

<p>Nama: Hashemi Ralf Koizumi</p> <p>NIM: 06400220032</p>		<h2>MODUL 3</h2> <p>Nama Dosen: Dedy Sugiarto</p> <p>Nama Asisten Labratorium:</p> <ol style="list-style-type: none">Elen Fadilla Estri 064002000008Rukhy Zaifa Aduhalim 064002000041
<p>Hari/Tanggal: Hari, 17/07/ 2023</p>	<p>Praktikum Statistika</p>	

Pengelolaan Data pada Data Frame

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

a. Latihan pertama – Merge Data

1. Jalankan souce code berikut. Ganti nama variable (seperti data_nama) menjadi variable dengan nama kalian masing - masing. Data yang digunakan adalah data houseprice.csv. Melakukan Read CSV dengan cara menginput data houseprice, sesuaikan dengan lokasi direktori dimana kalian menyimpan file csvnya.

```
data_nama = read.delim("clipboard")
View(data_nama)
```

Output:



The screenshot shows the RStudio interface. In the top-left pane, there is a data grid titled 'Untitled1' containing 12 rows of data. The columns are labeled: Price, SqFt, Bedrooms, Bathrooms, Offers, Brick, and Neighborhood. The data includes various house prices and characteristics. In the bottom-left pane, the 'Console' tab is active, displaying R code:

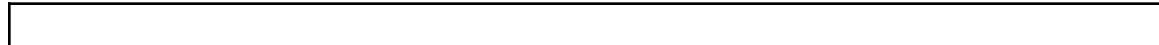
```
R 4.3.1 - /~>
> data_shemi = read.delim("clipboard")
> View(data_shemi)
> |
```

. The right side of the interface includes the 'Environment', 'History', 'Connections', and 'Tutorial' tabs, and a 'Data' pane showing the same data frame 'data_shemi' with 128 observations and 7 variables.

2. Lakukan subset data dengan cara sub set data frame khusus yang memiliki nilai variabel Bedrooms =2

```
nama <- data_nama[data_nama$Bedrooms==2,]
View(nama)
```

Output:

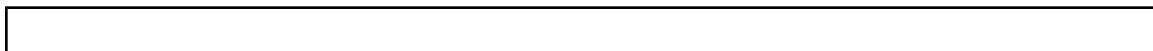


	Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood
1	114300	1790	2	2	2	No	East
12	123000	1870	2	2	2	Yes	East
19	111400	1700	2	2	1	Yes	East
29	69100	1600	2	2	3	No	North
32	112300	1930	2	2	2	Yes	North
35	117800	2000	2	2	3	No	North
37	117500	1880	2	2	2	No	North
41	106600	1560	2	2	1	No	East
43	105600	1990	2	2	3	No	East
47	129800	1990	2	3	2	No	North
49	115900	1980	2	2	2	No	East
52	91100	1860	2	2	3	No	North

3. Selanjutnya ialah mengubah nama variabel. Berikut adalah tahapan untuk mengubah nilai dalam kolom Bathrooms dalam kondisi jika Bathrooms >2 , maka akan diganti dengan nilai large, jika tidak maka small.

```
nama$Bathrooms <- ifelse(nama$Bathrooms > 2, c("large"), c("small"))
View(nama)
```

Output:



	Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood
1	114300	1790	2	small	2	No	East
12	123000	1870	2	small	2	Yes	East
19	111400	1700	2	small	1	Yes	East
29	69100	1600	2	small	3	No	North
32	112300	1930	2	small	2	Yes	North
35	117800	2000	2	small	3	No	North
37	117500	1880	2	small	2	No	North
41	106600	1560	2	small	1	No	East
43	105600	1990	2	small	3	No	East
47	129800	1990	2	large	2	No	North
49	115900	1980	2	small	2	No	East
52	91100	1860	2	small	3	No	North

4. Lalu langkah selanjutnya adalah menambah variabel. Tahap 1 □ Membuat variable baru dari dataku1 dengan nilai sesuai kondisi ifelse yang ditentukan. Sesuaikan nama variable baru dengan nama praktikan.

```
nama$newvariable <- ifelse(nama$Offers > 2, c("large"), c("small"))
View(nama)
```

