

**İSTATİSTİKSEL ANALİZ DERSİ FİNAL RAPORU**

**HAVA KALİTESİ HAKKINDA ÇALIŞMA**

**DERSİN SORUMLUSU: ARŞ. GÖR. DR. ONUR TOKA**

**HAZIRLAYAN: BERKE YAĞMUR 21322361**

**Kullanılan Paketler**

**ggplot2:** ggplot2, grafiksel dilbilgisine dayanan, grafik oluşturma için kullanılan bir sistemdir. Ggplot2'de değişkenleri estetiğe göre nasıl haritalandıracağınızı, hangi grafiksel temel öğeleri kullanacağınızı ve ayrıntıların nasıl ele alındığını gözlemlersiniz.

**dplyr:** R dilinin en populer paketlerinden biri dplyr’dır. Dplyr paketinin SQL’e benzer komutları vardır ve bu paketi kullanarak veri üzerinde işlemler yapabiliriz.

# Hava Kalitesi Araştırmasına Giriş

Hazırlamış olduğum bu rapor New York’ta 1973 senesinde mayıs ayından eylül ayına kadar hava kalitesini yorumlamak için günlük ölçüm yapılarak alınan verilerden yararlanılarak oluşturulmuştur. Araştırma yaptığım veri seti toplamda 9 değişken ve 100 gözlemden oluşmaktadır. 9 değişkenimiz içinde açıklanacak değişkenler sırasıyla;

Ozon: Roosevelt Island üzerinde saat 13.00 ila 15.00 arası ölçülen ortalama ozon değerlerine sahip gözlemlerin bulunduğu değişken.

Solar Radyasyon: Central Park'ta sabah 08.00 ila öğlen 12.00 saatleri arasında 7-334 Angstroms frekans bandında ölçümü yapılan güneş radyasyonu gözlemlerinin bulunduğu değişken.

Rüzgar: La Guardia Havaalanı’nda sabah 07.00 ila 10.00 saatleri arasında rüzgar hızı hakkında yapılan saatlik ölçüm gözlemlerinden oluşan değişken.

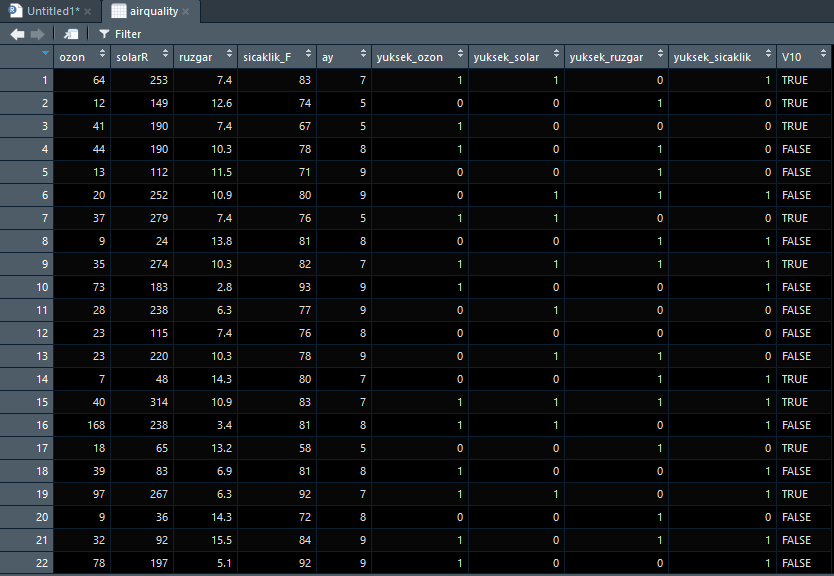
Sıcaklık: La Guardia Havalimanı'nda o gün içerisinde maksimum Fahrenheit derecesinde gözlemlenen sıcaklık.

Ay: Ay değişkenim ise ilk paragrafta belirttiğim gibi araştırmanın yapıldığı mayıs ayından eylül ayına kadar olan kısmın veri setimizde sayısal olarak ifade edildiği değişken.

Verilerimiz New York Eyaleti Koruma Bölümü'nden (ozon verileri) ve Ulusal Hava Durumu Servisi'nden elde edildi.

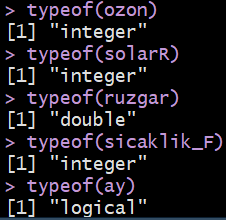
Yapılan bu araştırma hava kalitesinin durumuna dikkat çekmek için yapılmıştır. Araştırma genelinde istatistiksel olarak getirdiğimiz yorumlar durumu açıklar niteliktedir. İstatistiksel çalışmaların bir amacı da bilindiği üzere sorunun yapısını analiz edip ona çözümler üretmek üzerinedir. Raporun ilerleyen bölümlerinde bunlara değinilecektir.

## Veri Seti Hakkında Gerekli Açıklamalar

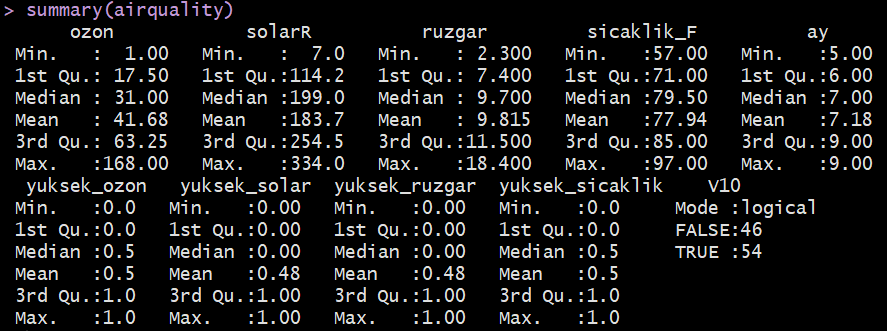


Elimizde temelde 5 esas değişken ve bu 5 esas değişkenin 4’ünden türetilmiş 9 değişkene sahip bir veri seti bulunmaktadır. “V10” adıyla listelenen değişkenimiz ise 5. Değişken olan ay değişkeninin kategorik değişkene dönüştürülmesinden elde edilmiştir. Kategorik değişkene dönüştürürken baz aldığımız 7. Ay ve daha önceki aylar “TRUE” 7. Aydan daha sonraki aylar ise “FALSE” olarak gösterilmiştir. Ay değişkenini kategorik değişkene çevirme amacımız ise veri setim üzerinde yapacağım analizlere yorum yaparken kolaylık sağlamak üzerinedir. Böylelikle ele alacağım analizlerde 7. Aydan önce veya 7. Aydan sonra gözlemlenen verilerin değişimlerini daha kolay yorumlayabiliriz.

Veri setimdeki değişkenlerimin açıklaması ise aşağıdaki R çıktısında görüldüğü üzeredir;



Son olarak ise veri setimin değişkenlerinin istatistiksel tablosunu verecek summary() komutunu kullanıyorum. Değişkenlerin ortalaması, çeyreklik değerleri, ortancası, minimum ve maksimum değerleri hakkında bilgiye sahip oluyorum. Analizlerimde yorumlama yaparken kullanabileceğim genel bir bakış açısına sahip olmak adına değişkenlerimizi bu şekilde tanıyoruz.



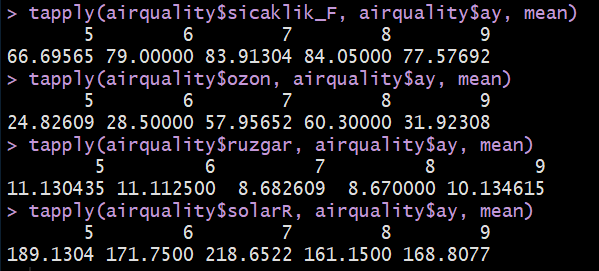
## Tapply() komutu ile aylara göre belirlediğim değişkenlerin tablolarını oluşturmak.

