 1

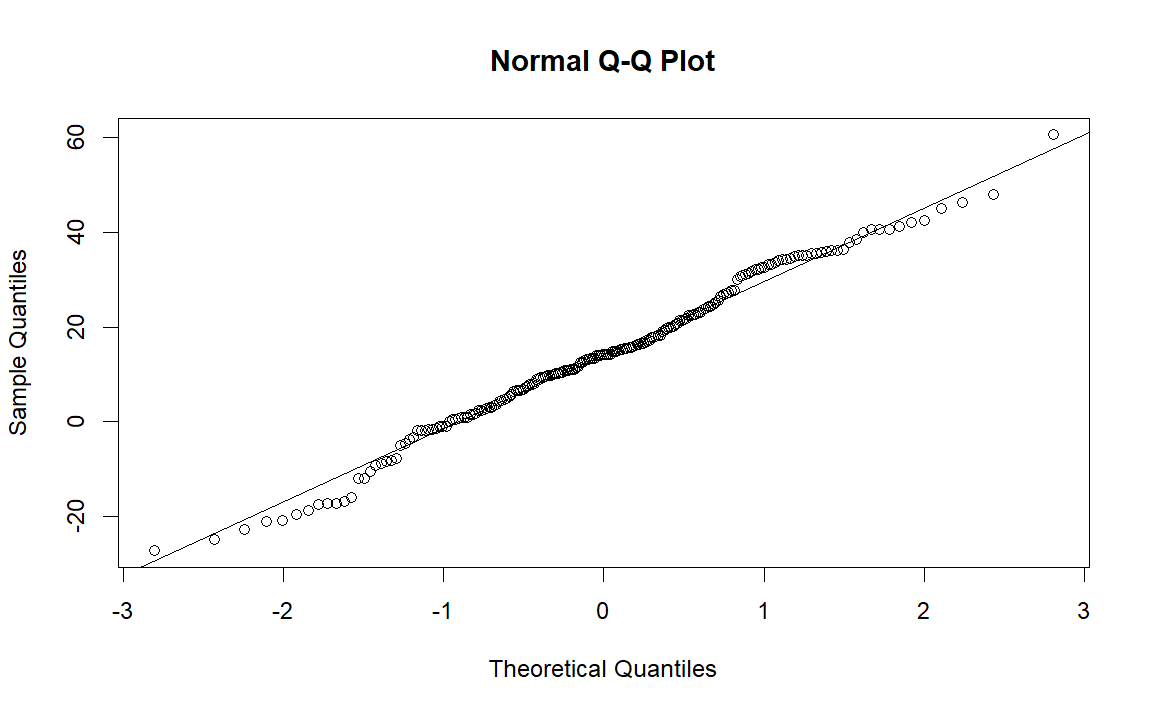
print(outlier\_values)

y x1 x2 x3\_categories x4 x5

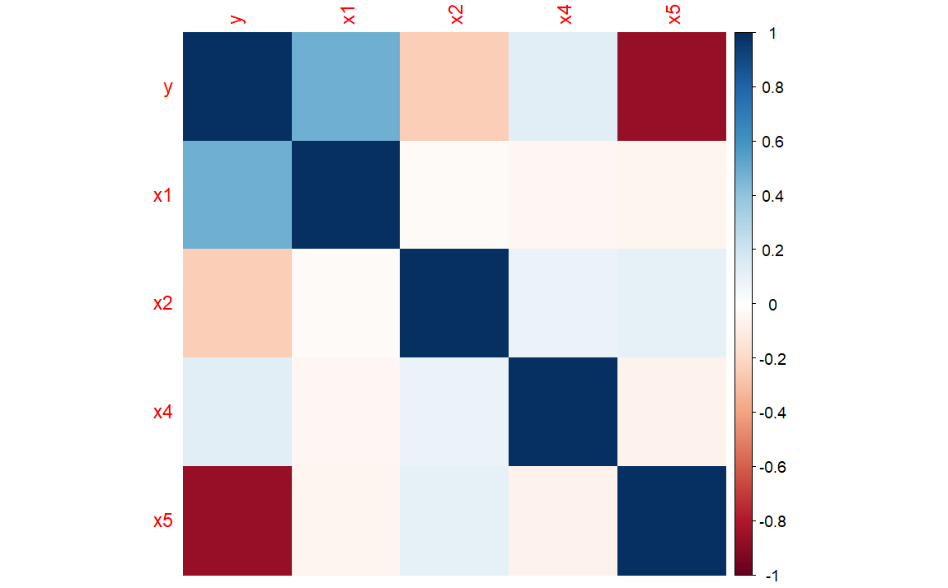
11 60.64595 91.12041 103.4309 C 5.205939 46.44458