Output



	Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood	newvariable
1	114300	1790	2	small	2	No	East	small
12	123000	1870	2	small	2	Yes	East	small
19	111400	1700	2	small	1	Yes	East	small
29	69100	1600	2	small	3	No	North	large
32	112300	1930	2	small	2	Yes	North	small
35	117800	2000	2	small	3	No	North	large
37	117500	1880	2	small	2	No	North	small
41	106600	1560	2	small	1	No	East	small
43	105600	1990	2	small	3	No	East	large
47	129800	1990	2	large	2	No	North	small
49	115900	1980	2	small	2	No	East	small
52	91100	1860	2	small	3	No	North	large

Showing 1 to 12 of 30 entries, 8 total columns

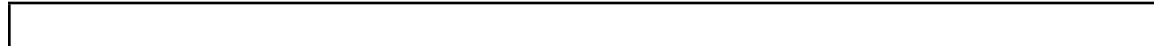
```
R 4.3.1 -- / 
> shemi$Bathrooms <- ifelse(shemi$Bathrooms > 2, c("large"), c("small"))
> View(shemi)
> shemi$newvariable <- ifelse(shemi$Offers > 2, c("large"), c("small"))
> View(shemi)
>
```

Files Plots Packages Help Viewer Presentation 13:48 17/07/2023

Tahap 2 □ Mengubah isi baris data dari kolom baru yang telah dibuat

```
nama$newvariable <- nama$Price/nama$SqFt
View(nama)
```

Output:

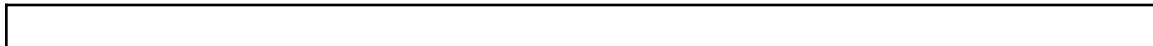


	Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood	newvariable
1	114300	1790	2	small	2	No	East	63.85475
12	123000	1870	2	small	2	Yes	East	65.77540
19	111400	1700	2	small	1	Yes	East	65.52941
29	69100	1600	2	small	3	No	North	43.18750
32	112300	1930	2	small	2	Yes	North	58.18653
35	117800	2000	2	small	3	No	North	58.90000
37	117500	1880	2	small	2	No	North	62.50000
41	106600	1560	2	small	1	No	East	68.33333
43	105600	1990	2	small	3	No	East	53.06533
47	129800	1990	2	large	2	No	North	65.22613
49	115900	1980	2	small	2	No	East	58.53535
52	91100	1860	2	small	3	No	North	48.97849

5. Delete Variabel. Selain bisa menambah, kita juga bisa menghapus variable. Dalam percobaan ini kita akan menghapus variable yang baru saja kita buat.

```
nama$newvariable<-NULL
```

Output:

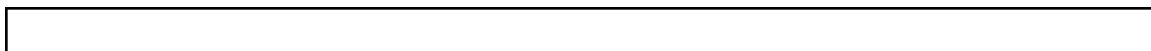


The screenshot shows the RStudio interface. In the top-left pane, there is a data grid titled 'shemi' with columns: Price, SqFt, Bedrooms, Bathrooms, Offers, Brick, and Neighborhood. The data consists of 128 observations. In the bottom-left pane, the 'Console' tab is active, displaying R code for creating new variables based on existing ones. The right-hand sidebar shows the 'Environment' tab with two objects: 'data_shemi' (128 obs. of 7 variables) and 'shemi' (30 obs. of 7 variables). The bottom right corner shows the system tray with the date and time.

6. Merge Data Frame. Tahap 1 □ Merge kolom Artinya memisahkan data sesuai kolom yang diinginkan lalu kemudian digabungkan.

```
data_nama.kolom1dan2 <- data_nama[,1:2]
View(data_nama.kolom1dan2)
```

Output:

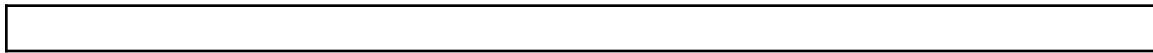


	Price	SqFt
1	114300	1790
2	114200	2030
3	114800	1740
4	94700	1980
5	119800	2130
6	114600	1780
7	151600	1830
8	150700	2160
9	119200	2110
10	104000	1730
11	132500	2030
12	123000	1870

