Nama: Adrian Maulana Y

> NIM: 065002100036

Hari/Tanggal: Kamis, 22 September 2022



Praktikum Statistika

MODUL 2

Nama Dosen: Dedy Sugiarto

Nama Asisten Labratorium:

- 1. Azhar Rizki Zulma 065001900001
- 2. Arfa Maulana 064001900039

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

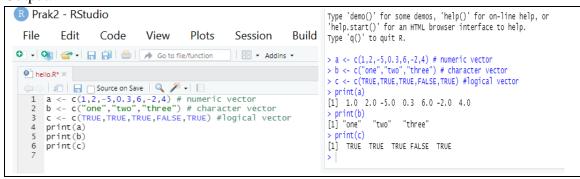
3. Elemen Kompetensi

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

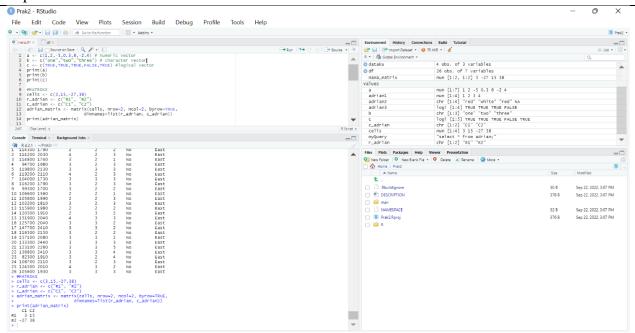
1. Tuliskan Perintah berikut ini di window R Script:

```
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

Output:



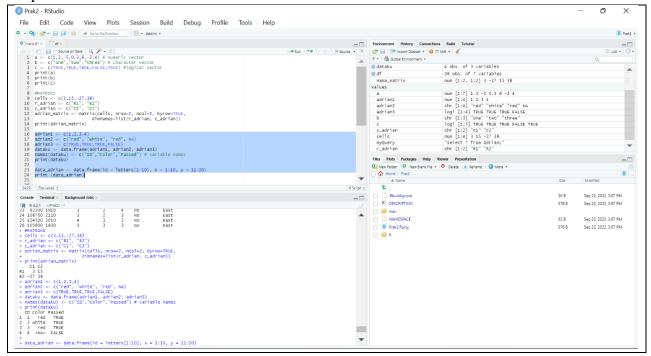
- c. Latihan Ketiga Data Frame
 - 1. Mengubah data input menjadi data frame
 - *gunakan nama variable dengan nama anda masing-masing

