Nama: Mendari pertiwi

NIM: 064002200037

Hari/Tanggal: Hari, 12 juli 2023



Praktikum Statistika

MODUL 2

Nama Dosen: **Dedy Sugiarto**

Nama Asisten Labratorium:

- 1. Elen Fadilla Estri 064002000008
- 2. Rukhy Zaifa Aduhalim 064002000041

Tipe Data, Filter Data & Koneksi ke Database

1. Teori Singkat

R (juga dikenal sebagai GNU S) adalah bahasa pemrograman dan perangkat lunak untuk analisis statistika dan grafik. R dibuat oleh Ross Ihaka dan Robert Gentleman di Universitas Auckland, Selandia Baru, dan kini dikembangkan oleh R Development Core Team, di mana Chambers merupakan anggotanya. R dinamakan sebagian setelah nama dua pembuatnya (Robert Gentleman dan Ross Ihaka), dan sebagian sebagian dari permainan nama dari S.

Bahasa R kini menjadi standar de facto di antara statistikawan untuk pengembangan perangkat lunak statistika, serta digunakan secara luas untuk pengembangan perangkat lunak statistika dan analisis data. R merupakan bagian dari proyek GNU. Kode sumbernya tersedia secara bebas di bawah Lisensi Publik Umum GNU, dan versi biner prekompilasinya tersedia untuk berbagai sistem operasi. R menggunakan antarmuka baris perintah, meski beberapa antarmuka pengguna grafik juga tersedia.

R menyediakan berbagai teknik statistika (permodelan linier dan nonlinier, uji statistik klasik, analisis deret waktu, klasifikasi, klasterisasi, dan sebagainya) serta grafik. R, sebagaimana S, dirancang sebagai bahasa komputer sebenarnya, dan mengizinkan penggunanya untuk menambah fungsi tambahan dengan mendefinisikan fungsi baru. Kekuatan besar dari R yang lain adalah fasilitas grafiknya, yang menghasilkan grafik dengan kualitas publikasi yang dapat memuat simbol matematika. R memiliki format dokumentasi seperti LaTeX, yang digunakan

untuk menyediakan dokumentasi yang lengkap, baik secara daring (dalam berbagai format) maupun secara cetakan.

RStudio merupakan integrated development environment (IDE) khusus bagi bahasa pemrograman R. Software ini menyediakan R console, code editor dengan syntax highlighting, code completion dan direct execution, environment, history, connections, dan fitur-fitur tambahan lainnya seperti file manager, packages manager, help, plot viewer, hingga project versioning menggunakan git. RStudio sebenarnya memiliki dua versi, yaitu open source (gratis) dan commercial edition (berbayar). RStudio juga tidak hanya terbatas dalam bentuk aplikasi dekstop, melainkan terdapat versi RStudio Server, yaitu RStudio yang dapat diakses melalui browser yang terhubung dengan suatu jaringan komputer. Untuk saat ini, versi RStudio yang akan dijelaskan hanyalah RStudio open source berbasis dekstop saja.

2. Alat dan Bahan

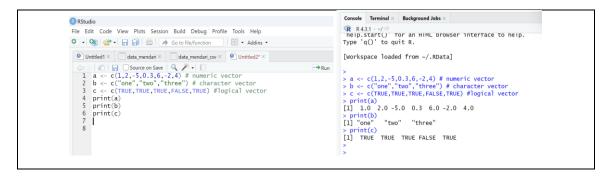
Hardware: Laptop/PC Software: R Studio

3. Elemen Kompetensi

Terdapat beberapa tipe data di R antara lain vektor, matriks dan data frame. Cantumkan setiap output yang dihasilkan dari console Rstudio, ke kolom yang sudah disediakan.

- a. Latihan pertama Vektor
 - 1. Tuliskan Perintah berikut ini di window R Script:

```
a < c(1,2,-5,0.3,6,-2,4) # numeric vector
b <- c("one","two","three") # character vector
c <- c(TRUE,TRUE,TRUE,FALSE,TRUE) #logical vector
print(a)
print(b)
print(c)
```

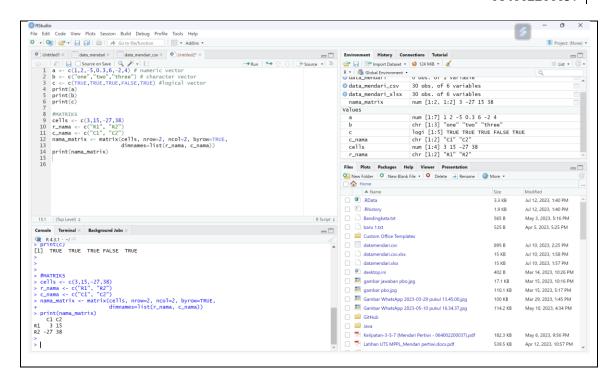


b. Latihan Kedua – Matriks

1. Seluruh kolom dalam sebuah matriks harus memiliki tipe yang sama (numerik semua, karakter semua, dll) dan memiliki panjang yang sama.