> qqnorm(newdata$y)

> qqline(newdata$y)



> corrplot(correlation\_matrix, method = "color")



> # Bağımsız değişkenlerin grafiğini çizme

> par(mfrow=c(3,2)) # Grafiklerin yan yana yerleştirilmesi için ayar

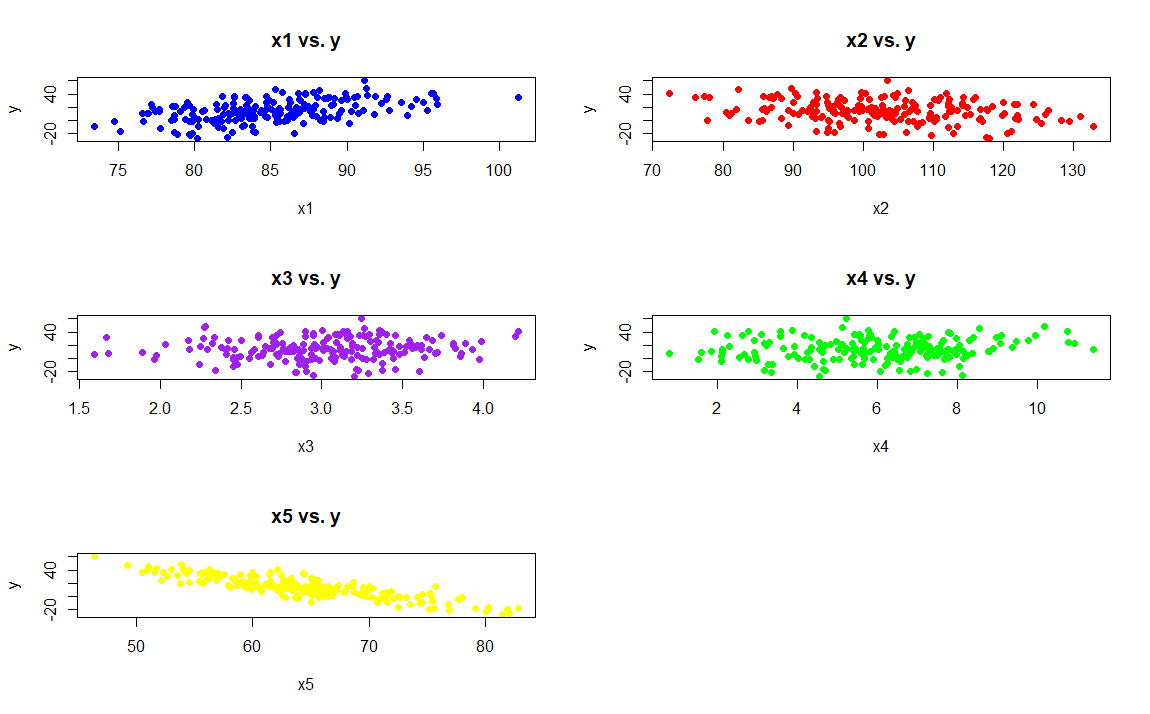
> plot(x1, y, main = "x1 vs. y", xlab = "x1", ylab = "y", col = "blue", pch = 16)

> plot(x2, y, main = "x2 vs. y", xlab = "x2", ylab = "y", col = "red", pch = 16)

> plot(x3, y, main = "x3 vs. y", xlab = "x3", ylab = "y", col = "purple", pch = 16)

> plot(x4, y, main = "x4 vs. y", xlab = "x4", ylab = "y", col = "green", pch = 16)

> plot(x5, y, main = "x5 vs. y", xlab = "x5", ylab = "y", col = "yellow", pch = 16)



data <- read.csv("crop.csv")

> head(data)

Nitrogen phosphorus potassium temperature humidity ph rainfall label X X.1

1 90 42 43 20.87974 82.00274 6.502985 202.9355 rice NA NA

2 85 58 41 21.77046 80.31964 7.038096 226.6555 rice NA NA

3 60 55 44 23.00446 82.32076 7.840207 263.9642 rice NA NA

4 74 35 40 26.49110 80.15836 6.980401 242.8640 rice NA NA

5 78 42 42 20.13017 81.60487 7.628473 262.7173 rice NA NA

6 69 37 42 23.05805 83.37012 7.073454 251.0550 rice NA NA

unique\_values <- unique(data$label)

> print(unique\_values)

[1] "rice" "maize" "chickpea" "kidneybeans" "pigeonpeas" "mothbeans" "mungbean"

[8] "blackgram" "lentil" "pomegranate" "banana" "mango" "grapes" "watermelon"

