



SUCUL TÜRLERDE İSTİLACILIK  
POTANSİYELİ BELİRLEME

# TD27: PROGRAMLAMA DİLİ OLARAK R VE R PROGRAMINA VERİ YÜKLEME

Dr. Öğr. Üyesi Burcu MESTAV

11:00- 11:45

# R Programlama Dili Nedir?

R,

kullanıcının algoritmaları programlamasına ve başkaları tarafından programlanmış araçları kullanmasına izin veren istatistiksel bir programlama dili olup esasında günümüzde veri bilimi olarak tanımlanan alanda kullanılan bir programlama dili ve yazılım ortamıdır.

'Açık kaynak' özelliğine sahiptir.

Açık kaynak, yazılımlarının dayandığı kodu korumalı olarak gizleyen ticari yazılım şirketlerinin aksine, R'yi geliştiren kişilerin herkesin kodlarına erişmesine izin verdiği anlamına gelir. Bu açık kaynaklı felsefe, herhangi bir yerdeki herhangi birinin, yazılıma katkıda bulunmasına izin verir. Sonuç olarak, R'nin yetenekleri, dünyanın dört bir yanından insanlar ekledikçe dinamik olarak genişlemektedir.

R,

GNU Genel Kamu Lisansı uyarınca ücretsiz şekilde kullanılabilir,

herhangi bir standart bilgi işlem platformunda veya işletim sisteminde çalışabilir; Mac, Windows, Linux ve hatta PlayStation 3'ünüzde bile nerde isterseniz isteyin çalışabilir.

R

# R Programlama Dili Nedir?

Tanıtımı yapıldığı zamandan bu yana R, veri analizi için baskın olan yazılım ortamlarından biri haline gelmiştir ve çeşitli bilimsel disiplinlerle kullanılmaktadır. R, özellikle grafik yetenekleri açısından popüler olmasının yanısıra, uydu görüntülerini gibi raster tabanlı modeller oluşturmayı nispeten kolay hale getiren yetenekleri nedeniyle de dikkat çekmektedir.

R ile fonksiyonlar yazabilir, hesaplamalar yapabilir, mevcut olan bir çok işlevi uygulayabilirsiniz.

R ve kütüphaneleri sayesinde doğrusal(linear) ve doğrusal olmayan(non-linear) modelleme, klasik istatistiksel testler, zaman serisi analizi, sınıflandırma, kümeleme ve diğerleri de dahil olmak üzere çok çeşitli istatistiksel teknikleri çalışabilir, basit veya karmaşık grafikler oluşturabilir ve hatta kendi kütüphane işlevlerinizi yazabilirsiniz.



# R Programlama Ne değildir?

## R Ne Değildir?

Bir veri tabanı değildir ama veri tabanlarına bağlanabilir.

Kullanıcı dostu olmasa da java gibi diller aracılığıyla ara yüz desteğine sahip bir yazılım geliştirme ortamıdır.

Tablolardan oluşan yazılım paketi (Excel, Minitab gibi) değildir ama bunlara bağlanabilir. Profesyonel veya ticari desteği tabi bir yazılım değildir. Kapalı kutu yazılımlardan oluşan bir yazılım değildir.