**#tapply(airquality$sicaklik\_F, airquality$ay, mean)**

**#tapply(airquality$ozon, airquality$ay, mean)**

**#tapply(airquality$ruzgar, airquality$ay, mean)**

**#tapply(airquality$solarR, airquality$ay, mean)**



Tabloda görüldüğü üzere 5. Aydan 9. Aya kadar olan sıcaklık değişkenim, ozon değişkenim, rüzgar değişkenim ve güneş ışığı değişkenim için ayrı ayrı ortalamaları hesaplattım. Sıcaklık ortalama olarak 7. Ve 8. Ayda maksimum ortalamada seyrederken, ozon konusunda da 7. Ve 8. Ayda ortalama maksimum seyretmiştir. Rüzgar hızı değişkenimin ortalaması ise 5. Ve 6. Ayda maksimum noktalarındadır. Güneş ışığı değişkenim 7. Ayda ortalama olarak maksimum değere ulaşmıştır. İstatistikte ortalama yorumlamanın etkisi çok önemlidir.

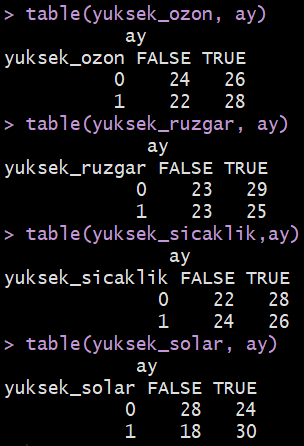
## Table() komutunu kullanarak tablolar oluşturup fikir elde ediyorum.

**#table(yuksek\_ozon, ay)**

**#table(yuksek\_ruzgar, ay)**

**#table(yuksek\_sicaklik,ay)**

**#table(yuksek\_solar, ay)**



Yüksek ozon değişkenimin ortancadan büyük ve küçük olduğu durumları kategorize ettiğim ay değişkenine göre değerlendirip 7. Aydan önce ve 7. Aydan sonra ne kadar gözlemlendiğine bakıyorum.

Yüksek ozon değerleri için ortancadan küçük olan gözlemler 7. Aydan önce 26 iken 7. Aydan sonra 24 kere gözlenmiştir.

Yüksek ozon değerleri için ortancadan büyük olan gözlemler ise 7. Aydan önce 28, 7. Aydan sonra 22 olarak gözlenmiştir.

Yüksek rüzgar değişkenimin ortancadan büyük değerleri için 7. Aydan önce 25 7. Aydan sonra 23 olarak gözlemlenmiştir.

Bu şekilde yorumlar tabloda gözlendiği gibi çıkarım yapılarak bulunabilir.

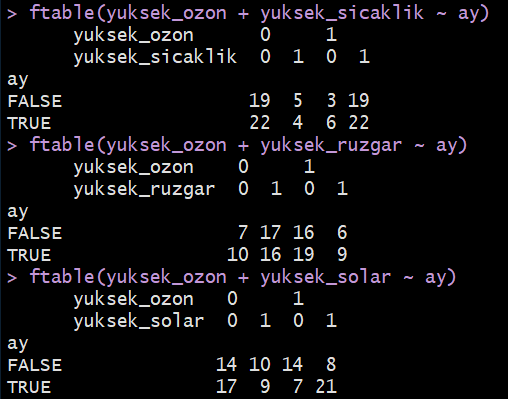
## Ftable() komutu kullanarak tabloları yorumlama

Biraz daha geliştirip iki değişkenimin birbiri içindeki tablosunu ftable komutuyla oluşturacağım.

**#ftable(yuksek\_ozon + yuksek\_sicaklik ~ ay)**

**#ftable(yuksek\_ozon + yuksek\_ruzgar ~ ay)**

**#ftable(yuksek\_ozon + yuksek\_solar ~ ay)**



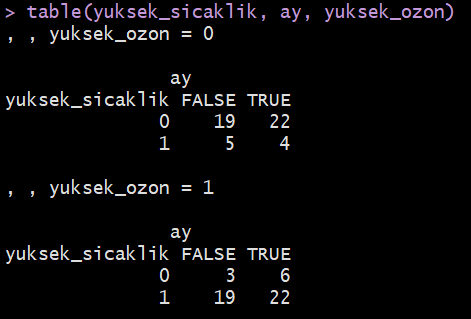
Yüksek ozon değişkenimin ortancadan düşük olduğu durumlarda yüksek sıcaklık değişkenimin ortancadan düşük ve büyük değerleri içinde kaç kere gözlemlendiği gösteren bir tablo oluşturdum. Burdan yüksek ile başlayan bütün değişkenlerim için türetilebilir ftable komutlarıyla bir izlenim sahibi olabilirim. Örneğin 1. Tabloda görülen ortancadan düşük ozon değerlerim için ortancadan büyük yüksek sıcaklık değerlerimin tam tersi duruma göre çok az gözlendiği söylenebilir.

## İkili grupları karşılaştırıp tablolarına ulaşmak

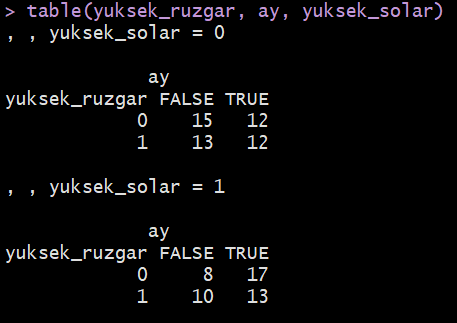
**#table(yuksek\_sicaklik, ay, yuksek\_ozon)**

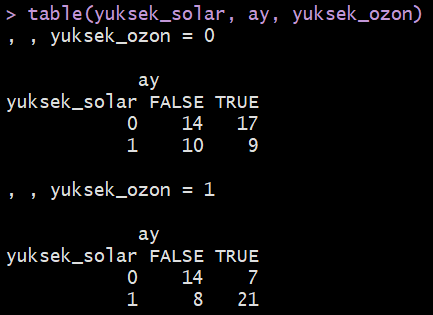
**#table(yuksek\_ruzgar, ay, yuksek\_solar)**

**#table(yuksek\_solar, ay, yuksek\_ozon)**



Yüksek ozon değişkenimin ortancadan küçük değerleri için yüksek sıcaklığın aylara göre nasıl gözlemler ürettiğini gösteren tablodur. İkili grupların karşılaştırılmasında kullanılır.

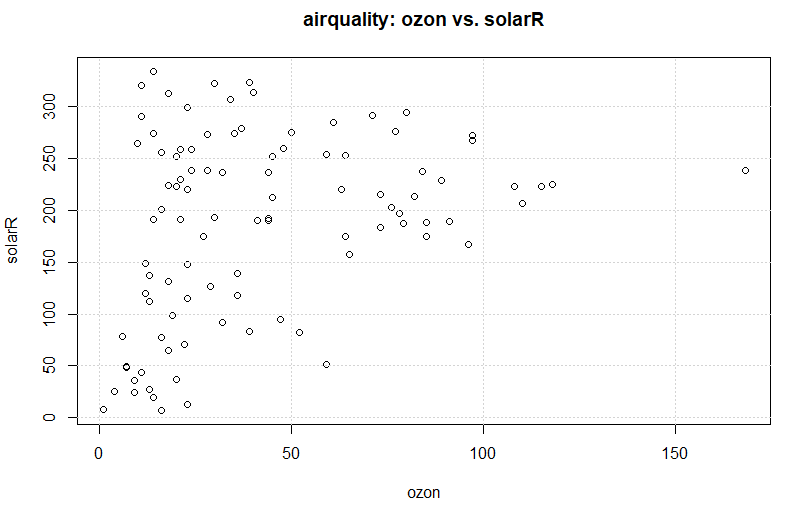


****

# Veri Setini Analiz Etme

### **1-) plot(x,y)** komutunu kullanarak iki değişkeni karşılaştırıp bir fikir edinelim.

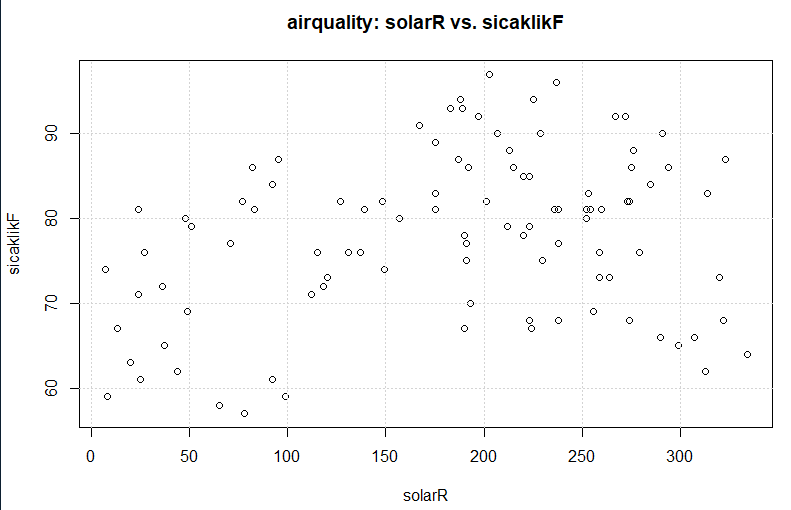
**#plot(airquality$ozon, airquality$solarR ,main="airquality: ozon vs. solarR", xlab="ozon", ylab="solarR")**



