

:: Praktikum Statistika menggunakan R ::

01. Pengenalan R

Pengenalan R (1)

MA2181 Analisis Data / MA2081
Statistika Dasar / MA2082 Biostatistika

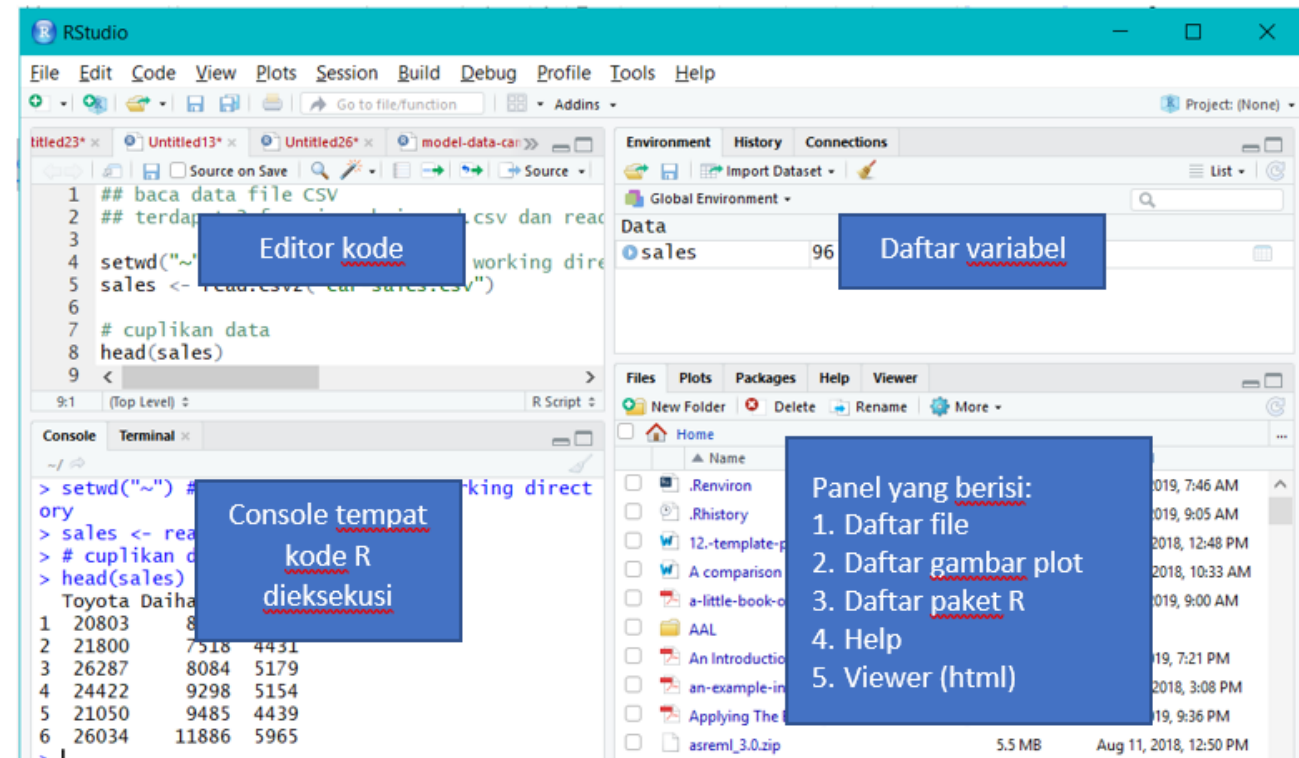
Kelompok Keilmuan Statistika

Laboratorium Statistika dan Komputasi Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam



R dan RStudio

- **R** adalah lingkungan perangkat lunak gratis untuk komputasi statistika dan grafik. R dapat diunduh pada situs <https://CRAN.R-project.org>
- **RStudio** adalah lingkungan pengembangan yang digunakan untuk mengelola dan menjalankan R sehingga mudah dioperasikan. RStudio berisi editor kode, kakas pelacak kesalahan, dan visualisasi. Unduh RStudio pada situs <https://rstudio.com/products/rstudio/download/>.
- Secara singkat R adalah bahasa pemrograman, dan Rstudio adalah aplikasi editor kode yang membuat penggunaan R menjadi lebih mudah.



