

1. Data Seti - PlantGrowth

```
library(readxl)
PlantGrowth <- read_excel("2.SINIF/2. SINIF 2. DÖNEM/isttistik/BigData24/PlantGrowth.xlsx")
```

New names:

- `` -> `...1`

```
View(PlantGrowth)
```

Tanımlayıcı İstatistik - 1

"Plantgrowth" veri seti, bitki büyüme oranlarını ölçmek için yapılmış bir deneyde elde edilen verileri içeren bir R veri setidir. Bu veri seti, R programlama dilinde örnek veri seti olarak sıklıkla kullanılmaktadır. Ancak biz bu data setini R'dan direkt almak yerine veri seti üzerinde çalışma yapmış bazı kaynakların elde ettiği başka değişken verilerini de projemize dahil etmek için veri setimizi genişlettik.

"Plantgrowth" veri setinde, 30 (aynı türde) bitki örneği ve üç farklı gübre cinsi kullanılarak deneydeki değişkenlerin bitkiler üzerindeki büyüme etkileri incelenmiştir.

1. "ctrl" (kontrol): Hiçbir gübre uygulanmamış bitkiler (kontrol grubu).
2. "trt1" (tedavi 1): Verilen bir gübre türü ile beslenmiş bitkiler.
3. "trt2" (tedavi 2): Verilen diğer bir gübre türü ile beslenmiş bitkiler.
4. "Water"(Su): Bitkilere verilen toplam su miktarları.
5. "Leaves"(Yaprak sayısı): Çiçeklerin yaprak sayıları.

Herbir bitki örneği için büyüme oranı (g) ölçülmüştür ve veri setinde bu büyüme oranları yer almaktadır.

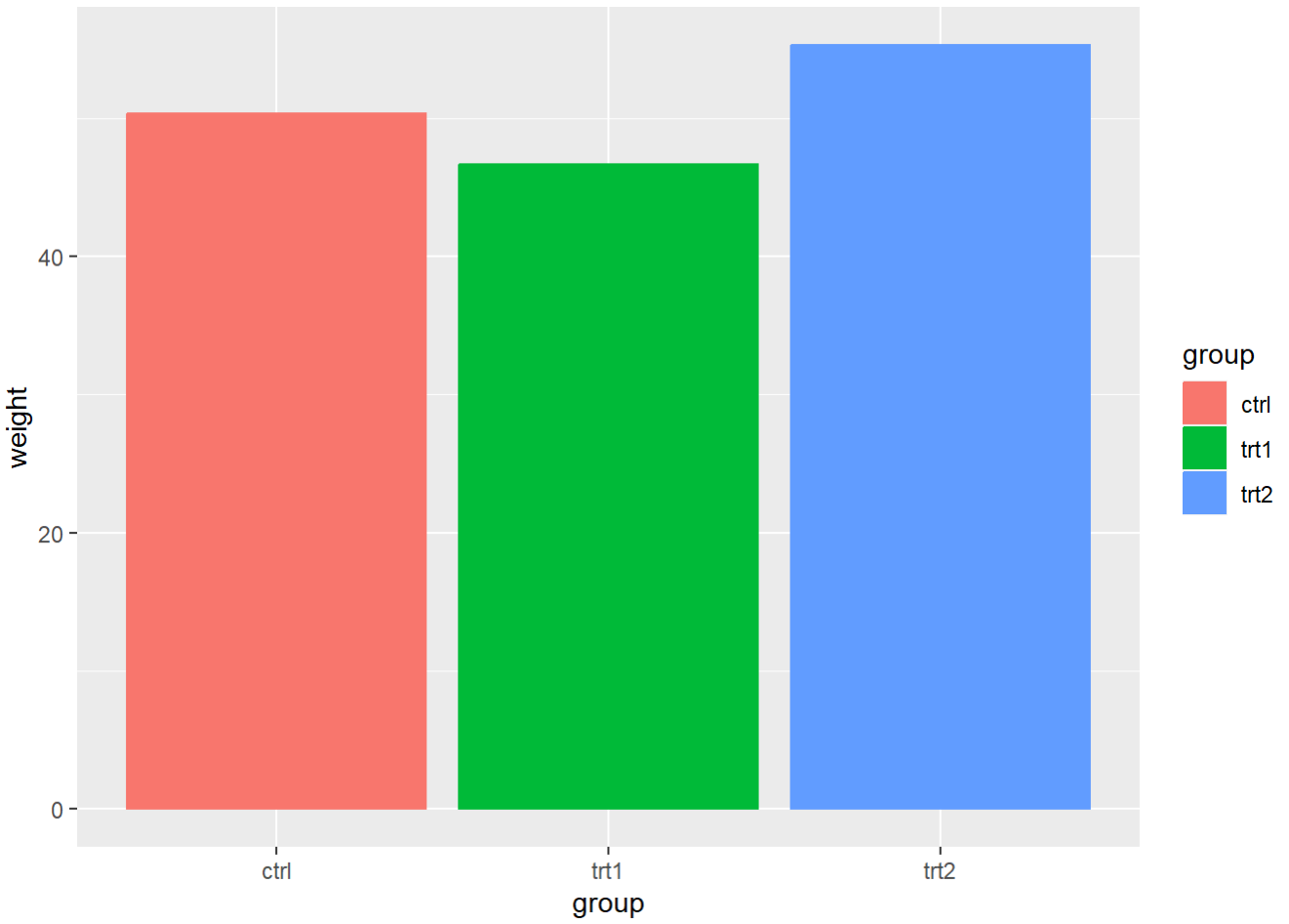
```
str(PlantGrowth)
```

```
tibble [30 × 5] (S3: tbl_df/tbl/data.frame)
 $ ...1  : num [1:30] 1 2 3 4 5 6 7 8 9 10 ...
 $ weight: num [1:30] 4.17 5.58 5.18 6.11 4.5 4.61 5.17 4.53 5.33 5.14 ...
 $ leaves: num [1:30] 1 2 3 4 5 6 7 8 9 10 ...
 $ group : chr [1:30] "ctrl" "ctrl" "ctrl" "ctrl" ...
 $ water : num [1:30] 3.63 3.56 2.23 3.15 2.38 ...
```

