

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

## İçindekiler

R Kodları .....	1
R Shiny Web Application Araştırma Süreci Hakkında .....	5
Veri68 Analiz Raporu .....	6
Veri Tiplerine Dair Analiz.....	7
Kutu Grafiği Yorumu .....	8
State Değişkeni.....	8
Time Değişkeni .....	9
Net Değişkeni .....	9
Age Değişkeni .....	10
Sex Değişkeni.....	10
Normallik Grafiği .....	11
State Değişkeni.....	11
Sex Değişkeni.....	11
Time Değişkeni .....	12
Net Değişkeni .....	12
Age Değişkeni .....	13
Kaynakça.....	14
Kitaplar.....	14
Web Siteler.....	14

## R Kodları

```
library(shiny)
```

```
//Kütüphaneden shiny iskeletini çağırmaı sağlayan komut
```

```
library(data.table)
```

```
veri68<- read.csv(file ="C:/Users/Victory/OneDrive/Belgeler/veri68yeni.csv", header =  
FALSE, sep = ";")
```

```
veri68s <- data.matrix(veri68)
```

```
// data.frame sınıfında olan verin matris durumuna getirilmesi
```

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

```
ui <- fluidPage(
```

```
  titlePanel("Web Ortamında Veri Analizi"),
```

```
  //Web uygulamasının ana başlığı
```

```
  sidebarLayout(
```

```
    sidebarPanel(
```

```
      sliderInput(inputId = "aralik",
```

```
        label = "Bir aralik belirleyiniz.",
```

```
        min = 1,
```

```
        max =300,
```

```
        value = 150),
```

```
      //Analizlerde ihtiyaç duyulacak girdilerin aralığı
```

```
      selectInput(inputId = "degisken",
```

```
        label = "Bir degisken seciniz:",
```

```
        choices = c("state"=1,
```

```
                    "time"=2,
```

```
                    "net"=3,
```

```
                    "age"=4,
```

```
                    "sex"=5)),
```

```
      // Analizlerin değişken belirleyerek yapılmasını sağlayan komut dizisi
```

```
      radioButtons(inputId = "renk",
```

```
        label = "Grafik rengi belirleyiniz:",
```

```
        choices = c("green", "red", "yellow", "purple"),
```

```
        selected = "red")
```

```
    // Renk butonlarının oluşmasını sağlayan komut dizini. (Kullanıcıların grafik renklerini belirlemesini sağlar.)
```

```
  ),
```

## İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

```
mainPanel(  
  // Sekmelerin arayüzde görünmesini sağlayacak olan kod bölümü  
  tabsetPanel(  
  
    tabPanel(title = "Summary",  
              verbatimTextOutput("summary")),  
  
    tabPanel(title = "Str",  
              verbatimTextOutput("str")),  
  
    tabPanel(title = "Veri",  
              tableOutput("data_out")),  
  
    tabPanel(title = "Histogram",  
              plotOutput("histogram")),  
    //Grafik çıktısının sekmede görünmesini sağlar. (plotOutput)  
    tabPanel(title = "Kutu Grafiği",  
              plotOutput("boxplot")),  
  
    tabPanel(title = "Normallik Grafiği",  
              plotOutput("qqnorm"))  
  
  )  
  
)  
  
)
```

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

)

```
server <- function(input, output, session) {
```

**//Kullanıcının etkileşimli ortamda girdiği komutların anlamlı bir halde çıktıya dönüşebilmesini sağlayan ana kod bölümü.**

```
  output$summary <- renderPrint({
```

```
    summary(veri68s)
```

**//Summary çıktısının alınmasını sağlayan komut**

```
  })
```

```
  output$str <- renderPrint({
```

```
    str(veri68)
```

```
  })
```

```
  output$data_out <- renderTable({
```

```
    veri68
```

```
  })
```

```
  output$histogram <- renderPlot({
```

```
    colm <- as.numeric(input$degisken)
```

```
    hist(veri68s[,colm],
```

```
    breaks = seq(1,300),
```

```
    col = input$renk,
```

```
    border = "red",
```

```
    xlab = "aralik")
```

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

```
})
```

```
output$boxplot<- renderPlot({
```

```
  colm <- as.numeric(input$degisken)
```

```
  boxplot(veri68s[, colm], col = input$renk)
```

```
})
```

```
output$qqnorm <- renderPlot({
```

```
  colm <- as.numeric(input$degisken)
```

```
  qqnorm(veri68s[,colm])
```

```
  qqline(veri68s[,colm])
```

```
//Veri68 isimli veri kümesinin normal dağılıp dağılmadığını gösteren grafik
```

```
  })
```

```
}
```