[15] "muskmelon" "apple" "orange" "papaya" "coconut" "cotton" "jute"

[22] "coffee"

clean\_data %>% count(label)

label n

1 apple 100

2 banana 100

3 blackgram 100

4 chickpea 100

5 coconut 100

6 coffee 100

7 cotton 100

8 grapes 100

9 jute 100

10 kidneybeans 100

11 lentil 100

12 maize 100

13 mango 100

14 mothbeans 100

15 mungbean 100

16 muskmelon 100

17 orange 100

18 papaya 100

19 pigeonpeas 100

20 pomegranate 100

21 rice 100

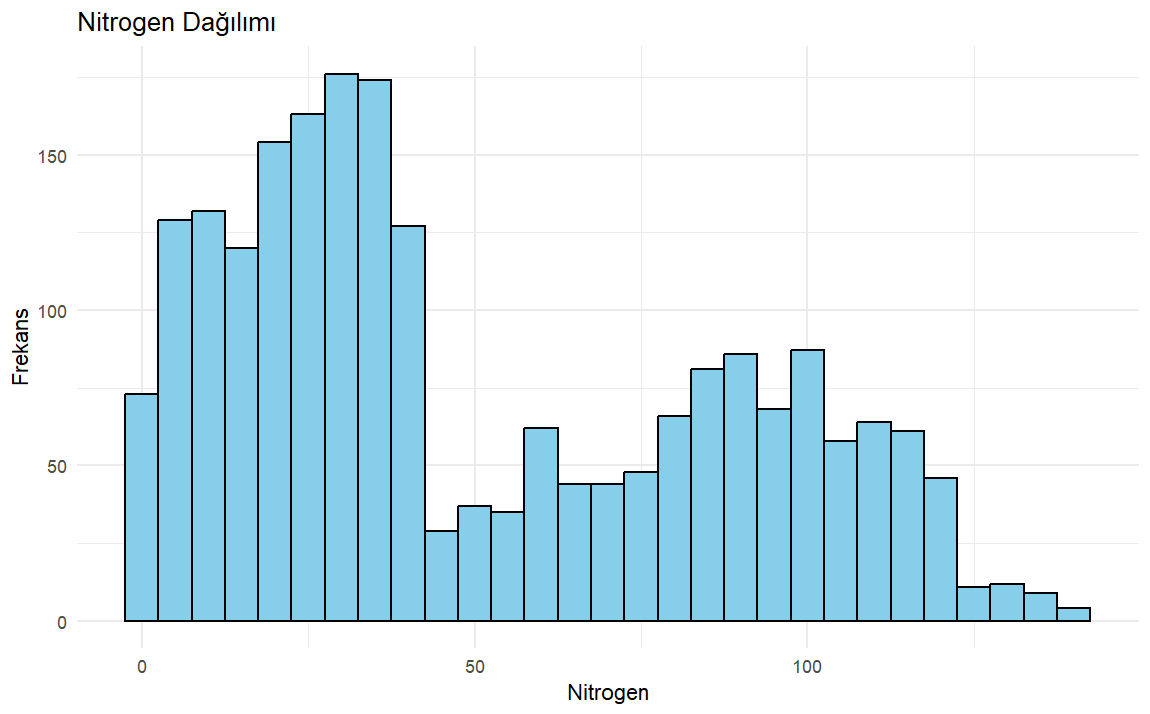
22 watermelon 100

ggplot(clean\_data, aes(x = Nitrogen)) +

+ geom\_histogram(binwidth = 5, fill = "skyblue", color = "black") +

+ labs(x = "Nitrogen", y = "Frekans", title = "Nitrogen Dağılımı") +

+ theme\_minimal()