**---** Grafikten görüldüğü üzere ozon değerinin az olduğu bölgede güneş ışığı radyasyonunda bir artış olduğu gözüküyor. Bunun aksi durumunda yani ozon değerinin yüksek gözlemlendiği kısımlarda (130’dan büyük olduğu) güneş ışığı radyasyonu neredeyse gözlemlenmemiştir. Ozon miktarı ile güneş radyasyonu değerlerimiz arasında bir ters ilişki mevcuttur.

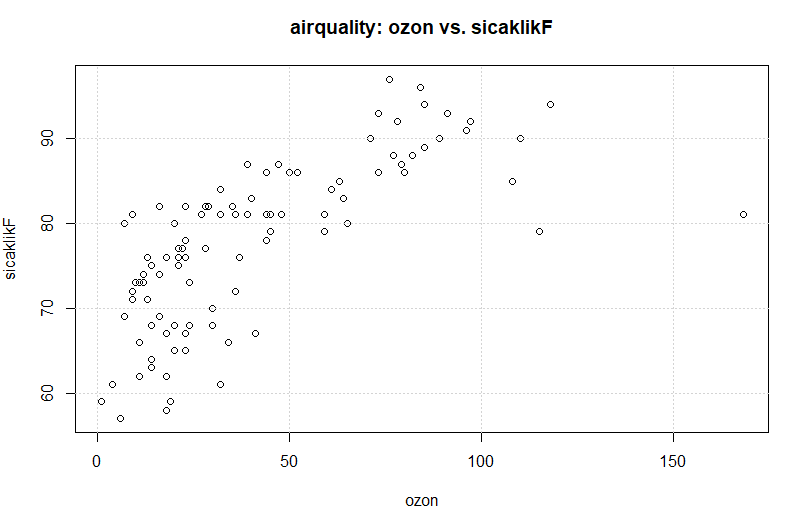
**#plot(airquality$solarR, airquality$sicaklik\_F, main="airquality: solarR vs. sicaklikF", xlab = "solarR", ylab = "sicaklikF")**

**grid()**



**---** Solar radyasyon ve sıcaklık değişkenlerini karşılaştırdığımda saçılımın dağılarak gittiğini gözlemliyorum. Aykırı değerlerin yoğunlukta olduğu kanısına varılabilir. Ayrıca 100 birim solar radyasyondan büyük olan birimlerde sıcaklık düşükse (60 derece altında) güneş radyasyonunun fazla gözlenmediğine kanaat getirdik.

**---** İlk iki saçılım grafiğimde güneş radyasyon değerinin ozona ve sıcaklığa olan ilişkisini gözlemledim. Bu bölümde son olarak yapacağım işlem ise ozon değişkeninin sıcaklıkla olan dağılım grafiğini oluşturmak olacak.



Ozon biriminin düşük olduğu yerlerde (0 ile 100 arası) gözlemlerin çok olduğu görülüyor. Ozon miktarı ile sıcaklık arasında ters bir ilişki vardır.

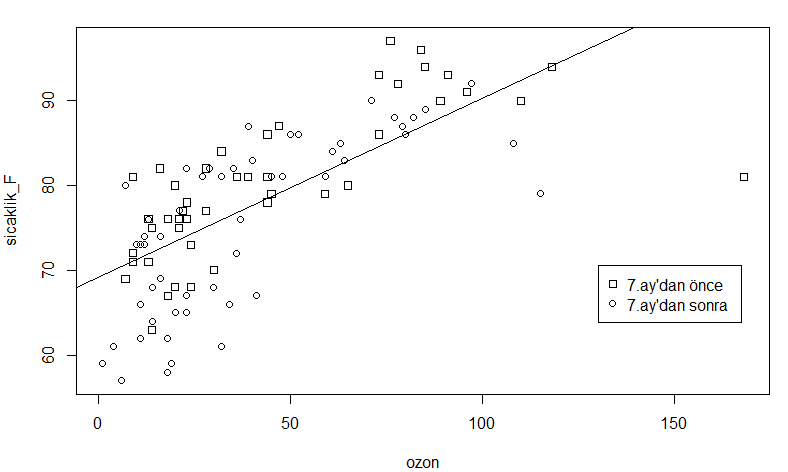
**---** Saçılım grafiğini biraz daha geliştirip legend komutunu ekliyorum.

**#plot(ozon, sicaklik\_F, pch=as.integer(ay))**

**legend(130.140, 70.60, c("7.ay'dan önce" , "7.ay'dan sonra"), pch=0:1)**

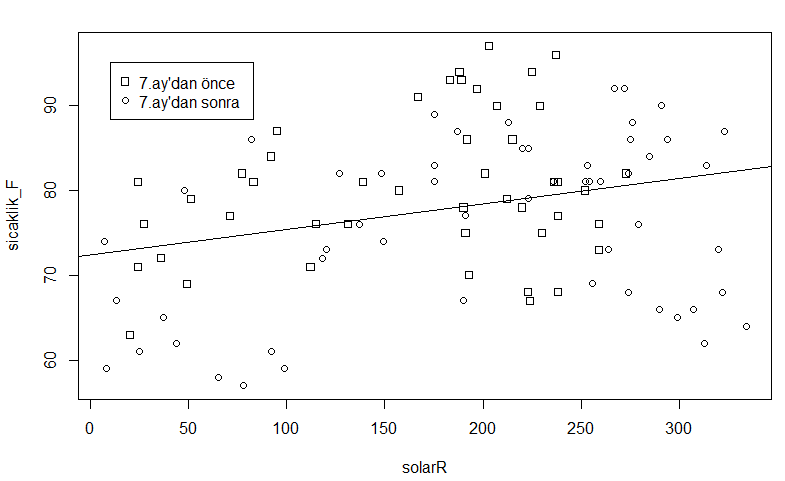
**a <- lm(sicaklik\_F ~ ozon)**

**abline(a)**

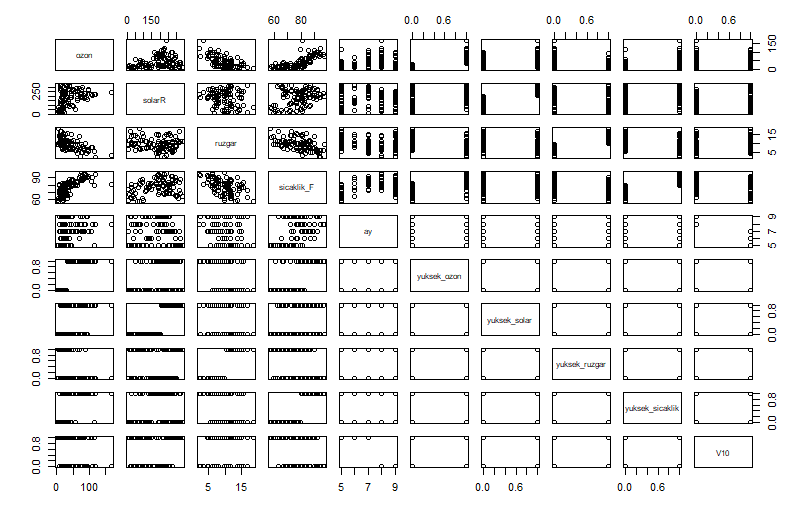


**---** Saçılım grafiğimi 7. Aydan önce ve 7. Aydan sonra diye özelleştirdim. Abline komutuyla **regresyon** çizgisini çizdirdim.