[This Photo](#) by Unknown Author is licensed under [CC BY-SA-NC](#)

Kode Pemrograman R

Direktori kerja

```
dir()           #menampilkan isi direktori kerja  
getwd()         #menampilkan lokasi direktori  
setwd("D:/praktikum") #menetapkan lokasi direktori
```

Mencari informasi/bantuan tentang suatu fungsi/topik tertentu

```
help(base) atau ?base    #menampilkan bantuan informasi tentang "base"  
help.start()             # dalam bentuk html (browser)  
??topik                  # adalah alternatif dari help.search('topik')
```

Menjalankan file R yang berada dalam working directory, misalnya tugas.R:

```
source("tugas.R")
```

Tipe Data

Scalar

- **Numeric** (ex : 12.3, 5, 999)
- **Integer / Double** (ex: 2L, 34L, 0L)
- **Complex** (ex: 3 + 2i)
- **Character** (ex : 'a', "baik", '23.4')
- **Logical** (ex : TRUE, FALSE)

Matrix dan Vector

- **Matrix** (ex : `matrix(1:20, nrow=5, ncol=4)`)
- **Vector** (ex : `c(2,3,5,6)`)

Data frame

- `data.frame(df, stringsAsFactors = TRUE)`

List

- `l <- list(1, "Satu", TRUE)`

Factor

- `factor_apple <- factor('green','green','yellow')`

Catatan:

- Tipe data dapat diperiksa dengan menggunakan fungsi `class()` dan `typeof()`.
- Nilai khusus (special value): NA (not applicable/not available), NaN (Not a Number), Inf (Infinity)

Matriks dan Vektor

- **Matriks**, properti `matrix(data, nrow, ncol, byrow = FALSE)`

Keterangan argumen: **data** (nilai skalar (sebuah nilai) atau kumpulan nilai skalar (vektor)), **nrow** (banyaknya baris), **ncol** (banyaknya kolom), **byrow** (cara menyusun elemen vektor dalam matriks, default nilainya `byrow = FALSE` artinya elemen-elemen vektor disusun dengan mengisi kolom pertama, kedua dst).

- Vektor adalah array berdimensi satu dan setiap elemennya punya tipe data yang sama. Tipe datanya harus skalar.

Operasi yang sering dilakukan terhadap vektor dan matriks

Kegunaan	Matriks	Vektor
Nama data	<code>rownames()</code> , <code>colnames()</code>	<code>names()</code>
Ukuran data	<code>nrow()</code> , <code>ncol()</code> , <code>dim()</code>	<code>length()</code>
Membuat data (menggabungkan)	<code>rbind()</code> , <code>cbind()</code>	<code>c()</code> , <code>:</code> (titik dua), atau <code>seq()</code>
Transpose	<code>t()</code>	—
Mengecek jenis data	<code>is.matrix()</code>	<code>is.null(dim())</code>

Contoh Vektor

```
1:10
## [1] 1 2 3 4 5 6 7 8 9 10
typeof(1:10)
## [1] "integer"
(da <- c(1,2,3))
## [1] 1 2 3
c(3:7, 5, 2, 1)
## [1] 3 4 5 6 7 5 2 1
(v <- c(12,34,56,78,23:30))
## [1] 12 34 56 78 23 24 25 26 27 28 29 30
length(v)
## [1] 12
is.null(dim(v))
## [1] TRUE
seq(3, 12, by = 2)
## [1] 3 5 7 9 11
```

operasi pada vektor

```
v + 1:2 #
## [1] 13 36 57 80 24 26 26 28 28 30 30 32
v + rep(1:2, 6)
## [1] 13 36 57 80 24 26 26 28 28 30 30 32
v * 1:2
## [1] 12 68 56 156 23 48 25 52 27 56
29 60
v * rep(1:2, 6)
## [1] 12 68 56 156 23 48 25 52 27 56
29 60
```

Contoh Matriks

```
# - matrix
matrix(10, ncol = 3, nrow = 2)

##      [,1] [,2] [,3]
## [1,]  10  10  10
## [2,]  10  10  10

matrix(10, 2, 3)

##      [,1] [,2] [,3]
## [1,]  10  10  10
## [2,]  10  10  10

matrix(c(10, 2), nrow = 2, ncol = 3)

##      [,1] [,2] [,3]
## [1,]  10  10  10
## [2,]   2   2   2

rbind(da,da)

##      [,1] [,2] [,3]
## [1,]   1   2   3
## [2,]   1   2   3
```

```
(a<-matrix(c(10, 2, 5), nrow = 2, ncol = 3))

##      [,1] [,2] [,3]
## [1,]  10   5   2
## [2,]   2  10   5

rownames(a) <- c("r1","r2")

colnames(a) <- colnames(a, do.NULL = F, prefix = "Var ")

##   Var 1 Var 2 Var 3
## r1   10   5   2
## r2    2  10   5

t(a)

##           r1    r2
## Var 1    10     2
## Var 2     5    10
## Var 3     2     5

is.matrix(a)

[1] TRUE
```

:: Praktikum Statistika menggunakan R ::

01. Pengenalan R

Pengenalan R (2)