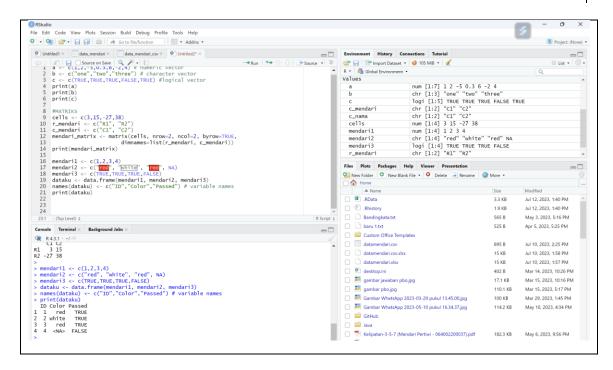
*gunakan nama variable dengan nama anda masing-masing

```
#MATRIKS
cells <- c(3,15,-27,38)
r_nama <- c("R1", "R2")
c_nama <- c("C1", "C2")
nama_matrix <- matrix(cells, nrow=2, ncol=2, byrow=TRUE,
           dimnames=list(r_nama, c_nama))
print(nama_matrix)
```



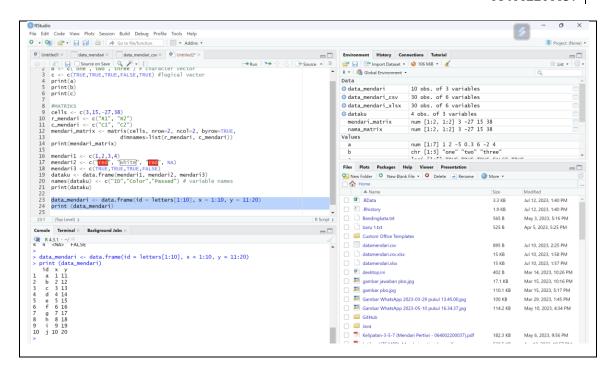
- c. Latihan Ketiga Data Frame
 - 1. Mengubah data input menjadi data frame
 - *gunakan nama variable dengan nama anda masing-masing

```
nama1 <- c(1,2,3,4)
nama2 <- c("red", "white", "red", NA)
nama3 <- c(TRUE,TRUE,TRUE,FALSE)
dataku <- data.frame(nama1, nama2, nama3)
names(dataku) <- c("ID", "Color", "Passed") # variable names
print(dataku)
```



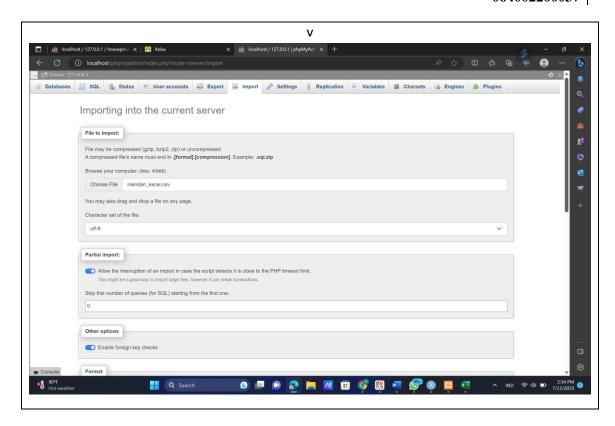
2. Selanjutnya ketikkan perintah dibawah ini

```
data_nama <- data.frame(id = letters[1:10], x = 1:10, y = 11:20)
print (data_nama)
```