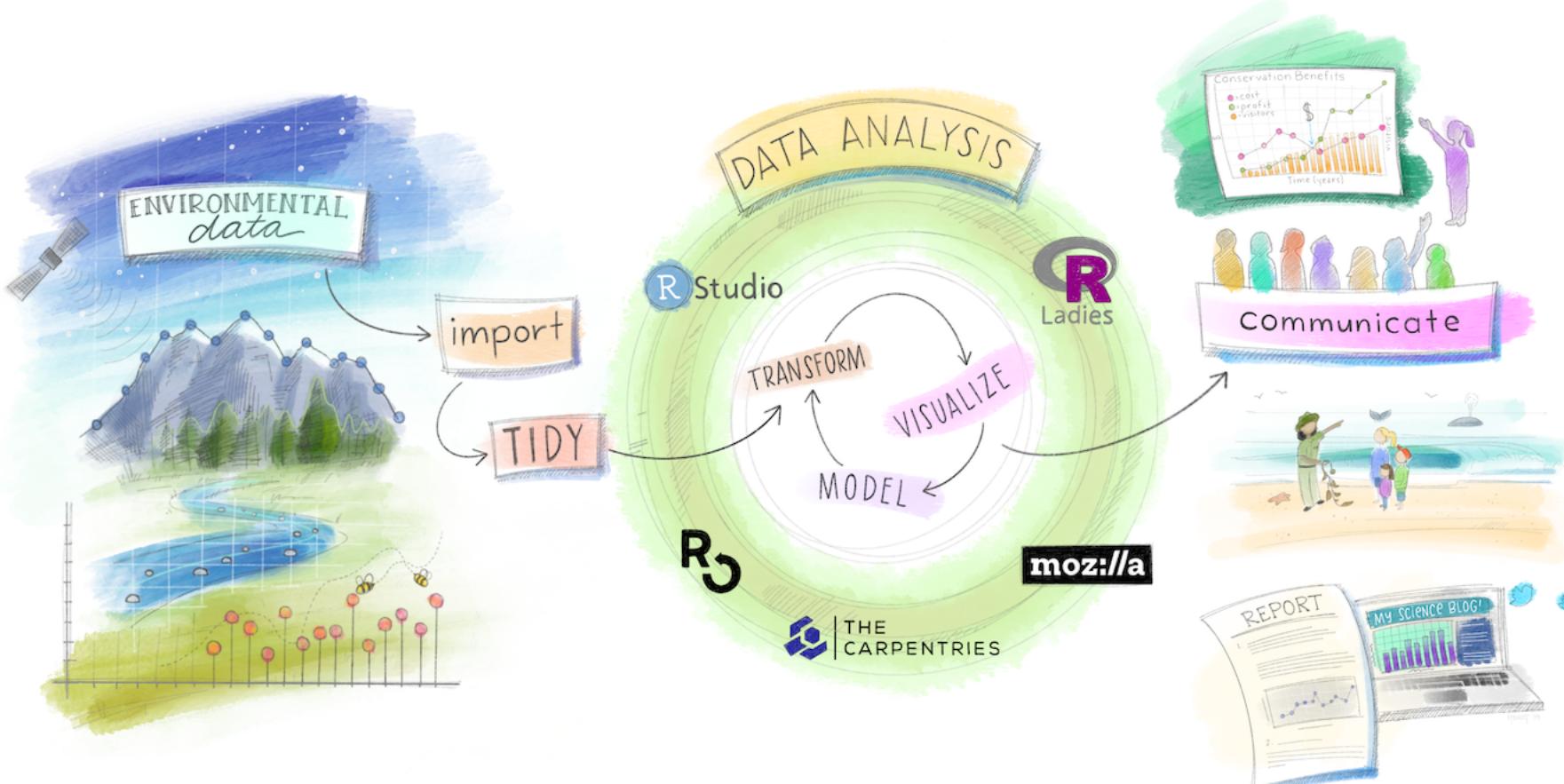
R

# R nerede kullanılır?



Veri düzenleme  
İstatistiksel analiz  
Web sayfası hazırlama (statik/dinamik)  
Sunum hazırlama  
Programlama  
Otomatik, periyodik ve tekrarlabılır rapor hazırlama  
pdf ve html oluşturma  
Tez, kitap yazma ve CV oluşturma  
Poster hazırlama  
Rapor şablonu oluşturma  
...

# R nerede kullanılır?

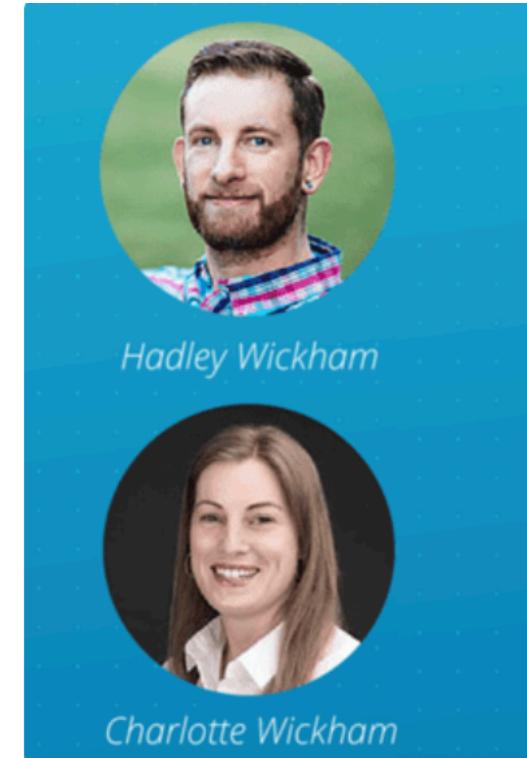


R

**Robert Gentleman**



**Ross Ihaka**



R

# R-chitecture

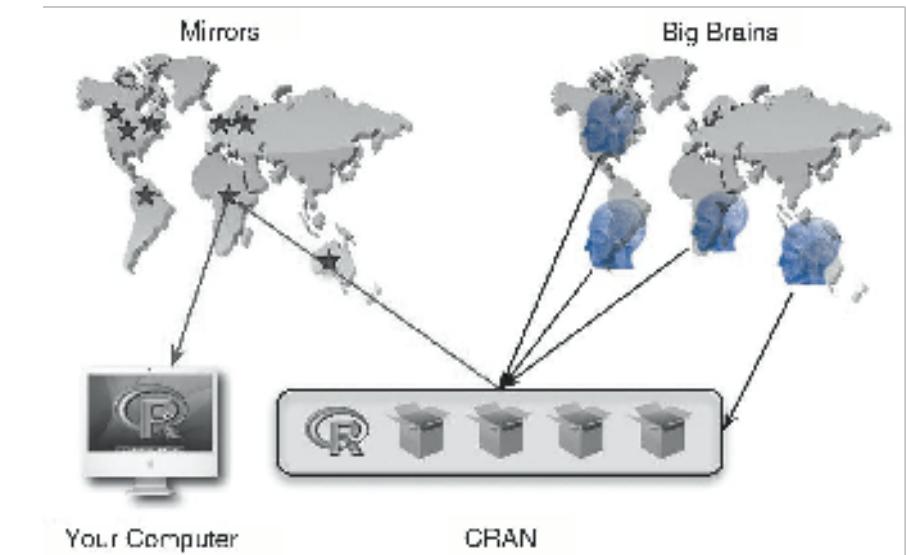
Temel olarak, R makul miktarda işlevselligi sahip temel bir paket olarak mevcuttur. Programa özel işlevler ekleyen bu paketleri indirerek genişletilebilmesidir.

Bu paketler, yazılımın yanı sıra, CRAN\* (Kapsamlı R Arşiv Ağı) olarak bilinen merkezi bir yerde saklanır. Bir paket CRAN'a kaydedildikten sonra, Internet bağlantısı olan herhangi biri, onu CRAN'dan indirebilir ve kendi R kopyasında kullanmak üzere yükleyebilir.

\*CRAN, R'yi kullanmak için bir merkezdir. Kısacası yazılımı ve yüklemek istediğiniz paketleri indirdiğiniz yerdir. Herkesin erişebilecegi tek bir CRAN konumuna sahip olmak yerine, CRAN dünyadaki farklı yerlerde ‘yansıtılmaktadır’. “Yansıtılmış”, dünyaya dağılmış CRAN'ın özdeş sürümleri olduğu anlamına gelir.

Temel felsefe, size coğrafi olarak en yakın bir CRAN'ı seçmektir.

R, temel olarak, üretim hedefine katkıda bulunan kabarık fedakar insanların büyük bir küresel ailesidir.



