Nama: (Isi Nama Anda)

NIM: (Isi NIM Anda)

Hari/Tanggal: Hari, Tanggal Bulan 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

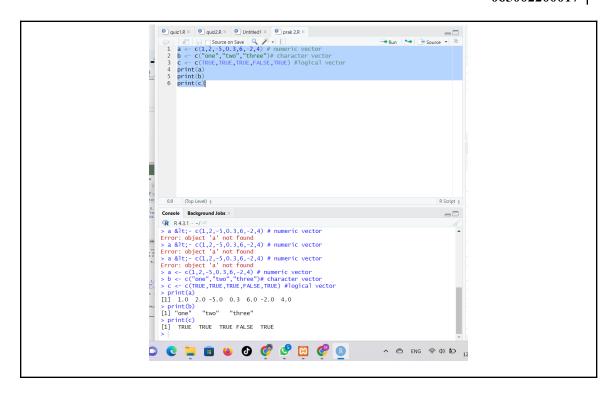
Terdapat beberapa tipe data di R antara lain vektor, matriks dan data frame. Cantumkan setiap output yang dihasilkan dari console Rstudio, ke kolom yang sudah disediakan.

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)
```





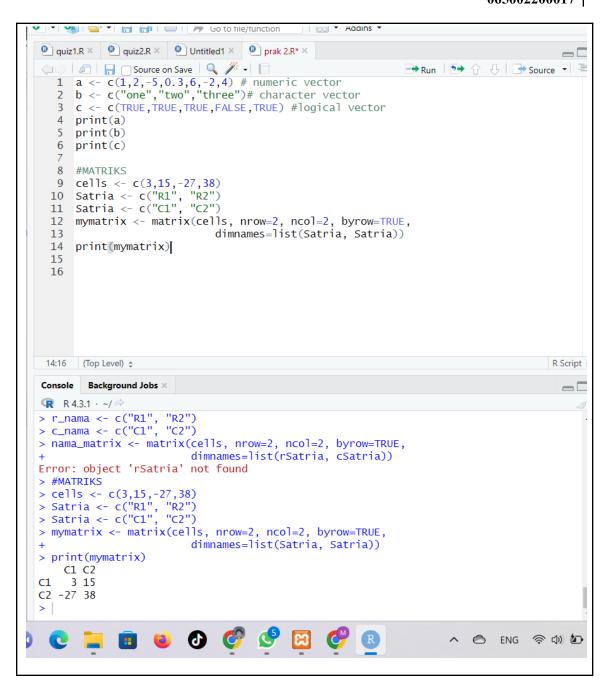
b. Latihan Kedua – Matriks

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)
```

```
quiz1.R × Q quiz2.R × Q Untitled1 × Q prak 2.R* ×

↓□□ | □□ | □ Source on Save | □ | ▼ ▼ | □ |

                                                                   ➡ Run | ➡ 介 ⇩ | ➡ Source ▾ | ਵ
   1 a <- c(1,2,-5,0.3,6,-2,4) # numeric vector
2 b <- c("one", "two", "three") # character vector
   3 c <- c(TRUE, TRUE, FALSE, TRUE) #logical vector
   4 print(a)
   5 print(b)
   6 print(c)
   8 #MATRIKS
  9 cells <- c(3,15,-27,38)
10 Satria <- c("R1", "R2")
11 Satria <- c("C1", "C2")
  12 mymatrix <- matrix(cells, nrow=2, ncol=2, byrow=TRUE,
  13
                                dimnames=list(Satria, Satria))
  14 print(mymatrix)
  15
  16
      satria1 <- c(1,2,3,4)
      satria2 <- c("red", "white", "red", NA)
satria3 <- c(TRUE,TRUE,TRUE,FALSE)</pre>
  17
  18
      dataku <- data.frame(satria1, satria2, satria3)</pre>
      names(dataku) <- c("ID","Color","Passed") # variable names
  20
  21
      print(dataku)
  22
  16:1
      (Top Level) $
                                                                                              R Script $
Console Background Jobs
                                                                                                R 4.3.1 · ~/ ≈
> print(mymatrix)
    C1 C2
     3 15
C2 -27 38
> satria1 <- c(1,2,3,4)
> satria2 <- c("red", "white", "red", NA)</pre>
> satria3 <- c(TRUE,TRUE,TRUE,FALSE)</pre>
> dataku <- data.frame(satria1, satria2, satria3)</pre>
> names(dataku) <- c("ID","Color","Passed") # variable names
  print(dataku)
  ID Color Passed
1
        red
              TRUE
   2 white
               TRUE
3
   3
       red
              TRUE
       <NA> FALSE
```