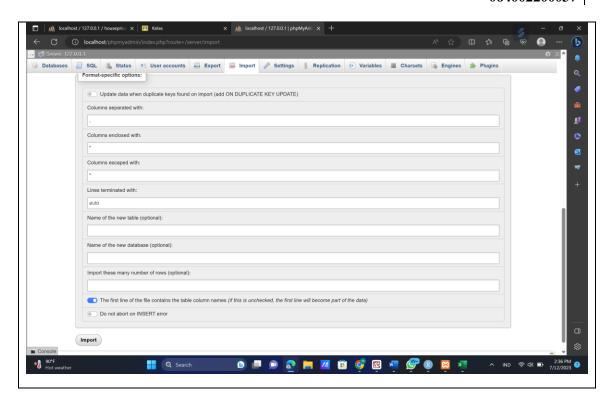


Beberapa fungsi penting terkait data frame:

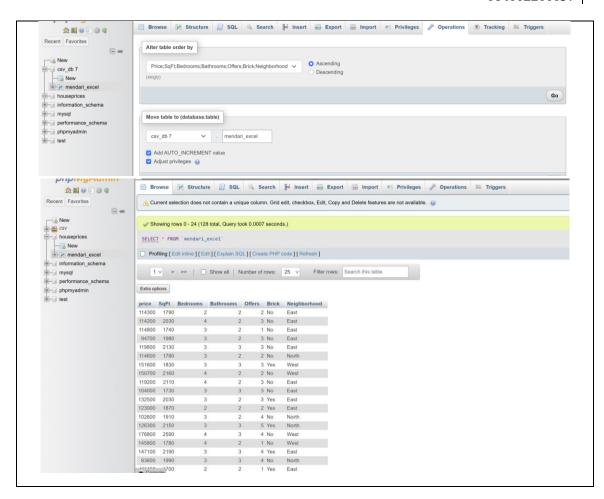
- head() shows first 6 rows
- tail() shows last 6 rows
- dim() returns the dimensions of data frame (i.e. number of rows and number of columns)
- nrow() number of rows
- ncol() number of columns
- str() structure of data frame name, type and preview of data in each column
- names() shows the names attribute for a data frame, which gives the column names.
- d. Latihan Keempat Koneksi ke Database
 - 1. Buat sebuah nama database terlebih dahulu dengan nama houseprices di phpmyadmin, Lalu klik menu import



2. Pilih file yang ingin di import ke database (untuk file nama_excel.csv), Ceklis the first line of the file contains the table column name untuk membuat baris pertama pada file excel tersebut menjadi nama atributnya atau nama kolomnya pada database.



3. Klik go, Rename nama tabel sesuai dengan nama anda



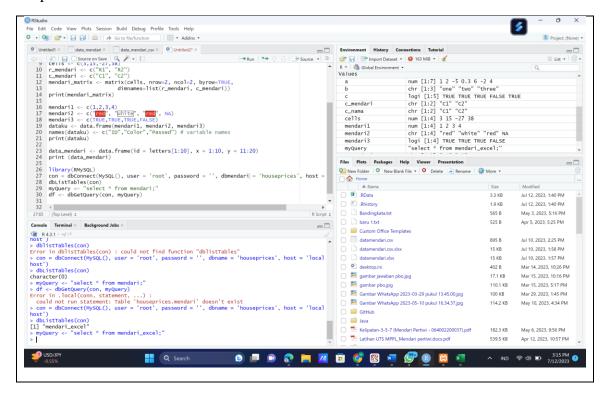
4. Kembali ke R Studio, lalu instal dahulu library yang dibutuhkan pada RStudio. Jika belum tersedia, maka lakukan instruksi:

install.packages("RMySQL")

```
'lib' is unspecified)
also installing the dependency 'DBI'
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.3/DBI_1.1.3.zip'
Content type 'application/zip' length 774701 bytes (756 KB)
downloaded 756 KB
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.3/RMySQL_0.10.25.zip'
Content type 'application/zip' length 475309 bytes (464 KB)
downloaded 464 KB
package 'DBI' successfully unpacked and MD5 sums checked
package 'RMySQL' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
        C:\Users\mendari pertiwi\AppData\Local\Temp\RtmpIPvZc3\downloaded_packages
```