```
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)
```



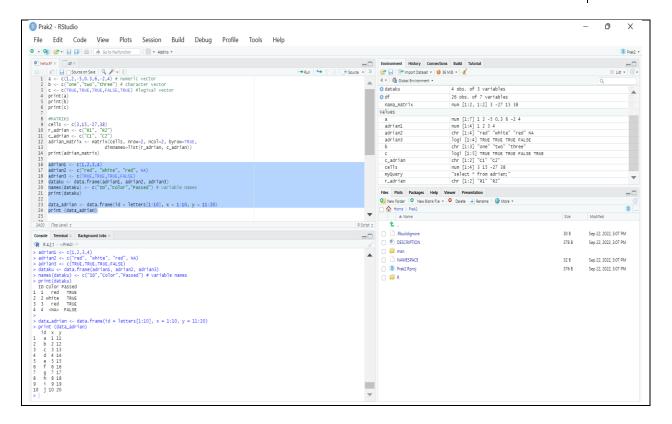
Jurusan Teknik Informatika & Sistem Informasi

Output:



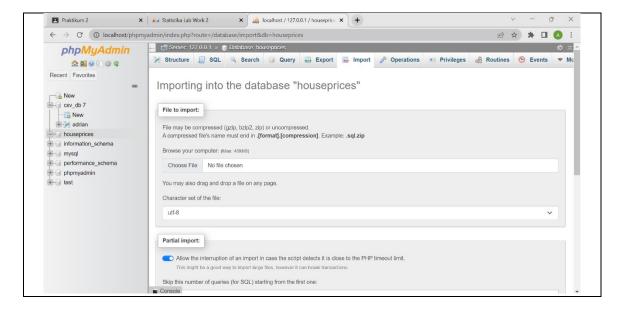
2. Selanjutnya ketikkan perintah dibawah ini

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

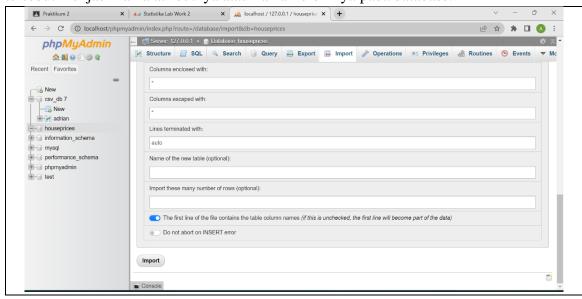


Beberapa fungsi penting terkait data frame:

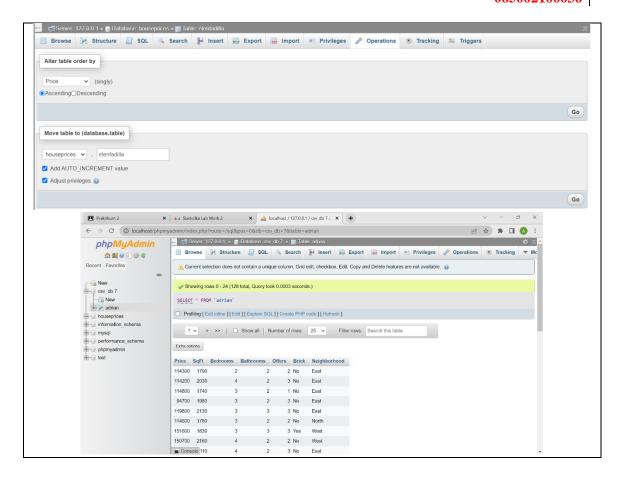
- head() shows first 6 rows
- tail() shows last 6 rows
- dim() returns the dimensions of data frame (i.e. number of rows and number of columns)
- nrow() number of rows
- ncol() number of columns
- str() structure of data frame name, type and preview of data in each column names() shows the names attribute 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 houseprices.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") install.packages("dplyr")

```
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/dplyr_1.0.10.zip'
Content type 'application/zip' length 1305159 bytes (1.2 MB)
downloaded 1.2 MB