Grafikler - 1

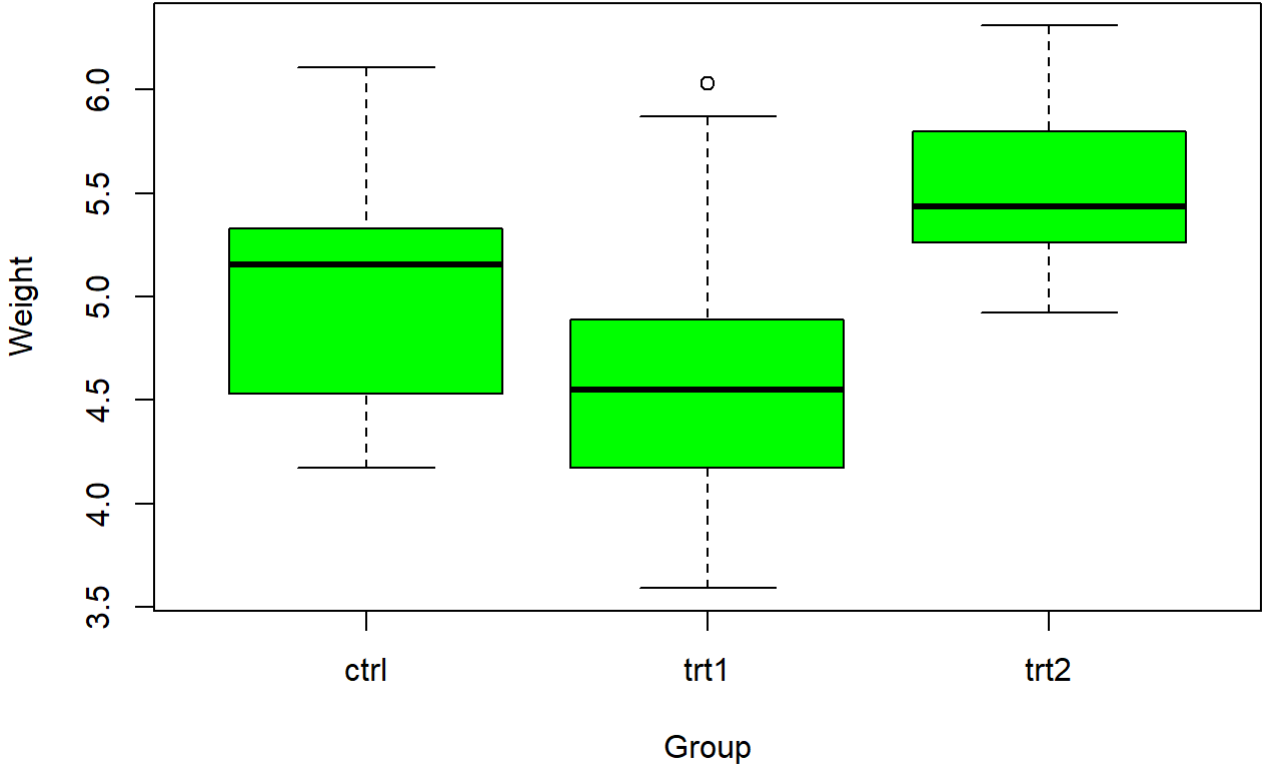
İlk olarak bitkilerin büyüme oranları ile gübreleme arasındaki bağlantıyı incelemek için gübre tipleri ve bitki uzunluklarını bir çubuk grafiği ile görselleştirdik. Bu grafikten de anlaşılacağı üzere trt2 gübresi bitkilerin uzamasını pozitif yönde etkilemektedir.

```
library(ggplot2)
ggplot(PlantGrowth, aes(x=group, y= weight)) +
  geom_col(aes(color=group, fill = group))
```