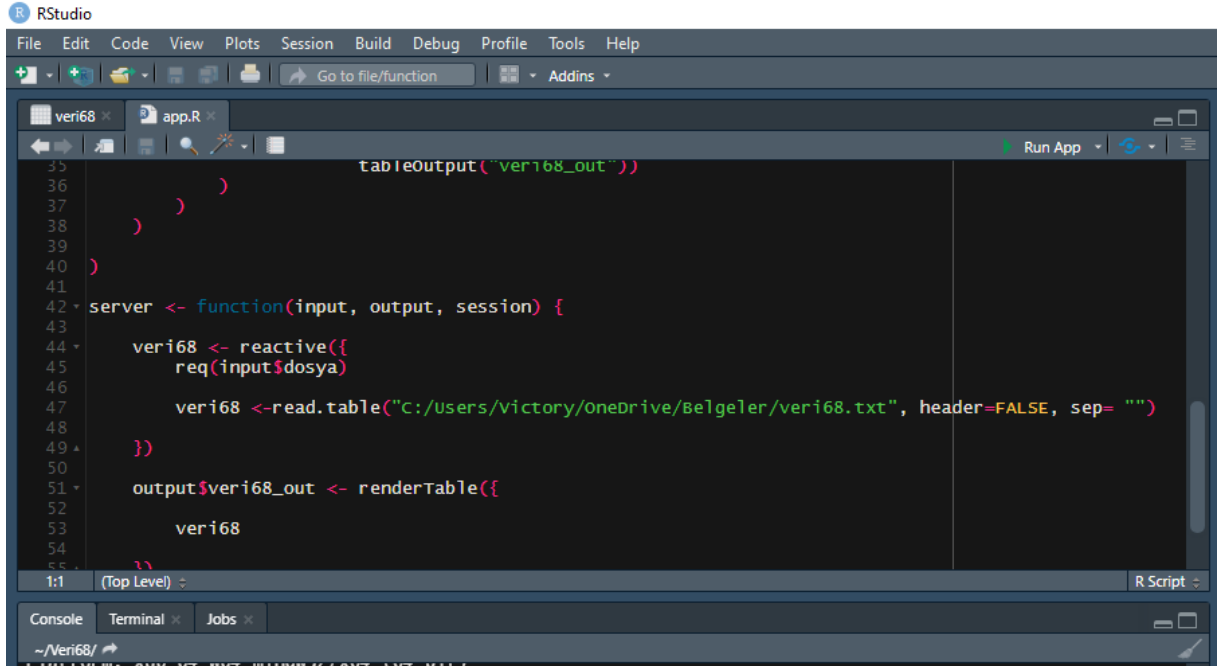
```
shinyApp(ui, server)
```

## R Shiny Web Application Araştırma Süreci Hakkında

R Shiny Web Application araştırma ödevine, bu konuyla alakalı olarak yapılan uygulamaları incelemekle başladım. Araştırma sürecinde ilgimi çeken en önemli nokta ise yazılacak olan kodlarla özgür bir etkileşim ortamı yaratılmasıdır. İstediğim analizleri özellikle web ortamında görmek, bunları tabandan tavana inşa etmek oldukça verimli bir süreçti.

Ödevin hazırlık sürecinin ilk iki haftasında verilerimi (veri68) web ortamına çekmekte oldukça zorlandım. Hatta ilk günlerde verinin server içerisinde yer alması gerektiğini ön gördüğümünden, bir süre boyunca veri çekme kodlarımı sürekli server alanına yazdım.

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi



```
1:1 (Top Level) - R Script
library(data.table)
library(shiny)
library(readr)

tableoutput("veri68_out")

server <- function(input, output, session) {
  veri68 <- reactive({
    req(input$dosya)
    veri68 <- read.table("C:/Users/Victory/OneDrive/Belgeler/veri68.txt", header=FALSE, sep="")
  })
  output$veri68_out <- renderTable({
    veri68
  })
}
```

Bu hatadan dönmemi sağlayan ise yabancı kaynaklardaki araştırmalarım oldu. Verilerimi web ortamına taşımak için ise, klasik shiny iskeletine başlamadan önce library komutunu kullandım.

➤ Library(data.table)

Komutunun ardından belirlediğim konumda yer alan veriyi rahatlıkla çekmeyi başardım.