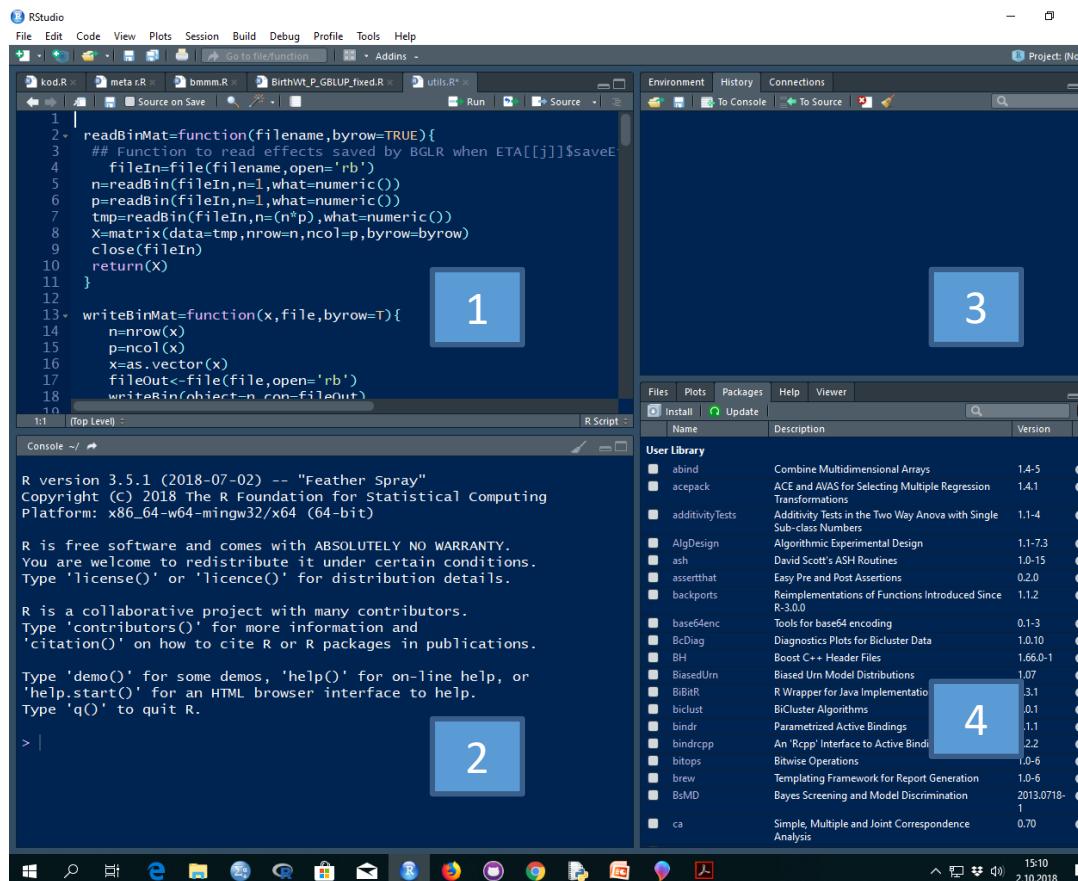
R

## R'ın kullanıcı dostu arayüzler

R programının bizlere sunduğu arayüzün yanı sıra kullanıcıya kolaylık sağlama amacıyla geliştirilmiş olan editör + arayüzler de bulunmaktadır. En yaygın olanlar R commander ve RStudio'dur.

RStudio, R için güçlü ve açık kaynaklı bir entegre geliştirme ortamıdır. 2010 yılında geliştirmeye başlanmış ve ilk beta sürümü Şubat 2011'de piyasaya sürülmüştür. RStudio, R ile daha kolay etkileşime girmenize izin veren entegre bir geliştirme ortamıdır. RStudio oldukça kullanıcı dostudur. Açıılır menülere, birden çok sekme içeren pencerelere ve birçok özelleştirme seçeneğine sahiptir.





# R'ın kullanıcı dostu arayüzler

## 1- R script (betik):

R programlama kodlarının yazıldığı kısımdır. Komut dosyalarını, fonksiyonları veya diğer dökümanyoların hazırlandığı yerdir.

## 2- Konsol (Console):

Temel işlemlerin ve komutların girildiği kısımdır. Başka bir deyişle R scripte yazılan kodların derlendiği yanı kodların çalıştırıldığı bölümdür

## 3- “Environment ve History” sekmleri:

R scripte ve consola yazılan ve oluşturulan değişkenler ve veri kümeleri **Environment** sekmesinde,

Yazılmış olan kodlar ise **History** sekmesinde tutulur.

## 4-Files”, “Plots”, “Packages”, “Help” ve “Viewer” sekmleri:

Analiz için çizdirdiğiniz grafiklerin çıktıları **Plots** sekmesinde,

Fonksiyon ve paketlerle ilgili olarak çağrıdığımız yardım belgeleri **Help** sekmesinde,

Rstudio programına yüklediğiniz paketlerin listesi **Package** sekmesinde yer almaktadır

# R'ın kullanıcı dostu arayüzler

- Mac OS için R studio'yu kurduktan sonra
- <https://www.xquartz.org/> ve [https://support.apple.com/kb/dl1572?locale=tr\\_TR](https://support.apple.com/kb/dl1572?locale=tr_TR) kurarak R studionun bazı özelliklerini çalıştırmanıza katkı sağlar.
  - Diğer arayüzler
    - Geany
    - <http://www.geany.org/>
    - Notepad++
    - <http://notepad-plus-plus.org/>
    - RWinEdt
  - <http://cran.r-project.org/web/packages/RWinEdt/index.html>
    - Tinn-R
    - <http://sourceforge.net/projects/tinn-r/>
      - JGR (R için Java ara yüzü)
        - <http://www.rforge.net/JGR/>
        - Emacs + ESS
      - <http://www.gnu.org/software/emacs/>
        - <http://ess.r-project.org/>
        - Rattle
        - <http://rattle.togaware.com/>
      - Playwith (grafikler için)
      - <https://code.google.com/p/playwith/>



# R Paketleri (R package)

Paketler isteğe bağlı olarak yüklenebilen ek işlev koleksiyonlarıdır. Genellikle, bu işlevleri göstermek için kullanılabilecek örnek veriler içerir. R birçok ortak istatistiksel işlev ve model ile dolu olmasına rağmen, çalışmalarınızın çoğunda ek paketler gereklidir. Bir paketi kullanmak için önce paketi kurup yüklemelisiniz.