**---** Şimdi ise solarR değişkenim ile sıcaklık değişkenimi karşılaştırıyorum.



**---** #**plot(airquality[,1:10])**



Bütün değişkenlerimin birbirine göre saçılım grafiklerini çizdirdim.

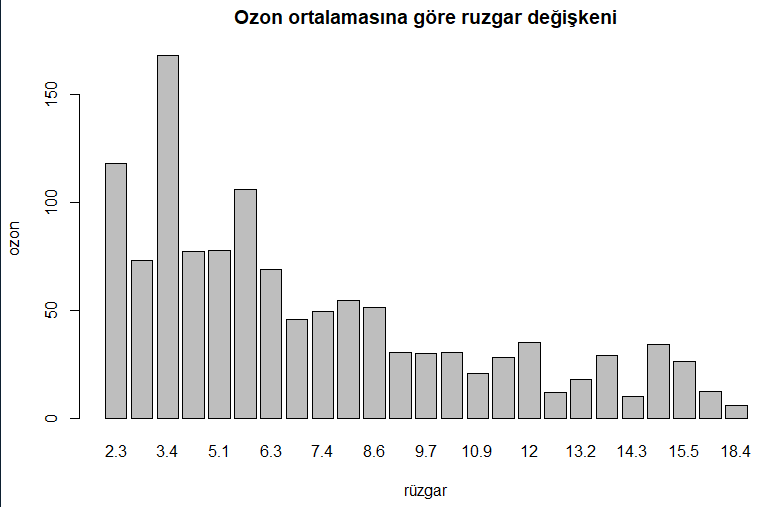
Amaca yönelik grafik istenmediğinde yukarıdaki gibi yapıp değişkenlerimi yorumlamam bana kolaylık sağlayacaktır. Genel bakış açısı için bu bahsettiğim geçerlidir. Abline komutuyla regresyon çizgisi çekip yorumlamak için özel olarak o grafiği büyük fontta çıkarmak her zaman daha avantajlıdır.

### 2-) barplot() komutunu kullanarak çubuk grafiği çizdirmek

**#agirlikli <- tapply(airquality$ozon, airquality$ruzgar, mean)**

**barplot(agirlikli)**

**barplot(agirlikli, main="Ozon ortalamasına göre ruzgar değişkeni",xlab = "rüzgar", ylab="ozon")**

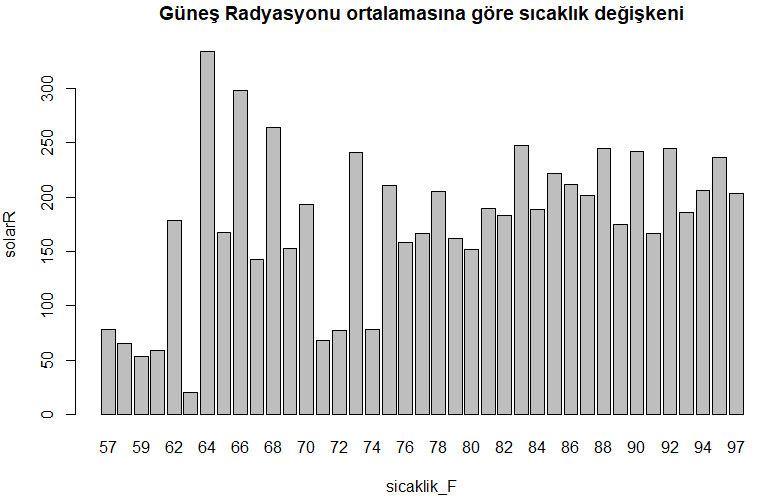


**---**  Ozon miktarının en yüksek olduğu gözlem rüzgar hızının 3.4 olduğu anında görülmüştür. Ozon miktarının en düşük olduğu gözlemde ise rüzgar hızı 18.4’tür. Rüzgar hızının 8.7 ila 12 arasında ortalama ozon değeri bakımından aynı seyrettiği görülmüştür.

**agirlikli2 <- tapply(airquality$solarR, airquality$sicaklik\_F, mean)**

**barplot(agirlikli2)**

**barplot(agirlikli2, main="Güneş Radyasyonu ortalamasına göre sıcaklık değişkeni",xlab = "sicaklik\_F", ylab="solarR")**

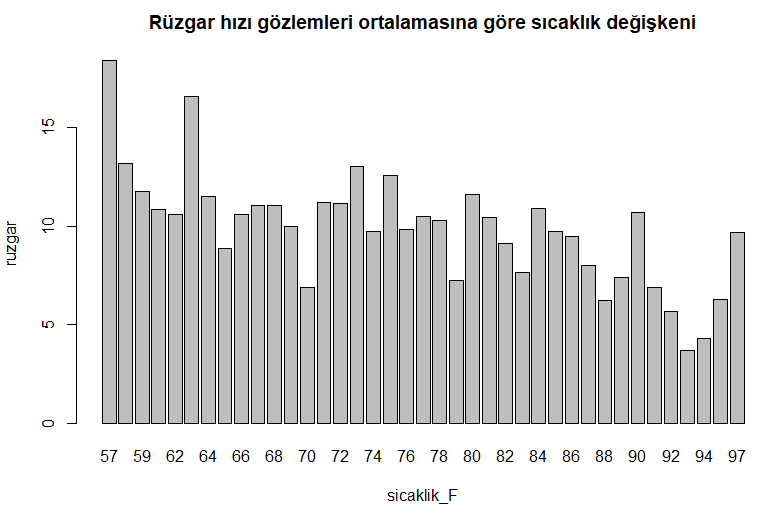


**---** Sıcaklığın 64F olduğu gözlem değerinde solarR en fazla değerine ulaşmıştır. Bu tabloda ilginç olan ise sıcaklığın 1F azaldığında solarR’nin gözlem değerinin en küçükte olmasıdır. Sıcaklık değişkeniyle güneş radyasyonu değişkeninin arasında anlamlı bir ilişki olmadığını söyleyebiliriz. Birbirlerini açıklamada yetersizdirler.

**#agirlikli3 <- tapply(airquality$ruzgar, airquality$sicaklik\_F, mean)**

**barplot(agirlikli3)**

**barplot(agirlikli3, main="Rüzgar hızı gözlemleri ortalamasına göre sıcaklık değişkeni",xlab = "sicaklik\_F", ylab="ruzgar")**



Rüzgar hızının en düşük gözlemlendiği değer 93F’dır. Onun haricinde düşük sıcaklıklarda kısmen rüzgar hızının daha fazla olduğunu görüyoruz.