Ödevin ilerleyen sürecinde karşılaştığım bir diğer zorluk ise, verilerimi doğru bir şekilde sınıflandırmamamdı. Bilgi eksikliğine dayalı bir hata olmasa da karşılaştığım bu hata dolayısıyla bir süre verilerimden hatalı sonuç aldım. Yaptığım hata ise, belirlediğim konumdaki verileri web ortamına çağırırken

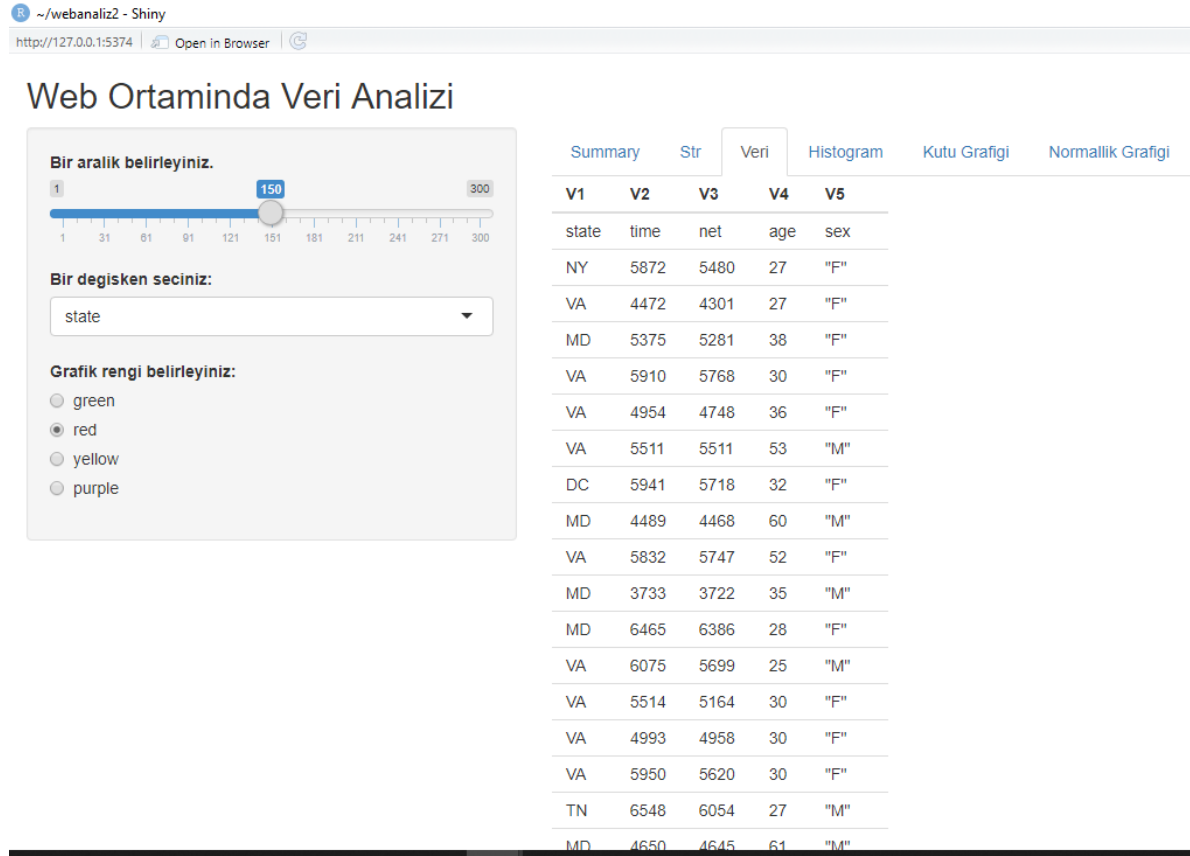
➤ Header  
➤ Sep

Gibi belirteçleri kullanmamaktı. Verilerim “state.time.net.age.sex” şeklinde noktalarla birbirinden ayrıldığından detaylı analiz yapmak ve histogram grafiği çizdirmek mümkün olmadı. Bu sorunu da keşfetmenin ardından elimde bulunan hatasız veri ile işler daha ilgi çekici bir hale geldi. Belirlediğim sekmelerle etkileşimli bir web ortamını yaratmak, R dili öğrenimin açısından oldukça verimli ve bir o kadar da keyifli oldu.

## Veri68 Analiz Raporu

R üzerinden web ortamı yaratma amacıyla başladığım ödevde, veri68 isimli veri kaynağından yaralandım.