R paketleri sadece bir kez kurulmalıdır (R yükseltilene veya yeniden kurulana kadar). Bununla birlikte, her yeni bir R oturumu başlattığınızda, o oturumda kullanmak istediğiniz her paketi yüklemeniz gereklidir.

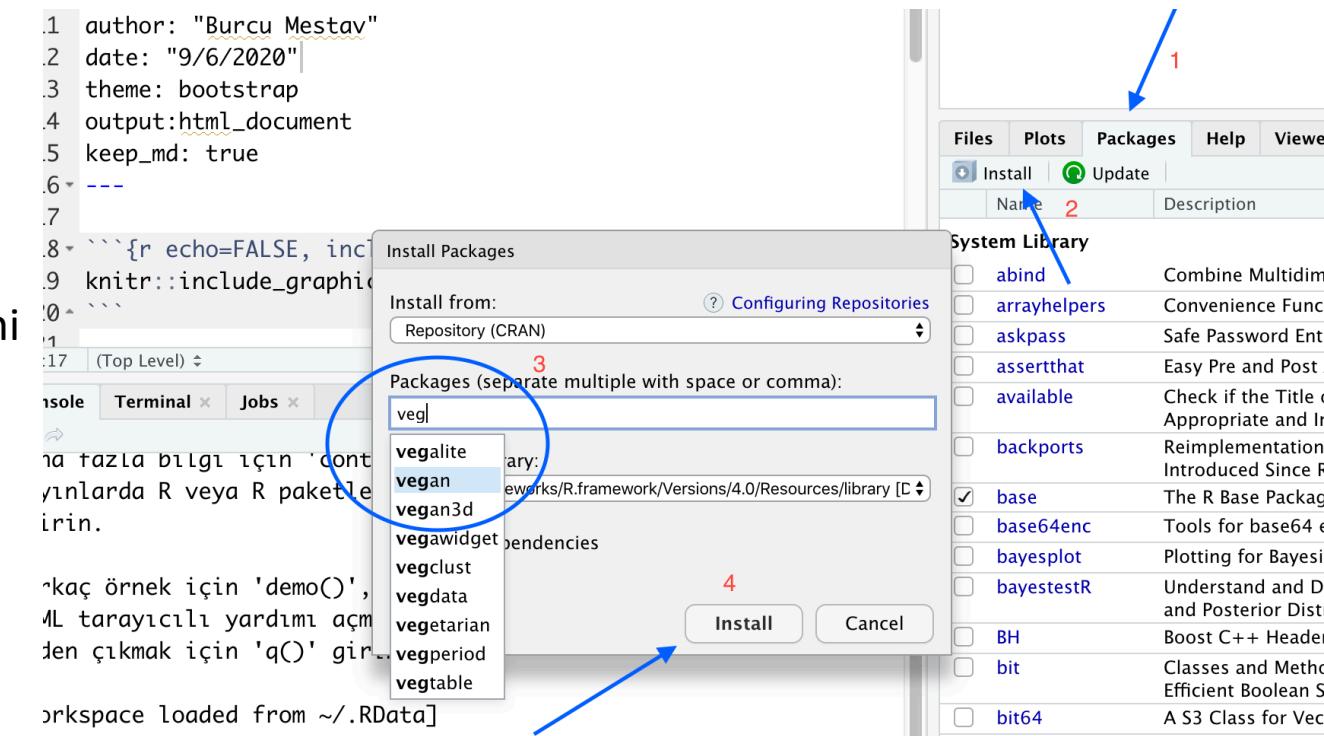
komut satırında

```
> install.packages("name of the package")
```

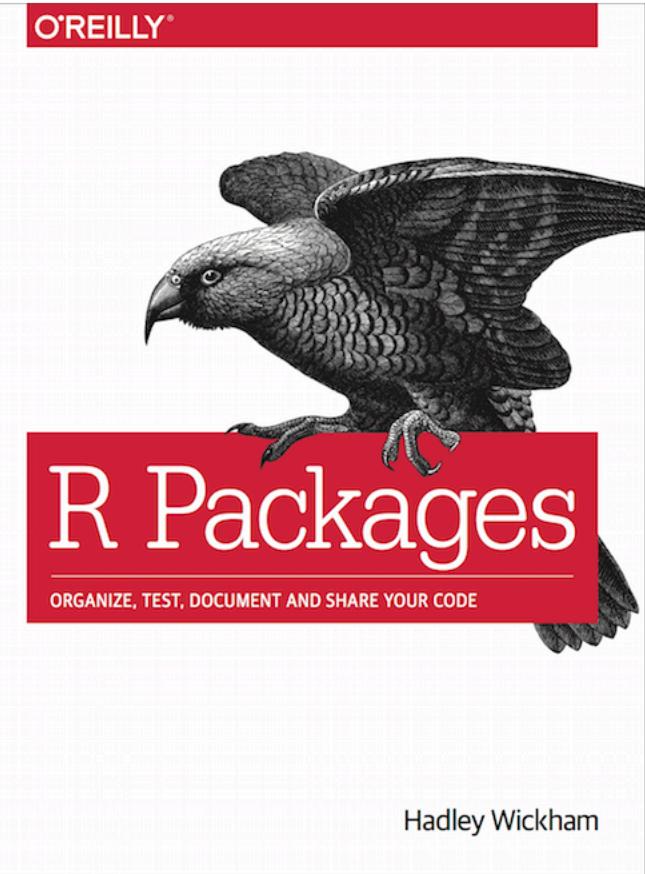
```
➤ install.packages("vegan")
```

```
➤ install.packages(c("ncdf4", "RNetCDF"))
```

Paketler Sekmesi kullanılarak yapılabilir.



