

Data Frame

Data frame'ler R'in temeli ve R'ı güzelleştiren, onu çalışması çok kolay bir dil hale getiren bir yapıdır. Dataframe'ler vektörler kümelerinde oluşur. Bir kaç (aynı uzunlukta) vektörün yan yana gelmesi ile bir data frame oluşturabiliriz.

```
v1 <- c(1:10)
```

v2 <- c(11:20) diyelim. v1 ve v2 şu an iki ayrı vektör durumunda. Bunları bir data frame'de birleştirelim.

```
df1 <- data.frame(Sutun1 = v1, Sutun2 = v2)
```

oluşan data frame aşağıdaki gibidir.

	Sutun1	Sutun2
1	1	11
2	2	12
3	3	13
4	4	14
5	5	15
6	6	16
7	7	17
8	8	18
9	9	19
10	10	20

Bir vektör tek başına da data frame olabilir. Yani `v1 <- as.data.frame(v1)` dersek `v1` değişkenimiz artık bir vektör değil data frame sınıfına aittir. Vektörün içindeki değişkenler bu dönüşümden etkilenmezler, sadece değişkenin sınıfı değişir.

Sapply ve benzer yöntemler

Bu kısımda `apply` fonksiyonlarını bir örnek üzerinden tekrar edeceğiz. Diyelimki elimizde kategorik bir değişken olan ‘kalite’ değişkeni ve bu kaliteye bağlı değişen bir sayısal değişken olan ‘değer’ değişkeni var. Bu değişkenlerde bir data frame’in içinde bulunuyor. Biz de bu kategorik değişken olan kalite’ye göre ortalamalar bulmak istiyoruz. Eğer kalite faktöründe iyi - orta - kötü şeklinde seviyeler varsa bize 3 adet ortalama dönecek. İyi derecede kaliteye sahip verilerin değerlerinin ortalaması, orta derecede kaliteye sahip verilerin değerlerinin ortalaması ve kötü derecede kaliteye sahip verilerin değerlerinin ortalaması.

Bu işlemi aşağıda sıralanan yöntemlerle yapabiliriz.

tapply ile

```
tapply(df$deger, df$kalite, mean)
```

aggregate ile

```
aggregate(deger ~ kalite, df, mean)
```

dplyr ile

```
library(dplyr)
```

```
df %>% group_by(kalite) %>% summarise(x = mean(deger))
```

sapply kullanmak için

```
splitmean <- function(df){  
  s <- split(df, df$kalite)  
  sapply(s, function(x) mean(x$speed))  
}  
splitmean(df)
```

Vektörler ve Döngüler

İki vektörü for döngüsü içinde nasıl ekleriz?

```
v1 <- c(1:10)  
v2 <- c(1:10)  
v3 <- rep(0, length(v2))  
for(i in length(v1)){  
  v3[i] <- v1[i] + v2[i]  
}  
v4 <- v1 + v2
```