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

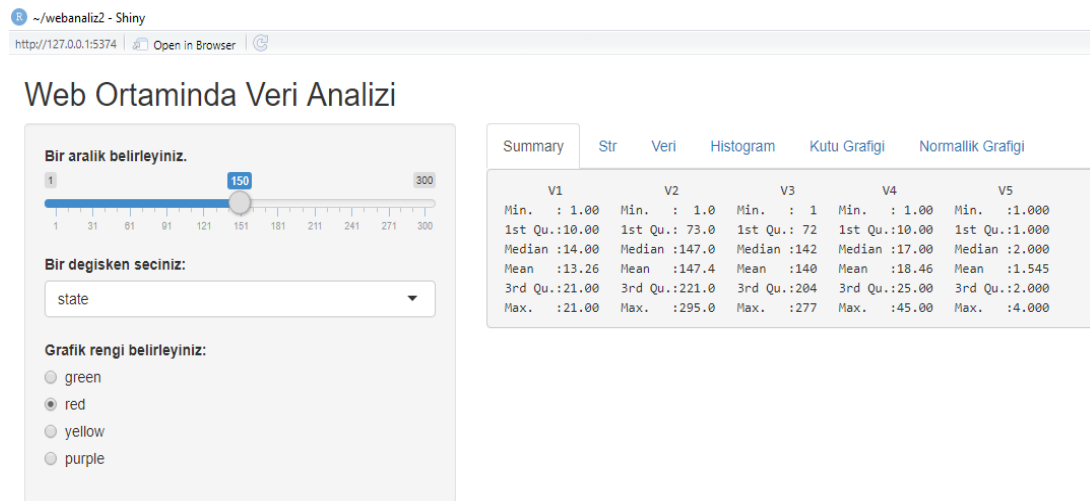


## Veri Tiplerine Dair Analiz

Veri setimde 5 değişken olmak üzere toplam 300 adet veri bulunmaktadır. “State” ve “Sex” isminde iki adet char sınıfında değişken bulunmaktadır.

➤ `Class(veri68) //` komutundan da görülebileceği üzere veri `data.frame` sınıfındadır.

Analiz raporunda olumlu sonuçlarla karşılaşabilmek adına belli noktalarda (histogram, boxplot) verimi `data.matrix` sınıfına çevirdim. (numeric olmasından dolayı)



Time değişkeninde yer alan verilerin ortalaması (mean) = 147.4

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

Net değişkeninde yer alan verilen ortalaması (mean) = 140

Age değişkeninde yer alan verilerin ortalaması (mean) = 18.46

## Web Ortamında Veri Analizi

**Bir aralık belirleyiniz.**

1 150 300

1 31 61 91 121 151 181 211 241 271 300

**Bir degisken seciniz:**

state ▼

**Grafik rengi belirleyiniz:**

☐ green

☒ red

☐ yellow

☐ purple

Summary Str Veri Histogram Kutu Grafiği Normallik Grafiği

```
'data.frame': 301 obs. of 5 variables:
 $ V1: chr "state" "NY" "VA" "MD" ...
 $ V2: chr "time" "5872" "4472" "5375" ...
 $ V3: chr "net" "5480" "4301" "5281" ...
 $ V4: chr "age" "27" "27" "38" ...
 $ V5: chr "sex" "\"F\"" "\"F\"" "\"F\"" ...
```

## Kutu Grafiği Yorumu

### State Değişkeni

## Web Ortamında Veri Analizi

**Bir aralık belirleyiniz.**

1 150 300

1 31 61 91 121 151 181 211 241 271 300

**Bir degisken seciniz:**

state ▼

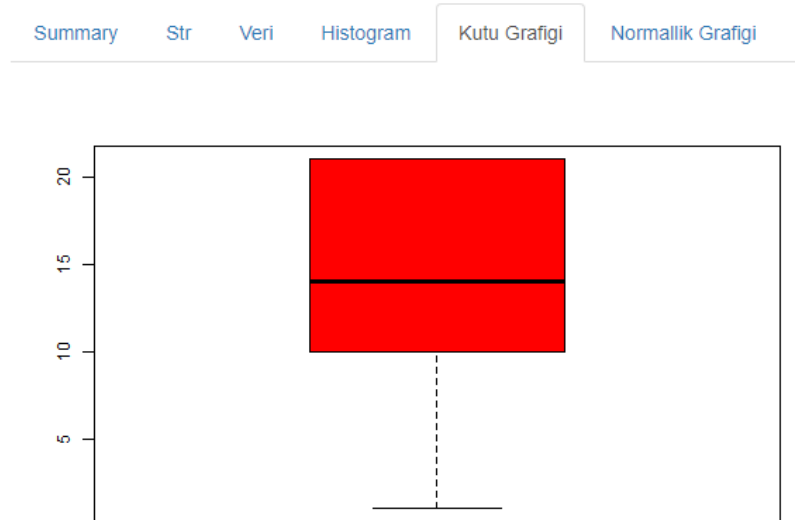
**Grafik rengi belirleyiniz:**

☐ green

☒ red

☐ yellow

☐ purple



Değişkene ait verilerin maksimum değeri ile 3.çeyreklik değeri aynıdır. Bu nedenle grafikte üst bıyık bulunmamaktadır. Medyan 14 ve 1. Çeyreklik 10 değerindedir. Q1 ile Q2 arasındaki uzaklık Q2 ile Q3 arasındaki uzaklıktan daha azdır. Veri kümesindeki değerler, küçük değerlere daha eğilimlidir.



# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

## Time Değişkeni

### Web Ortamında Veri Analizi

**Bir aralık belirleyiniz.**

1 150 300

1 31 61 91 121 151 181 211 241 271 300

**Bir değişken seciniz:**

time

**Grafik rengi belirleyiniz:**

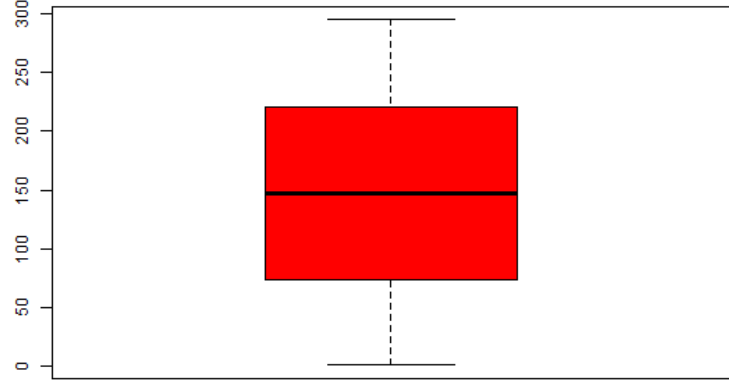
☐ green

☒ red

☐ yellow

☐ purple

Summary Str Veri Histogram Kutu Grafiği Normallik Grafiği



State değişkenine ait veri değerleri normal dağılım gösterir.

## Net Değişkeni

### Web Ortamında Veri Analizi

**Bir aralık belirleyiniz.**

1 150 300

1 31 61 91 121 151 181 211 241 271 300

**Bir değişken seciniz:**

net

**Grafik rengi belirleyiniz:**

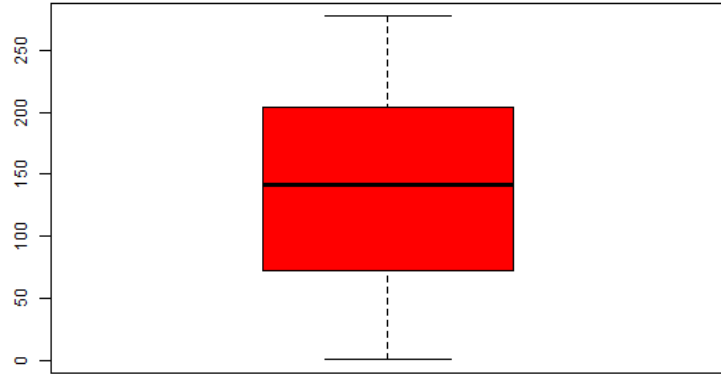
☐ green

☒ red

☐ yellow

☐ purple

Summary Str Veri Histogram Kutu Grafiği Normallik Grafiği



Net değişkeninin kutu grafiği time değişkeninin kutu grafiğine göre daha basıktır. Bunun nedeni Net isimli değişkenin maksimum değeri 277 olmasıdır. Time değişkenine göre daha basıktır ancak, yine de veri değerleri normal dağılım gösterir.

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

## Age Değişkeni

### Web Ortamında Veri Analizi

Bir aralık belirleyiniz.

1 150 300

Bir değişken seçiniz:

age

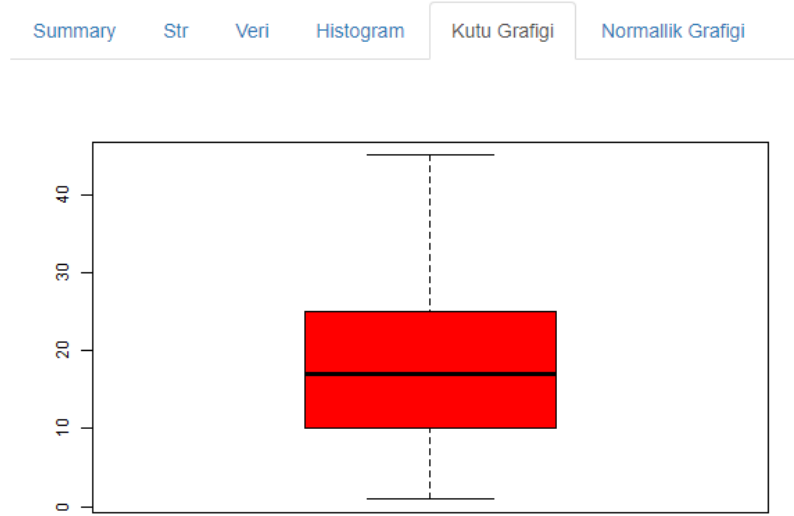
Grafik rengi belirleyiniz:

☐ green

☒ red

☐ yellow

☐ purple



1.çeyreklik=10

3.çeyreklik=25

Ortalama (mean)=18.46

Ortanca (median) = 17.00

Maksimum (max) =45.00

Age değişkeninin veri kümesinde değer aralığı küçüktür. Ancak değerlerin ortalaması küçük değerlere daha eğilimli durumdadır.

## Sex Değişkeni

### Web Ortamında Veri Analizi

Bir aralık belirleyiniz.

1 150 300

Bir değişken seçiniz:

sex

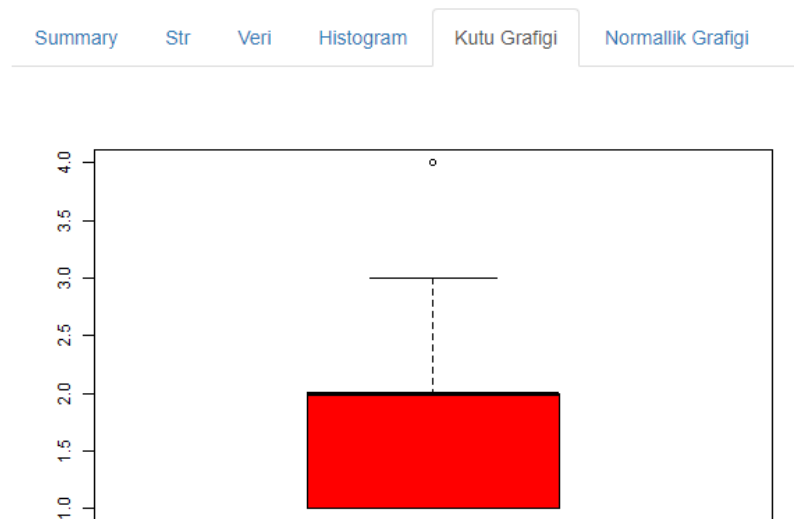
Grafik rengi belirleyiniz:

☐ green

☒ red

☐ yellow

☐ purple



Minimum değer ile 1.çeyreklik değeri eşittir. (1.00) Bu nedenle grafikte alt bıyık bulunmamaktadır. State değişkeninin kutu-bıyık grafiğinin tam tersidir.

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

## Normallik Grafiği

Veri68 isimli veri setinde “State” ve “Sex” isimli 2 adet char değişeni bulunduğundan, verilerin normal dağılım grafiğinde sapmaların yaşanması oldukça doğaldır.

### State Değişkeni

## Web Ortamında Veri Analizi

Bir aralık belirleyiniz.

1 150 300

Bir degisken seciniz:

state

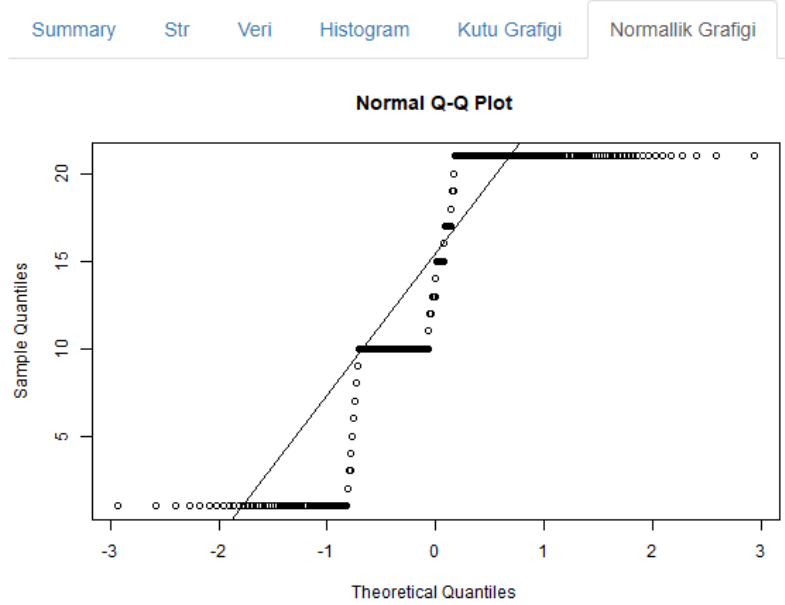
Grafik rengi belirleyiniz:

☐ green

☒ red

☐ yellow

☐ purple



### Sex Değişkeni

## Web Ortamında Veri Analizi

Bir aralık belirleyiniz.