2. Selanjutnya ketikkan perintah dibawah ini

```
data nama \leftarrow data.frame(id = letters[1:10], x = 1:10, y = 11:20)
print (data nama)
```

```
Quiz1.R × Quiz2.R × Qui
                                                                                                                                                                                                                                                                              4 print(a)
         5 print(b)
         6 print(c)
         8 #MATRIKS
         9 cells <- c(3,15,-27,38)
     10 Satria <- c("R1", "R2")
11 Satria <- c("C1", "C2")
                  mymatrix <- matrix(cells, nrow=2, ncol=2, byrow=TRUE,</pre>
      12
                                                                                          dimnames=list(Satria, Satria))
     13
     14 print(mymatrix)
     15
      16 satria1 <- c(1,2,3,4)
     17 satria2 <- c("red", "white", "red", NA)
18 satria3 <- c(TRUE, TRUE, TRUE, FALSE)
                 dataku <- data.frame(satria1, satria2, satria3)
names(dataku) <- c("ID","Color","Passed") # variable names</pre>
      19
       20
      21
                  print(dataku)
      22
      23
                  data_satria \leftarrow data.frame(id = letters[1:10], x = 1:10, y = 11:20)
       24
                  print (data_satria)
      25
    18:35
                 (Top Level) $
                                                                                                                                                                                                                                                                        R Script $
                         Background Jobs X
 Console
                                                                                                                                                                                                                                                                              R 4.3.1 · ~/ ≈
 3 3
                 red TRUE
4 4 <NA> FALSE
     data_satria \leftarrow data.frame(id = letters[1:10], x = 1:10, y = 11:20)
> print (data_satria)
          id x y
             a 1 11
1
            b
3
            c 3 13
            d 4 14
5
             e 5 15
6
                      6 16
             g 7 17
           h 8 18
9
           i 9 19
10
         j 10 20
                                                                            3 🚱 👺 🔞 🧬 📵
                                                                                                                                                                                                                    ^ ♠ ENG ♠ ♠ ♠ 🐿
```

Beberapa fungsi penting terkait data frame:

head() - shows first 6 rows

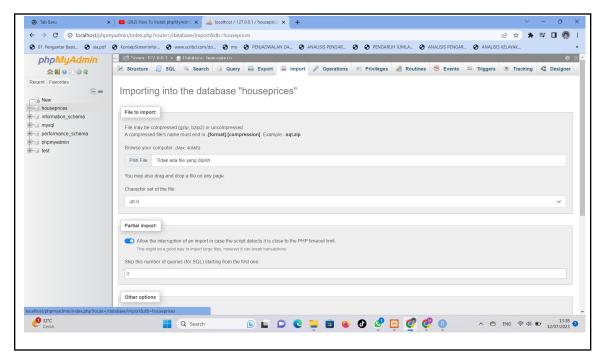
tail() - shows last 6 rows

dim() - returns the dimensions of data frame (i.e. number of rows and number of columns)

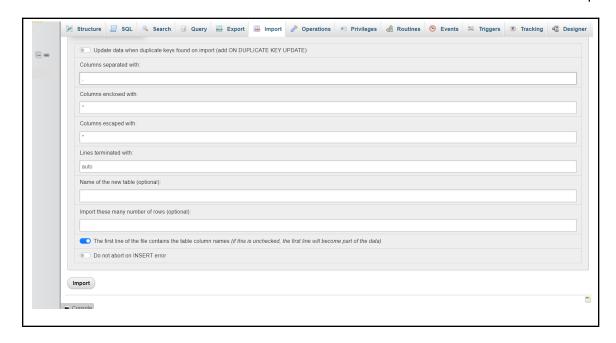
nrow() - number of rows

ncol() - number of columns

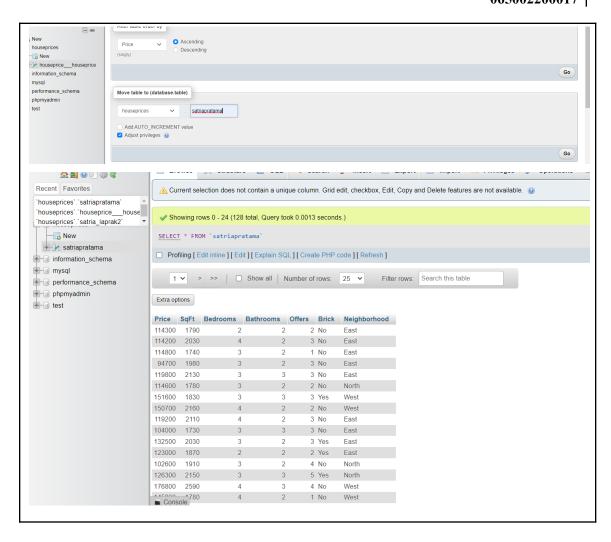
- str() structure of data frame name, type and preview of data in each column names() - shows the names attribute for a data frame, which gives the column names.
- d. Latihan Keempat Koneksi ke Database
 - 1. Buat sebuah nama database terlebih dahulu dengan nama houseprices di phpmyadmin, Lalu klik menu 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