Tahap 2 □ Merge kolom Artinya memisahkan data sesuai kolom yang diinginkan lalu kemudian digabungkan.

```
data_elen.kolom3dan4 <- data_elen[,3:4]
View(data_elen.kolom3dan4)
```

Output:



The screenshot shows the RStudio interface. In the top-left, there's a preview of a data frame with columns 'Bedrooms' and 'Bathrooms'. The top-right panel shows the 'Data' pane with several objects listed: 'data_shemi' (128 obs. of 7 variables), 'data_shem...' (128 obs. of 2 variables), 'data_shem...' (128 obs. of 2 variables), and 'shemi' (30 obs. of 7 variables). The bottom-left is the 'Console' tab, which contains R code for creating new variables ('newvariable') based on existing ones ('Price' and 'SqFT'). The bottom-right shows the system tray with the date and time (17/07/2023, 13:51).

```
R 4.3.1 / 
> View(shemi)
> shemi$newvariable <- shemi$Price/shemi$SqFT
> View(shemi)
> shemi$newvariable<-NULL
>
> data_shemi.kolomidan2 <- data_shemi[,1:2]
> View(data_shemi.kolomidan2)
> data_shemi.kolom3dan4 <- data_shemi[,3:4]
> View(data_shemi.kolom3dan4)
> |
```



Tahap 3 □ Merge kolom Artinya memisahkan data sesuai kolom yang diinginkan lalu kemudian digabungkan.

```
data_nama.kolom1sd4<-cbind(data_nama.kolom1dan2, data_nama.kolom3dan4)
View(data_nama.kolom1sd4)
```

Output:

The screenshot shows the RStudio interface. In the top-left, there's a data grid displaying a subset of the data. The columns are labeled 'Price', 'SqFt', 'Bedrooms', and 'Bathrooms'. The data shows 12 rows of house prices and features. Below the grid, the status bar indicates 'Showing 1 to 12 of 128 entries, 4 total columns'. In the bottom-left, the console tab is active, showing R code and its execution. The code includes loading data frames 'data_shemi' and 'shemi', creating a new variable 'newvariable', and performing a cbind operation to merge 'kolom1dan2' and 'kolom3dan4'. The status bar at the bottom right shows the date '17/07/2023' and time '13:53'.

7. Merge Baris artinya memisahkan data sesuai baris yang diinginkan dengan menggunakan range baris. Lalu kemudian digabungkan.

```
data_nama.baris1sd3 <- data_nama[1:3,]
data_nama.baris4sd6 <- data_nama[4:6,]
data_nama.baris1sd6 <- rbind(data_nama.baris1sd3, data_nama.baris4sd6)
View(data_nama.baris1sd6)
```

Output:

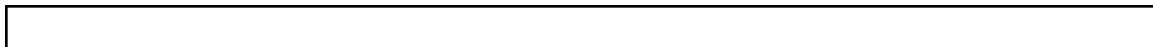
	Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood
1	114300	1790	2	2	2	No	East
2	114200	2030	4	2	3	No	East
3	114800	1740	3	2	1	No	East
4	94700	1980	3	2	3	No	East
5	119800	2130	3	3	3	No	East
6	114600	1780	3	2	2	No	North

8. Sort data frame. Apa yang terjadi dengan data setelah di sort?

Jawaban: ?

```
data_nama.sort<-data_nama[order(data_nama$Price),]  
View(data_nama.sort)
```

Output:



The screenshot shows the RStudio interface with the following components:

- Data Viewer:** Displays a table of 128 rows and 7 columns from the dataset `shemi`. The columns are: Price, SqFt, Bedrooms, Bathrooms, Offers, Brick, and Neighborhood. Rows are numbered 29 to 28.
- Environment Browser:** Shows the global environment with objects like `data_shemi.kolom1dan2`, `data_shemi.kolom3dan4`, `data_shemi.kolom1sd4`, `data_shemi.baris1sd4`, `data_shemi.baris4sd6`, `data_shemi.baris4sd6`, and `shemi`.
- Console:** Displays the R session history, including code for reading data frames and creating sorted versions of the dataset.
- Bottom Bar:** Includes icons for file operations, plots, packages, help, viewer, and presentation.

b. Latihan Kedua – Tugas

Gunakan dataset pada tugas 1 yang telah ditambah lagi datanya dengan 10 mhs TIF/SI. Sehingga total baris data pada file tersebut berjumlah 30. Ulangi kembali perintah-perintah di atas dan sesuaikan dengan data anda. Dan lampirkan Screen Capture untuk tiap poin yang ada.