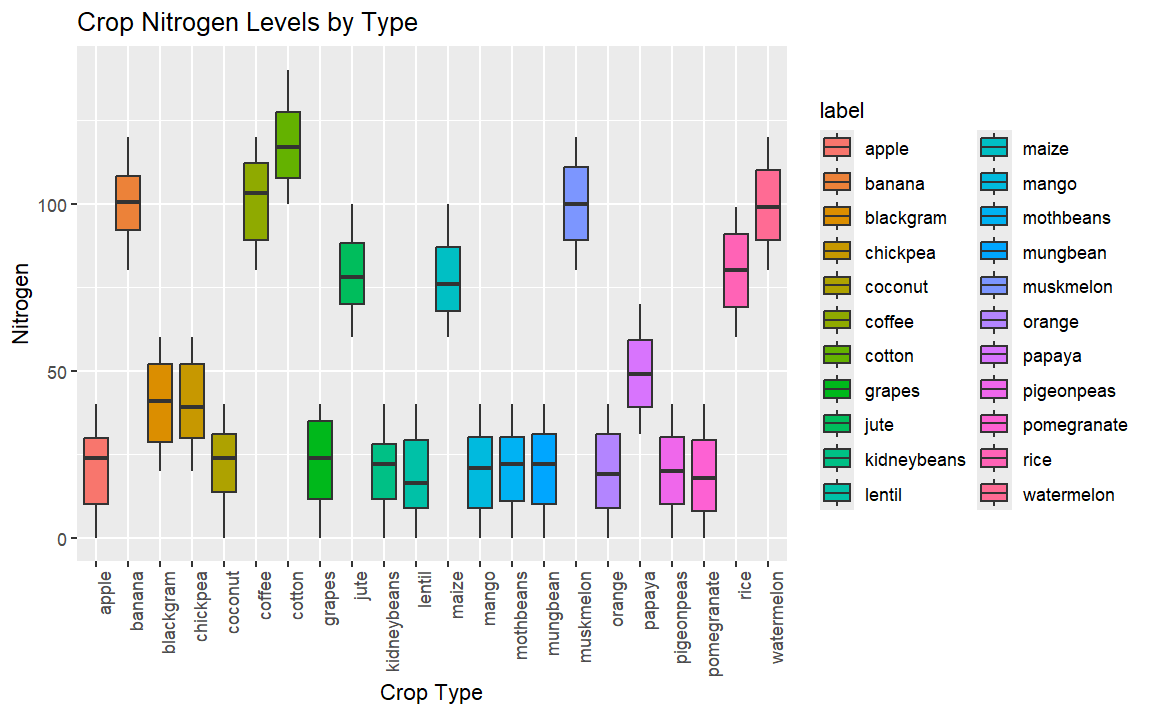


ggplot(clean\_data, aes(x = label, y = Nitrogen, fill = label)) +

+ geom\_boxplot() +

+ theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +

+ labs(x = "Crop Type", y = "Nitrogen", title = "Crop Nitrogen Levels by Type")