### 3-) boxplot() komutunu kullanarak kutu grafiği çizdirmek

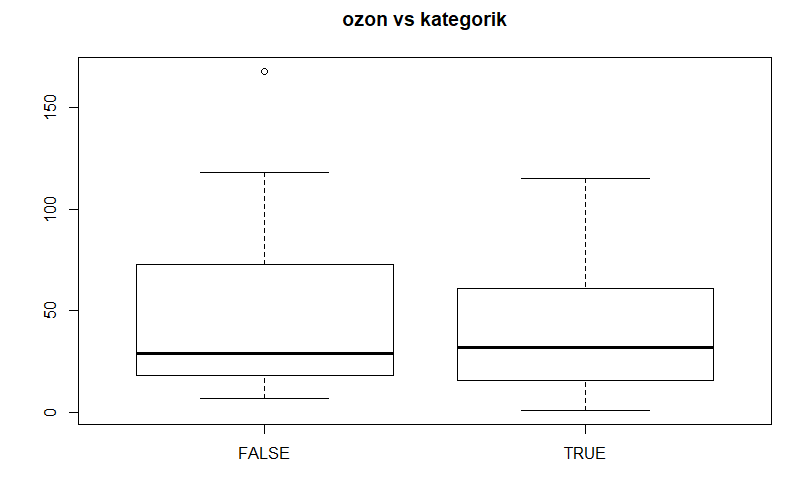
**#boxplot(ozon ~ V10, data=airquality, main="ozon vs kategorik")**

**#boxplot(solarR ~ V10, data=airquality, main= "solarR vs kategorik")**

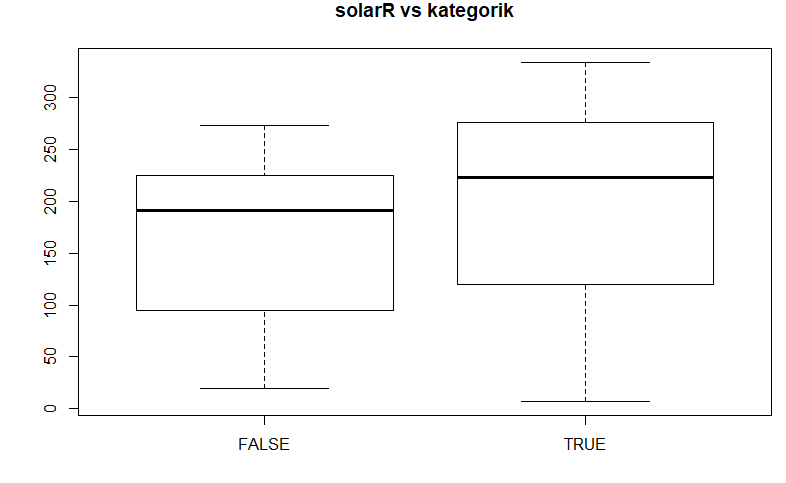
**#boxplot(ruzgar ~ V10, data=airquality, main= "ruzgar vs kategorik")**

**#boxplot(sicaklik\_F ~ V10, data=airquality, main= "sicaklik vs kategorik")**

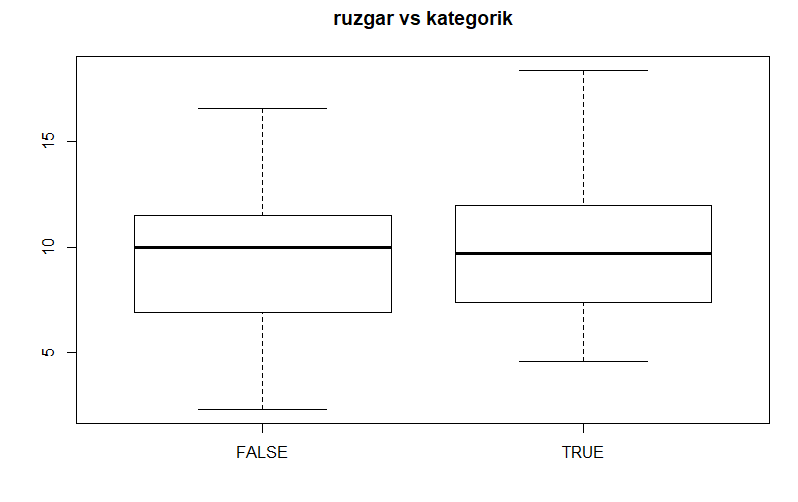
Sırasıyla kutu grafiklerini çizdirdiğim değişkenler.



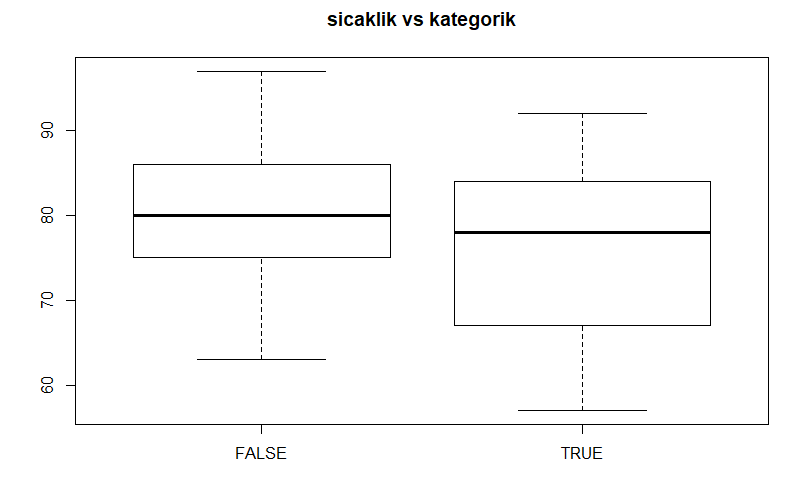
**---** Ozon değişkeninin ortalamasının 7. Aydan sonra ve 7. Aydan önce hemen hemen aynı olduğunu ve 40 civarlarında olduğunu söyleyebiliriz. Ozon değişkenim kategorize ettiğim aylara göre hemen hemen eşit dağılmış diyebilirim.



**---** SolarR değişkenimin ise 7. Aydan sonra ortalamasının 190 civarlarında, 7. Aydan önceyse 220 civarlarında olduğunu söyleyebilirim.



**---** Rüzgar değişkenimin ise kategorize ettiğim aylara göre ortalamada farklılık göstermediğini aynı zamanda çeyreklik değerlerde de hemen hemen aynı çizgide seyrettiğini söyleyebilirim.



**---** Sıcaklık değişkenim ise 7. Aydan sonra ortalama 80 F civarında iken 7. Aydan önce 77F civarlarında seyretmiştir. Ortalamalar bakımın çok fark olmasa da çeyreklik değerler arasında bir fark olduğunu söyleyebiliriz.

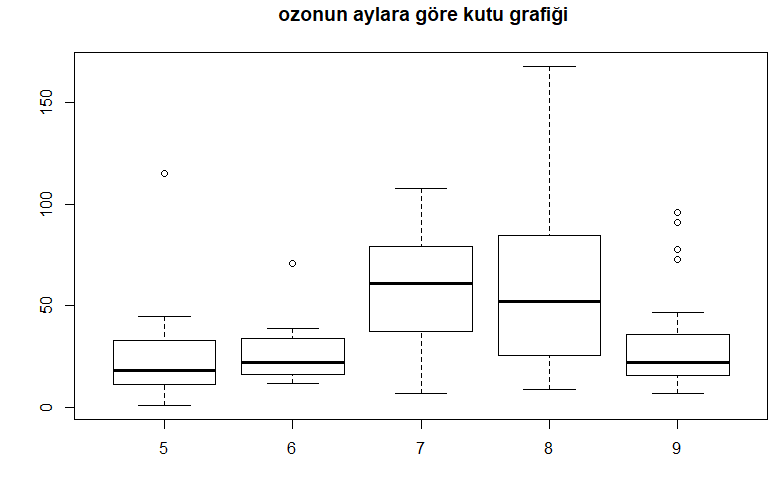
### 4-) boxplot grafiğini oluşturup 5 ay için ayrı ayrı kutu grafikleri çıkarmak

**# boxplot(ozon ~ ay , data=airquality, main = " ozonun aylara göre kutu grafiği")**

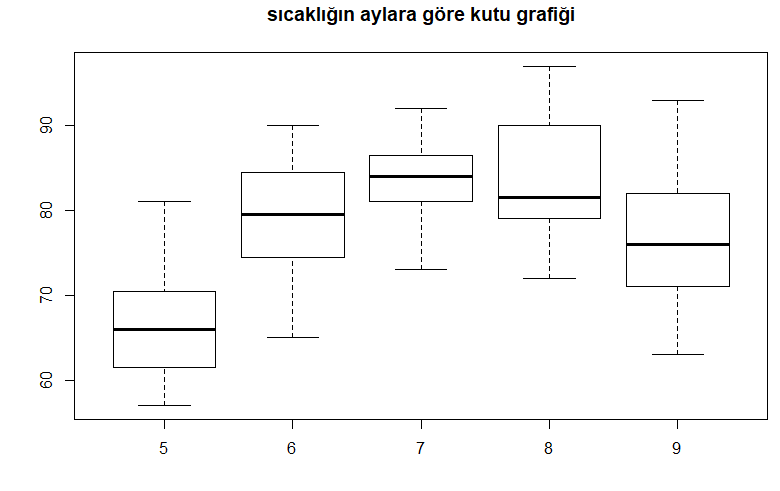
**#boxplot(sicaklik\_F ~ ay , data=airquality, main = "sıcaklığın aylara göre kutu grafiği")**