1. Read CSV

```
data_shemi_csv = read.csv("C:/Users/ACER/OneDrive/Documents/Praktikum  
Statistika/amoqus.csv")
```

Output:



The screenshot shows the RStudio interface with the following details:

- Environment View:** Shows three objects:
 - `data_shemi`: 30 obs. of 6 variables
 - `data_shemi_csv`: 30 obs. of 1 variable
 - `shemi`: 5 obs. of 6 variables
- Data View:** Displays the contents of the `shemi` data frame, which includes 30 rows of data with columns: Nama, Gender, Angkatan, Tinggi, Badan, Waktu, Perjalanan, Wilayah, and Tinggi.
- Console View:** Shows the R session history:

```
R 4.3.1 -- > /Praktikum Statistika/prak3/ >
> data_shemi_csv = read.csv("C:/Users/ACER/OneDrive/Documents/Praktikum Statistika/amogus.csv")
> view(data_shemi_csv)
> |
```

2. Subset Data Gunakan Kolom Tinggi Badan untuk bagian ini

```
shemi <- data_shemi[data_shemi$Tinggi.Badan==165,]
```

Output:



The screenshot shows the RStudio interface. In the top-left pane, there is a data grid titled 'shemi' with columns: Nama, Gender, Angkatan, Tinggi.Badan, Waktu.Perjalanan, and Wilayah.Tinggal. The data includes entries for Evan, Grid, Messi, Ojan, and Riska. In the top-right pane, the 'Environment' tab shows two objects: 'data_shemi' (30 obs. of 6 variables) and 'shemi' (5 obs. of 6 variables). The bottom-left pane is the 'Console' tab, which contains the command: > shemi <- data_shemi[data_shemi\$Tinggi.Badan==165,]. The bottom-right pane shows the system tray with various icons and the date/time: 14:23 17/07/2023.

3. Mengubah nilai suatu variable □ Ubah isi kolom tinggi badan untuk tinggi > 160 menjadi “Tinggi” dan jika bukan berarti “Pendek”

```
> shemi$Tinggibadan <- ifelse(shemi$TinggiBadan > 160, c("tinggi"),
  c("pendek"))

> View(shemi)
```

Output:



4. Menambah Variabel

- Buat variabel kolom baru dengan nama “Jurusan” dengan isi baris datanya adalah “Infor20”
- Buat varibel kolom baru kedua dengan nama “Fakultas” dengan isi baris datanya adalah “FTI”

```
shemi$jurusan <- ("Infor20")
View(shemi)
shemi$fakultas <- ("FTI")
View(shemi)
```

Output:



The image contains two side-by-side screenshots of the RStudio interface. Both screenshots show a data frame named 'shemi' with 30 rows and 9 columns. The columns are: Nama, Gender, Angkatan, TinggiBadan, WaktuPerjalanan, WilayahTinggal, Tinggibadan, jurusan, and fakultas.

Screenshot 1 (Top):

- Data View:** Shows the first 10 rows of the 'shemi' data frame.
- Console View:** Displays the following R code and its execution errors. The code attempts to update the 'jurusan' column based on 'Angkatan' and 'WilayahTinggal'.

```
R 4.3.1 --> /Praktikum Statistika/prak3/
> shemi$newvariable <- ifelse(shemi$Jurusan == "Infor20" & shemi$WilayahTinggal == "Bandung", 1, 0)
Error: unexpected '=' in "shemi$newvariable <- ifelse(shemi$Jurusan == "
> shemi$newvariable <- ifelse(shemi$Jurusan > 2022, c("Infor20"))
Error in <-- data.frame (*tmp*, newvariable, value = logical(0)) :
  replacement has 0 rows, data has 30
> shemi$newvariable <- ifelse(shemi$Jurusan > 2022, c("Infor20"))
Error in <-- data.frame (*tmp*, newvariable, value = logical(0)) :
  replacement has 0 rows, data has 30
> shemi$newvariable <- ifelse(shemi$Jurusan > 2022, c("Infor20"))
Error in <-- data.frame (*tmp*, newvariable, value = logical(0)) :
  replacement has 0 rows, data has 30
> shemi$Jurusan <- ("Infor20")
> |
```

Screenshot 2 (Bottom):

- Data View:** Shows the first 10 rows of the 'shemi' data frame.
- Console View:** Displays the same R code as Screenshot 1, but it has been successfully executed. The 'jurusan' column now contains values 1 or 0 based on the conditions specified in the code.