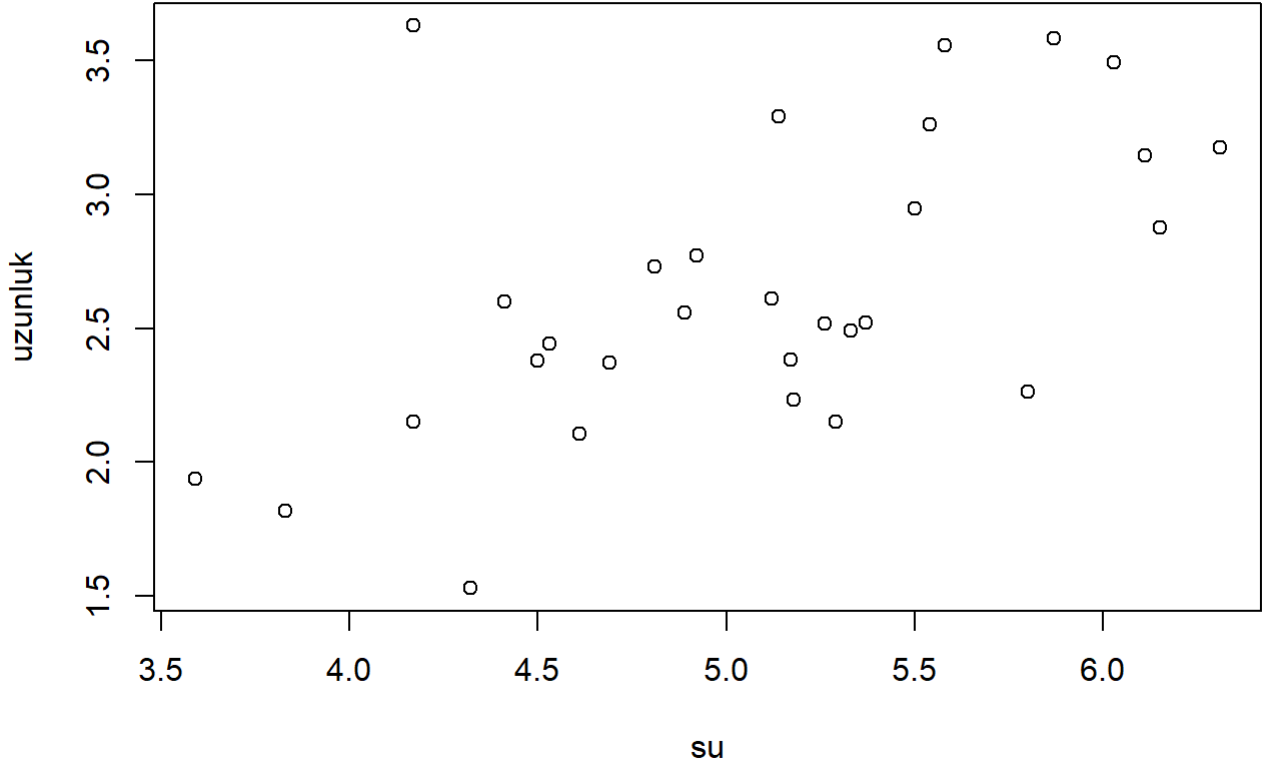
Aşağıdaki kutu grafiğinde de görüleceği üzere trt2 gübresi verilen bitkilerde (diğer değişkenlerden bağımsız olarak) daha çok büyüme gözlemlenmiştir. Bununla birlikte trt1 gübresi verilen bitkilerin boyu diğerlerinden kısa kalmıştır.

```
boxplot(weight~group, data=PlantGrowth, col="green",xlab = "Group",ylab = "Weight")
```



Aşağıdaki serpmme grafiğinde su ve uzunluk değişkenleri arasındaki bağlantıyı inceledik. Grafikte de görüleceği üzere 2 değişken arasında pozitif ve nispeten güçlü bir korelasyon vardır. Yani diğer değişkenlerden bağımsız incelendiğinde daha çok sulanan bitki daha çok uzamaktadır sonucuna varılabilir.

```
plot(PlantGrowth$water~PlantGrowth$weight,xlab="su",ylab="uzunluk")
```



Yukarıdaki grafiğin korelasyon değeri aşağıdaki gibidir.