# R Paketleri (R package)



# Paketleri nereden bulabiliyoruz?

- CRAN
  - [https://cran.r-project.org/web/packages/available\\_packages\\_by\\_name.html](https://cran.r-project.org/web/packages/available_packages_by_name.html)
- CRAN task views (ana başlık halinde veriyor)
  - <https://cran.r-project.org/web/views/>
- Bioconductor (bioinformatik paketler için)
  - <http://bioconductor.org/>
    - pkgsearch
  - <https://rdrr.io/github/metacran/pkgsearch/>
    - AWESOME R
  - <https://awesome-r.com/>



# R ile programlamaya giriş

R'yi Hesap makinesi olarak kullanabilirsiniz.

```
> 10 + 100 # Toplama  
[1] 110  
> 12 * 6 # Çarpma  
[1] 72  
> 100 ^ 2 # Üstlü işlemler  
[1] 10000  
# İşlemlerin önceliğini belirlemek için  
parantez kullanabilirsiniz.  
> 3 * (13 * 2)  
[1] 78  
> (11 ^ 2) * (1.4 * 10)  
[1] 1694
```

Temel Matematik İşlevleri gerçekleştirebilirsiniz

```
# Karekök almak  
> sqrt(4)  
[1] 2  
# Sinüs değerini hesaplamak  
> sin(pi / 2)  
[1] 1  
# Mutlak değer hesaplamak  
> abs(-4)  
[1] 4  
# Logaritmik hesaplamalar  
> log10(100)  
[1] 2  
exp(1)  
[1] 2.71828
```

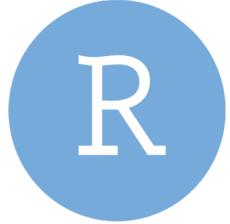


# R ile programlamaya giriş

## Değişkenlere Değer Atayarak Kullanmak

R kullanırken işlemleriniz biraz karmaşık hale gelmeye başladığında, önceki işlem sonuçlarınızı kaydederek sonra yapacağınız işlemlerde kullanmak isteyeceksiniz. Ve belirli bir seyiye geldikten sonra sizi temin ederim ki bu bir istek olmaktan çıkacak ve zorunluluk haline gelecek.

Bu bilgi saklama görevini yerine getirmek için nesneler (object) kullanılır. Nesne denen şey özünde herhangi bir **değişkeni** veya **fonksiyonu** tutan bir isim olarak tanımlanabilir. Bilgisayarın hafızasında (bellek de denir) duran bir değere ulaşmanızı imkan veren bir isim.



R

# R ile programlamaya giriş

R'da nesneler ve fonksiyonlar “-dan yaratıldığı” anlamını düşünebileceğiniz “`<-`” ile veya “`=`” ayrıılır. Bu nedenle, bir komutun genel şekli şöyledir:

## **Object<-function veya Object = function**

Bir nesne R'da herhangi birşeyle (değişken, bir değişkenler topluluğu, bir istatistiksel model vb. olabilir.) oluşturulabilir.

Nesneler tek değerler (bir puan kümesinin ortalaması gibi) veya bilgi koleksiyonları olabilir;

örneğin, bir analizi çalıştırığınızda, bu analizin çıktısını içeren bir nesne yaratırsınız; bu, nesnenin birçok farklı değer ve değişken içerdiği anlamına gelir.

İşlevler ise nesnelerinizi oluşturmak için R'de yaptığınız şeylerdir.



# R ile programlamaya giriş

```
x1 = 10
x2 = 100
y5 = 232132
# Değişkenlere kişilerle ilgili bilgileri atayabilirsiniz.
ali = 25
# Değişken isimleri büyük ve küçük harfler arasındaki ayrımı dikkat
eder.
# Siz de dikkat edin.
> mehmet = 79
> ahmet = 45
> murat = 79+45
> murat [1] 124
# Değişken isimlerinde Türkçe karakterler ve boşluk kullanamazsınız.
> bu_da_baska_bir_degisken = 12.1
# Değişkenleri metin değerlerini tutmak için de kullanabilirsiniz.
> degisken = "Metin değerlerini tutan bir değişkenin içeriği."
```



R

# R'da veri türleri ve yapıları

- Nesnelerin hepsine birden veri yapıları adı verilir ve iki başlık altında toplanır

- **Atomik veya basit veri türleri**

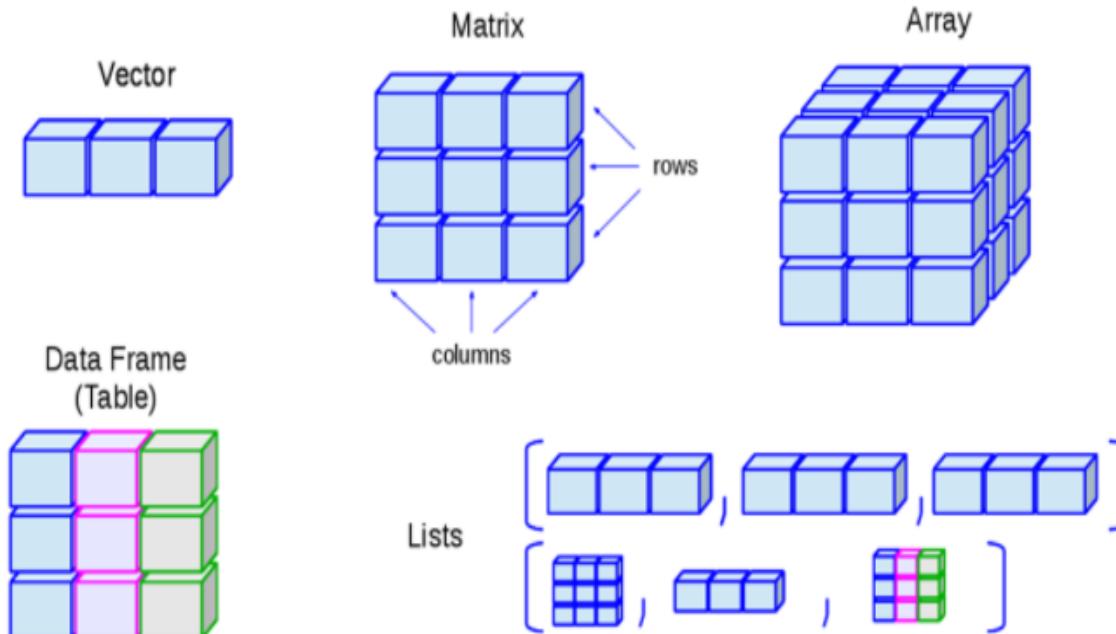
- Tamsayı (integer) : 2L, 199L, 0L, -14L
- Sayı (double) : 32, 45.2, 0, -0.456773, 6.2e+19
- Karmaşık sayı (complex) : 4+6i, 2.3+6.2i
- Metin (character) : "a", "Çanakkale", "Ahmet", "345", "17GL269", "TRUE"
- Ham veri (raw) : "Ada" için 41 64 61
- Mantıksal (logical) : TRUE, FALSE