package 'fansi' successfully unpacked and MD5 sums checked
package 'pkgconfig' successfully unpacked and MD5 sums checked
package 'ellipsis' successfully unpacked and MD5 sums checked
package 'cli' successfully unpacked and MD5 sums checked
package 'cli' successfully unpacked and MD5 sums checked
package 'utf8' successfully unpacked and MD5 sums checked
package 'glue' successfully unpacked and MD5 sums checked
package 'glue' successfully unpacked and MD5 sums checked
package 'iffecycle' successfully unpacked and MD5 sums checked
package 'magrittr' successfully unpacked and MD5 sums checked
package 'rlang' successfully unpacked and MD5 sums checked
package 'rlang' successfully unpacked and MD5 sums checked
package 'tibble' successfully unpacked and MD5 sums checked
package 'titosecesfully unpacked and MD5 sums checked
package 'titosecesfully unpacked and MD5 sums checked
package 'pillar' successfully unpacked and MD5 sums checked
```

5. Lalu jalankan perintah dibawah ini

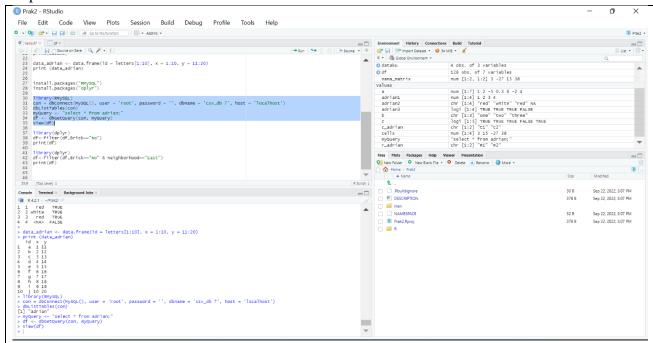
library(RMySQL)



Jurusan Teknik Informatika & Sistem Informasi Fakultas Teknologi Industri – Universitas Trisakti

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

Output:

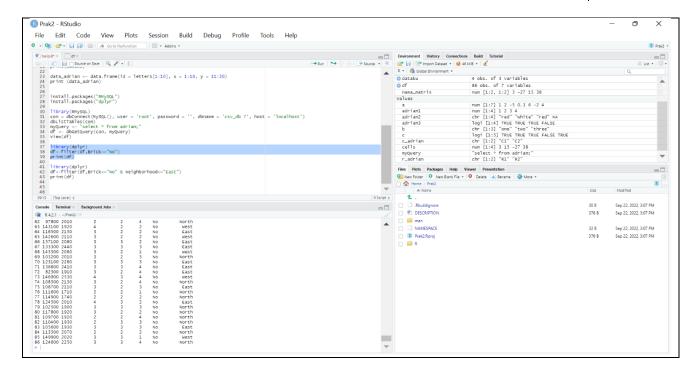


- e. Latihan Kelima Teknik Filter Data
 - 1. Jalankan perintah dibawah ini:

*Perintah ini akan menampilkan 86 baris data hasil filter.

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

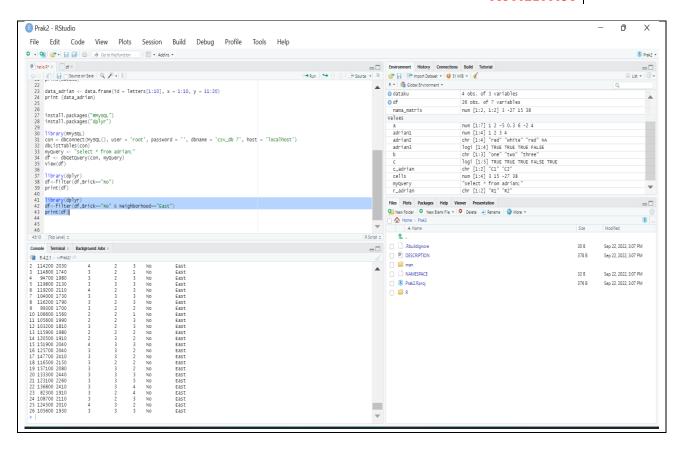