**#boxplot(solarR ~ ay , data=airquality, main = " Güneş radyasyonunun aylara göre kutu grafiği")**

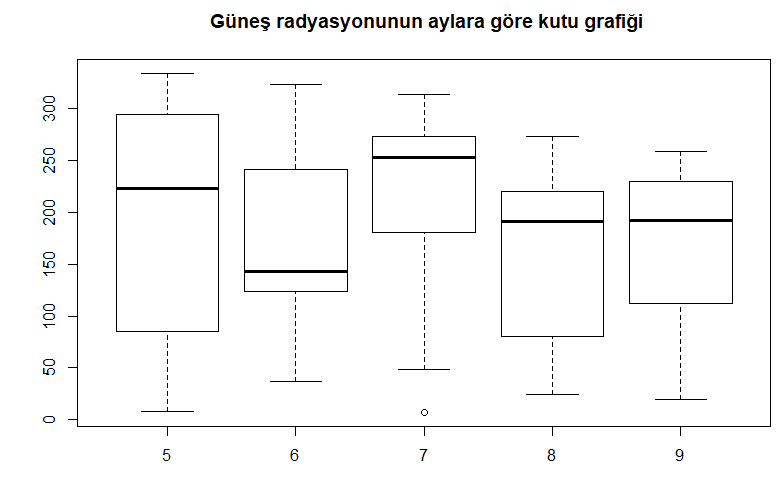
**#boxplot(ruzgar ~ ay , data=airquality, main = " rüzgar hızının aylara göre kutu grafiği")**



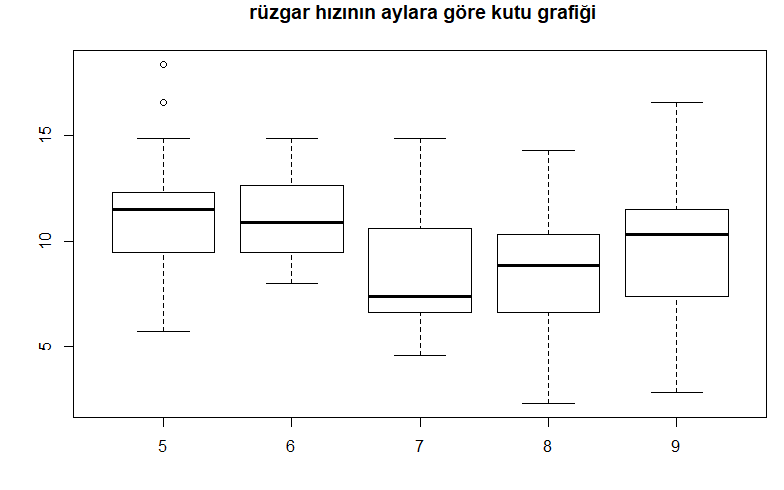
**---** Görüldüğü üzere 5. 6. Ve 9. Ayda aykırı değerler bulunmaktadır. Onun haricinde ozon ortalamasının en yüksek olduğu ay olarak 7. Ay söylenebilir. Çeyrek değerler kısmının uzunluğu hakkında ise 8. Ayda en büyük genişliğin gözlendiği söylenebilir. En çok 8. Ayda ozon miktarları arasında dalgalanma olduğu gözükmüştür.



**---** Aykırı değerlerin gözlenmediği görünmüştür. Sıcaklığın ortalaması yine 7. Ayda en çok olmuştur. 5. Aydan 7. Aya sıcaklık ortalamalarında bir artış gözükmüş, 7. Aydan 9. Aya sıcaklık ortalamasının düşüşe geçtiği gözlemlenmiştir. 8. Ayda gözlenen çeyrek değerler en geniş aralığa sahiptir.



**---** 7. ayın güneş radyasyonu değişkeninde bir aykırı değer gözükmektedir. Güneş radyasyonu değişkeni için ortalamanın en yukarda olduğu ay 7. Aydır. En düşük ortalama gözlendiği ay ise 6. Aydır. Çeyrek değerlerin genişliğinin en fazla gözlendiği ay ise 5. Aydır. En çok güneş radyasyonu değişimi dalgalanması 5. Ayda meydana gelmiştir.



**---** Rüzgar hızı ortalaması için en yüksek ortalamaya sahip ay 5. Ay olarak gözlemlenmiştir. Bunun yanı sıra çeyrek değerler arasındaki genişliğin en fazla olduğu ay 7. Aydır.