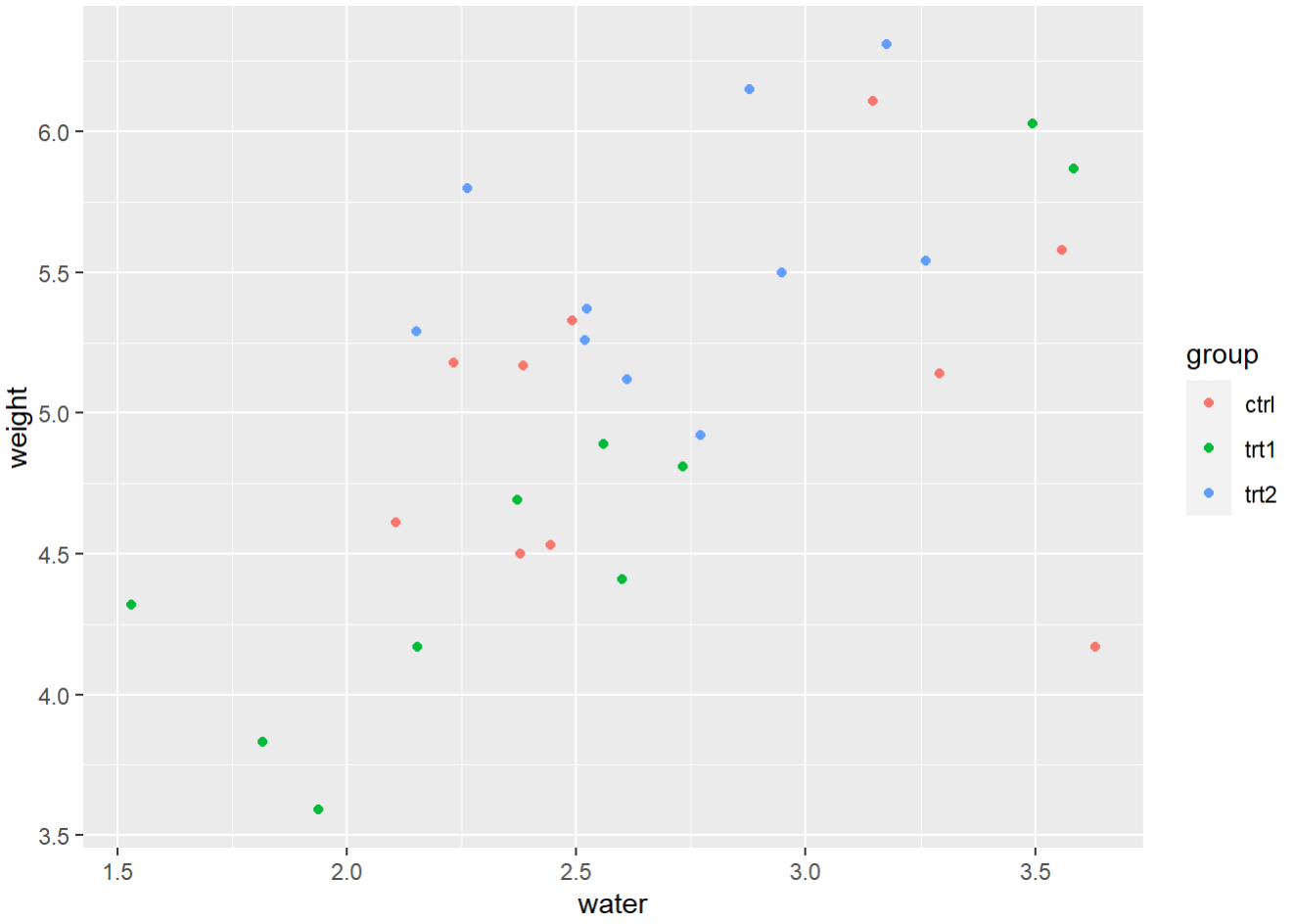
```
cor(PlantGrowth$water,PlantGrowth$weight)
```

```
[1] 0.5774683
```

Aşağıdaki grafikte su ve uzunluk değişkenlerini gübre değişkenini de hesaba katarak irdeledik. Bu grafikten şu sonuçları çıkardık:

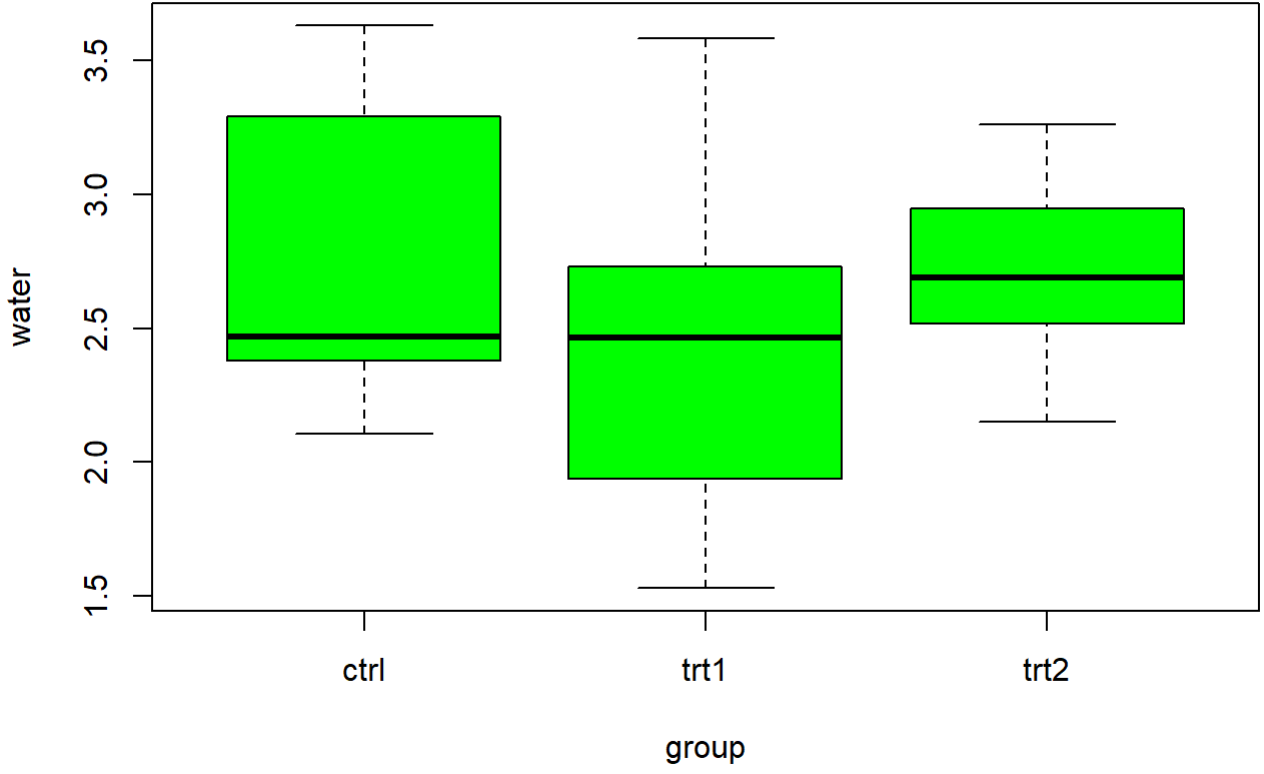
Diğer değişkenlerden bağımsız olarak bitkilerin büyümesinde maksimum verim sağlayan trt2 değişkeni, su değişkeni hesaba katıldığında da verimliliğini korumaktadır.

```
ggplot(data=PlantGrowth,aes(x=water,y=weight,color=group))+geom_point()
```