- **Bileşik veriler**

- Birden fazla basit verinin bileşiminden oluşan verilerdir. Tek boyutlu veya çok boyutlu olabilirler.
- 5 farklı veri yapısı bulunmaktadır.
- Vektör (vector) : burcu = c(3, 3.14)
- Matris (matrix) : burcu = matrix(x1, nrow=4, ncol=4, byrow=T)
- Dizi (array)
- Veri Çerçeveesi (data.frame)
- Liste (list)



# R'da veri türleri ve yapıları



**Vektör:** tek boyutlu ve aynı türden bir veya daha fazla değere sahip olan değişkendir.

**Matris :** iki boyutlu ve aynı türden bir veya daha fazla değere sahip olan değişkendir.

**Array (dizi):** üç boyutlu ve aynı türden bir veya daha fazla değere sahip olan değişkendir.

**Faktör (factor) :** tek boyutludur ancak metin (karakter) değerleri için uygundur.

**Veri Çerçeveşi (Data Frame) :** Matrise benzer ancak her sütunu farklı data tipi olabilir.

**Liste (List) :** kompleks data yapısıdır ancak oldukça esnektir. Tüm data yapılarını içinde bulundurabilir.

R

# R'da veri türleri ve yapıları

## Verilere erişim;

Veri Çerçeveleri / Matrisler / Diziler / Listeler'deki herhangi bir sütun ya da satırı ulaşmak isterseniz:

- Matris / Veri Çerçevesi: nesne [c ("rownames"), c ("colnames")]
- Dizi: array [c ("rownames"), c ("colnames"), c ("boyut")]
- Veri Çerçevesi / Liste: nesne \$ değişken
- Liste: [[„element“]] listeley veya \$ elementi listeley

R

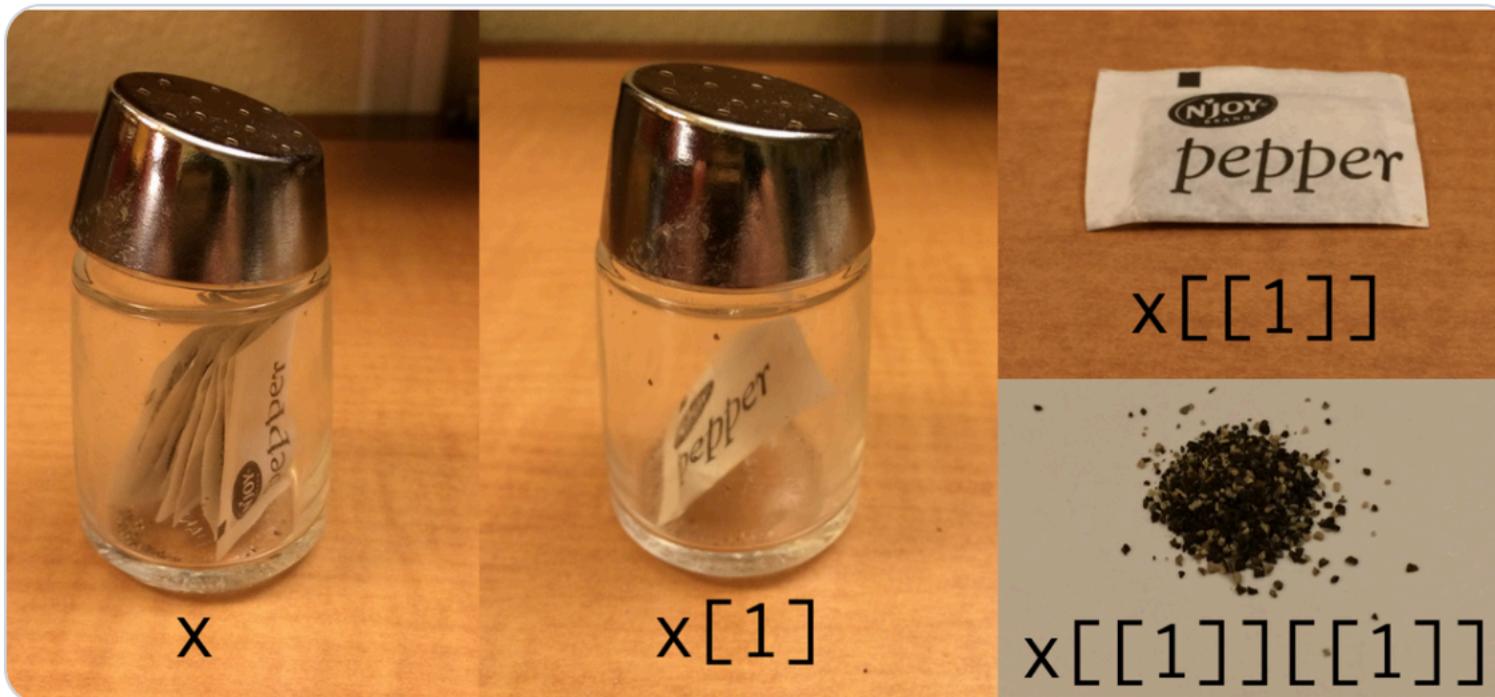
# R'da veri türleri ve yapıları



Hadley Wickham @hadleywickham



Indexing lists in #rstats. Inspired by the Residence Inn



R

# R'da veri yükleme

- Programı açığınızda ilk yapmanız gereken çalışma klasörünü (Working Directory) belirlemenizdir.

```
#### Program ilk açıldığında hangi çalışma alanında olduğunu gösterir.
```