1 150 300

Bir degisken seciniz:

sex

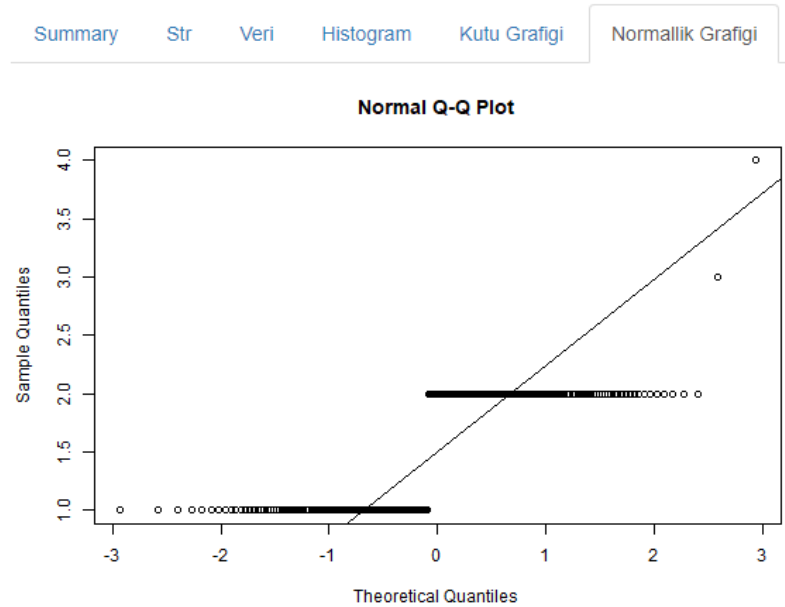
Grafik rengi belirleyiniz:

☐ green

☒ red

☐ yellow

☐ purple



# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

## Time Değişkeni

### Web Ortamında Veri Analizi

**Bir aralık belirleyiniz.**

1 150 300

**Bir degisken seciniz:**

time

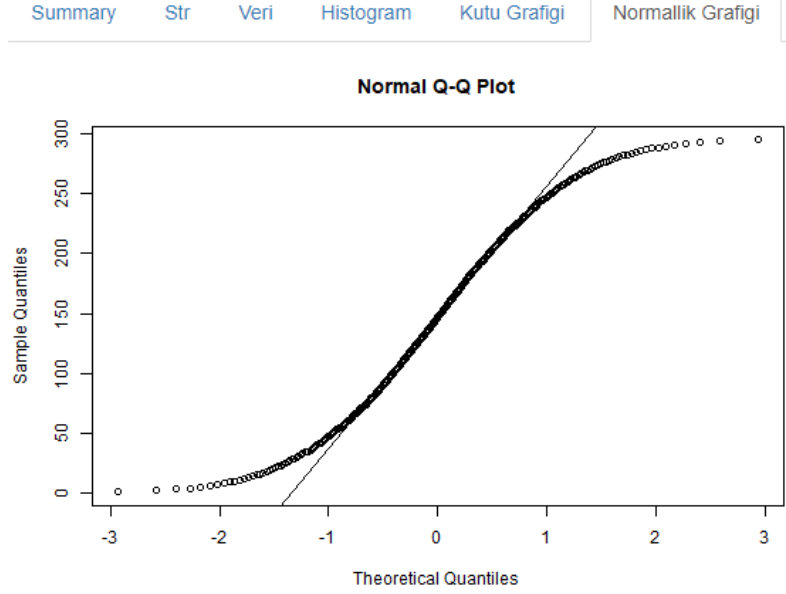
**Grafik rengi belirleyiniz:**

☐ green

☒ red

☐ yellow

☐ purple



## Net Değişkeni

### Web Ortamında Veri Analizi

**Bir aralık belirleyiniz.**

1 150 300

**Bir degisken seciniz:**

net

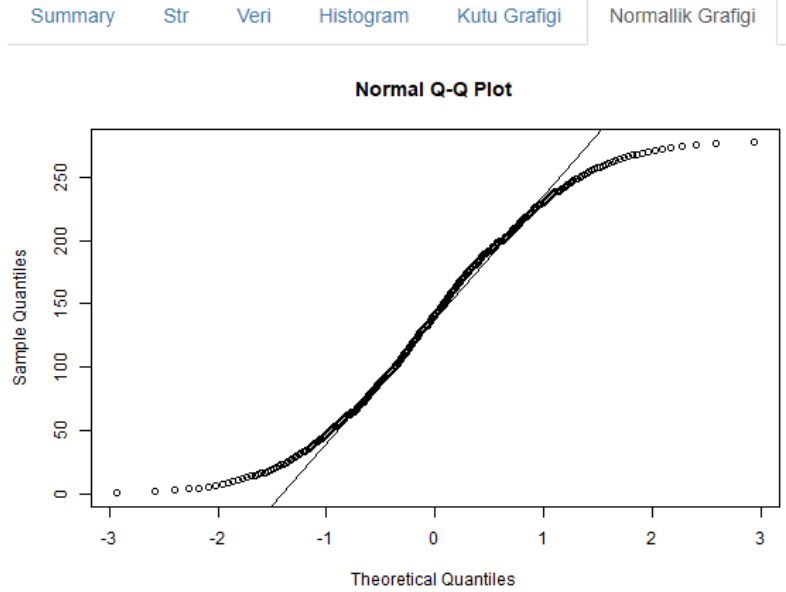
**Grafik rengi belirleyiniz:**

☐ green

☒ red

☐ yellow

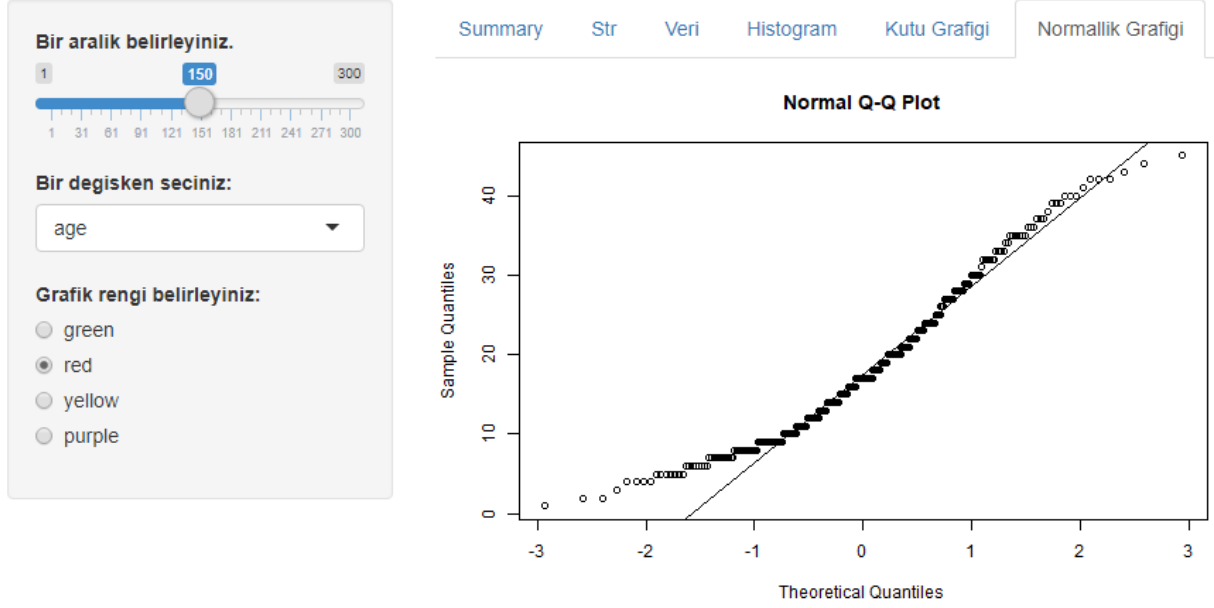
☐ purple



# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

## Age Değişkeni

### Web Ortamında Veri Analizi



Kutu grafiklerinin çıktı yorumlarında da görüldüğü üzere “Time” ve “Net” değişkenlerinin büyük bir kısmı normal dağılıma uygundur. Aynı şekilde “Age” değişkenine ait verilerin dağılımı da normal dağılıma oldukça yakındır.

# İST-347\_1 İstatistiksel Analiz Uygulamaları Araştırma Ödevi

## Kaynakça

### Kitaplar

R Yazılımına Giriş - Yrd. Doç. Dr. Özlem İlk

R İle Betimsel İstatistik – Necmi Gürsakal

### Web Siteler

Data Professor” isimli Youtube kanalından faydalandım.

[https://www.youtube.com/playlist?list=PLtqF5YXg7GLkxx\\_GGXDI\\_EiAvkhY9olbe](https://www.youtube.com/playlist?list=PLtqF5YXg7GLkxx_GGXDI_EiAvkhY9olbe)

Aynı zamanda araştırma sürecimin belli noktalarında Veri Bilimi Okulu isimli Youtube kanalından da faydalandım.

<https://www.youtube.com/channel/UC6sRvi8MsLRpeYY35EuDNuw>