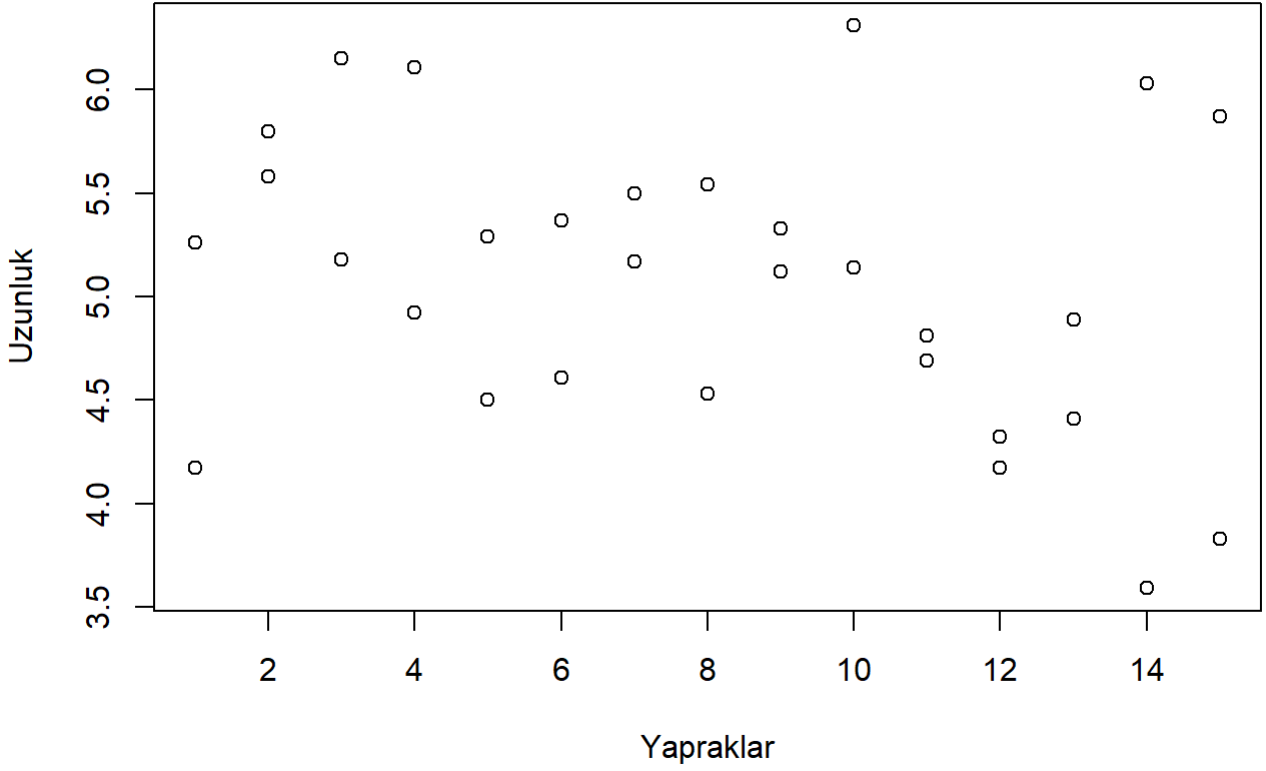
Aşağıdaki grafikte gübre miktarları ile uygulanan su miktarları arasındaki bağlantıyı inceledik. Bu grafikten elde ettiğimiz sonuçlara göre hiç gübre uygulanmayan bitkiler, gübre uygulananlara göre daha fazla suya ihtiyaç duyuyor. Bununla birlikte trt2 gübresi trt1 gübresine göre daha çok suya ihtiyaç duyuyor.

```
boxplot(water~group, data=PlantGrowth, col="green")
```



Aşağıdaki grafikte bitkilerin uzunlukları ile yaprak sayıları arasındaki bağlantıyı inceledik ve iki değişken arasında zayıf ve negatif bir korelasyon olduğunu gördük.

```
plot(PlantGrowth$weight~PlantGrowth$leaves,xlab="Yapraklar",ylab = "Uzunluk")
```



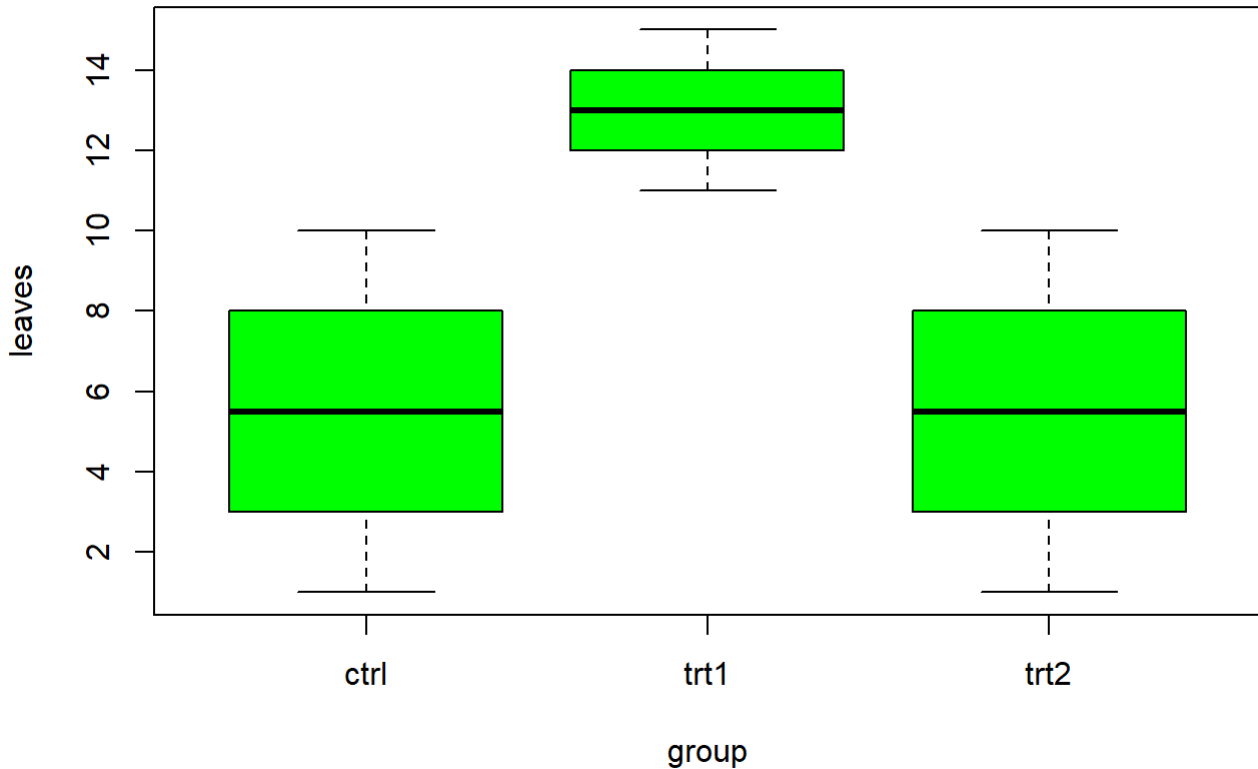
Grafiğin korelasyon değeri:

```
cor(PlantGrowth$weight,PlantGrowth$leaves)
```

```
[1] -0.3036127
```