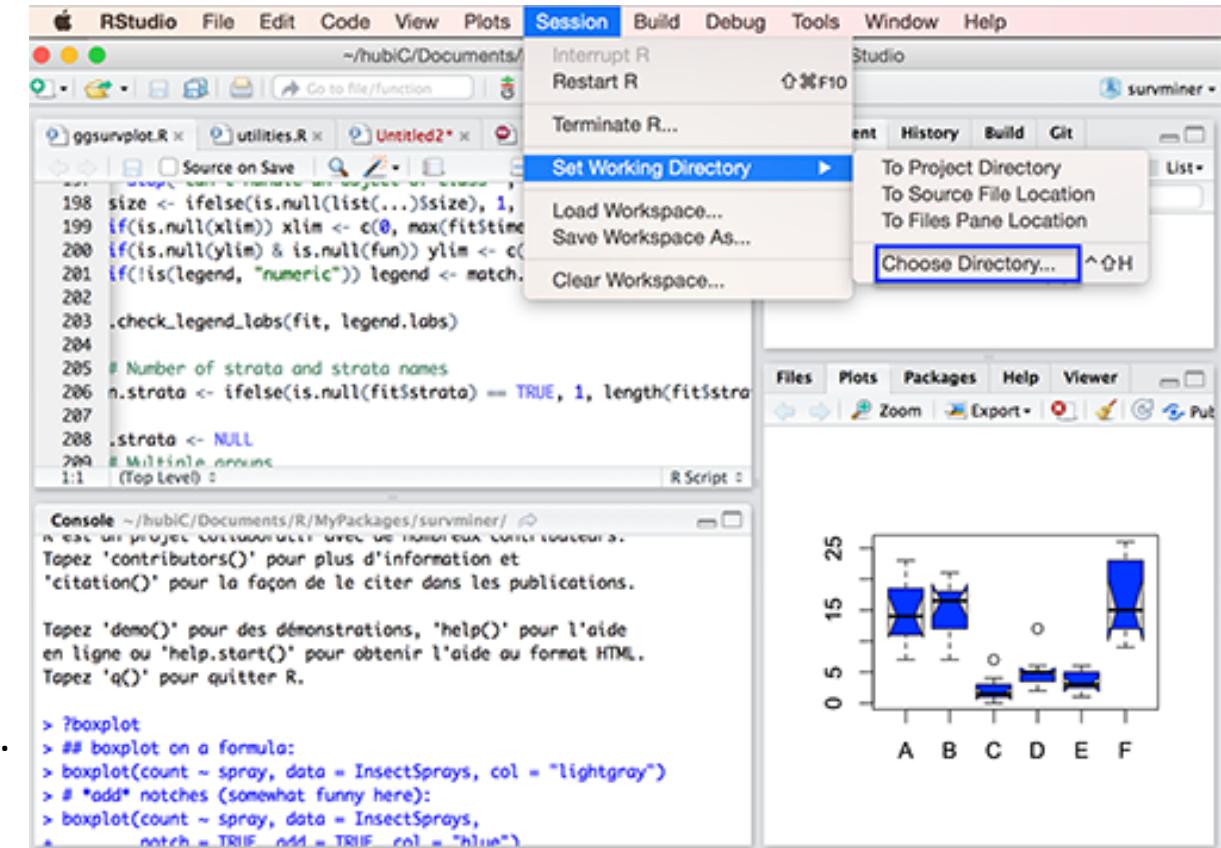
```
> getwd()
[1] "C:/Users/COMU/Desktop/..."
```

```
#### Çalışmak istediğiniz yolu belirlersiniz (windows için).
```

```
> setwd("C:/Users/COMU/Desktop/R_islem")
```

```
#### Çalışmak istediğiniz yolu belirlersiniz (MacOS için).
```

```
> setwd("~/Desktop/R_islem")
```



# R'da veri yükleme

country	year	cases	population
Afghanistan	1990	45	1537071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	216766	1280425583

variables

country	year	cases	population
Afghanistan	1990	45	1537071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	216766	1280425583

observations

country	year	cases	population
Afghanistan	1990	45	1537071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	216766	1280425583