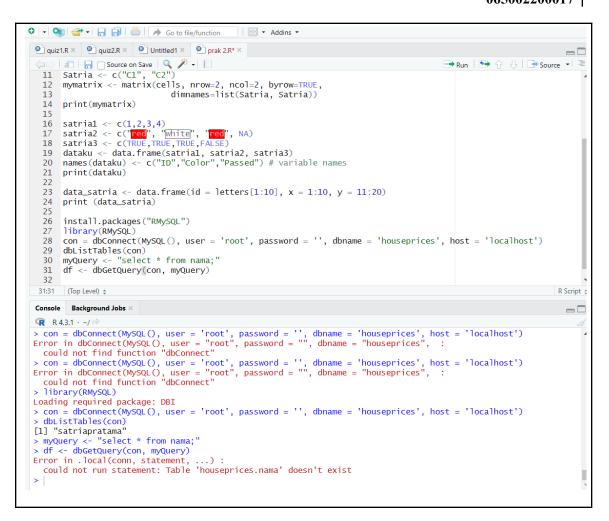
4. Kembali ke R Studio, lalu instal dahulu library yang dibutuhkan pada RStudio. Jika belum tersedia, maka lakukan instruksi:

install.packages("RMySQL")

```
Console Background Jobs ×
R 4.3.1 · ~/ ≈
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\satria\AppData\Local\Temp\RtmpULh3wt\downloaded_packages
```

5. Lalu jalankan perintah dibawah ini

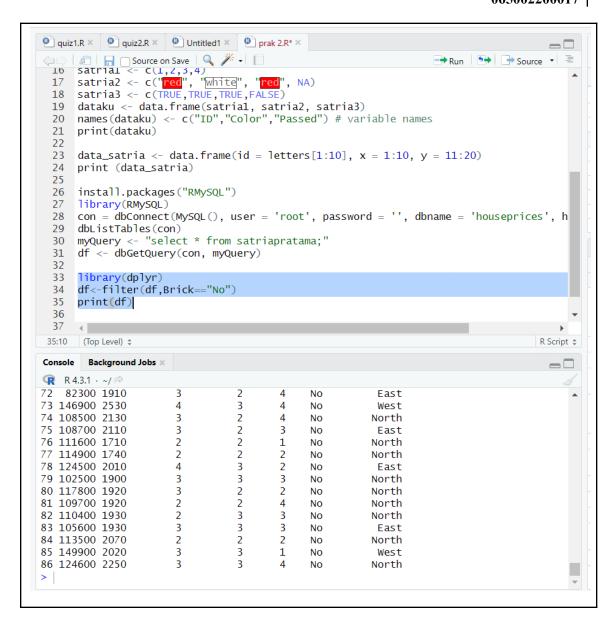
```
library(RMySQL)
con = dbConnect(MySQL(), user = 'root', password = ", dbname = 'houseprices', host =
'localhost')
dbListTables(con)
myQuery <- "select * from nama;"
df <- dbGetQuery(con, myQuery)</pre>
```