Aşağıdaki kutu grafiğinde hiç gübre verilmeyen bitkilerin yaprak sayılarını büyük bir aralık içerisinde oldukça düşük bir değerde olduğunu, bununla birlikte trt2 gübresi yaprak sayısında anlamlı bir değişim elde edilmesine sebep olamazken trt1 gübresi yaprak sayısında büyük bir etki yarattığını gözlemledik.

```
boxplot(leaves~group, data=PlantGrowth, col="green")
```



Hipotez Testleri - 1

```
ctrl_weight=PlantGrowth$weight[PlantGrowth$group=="ctrl"]  
trt1_weight=PlantGrowth$weight[PlantGrowth$group=="trt1"]  
trt2_weight=PlantGrowth$weight[PlantGrowth$group=="trt2"]
```

İlk olarak shapiro testi ile T testi uygulayacağımız değişkenlerin normal dağılıma uygunluklarını inceledik.

```
shapiro.test(trt2_weight)
```

Shapiro-Wilk normality test

```
data:  trt2_weight  
W = 0.94101, p-value = 0.5643
```

```
shapiro.test(ctrl_weight)
```

Shapiro-Wilk normality test

```
data:  ctrl_weight  
W = 0.95668, p-value = 0.7475
```



```
shapiro.test(trt1_weight)
```

Shapiro-Wilk normality test

```
data: trt1_weight  
W = 0.93041, p-value = 0.4519
```

```
var.test(x=ctrl_weight,y=trt1_weight,ratio = 1,alternative = "t")
```

F test to compare two variances

```
data: ctrl_weight and trt1_weight  
F = 0.53974, num df = 9, denom df = 9, p-value = 0.3719  
alternative hypothesis: true ratio of variances is not equal to 1  
95 percent confidence interval:  
 0.1340645 2.1730025  
sample estimates:  
ratio of variances  
 0.5397431
```

Tanımlı veri setimizin varyansları arasında anlamlı bir fark çıkmadığı için t-testimizde varyansları eşit kabul edebiliriz.

- Kontrol grubu(hiç gübre verilmeyen grup) ile trt1 grubunun boyları arasında %5 anlam düzeyinde, anlamlı bir fark var mıdır?

H0: kontrol grubunun boyları ile trt1 grubunun boyları arasında anlamlı bir fark yoktur.

H1: Kontrol grubu ile trt1 grubu boyları arasında anlamlı bir fark vardır.

```
t.test(ctrl_weight,trt1_weight,alternative = "t",conf.level=0.95,var.equal = FALSE,paired = TRUE)
```

Paired t-test

```
data: ctrl_weight and trt1_weight  
t = 0.99384, df = 9, p-value = 0.3463  
alternative hypothesis: true mean difference is not equal to 0  
95 percent confidence interval:  
 -0.4734609 1.2154609  
sample estimates:  
mean difference  
 0.371
```

Görüldüğü üzere **H0 reddedilemez durumdadır.**

Hipotezi biraz daha değiştirip doğruluğunu kontrol etmek istersek. Gübre verilen bitkilerin boylarında pozitif yönde bir değişim olmuş mudur diye sorarsak cevabımız netleşiyor.

- Anlam düzeyimiz %5 iken trt1 gübresinin uygulandığı bitkilerin boyları hiç gübre verilmeyen bitkilerin boylarından uzundur.

H0: trt1 grubunun boyları = kontrol grubu boyları.