MA2181 Analisis Data / MA2081
Statistika Dasar / MA2082 Biostatistika

Kelompok Keilmuan Statistika

Laboratorium Statistika dan Komputasi Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam



Data Frame

Tipe data frame adalah deretan vektor yang panjangnya sama. Perbedaannya dengan matriks adalah data frame membolehkan kolom yang satu dengan yang lainnya mempunyai tipe data yang berbeda (namun tetap bertipe skalar).

```
data.frame(df, stringsAsFactors = TRUE)
```

Keterangan argumen: df (dapat berupa matriks atau kumpulan beberapa variabel (vektor)), stringsAsFactors (mengkonversi string sebagai faktor (default))

```
v1 <- 1:5  
v2 <- c("Apel", "Belimbing", "Cherry", "Dukuh", "-")  
v3 <- c(50000, 20000, 35000, 18000, NA)  
  
dfHargaBuah <- data.frame(v1, v2, v3)  
colnames(dfHargaBuah) <- c("No", "Nama Buah", "Harga")
```

dfHargaBuah

##	No	Nama Buah	Harga
##	1	Apel	50000
##	2	Belimbing	20000
##	3	Cherry	35000
##	4	Dukuh	18000
##	5	-	NA

Subsetting

Subsetting adalah operasi untuk mengambil sebagian atau seluruh nilai dari suatu vektor, matriks atau list. Subsetting dapat dilakukan dengan menggunakan fungsi subset(). **Operator Subsetting** : [] ; [[]] ; \$

matriks

```
m <- matrix(c(10,2,5), nr
ow=2, ncol=3, byrow=T)
```

subsetting/indexing

```
m[1,1:3]
## [1] 10  2  5
m[1,]
## [1] 10  2  5
is.matrix(m[1:2,1:2])
## [1] TRUE
```

Vektor

```
(v <- c(c(12,34,56),78,23:25))
## [1] 12 34 56 78 23 24 25
```

subsetting/indexing

```
v[1]
## [1] 12
v[c(2,4)]
## [1] 34 78
v[seq(2, 7, by = 2)]
## [1] 34 78 24
v >= 30
## [1] FALSE TRUE TRUE TRUE
FALSE FALSE FALSE
```

Data Frame

```
dfHargaBuah
```

```
## No Nama Buah Harga
## 1      Apel 50000
## 2 Belimbing 20000
## 3      Cherry 35000
## 4      Dukuh 18000
## 5          -    NA
```

subsetting/indexing

```
dfHargaBuah$No
```

```
## [1] 1 2 3 4 5
```

```
dfHargaBuah$`Nama Buah`[3]
```

```
## [1] "Cherry"
```

```
dfBuah <- data.frame(dfHargaBuah[1:3,])
```

List

List adalah objek yang dapat berisi berbagai jenis elemen seperti vektor, fungsi, dan bahkan list lain di dalamnya. List dapat dilakukan dengan menggunakan fungsi `list()`.

```
l <- list(1, "Satu", TRUE)
ll <- list(vektor = v, matriks = m, l)

# subsetting
ll[[3]]
## [[1]]
## [1] 1
## [[2]]
## [1] "Satu"
## [[3]]
## [1] TRUE
ll[[3]][[2]]
## [1] "Satu"
```

Faktor

Faktor adalah tipe data dalam R yang terdiri dari sejumlah nilai yang berbeda-beda; data bertipe faktor sering disebut sebagai variabel kategori. Terdapat 2 jenis tipe kategori: nominal dan ordinal. Faktor dibuat dengan fungsi `factor()`.

```
factor(x = character(), levels, labels
= levels, ordered = is.ordered(x))
```

```
# Membuat vektor.
apple_colors <- c('green','green','yellow',
'red','red','red','green')

# Membuat objek faktor
(factor_apple <- factor(apple_colors))

## [1] green green yellow red red red green
## Levels: green red yellow
```



Function

Fungsi adalah sekumpulan perintah dan dapat menghasilkan suatu output tertentu. Fungsi boleh juga mempunyai input yang disebut argumen. Fungsi dibuat untuk menghindari pengulangan penulisan kode perintah yang sama atau mengurangi kerumitan.

Fungsi yang sering digunakan

str, diff, abs, log, exp, sqrt, factorial, runif, rnorm, sum, scale, eigen, sin, cos, tan, union, intersect, lm, summary, rownames, colnames, as.character.

Konstanta

pi, letters, LETTERS, month.abb, month.name

Membuat fungsi baru

```
function_name <- function(arg_1,  
arg_2, ...) {  
    Function body  
}
```

Contoh :

```
new.function <- function(a) {  
    for(i in 1:a) {  
        b <- i^2  
        print(b)  
    }  
}  
new.function(6)
```

Package

Paket dalam R adalah kumpulan fungsi dan data yang saling berkaitan serta dokumentasi yang menjelaskan fungsi dan data tersebut. Agar dapat digunakan, suatu paket harus diinstal terlebih dahulu. R mempunyai sumber tempat paket disimpan (repository), di antaranya CRAN (Comprehensive R Archive Network), Bioconductor, R-Forge. Paket R dapat diinstal baik secara online maupun offline.