2. Jalankan perintah dibawah ini:

*Perintah ini akan menampilkan 105 baris data hasil filter.

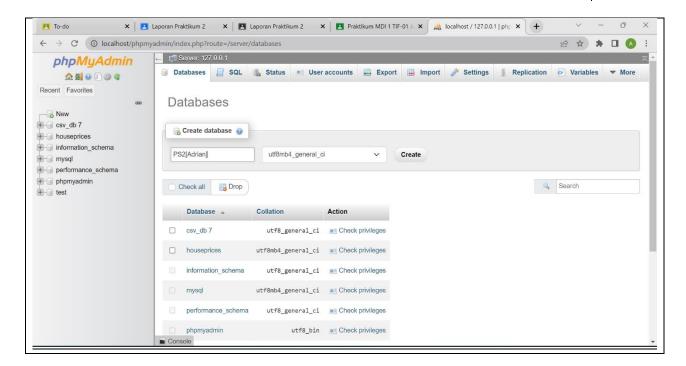
```
library(dplyr)
df<-filter(df,Brick=="No"|Neighborhood=="East")
print(df)
```



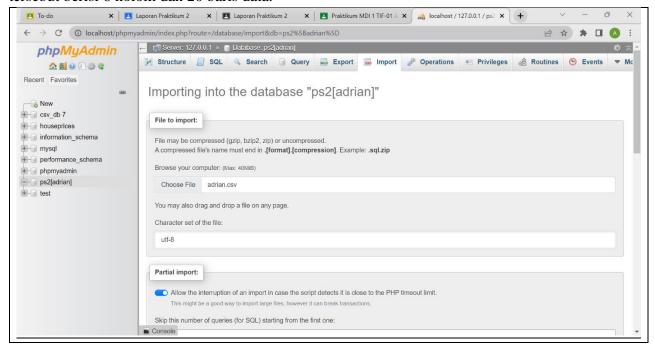
f. Latihan Keenam – Tugas

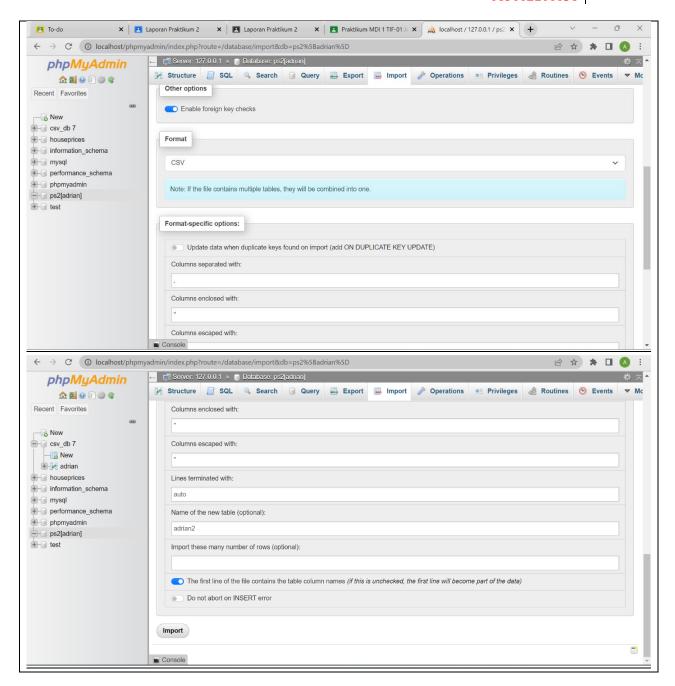
Buat sebuah database serta tabel di dalamnya (bisa gunakan data teman dipraktikum ke 1). Lakukan koneksi R ke database serta berikan beberapa filter data sesuai yang anda inginkan. Tampilkan data tersebut

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



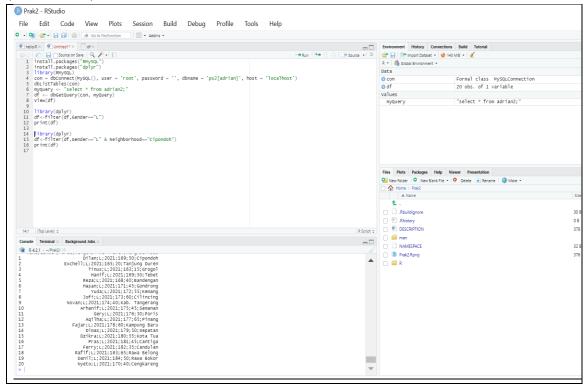
2. Import file berformat csv yang telah anda buat sebelumnya di Praktikum 1, dimana file tersebut berisi 6 kolom dan 20 baris data.





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.

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

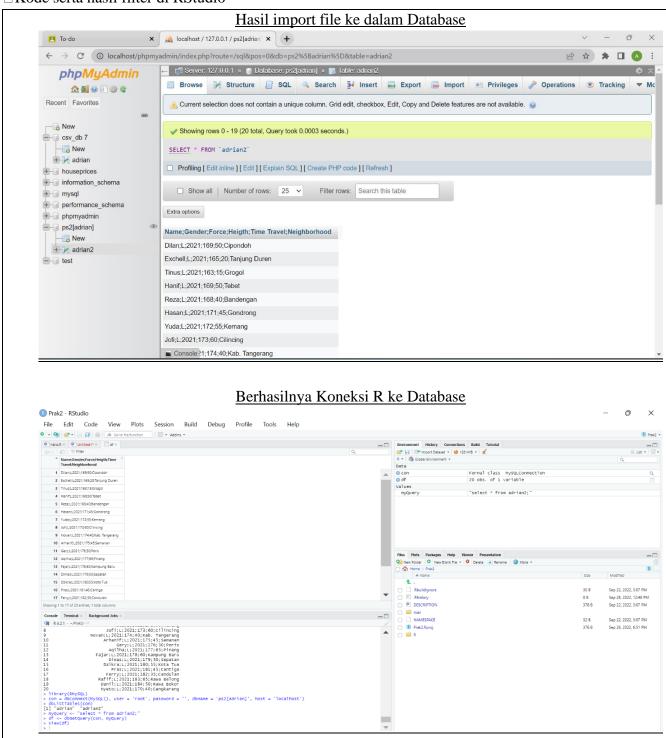


- 5. Lampirkan Screenshot
- ☐ Hasil import file kedalam Database

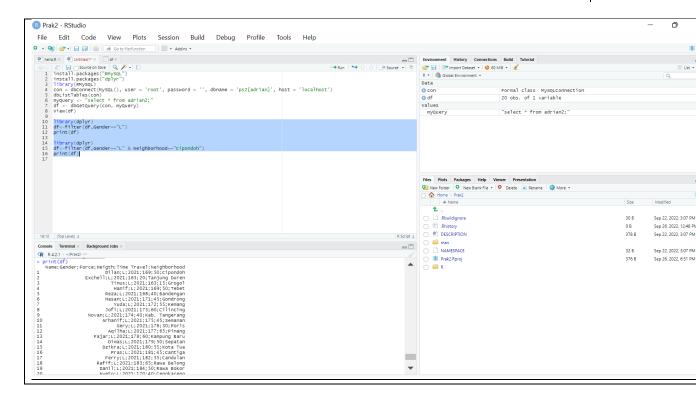


Jurusan Teknik Informatika & Sistem Informasi Fakultas Teknologi Industri – Universitas Trisakti

- □Berhasilnya koneksi R ke Database
- □Kode serta hasil filter di RStudio



Kode serta hasil filter di RStudio



4. File Praktikum

Github Repository:

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. Tipe data yang didekslarasikan oleh RStudio, diantaranya:
- -Character(Char), -Numeric, -Integer, -Logical, -Complex, -Vektor, -Matriks dan -Data frame