values

R

# R'da veri yükleme

## 5 types

- Flat files



- Data from Excel



- Databases



- Web



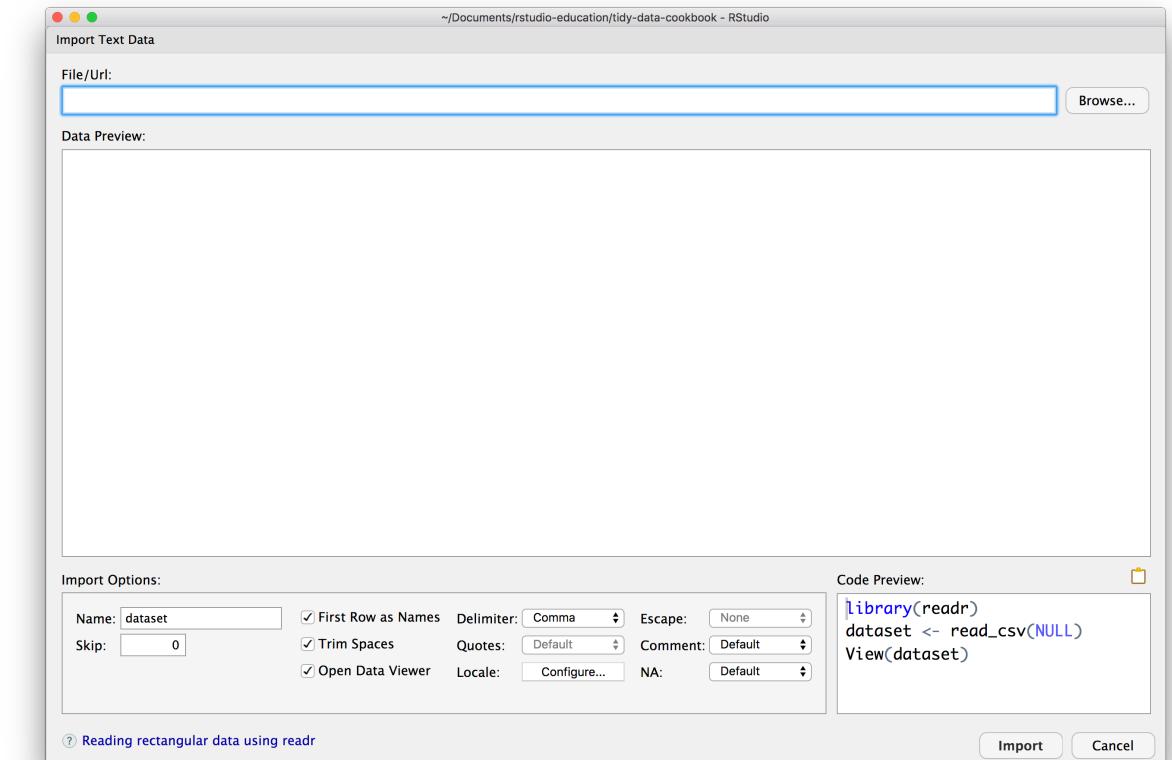
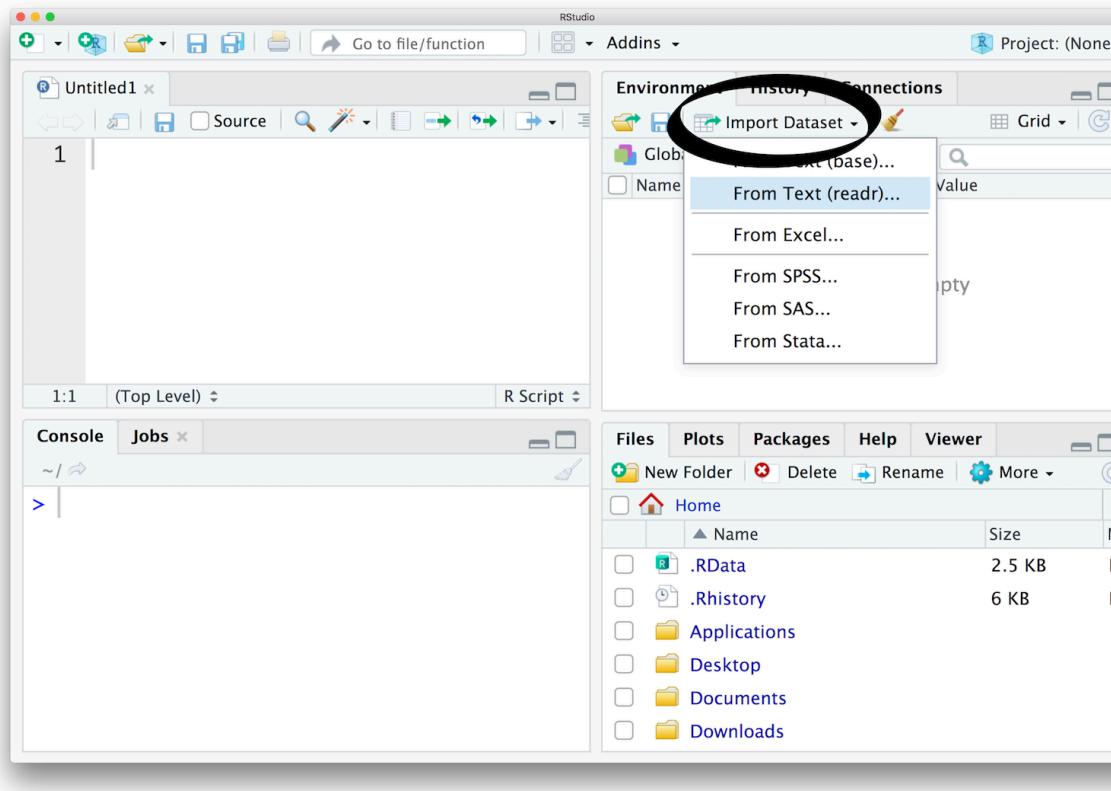
- Statistical software



R

# R'da veri türleri ve yapıları

Bir GUI (grafik kullanıcı arabirim ) ile verileri hızla içe aktarabilirsiniz.



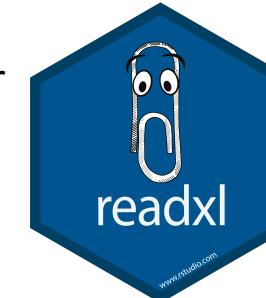
## R'da veri yükleme



# R'da veri türleri ve yapıları



- `read\_csv ()` - virgülle ayrılmış dosyalar  
my\_data <- read\_csv("file.csv")
- `read\_csv2 ()` - noktalı virgülle ayrılmış dosyalar (virgülün ondalık basamak olarak kullanıldığı ülkelerde yaygındır)
  - my\_data <- read\_csv2("file.csv")
- `read\_tsv ()` - sekmeye ayrılmış dosyalar
  - my\_data <- read\_tsv("file.tsv")
- `read\_delim ()` - herhangi bir sınırlayıcıya sahip dosyalarda okur
  - my\_data <- read\_delim("file.txt", delim = "|")
- `read\_excel ()` - excel dosyalarını okur
  - my\_data<- read\_excel("file.xlsx")



# R'da veri türleri ve yapıları

## ## Diğer veri türleri

- **\*\* googlesheets4: \*\*** Google E-Tablolar
- **\*\* haven \*\*:** SPSS, Stata ve SAS dosyaları
- **\*\* DBI \*\*:** veritabanına özgü bir arka uça birlikte (örn. RMySQL, RSQLite, RPostgreSQL vb.): Bir veritabanında SQL sorguları çalıştırmanıza ve bir veri çerçevesi döndürmenize olanak tanır
- **\*\* jsonline \*\*:** JSON
- **\*\* xml2 \*\*:** xml
- **\*\* rvest \*\*:** web scraping
- **\*\* httr \*\*:** web API'leri
- **\*\* sparklyr \*\*:** spark veriler