5. Delete Variabel □ Hapus kolom Fakultas

```
> shemi$fakultas<-NULL
> View(shemi)
```



Output:

The screenshot shows the RStudio interface with two data frames loaded: 'data_shemi' and 'shemi'. The 'data_shemi' frame contains 30 rows of student data with columns: Nama, Gender, Angkatan, TinggiBadan, WaktuPerjalanan, WilayahTinggal, Tinggibadan, and jurusan. The 'shemi' frame contains 30 rows of course data with columns: Jurusan, Fakultas, and Gabungan. A console window shows an R script running, which includes code to merge the data frames. An error message is displayed, indicating that the 'shemi\$jurusan' column has 0 rows while the data frame has 30 rows, and it cannot find the 'namafakultas' variable.

6. Merge Data Frame

- Gabung kolom Nama dan Gender
- Gabung kolom Angkatan dan Tinggi Badan
- Gabungkan 2 variabel kolom gabungan diatas

```
data_shemi <- mutate(data_shemi, Nama_Gender = paste(Nama, Gender, sep = " - "))
shemi <- mutate(shemi, Angkatan_Tinggi = paste0(Angkatan,
Tinggi.Badan))
shemi <- mutate(shemi, Gabungan = paste(Nama_Gender, Angkatan_Tinggi,
sep = " - "))
```

Output:



	Nama	Gender	Angkatan	TinggiBadan	WaktuPerjalanan	WilayahTinggal	Tinggibadan	jurusan
1	Tommy	L	2022	168	35	Bandung	tinggi	Infor20
2	Jamal	L	2022	185	75	Bogor	tinggi	Infor20
3	Agi	L	2022	174	5	Bengkulu	tinggi	Infor20
4	Yono	L	2022	190	2	Bukit tinggi	tinggi	Infor20
5	Shemi	L	2022	171	53	Kebayoran lama	tinggi	Infor20
6	Evan	L	2022	165	60	Tangerang	tinggi	Infor20
7	Syakir	P	2022	162	65	Jelambar	tinggi	Infor20
8	Grid	P	2022	165	75	Banten	tinggi	Infor20
9	Imam	L	2022	168	140	Serang	tinggi	Infor20
10	Denny	L	2022	167	150	Curug	tinggi	Infor20

```
R 4.3.1 -- /Praktikum Statistika/prak3/ 
could not find function "mutate"
> data_shemi <- mutate(data_shemi, Nama_Gender = paste(Nama, Gender, sep = " - "))
Error in mutate(data_shemi, Nama_Gender = paste(Nama, Gender, sep = " - ")) :
  could not find function "mutate"
> View(shemi)
> View(data_shemi)
> View(shemi)
> shemi <- mutate(shemi, Angkatan_Tinggi = paste0(Angkatan, Tinggi.Badan))
Error in mutate(shemi, Angkatan_Tinggi = paste0(Angkatan, Tinggi.Badan)) :
  could not find function "mutate"
> shemi <- mutate(shemi, Gabungan = paste(Nama_Gender, Angkatan_Tinggi, sep = " - "))
Error in mutate(shemi, Gabungan = paste(Nama_Gender, Angkatan_Tinggi, :
  could not find function "mutate"
```

7. Merge Baris

- Buat gabungan baris 1-5
- Buat gabungan baris 25-30
- Gabungkan 2 variabel diatas

```
> shemi.baris1sd5 <- shemi[25:30,]
> View(shemi.baris1sd5)
> shemi.baris1sd5 <- shemi[1:5,]
> shemi.baris25sd30 <- shemi[25:30,]
> shemi.baris1sd30<-rbind(shemi.baris1sd5, shemi.baris25sd30)
```

Output:



8. Sort Data Frame Lakukan sort berdasarkan waktu perjalanan

```
> shemi.sort<-shemi[order(shemi$WaktuPerjalanan),]  
  
> View(shemi.sort)
```