### 5-) hist() komutunu kullanarak histogram grafiği oluşturmak

**#hist(ozon, 30, prob=T)**

**#lines(density(ozon))**

**#hist(sicaklik\_F, 30, prob=T)**

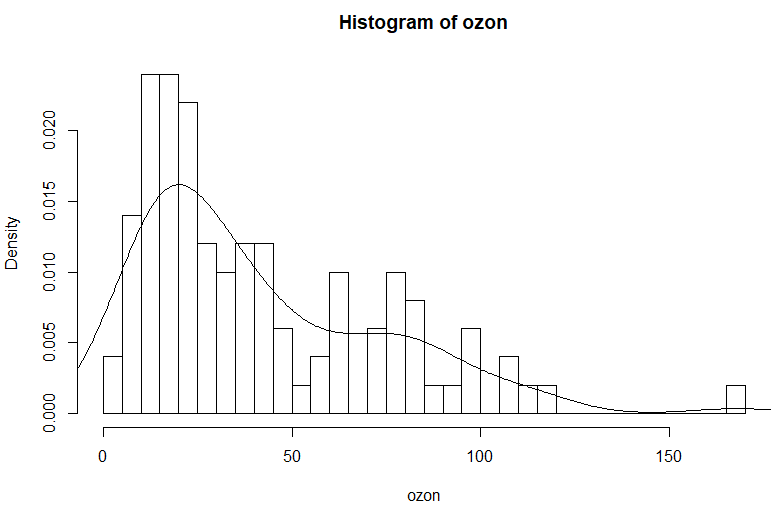
**#lines(density(sicaklik\_F))**

**#hist(ruzgar, 20, prob=T)**

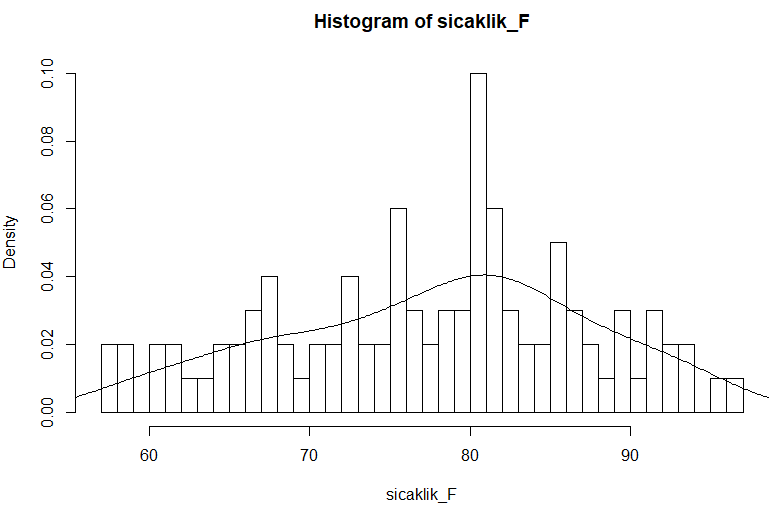
**#lines(density(ruzgar))**

**#hist(solarR, 30, prob=T)**

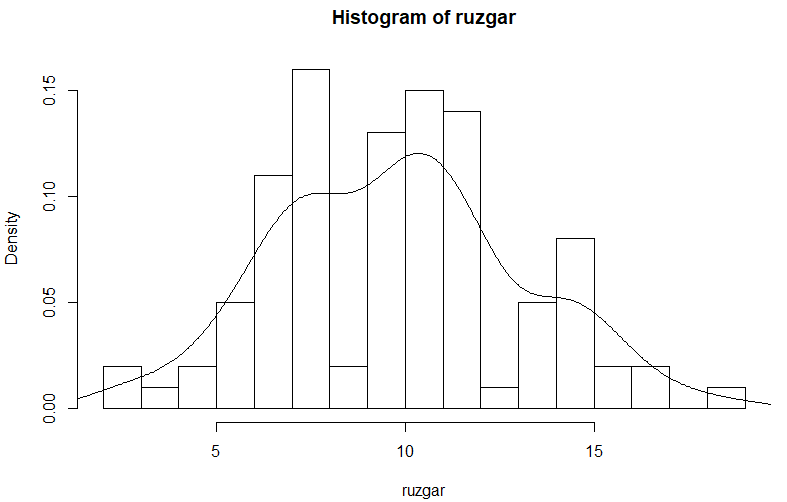
**#lines(density(solarR))**



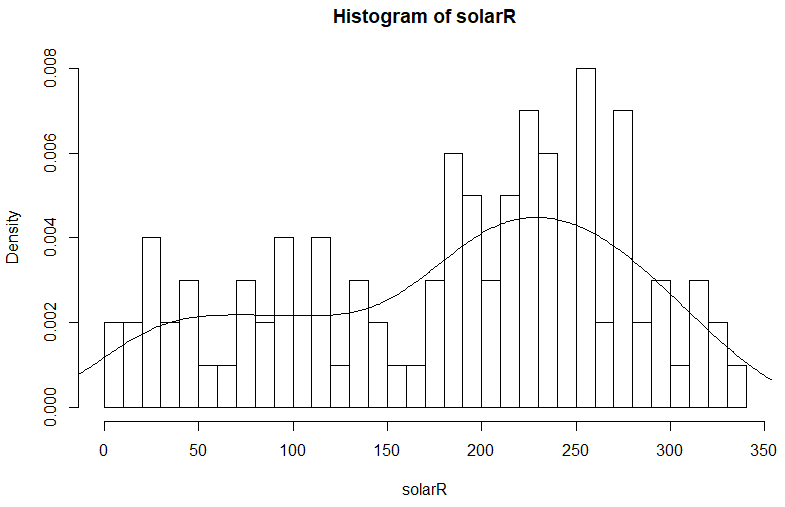
**---** Ozon değişkenin histogram grafiğini çizdirdiğimde ve yoğunluk çizgisi eklediğimde elde ettiğim sonuç yukarıdaki gibidir. Bu grafikte maksimum değerin 15-20 birimler arasında olduğu gözüküyor. Minimum değer ise 50’nin biraz üstünde seyrediyor. Yoğunluğun 10 ila 30 arasında en çok olduğunu söyleyebilirim.



**---** Sıcaklık değişkenin histogram grafiğini çizdirdiğimde ve yoğunluk çizgisi eklediğimde elde ettiğim sonuç yukarıdaki gibidir. Maksimum değerin 82F olduğunu görmekteyim. Yoğunluk çizgisinin en yüksek seyrettiği değer ise 80F ila 82F arasıdır.



**---** Rüzgar değişkenin histogram grafiğini çizdirdiğimde ve yoğunluk çizgisi eklediğimde elde ettiğim sonuç yukarıdaki gibidir. Maksimum gözlem değeri rüzgar hızının 8 civarında olduğu gözlemlerde gözlemlenmiştir. Minimum değer ise rüzgar hızının 3 ve 13 olduğu değerlerde gözlenmiştir. Yoğunluğun en çok 10 ve 11 değerleri arasında olduğunu söyleyebilirim.



**---** Güneş radyasyonu değişkenin histogram grafiğini çizdirdiğimde ve yoğunluk çizgisi eklediğimde elde ettiğim sonuç yukarıdaki gibidir. Maksimum değerin 260 solarR değerinde olduğunu, minimum değerin ise birden çok kısımda gözlendiğini görebiliyoruz. Yoğunluk olarak 200 ve 260 değerleri arasında en çok yoğunluk olduğu söylenebilir. Ayrıca 240 değerinde gözlem olmadığı grafiğe yansımıştır.

### 6-) Q-Q plot grafiği çizdirmek

Verilerimin normal olarak dağılıp dağılmadığını bilmek istediğimden verilerimin Q-Q grafiğini oluşturmak istiyorum.

qqnorm ve qqline komutlarını kullanarak test ediyorum.