Contoh paket R: base, stats, tools, utils, datasets, graphics, ggplot2, dll.

Berikut beberapa perintah yang berkaitan dengan paket R :

Paket	Kegunaan
<code>.packages(all.available = TRUE)</code>	Daftar paket yang telah diunduh ke dalam lingkungan R
<code>search()</code>	Daftar paket yang telah dimuat (<i>load</i>) ke dalam lingkungan R
<code>install.packages("<base>")</code>	Unduh paket yang bernama <base>
<code>library("<base>", lib.loc = "<lokasi>")</code>	Muat paket <base> yang diinstal di direktori <lokasi>
<code>remove.packages("<base>")</code>	Menghapus paket <base> dari direktori instalasi
<code>setRepositories()</code>	Memilih repository

Import Data csv (*Comma Delimited Text File*)

1. Import data dengan menggunakan syntax, berikut cara import data di R :

Syntax fungsi: `read.table(file, header = FALSE, sep = "", dec = ".", row.names, col.names)`

```
car.sales <- read.table("D:/car-sales.csv", header=TRUE, sep=";")
```

2. Import data dengan menggunakan menu

Pilih file csv yang akan diimport

Klik Import

Import Dataset

Name: car.sales

Input File: Toyota;Daihatsu;Honda
20803;8302;3755
21800;7518;4431
26287;8084;5179
24422;9298;5154
21050;9485;4439
26034;11886;5965
29448;7932;3684
25743;8423;4786
25725;10560;6354
22157;9600;4006

Encoding: Automatic

Heading: ☒ Yes ☐ No

Row names: Automatic

Separator: Semicolon

Decimal: Period

Quote: Double quote (")

Comment: None

na.strings: NA

☐ Strings as factors

Data Frame

Toyota	Daihatsu	Honda
20803	8302	3755
21800	7518	4431
26287	8084	5179
24422	9298	5154
21050	9485	4439
26034	11886	5965
29448	7932	3684
25743	8423	4786
25725	10560	6354
22157	9600	4006

Environment

Global Environment

Data

car.sales 10 obs. of 3 variables

Import Data xlsx (Excel)

1. Import data dengan menggunakan syntax, berikut cara import data di R :

Unduh dan muat paket "xlsx". Syntax fungsi : `read.xlsx (file, sheetName = NULL, as.data.frame = TRUE, header = TRUE)`

```
library(xlsx)
handphone_sales <- read.xlsx("D:/handphone-sales.xlsx", sheetName = "Sheet1")
```

2. Import data dengan menggunakan menu

Klik Browse → Pilih file excel yang akan diimport

Klik Import

Samsung (double)	Oppo (double)	Redmi (double)
100803	58302	33755
121800	67518	24431
116287	88084	55179
114422	49298	45154
131050	59485	34439
96034	91886	35965
109448	47932	43684
105743	68423	30786

Import Options:

Name: Max Rows:
 Sheet: Skip:
 Range: NA:
☒ First Row as Names ☐ Open Data Viewer

Code Preview:

```
library(readxl)
handphone_sales <- read_excel("D:/MA/PRAKTIKUM/ASPRAK GANJIL 2021/handphone-sales.xlsx")
```

Environment

Global Environment

Data

handphone_... 10 obs. of 3 variables

Tim Penyusun



Dr. Utriweni Mukhaiyar

Dosen KK Statistika

Kepala Laboratorium Statistika dan Komputasi Statistika



Fatia Amalia, S.Si

Asisten KK Statistika

Pengajar Semester I – 2020/2021



Dr. Udjianna S. Pasaribu

Dosen KK Statistika, MA2181 Analisis Data



Dr. Rr. Kurnia Novita Sari

Dosen KK Statistika, MA2181 Analisis Data



Dr. Sandy Vantika

Dosen KK Statistika,

MA2181 Analisis Data / MA2081 Statistika Dasar



Dr. Sapto Wahyu Indratno

Dosen KK Statistika, MA2082 Biostatistika



Yuli Sri Afrianti, S.Si., MT, MBA.

Dosen KK Statistika,

MA2181 Analisis Data / MA2081 Statistika Dasar



Dr. Utriweni Mukhaiyar

Dosen KK Statistika, MA2082 Biostatistika



Referensi

- KK Statistika ITB. 2019. *Modul R: Aplikasi Biostatistika dalam Industri Kelapa Sawit*. Bandung.

Selamat Praktikum!