5. Lalu jalankan perintah dibawah ini

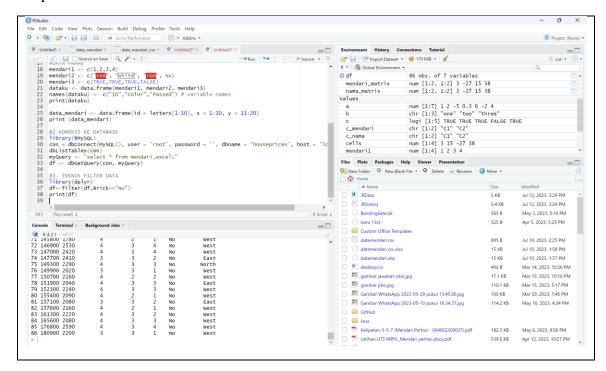
```
library(RMySQL)
con = dbConnect(MySQL(), user = 'root', password = ", dbname = 'houseprices', host =
'localhost')
dbListTables(con)
myQuery <- "select * from nama;"
df <- dbGetQuery(con, myQuery)
```



- e. Latihan Kelima Teknik Filter Data
 - 1. Jalankan perintah dibawah ini:
 - *Perintah ini akan menampilkan 86 baris data hasil filter.

```
library(dplyr)
df<-filter(df,Brick=="No")
print(df)
```

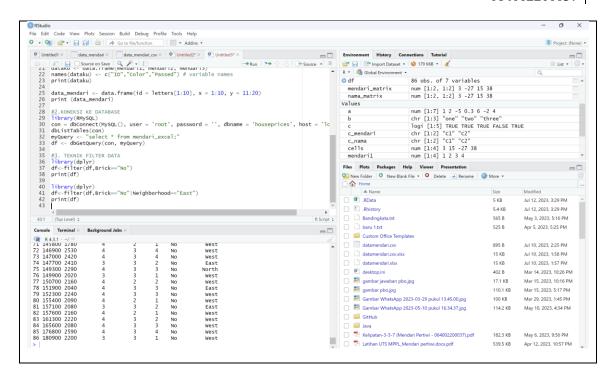
Output:



2. Jalankan perintah dibawah ini:

*Perintah ini akan menampilkan 105 baris data hasil filter.

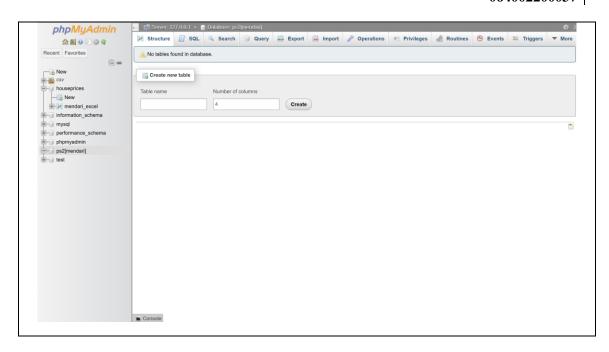
```
library(dplyr)
df<-filter(df,Brick=="No"|Neighborhood=="East")
print(df)
```



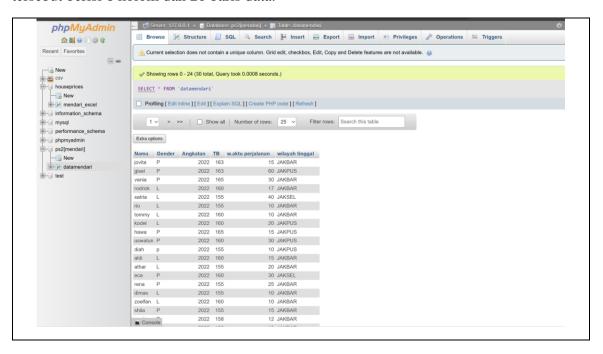
f. Latihan Keenam – Tugas

Buat sebuah database serta tabel di dalamnya (bisa gunakan data teman dipraktikum ke 1). Lakukan koneksi R ke database serta berikan beberapa filter data sesuai yang anda inginkan. Tampilkan data tersebut

1. Buatlah terlebih dahulu Database baru dengan format PS2[NamaAnda]



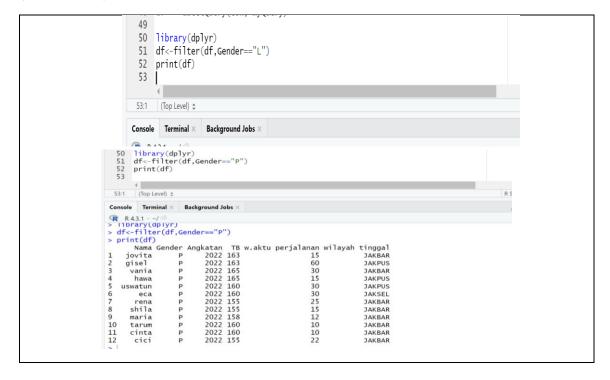
2. Import file berformat csv yang telah anda buat sebelumnya di Praktikum 1, dimana file tersebut berisi 6 kolom dan 20 baris data.