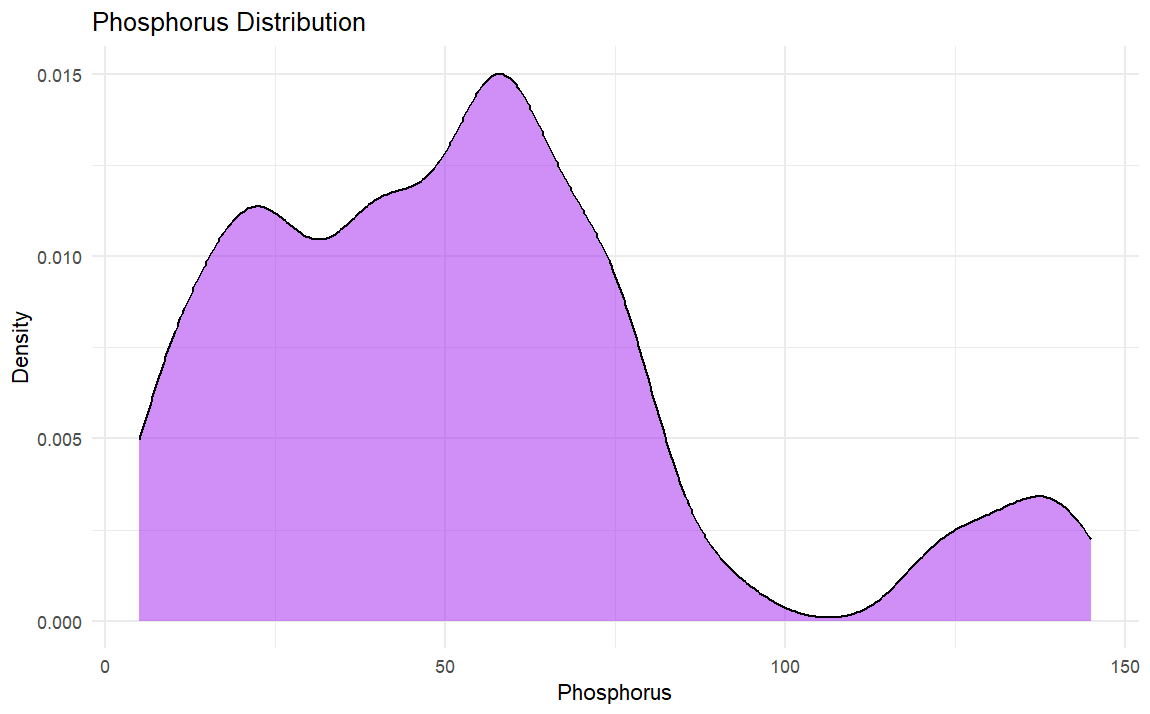


ggplot(clean\_data, aes(x = phosphorus)) +

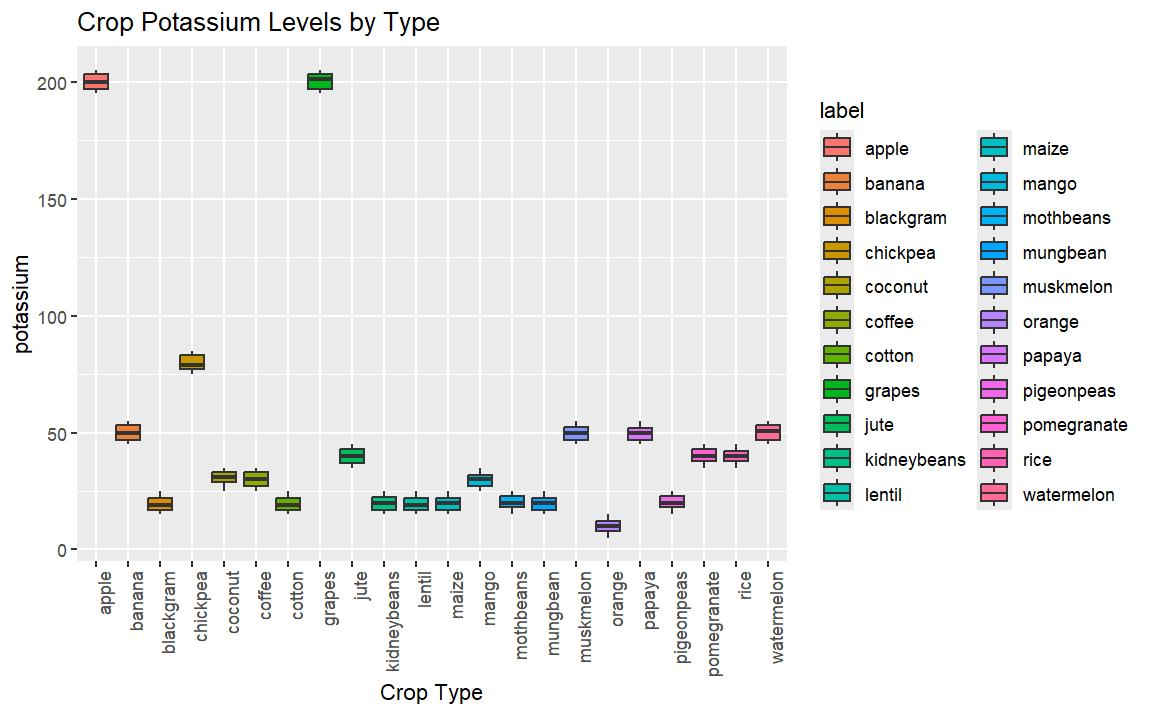
+ geom\_density(fill = "purple", alpha = 0.5) +

+ labs(x = "Phosphorus", y = "Density", title = "Phosphorus Distribution") +

+ theme\_minimal()



|  |
| --- |
| ggplot(clean\_data, aes(x = label, y = potassium,fill = label)) +  + geom\_boxplot() +  + theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  + labs(x = "Crop Type", y = "potassium", title = "Crop Potassium Levels by Type") |
|  |
| |  | | --- | | > | |