- 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)
```





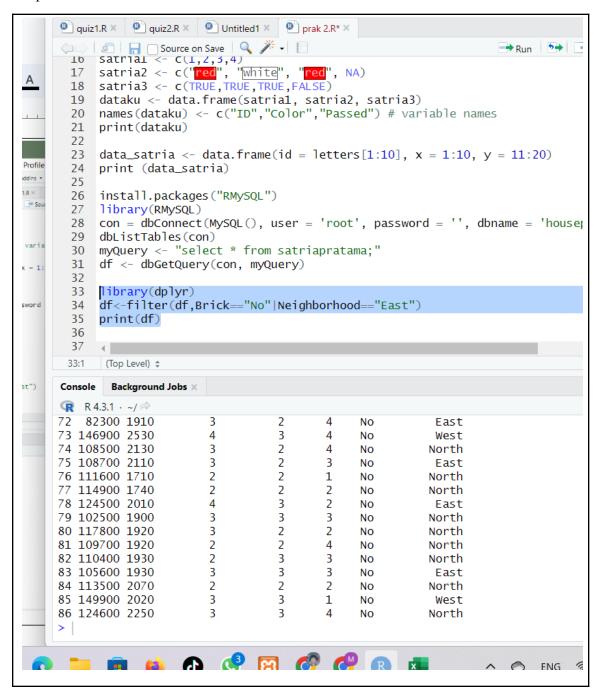
2. Jalankan perintah dibawah ini:

*Perintah ini akan menampilkan 105 baris data hasil filter.

library(dplyr)



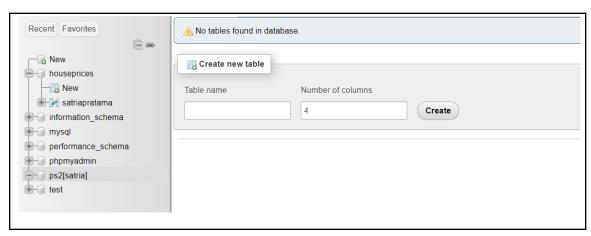
```
df<-filter(df,Brick=="No"|Neighborhood=="East")
print(df)
```



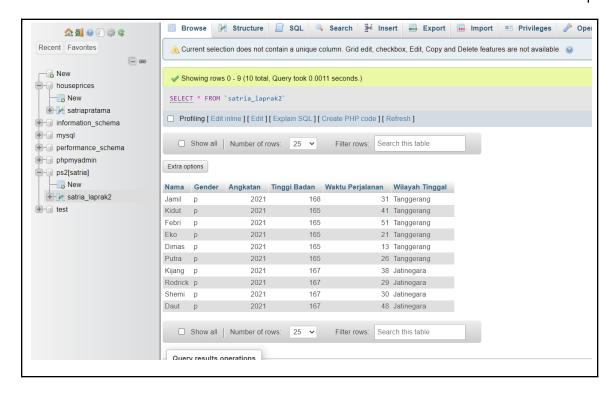
f. Latihan Keenam – Tugas

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]



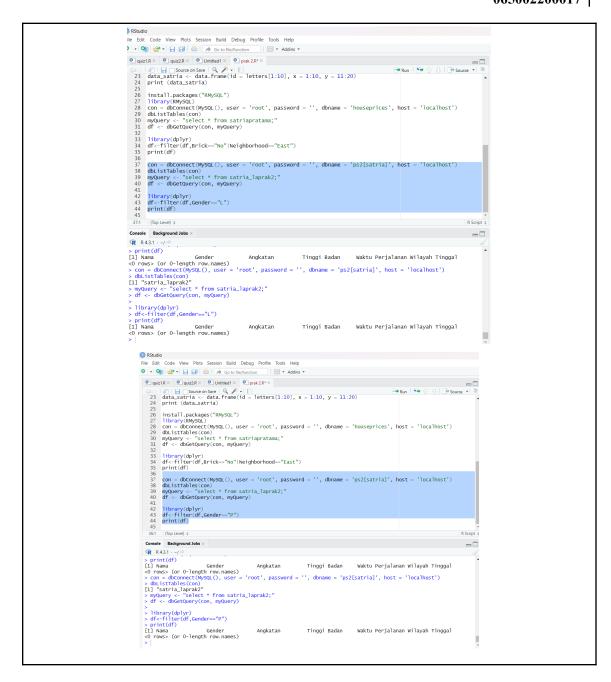
2. Import file berformat csv yang telah anda buat sebelumnya di Praktikum 1, dimana file tersebut berisi 6 kolom dan 20 baris data.



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.

```
con = dbConnect(MySQL(), user = 'root', password = '', dbname = 'ps2[satria]', host = 'localhost')
 dbListTables(con)
 myQuery <- "select * from satria_laprak2;"</pre>
 df <- dbGetQuery(con, myQuery)</pre>
 library(dplyr)
 df<-filter(df,Gender=="L")</pre>
 print(df)
  con = dbConnect(MySQL(), user = 'root', password = '', dbname = 'ps2[satria]', host = 'localhost')
38 dbListTables(con)
   myQuery <- "select * from satria_laprak2;"</pre>
10 df <- dbGetQuery(con, myQuery)</pre>
12 library(dplyr)
43 df<-filter(df,Gender=="P")</pre>
   print(df)
```

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



4.	File	Pra	ktikuı	m
----	------	-----	--------	---

Gitnub 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 dideklarasikan oleh RStudio meliputi numerik (angka), karakter (teks), faktor (kategorikal), logikal (benar/salah), dan kompleks (angka kompleks).
- 2. Untuk memanggil tabel dari database ke RStudio: a. Pastikan terhubung dengan database yang sesuai menggunakan paket R yang tepat. b. Gunakan fungsi yang sesuai untuk mengambil data dari database, seperti "dbGetQuery()" untuk MySQL. c. Simpan hasil pengambilan data ke dalam objek di RStudio. d. Tabel tersebut dapat digunakan untuk analisis atau manipulasi data di RStudio.

6. Kesimpulan

- a. Dalam pengerjaan praktikum Statistika, ...
- b. Kita juga dapat mengetahui...

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	Menit	
2.	Latihan Kedua	Menit	
3.	Latihan Ketiga	Menit	
4.	Latihan Keempat	Menit	
5.	Latihan Kelima	Menit	
6.	Latihan Keenam	Menit	

Keterangan:

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