3. Koneksikan R ke Database tersebut sebagaimana yang telah anda lakukan pada Elemen Kompetensi 1 dimodul kedua ini dengan menyesuaikan kembali nama Database baru yang sudah dibuat.

```
> library(RMySQL)
> con = dbConnect(MySQL(), user = 'root', password = '', dbname = 'PS2[mendari]', host = 'loca
lhost')
> dbListTables(con)
[1] "datamendari"
> myQuery <- "select * from datamendari;"</pre>
> df <- dbGetQuery(con, myQuery)</pre>
> install.packages("dplyr")
> iibrary(dpiyr)
> df<-filter(df,Gender=="P")</pre>
   df<-filterus.,
print(df)
Nama Gender Angkatan TB
jovita P 2022 163
gisel P 2022 163
- P 2022 165
                                                TB w.aktu perjalanan wilayah tinggal
                                                                                                    JAKBAR
JAKPUS
                                                                                30
                                                                                                    JAKBAR
                                     2022 165
2022 160
2022 160
2022 155
2022 155
2022 158
2022 160
2022 160
2022 155
                                                                                15
30
                                                                                                    JAKPUS
JAKPUS
     uswatun
                                                                                                    JAKSEL
JAKBAR
6
7
8
9
            eca
           rena
         shila
                                                                                                    JAKBAR
         maria
10
                                                                                10
10
22
         tarum
                                                                                                    JAKBAR
11
12
                                                                                                     JAKBAR
                                                                                                     JAKBAR
           cici
```

4. Lakukan filter data terhadap Kolom Gender, untuk melihat berapa baris data Pria/Wanita (Pilih salah 1).



4. File Praktikum

Github	Reposit	ory
--------	---------	-----

5. Soal Latihan

Soal:

- 1. Ada berapa tipe data yang data di deklrasikan oleh Rstudio?
- 2. Bagaimana cara memanggil sebuah tabel di dalam database ke Rstudio?

Jawaban:

- 1. Secara umum tipe data ada: Numeric, Integer, Character, Logical, factor, date, posixct dan posixit,complex,raw.
- 2. Instal dan aktifkan paket R yang sesuai untuk sistem database yang Anda gunakan. jika Anda menggunakan MySQL maka ketik: install.packages("RMySQL") lalu library(RMySQL)

```
con = dbConnect(MySQL(), user = 'root', password = ", dbname = 'PS2[mendari]', host =
'localhost')
dbListTables(con)
myQuery <- "select * from datamendari;"
```

6. Kesimpulan

df <- dbGetQuery(con, myQuery)</pre>

- a. Dalam pengerjaan praktikum Statistika, pada praktikum kali ini mempelajari tentang Tipe Data, Filter Data & Koneksi ke Database
- b. Kita juga dapat mengetahui tipe data yaitu Tipe Data: untuk menggolongkan nilai atau variabel berdasarkan jenis atau sifatnya. Dalam RStudio, ada beberapa tipe data umum seperti Numeric (bilangan desimal), Integer (bilangan bulat), Character (teks atau karakter), Logical (nilai kebenaran), dan lainnya. Memahami tipe data penting karena hal ini mempengaruhi cara Anda menyimpan, memanipulasi, dan menganalisis data. Filter data: roses memilih subset data berdasarkan kondisi tertentu. Dalam RStudio, Anda dapat menggunakan fungsi filter() atau subset() untuk melakukan filtering pada data frame.selanjutnya kita mengetahui koneksi ke data base: untuk terhubung ke database eksternal dari RStudio dan mengambil data dari tabel yang ada di dalamnya.jadi itulah yang dapat kita ambil dari praktikum ini.

7. Cek List (**✓**)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	√	
2.	Latihan Kedua	√	
3.	Latihan Ketiga	√	
4.	Latihan Keempat	√	
5.	Latihan Kelima	√	
6.	Latihan Keenam	√	

8. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	10 Menit	menarik
2.	Latihan Kedua	15 Menit	menarik
3.	Latihan Ketiga	20 Menit	menarik
4.	Latihan Keempat	20 Menit	menarik
5.	Latihan Kelima	20 Menit	menarik

6.	Latihan Keenam	20 Menit	menarik

Keterangan:

- 1. Menarik
- 2. Baik
- 3. Cukup
- 4. Kurang