**#qqnorm(ozon, main="Q-Q Plot: ozon")**

**#qqline(ozon)**

**#qqnorm(solarR, main="Q-Q Plot: güneş radyasyonu")**

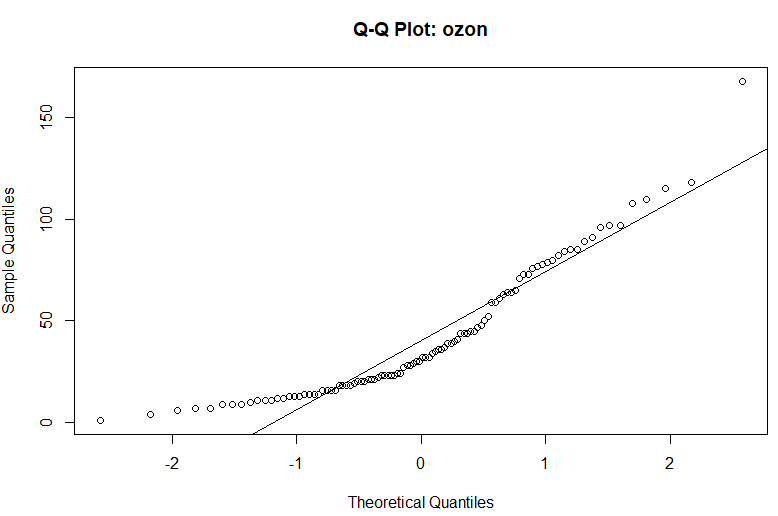
**#qqline(solarR)**

**#qqnorm(ruzgar, main="Q-Q Plot: rüzgar hızı")**

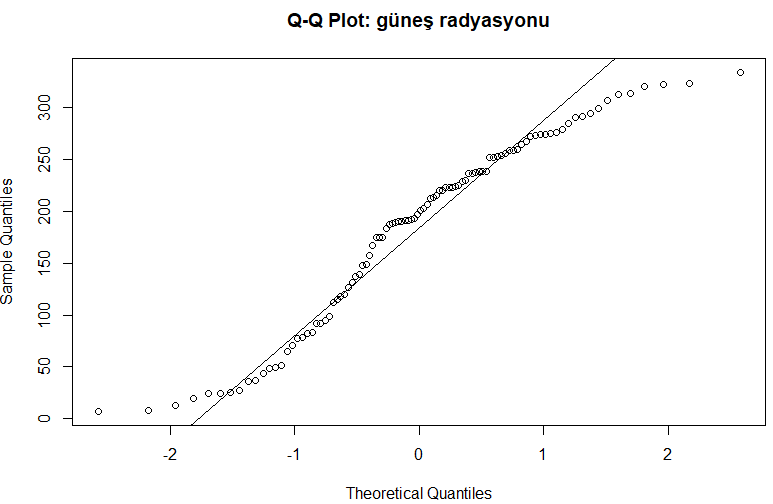
**#qqline(ruzgar)**

**#qqnorm(sicaklik\_F, main="Q-Q Plot: sıcaklık değeri")**

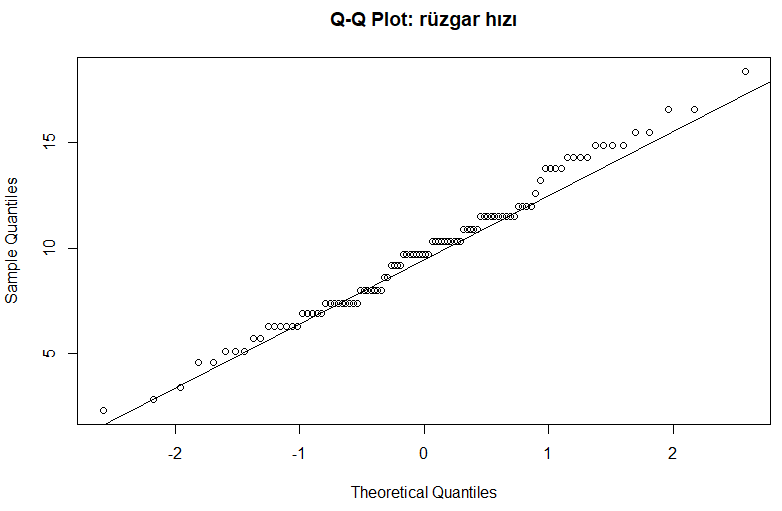
**#qqline(sicaklik\_F)**



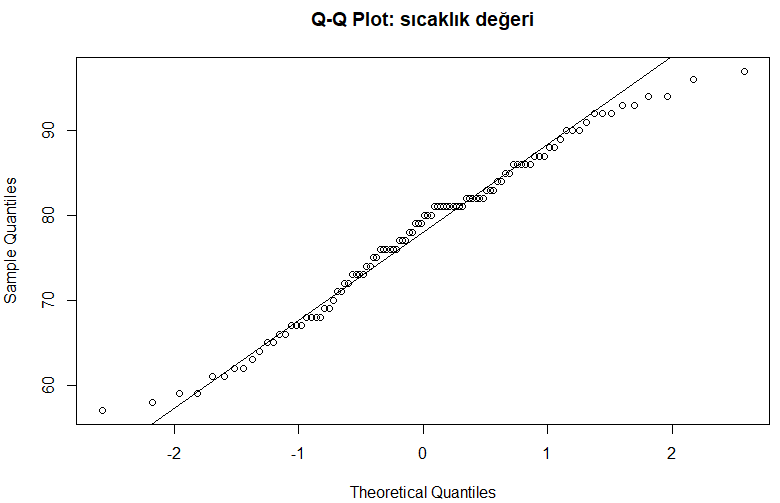
**---** Ozon değişkenim için verilerimin normal dağıldığını söyleyebilirim. Regresyon çizgimizin dağılım değerlerimize uygunluğu iyi olarak söylenebilir.



**---** SolarR değişkenimizinde normal dağıldığını söyleyebiliriz. Regresyon çizgisi dağılım değerlerinin çok fazla sapma yapmadığını gösteriyor.



**---** Rüzgar hızı değişkenimizinde sapma olmadan regresyon çizgisi üzerinde dağıldığını görebiliyoruz. Bu değişkenimiz içinde normal dağıldığı varsayımına ulaşabiliriz.



Sıcaklık değeri değişkenimizinde normal dağıldığı sonucuna varabiliriz.

**Kodlarım:**

**airquality <- read.csv("C:/airquality.txt", sep="")**

**attach(airquality)**

**index = ay <= 7**

**ay=airquality[,5] <=7**

**airquality[,10] = ay**

**install.packages("qqplot2")**

**install.packages("dplyr")**

**library(ggplot2)**

**library(dplyr)**

**typeof(ozon)**

**typeof(solarR)**

**typeof(ruzgar)**

**typeof(sicaklik\_F)**

**typeof(ay)**

**typeof(yuksek\_ozon)**

**summary(airquality)**

**tapply(airquality$sicaklik\_F, airquality$ay, mean)**

**tapply(airquality$ozon, airquality$ay, mean)**

**tapply(airquality$ruzgar, airquality$ay, mean)**

**tapply(airquality$solarR, airquality$ay, mean)**

**table(yuksek\_ozon, ay)**

**table(yuksek\_ruzgar, ay)**

**table(yuksek\_sicaklik,ay)**

**table(yuksek\_solar, ay)**

**ftable(yuksek\_ozon + yuksek\_sicaklik ~ ay)**

**ftable(yuksek\_ozon + yuksek\_ruzgar ~ ay)**

**ftable(yuksek\_ozon + yuksek\_solar ~ ay)**

**table(yuksek\_sicaklik, ay, yuksek\_ozon)**

**table(yuksek\_ruzgar, ay, yuksek\_solar)**

**table(yuksek\_solar, ay, yuksek\_ozon)**

**apply(airquality, 2, mean)**

**plot(airquality$ozon, airquality$solarR ,main="airquality: ozon vs. solarR", xlab="ozon", ylab="solarR")**

**grid()**

**plot(airquality$solarR, airquality$sicaklik\_F, main="airquality: solarR vs. sicaklikF", xlab = "solarR", ylab = "sicaklikF")**

**grid()**

**plot(airquality$ozon, airquality$sicaklik\_F, main="airquality: ozon vs. sicaklikF", xlab = "ozon", ylab = "sicaklikF")**

**grid()**

**plot(ozon, sicaklik\_F, pch=as.integer(ay))**

**legend(130.140, 70.60, c("7.ay'dan önce" , "7.ay'dan sonra"), pch=0:1)**

**a <- lm(sicaklik\_F ~ ozon)**

**abline(a)**

**plot(solarR, sicaklik\_F, pch=as.integer(ay))**

**legend(10.30, 95.105, c("7.ay'dan önce" , "7.ay'dan sonra"), pch=0:1)**

**b <- lm(sicaklik\_F ~ solarR)**

**abline(b)**

**plot(airquality[,1:10])**

**agirlikli <- tapply(airquality$ozon, airquality$ruzgar, mean)**

**barplot(agirlikli)**

**barplot(agirlikli, main="Ozon ortalamasına göre ruzgar değişkeni",xlab = "rüzgar", ylab="ozon")**

**agirlikli2 <- tapply(airquality$solarR, airquality$sicaklik\_F, mean)**

**barplot(agirlikli2)**

**barplot(agirlikli2, main="Güneş Radyasyonu ortalamasına göre sıcaklık değişkeni",xlab = "sicaklik\_F", ylab="solarR")**

**agirlikli3 <- tapply(airquality$ruzgar, airquality$sicaklik\_F, mean)**

**barplot(agirlikli3)**

**barplot(agirlikli3, main="Rüzgar hızı gözlemleri ortalamasına göre sıcaklık değişkeni",xlab = "sicaklik\_F", ylab="ruzgar")**

**boxplot(ozon ~ V10, data=airquality, main="ozon vs kategorik")**

**boxplot(solarR ~ V10, data=airquality, main= "solarR vs kategorik")**

**boxplot(ruzgar ~ V10, data=airquality, main= "ruzgar vs kategorik")**

**boxplot(sicaklik\_F ~ V10, data=airquality, main= "sicaklik vs kategorik")**

**boxplot(ozon ~ ay , data=airquality, main = " ozonun aylara göre kutu grafiği")**

**boxplot(sicaklik\_F ~ ay , data=airquality, main = "sıcaklığın aylara göre kutu grafiği")**

**boxplot(solarR ~ ay , data=airquality, main = " Güneş radyasyonunun aylara göre kutu grafiği")**

**boxplot(ruzgar ~ ay , data=airquality, main = " rüzgar hızının aylara göre kutu grafiği")**

**hist(ozon, 30, prob=T)**

**lines(density(ozon))**

**hist(sicaklik\_F, 30, prob=T)**

**lines(density(sicaklik\_F))**

**hist(ruzgar, 20, prob=T)**

**lines(density(ruzgar))**

**hist(solarR, 30, prob=T)**

**lines(density(solarR))**

**qqnorm(ozon, main="Q-Q Plot: ozon")**

**qqline(ozon)**

**qqnorm(solarR, main="Q-Q Plot: güneş radyasyonu")**

**qqline(solarR)**

**qqnorm(ruzgar, main="Q-Q Plot: rüzgar hızı")**

**qqline(ruzgar)**

**qqnorm(sicaklik\_F, main="Q-Q Plot: sıcaklık değeri")**

**qqline(sicaklik\_F)**

**colors <- ifelse(ozon >= 0, "green", "red")**

**plot(ozon, type="l", col=c("blue", "green"))**