Output:



4. File Praktikum

Github Repository:

<https://github.com/SRALFk/tugasstatistika.git>

5. Soal Latihan

Soal:

1. Apa saja kegunaan pengelolaan data pada data frame?
2. Dalam kasus apakah data perlu dihapus?

Jawaban:

1. Pengelolaan data pada data frame memiliki beberapa kegunaan penting, termasuk:
 1. Penyusunan dan pengaturan data: Data frame memungkinkan pengguna untuk mengatur data dalam format yang terstruktur, dengan baris dan kolom yang terorganisir. Ini memudahkan analisis dan pemrosesan data lebih lanjut.
 2. Manipulasi data: Data frame memungkinkan manipulasi data dengan berbagai cara. Pengguna dapat menambah, menghapus, atau mengubah nilai-nilai dalam data frame, serta melakukan operasi seperti pemfilteran, pengurutan, dan agregasi data.
 3. Pemrosesan statistik: Data frame menyediakan fungsionalitas statistik yang kaya. Pengguna dapat menghitung statistik ringkas, seperti rata-rata, median, atau deviasi standar



dari kolom-kolom data tertentu. Selain itu, data frame juga mendukung analisis regresi, pengujian hipotesis, dan visualisasi data.

4. Pemilihan dan penyaringan data: Data frame memungkinkan pengguna untuk memilih subset data berdasarkan kriteria tertentu. Misalnya, pengguna dapat memilih baris yang memenuhi kondisi tertentu atau memilih kolom-kolom tertentu untuk dianalisis lebih lanjut.

5. Integrasi dengan alat analisis data lainnya: Data frame sering digunakan dalam lingkungan pemrograman dan analisis data seperti Python dan R. Mereka memudahkan integrasi dengan alat dan paket analisis data lainnya, seperti Pandas, NumPy, atau tidyverse.

2. Ada beberapa kasus di mana penghapusan data diperlukan, termasuk:

1. Data yang tidak relevan: Jika data tidak relevan lagi untuk analisis atau keperluan saat ini, dapat dihapus untuk mengurangi kompleksitas dan ukuran dataset.

2. Data yang rusak atau tidak akurat: Jika terdapat data yang rusak, tidak akurat, atau tidak lengkap, dapat menjadi perlu untuk menghapusnya agar tidak mempengaruhi hasil analisis yang dapat menyesatkan.

3. Kebutuhan privasi dan kepatuhan: Ada situasi di mana data harus dihapus untuk mematuhi regulasi privasi atau kebijakan perusahaan. Misalnya, jika data pribadi atau informasi sensitif yang disimpan dalam dataset, harus dihapus sesuai dengan persyaratan hukum atau kebijakan privasi.

4. Pengelolaan kapasitas: Jika dataset menjadi terlalu besar untuk ditangani secara efisien atau melebihi kapasitas penyimpanan yang tersedia, maka beberapa data mungkin harus dihapus atau dipangkas agar dapat mengelola ukuran dataset.

5. Data yang telah usang: Dalam beberapa kasus, data yang tidak lagi relevan atau memiliki tanggal kedaluwarsa tertentu, seperti data historis yang telah melampaui periode kepentingannya, dapat dihapus untuk menjaga kebersihan dan keberlanjutan dataset.

6. Kesimpulan

a. Dalam pelajaran praktikum Statistika, kita membahas tentang cara bagaimana mengelola data berbentuk excel melalui Rstudio, seperti bagaimana menemukan sekumpulan data yang kita inginkan, mengubah kriteria data, memisah data menjadi beberapa kolom, menambahkan table dalam data dan menghapusnya



- b. Kita juga dapat mengetahui bagaimana cara untuk menggabungkan 2 buah file yang berisi beberapa kolom dan variable lalu memisahkan kolom dan baris sesuai yang dibutuhkan, kita juga dapat mengurutkan data tersebut menjadi apa yang dibutuhkan.

7. Cek List (✓)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	✓	
2.	Latihan Kedua	✓	

8. Formulir Umpam Balik

No	Elemen Kompetensi	Waktu Penggerjaan	Kriteria
1.	Latihan Pertama	30 Menit	2
2.	Latihan Kedua	65 Menit	1

Keterangan:

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