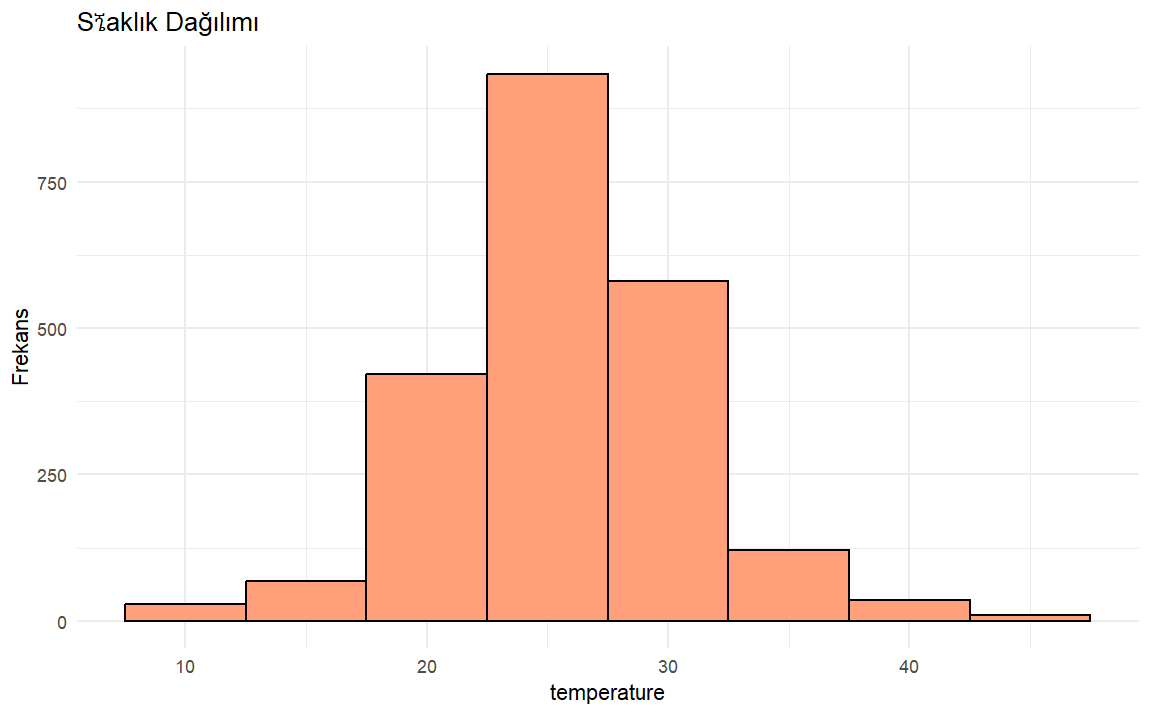


ggplot(clean\_data, aes(x = temperature)) +

+ geom\_histogram(binwidth = 5, fill = "lightsalmon", color = "black") +

+ labs(x = "temperature", y = "Frekans", title = "Sıcaklık Dağılımı") +

+ theme\_minimal()

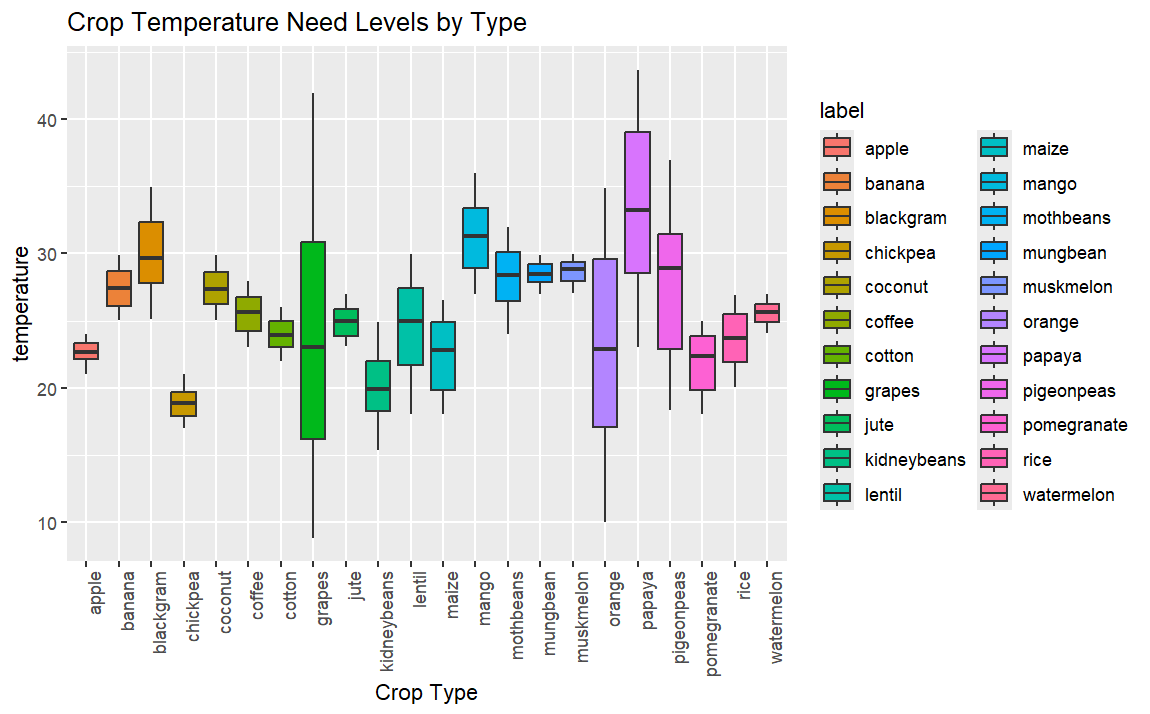


ggplot(clean\_data, aes(x = label, y = temperature,fill = label)) +

+ geom\_boxplot() +

+ theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +

+ labs(x = "Crop Type", y = "temperature", title = "Crop Temperature Need Levels by Type")