- 2. Cara memanggil sebuah table di dalam database ke RStudio menggunakan perintah

library(RMySQL)

con = dbConnect(MySQL(), user = 'root', password = ", dbname = 'ps2[adrian]', host = 'localhost')

dbListTables(con)

myQuery <- "select * from adrian2;"



Jurusan Teknik Informatika & Sistem Informasi Fakultas Teknologi Industri – Universitas Trisakti df <- dbGetQuery(con, myQuery)</pre> View(df)

Setelah berhasil menginstal package (RMySQL).

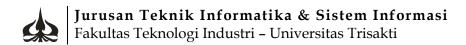
6. Kesimpulan

- a. Dalam pengerjaan praktikum Statistika, kita harus mengetahui apa saja tipe data yang didekslarasikan oleh RStudio, diantaranya yaitu Vektor, Matriks, dan Data Frame.
- b. Kita juga dapat mengetahui bahwa R dapat di koneksikan kedalam Database dari localhost/phpMyAdmin web, dengan membuat Database dan meng import file ke dalam database yang sudah dibuat, kemudian menggunakan sintaks/perintah koneksi ke database sebagaimana yang sudah tertera dalam modul.

7. Cek List (**✓**)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	$\sqrt{}$	
2.	Latihan Kedua	$\sqrt{}$	
3.	Latihan Ketiga	$\sqrt{}$	
4.	Latihan Keempat	$\sqrt{}$	
5.	Latihan Kelima	V	
6.	Latihan Keenam		

8. Formulir Umpan Balik



No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	20 Menit	1
2.	Latihan Kedua	20 Menit	1
3.	Latihan Ketiga	28 Menit	1
4.	Latihan Keempat	30 Menit	1
5.	Latihan Kelima	30 Menit	1
6.	Latihan Keenam	35 Menit	1

Keterangan:

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