H1: trt1 grubunun boyları > kontrol grubunun boyları.

```
t.test(trt1_weight,ctrl_weight,alternative = "g",conf.level = 0.95,var.equal = FALSE,paired =
```

Paired t-test

```
data: trt1_weight and ctrl_weight
t = -0.99384, df = 9, p-value = 0.8269
alternative hypothesis: true mean difference is greater than 0
95 percent confidence interval:
 -1.055299      Inf
sample estimates:
mean difference
      -0.371
```

T testimizin p değeri 0.82 olarak çıkar bu da H1 in reddedildiğini gösterir. Yani trt1 gübremiz bitkilerin boylarında anlamlı bir değişim sağlamamıştır.

Kutu grafiğimizde gördüğümüz boy uzunlukları trt1 ve trt2 grubu için anlamlı bir farka sahip gibi görülmektedir (trt2 grubunun boylarının daha uzun olduğu açıkça görülür). Bunu test etmek için bit t test oluşturalım. %5 anlam düzeyinde.

H0: trt2 gübresi uygulanan bitki boyları ile trt1 gübresi uygulanan bitki boylarına eşittir .

H1: trt2 grubunun boyları trt1 grubun boylarından büyük müdür?

```
var.test(x=trt1_weight,y=trt2_weight,ratio = 1,alternative = "t")
```

F test to compare two variances

```
data: trt1_weight and trt2_weight
F = 3.216, num df = 9, denom df = 9, p-value = 0.0968
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.7988084 12.9475894
sample estimates:
ratio of variances
      3.215998
```

```
t.test(trt2_weight,trt1_weight,alternative = "g",conf.level = 0.95,paired = TRUE,var.equal = F,
```

Paired t-test

```
data: trt2_weight and trt1_weight
t = 2.8464, df = 9, p-value = 0.009602
alternative hypothesis: true mean difference is greater than 0
95 percent confidence interval:
 0.307921      Inf
sample estimates:
mean difference
      0.865
```

H0 ret sonucunu görürüz. Buna göre trt2 grubu uygulanan bitkilerin uzunluğu daha fazladır.

```
ctrl_leaves=PlantGrowth$leaves[PlantGrowth$group=="ctrl"]
trt1_leaves=PlantGrowth$leaves[PlantGrowth$group=="trt1"]
9
```

```
[1] 9
```

Aşağıdaki testte, incelediğimiz veri setlerinin normalliğe uygun olduğunu gördük.

```
shapiro.test(trt1_leaves)
```

Shapiro-Wilk normality test

```
data: trt1_leaves
W = 0.91807, p-value = 0.3411
```

```
shapiro.test(ctrl_leaves)
```

Shapiro-Wilk normality test

```
data: ctrl_leaves
W = 0.97016, p-value = 0.8924
```

Aşağıda uyguladığımız F testi ile varyanslar arasında fark olduğunu tespit ettik.

```
var.test(x=trt1_leaves,y=ctrl_leaves,ratio = 1,alternative = "t")
```

F test to compare two variances

```
data: trt1_leaves and ctrl_leaves
F = 0.24242, num df = 9, denom df = 9, p-value = 0.04639
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.06021475 0.97599858
sample estimates:
ratio of variances
      0.2424242
```

Arařtırmalar sonucu trt2 gbresi uygulanan bitkilerde yaprak sayıları yaklaşık %30 oranında artış gözlemlenmiştir. Bu bulguya göre:

H0: trt1 grubunun yaprak sayısı kontrol grubunun yaprak sayısına eşittir.

H1: trt1 grubunun yaprak sayısı kontrol grubunun yaprak sayısından büyüktür.

```
t.test(trt1_leaves,ctrl_leaves,alternative = "g",conf.level = 0.95,paired = FALSE,var.equal =
```

Welch Two Sample t-test

data: trt1_leaves and ctrl_leaves

t = 7.0278, df = 13.121, p-value = 4.26e-06

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

5.611419 Inf

sample estimates:

mean of x mean of y

13.0 5.5

H0 ret. Buna göre kullandığımız gübre türü yaprak sayısını dolaylı olarak etkileyebilir. **H1 reddedilemez.**

1. DPlantGrowth Sonuçlar:

- Trt1 gübresi herhangi bir başka değişkende anlamlı bir farklılık yokken bitkiler üzerinde herhangi bir anlamlı sonuç oluşturmamaktadır ve kontrol değişkeni olarak kabul edilen gübre verilmeyen bitkiler ile trt1 gübresi verilen bitkiler arasında anlamlı bir boy farklı oluşmamaktadır.
- Trt2 gübresi uygulanan bitkilerin boyları trt1 gübresi uygulanan ve hiçbir gübre uygulanmayan bitkilerin boylarından uzundur, yani trt2 gübresi bitkiler üzerinde pozitif bir büyüme etkisine sebep olmaktadır.
- Kullanılan gübre türü yaprak sayısını dolaylı olarak etkileyebilir.