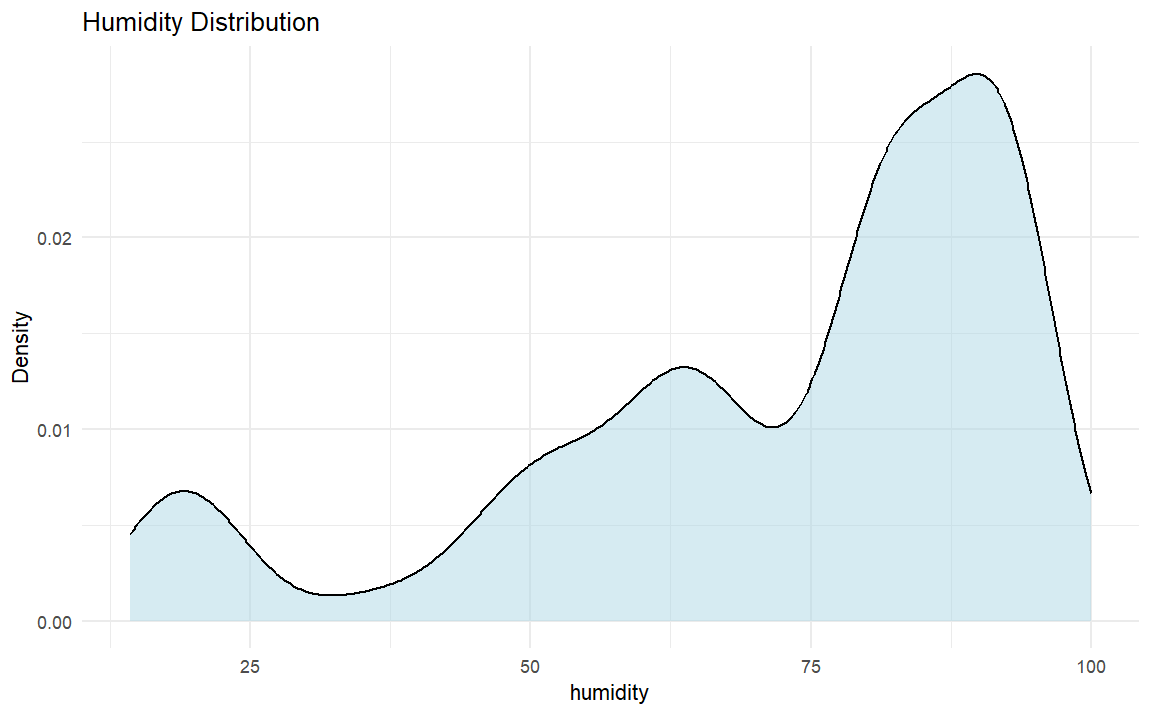


ggplot(clean\_data, aes(x = humidity)) +

+ geom\_density(fill = "lightblue", alpha = 0.5) +

+ labs(x = "humidity", y = "Density", title = "Humidity Distribution") +

+ theme\_minimal()

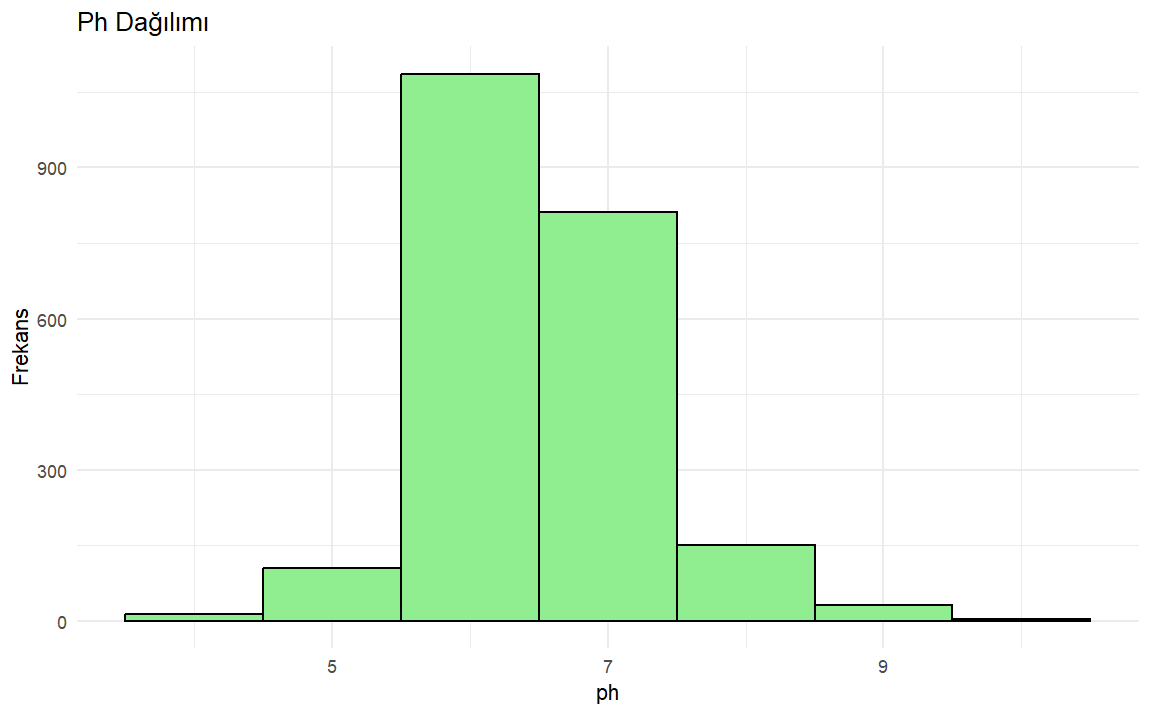


ggplot(clean\_data, aes(x = ph)) +

+ geom\_histogram(binwidth = 1, fill = "lightgreen", color = "black") +

+ labs(x = "ph", y = "Frekans", title = "Ph Dağılımı") +

+ theme\_minimal()

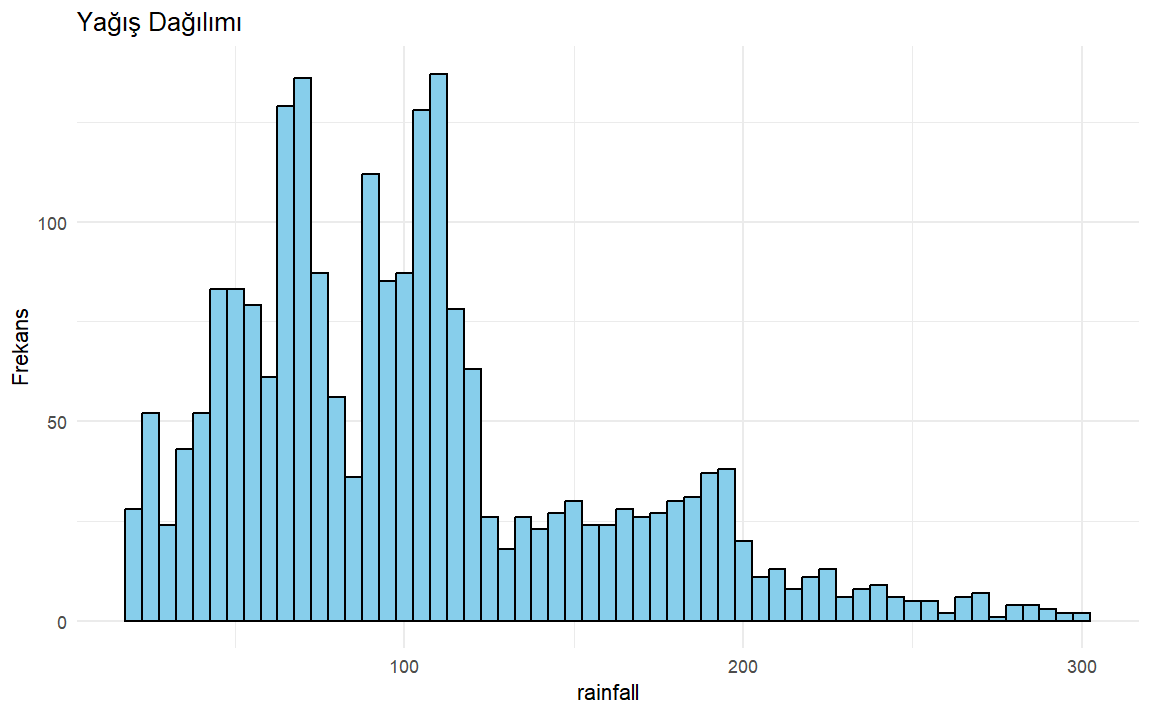


ggplot(clean\_data, aes(x = rainfall)) +

+ geom\_histogram(binwidth = 5, fill = "skyblue", color = "black") +

+ labs(x = "rainfall", y = "Frekans", title = "Yağış Dağılımı") +

+ theme\_minimal()



> corrplot(correlation\_matrix\_crop, method = "color")

