


<p><b>Nama:</b> <b>Vania Rahma Dewi</b></p> <p><b>NIM:</b> <b>064002200030</b></p>	 <p><b>Praktikum Statistika</b></p>	<p><b>MODUL 3</b></p> <p><b>Nama Dosen:</b> <b>Dedy Sugiarto</b></p>
<p><b>Hari/Tanggal:</b> <b>Hari, 17 Juli 2023</b></p>		<p><b>Nama Asisten Labratorium:</b></p> <ol style="list-style-type: none"> <li><b>Elen Fadilla Estri</b> <b>064002000008</b></li> <li><b>Rukhy Zaifa Aduhalim</b> <b>064002000041</b></li> </ol>

## 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 penggunaannya 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

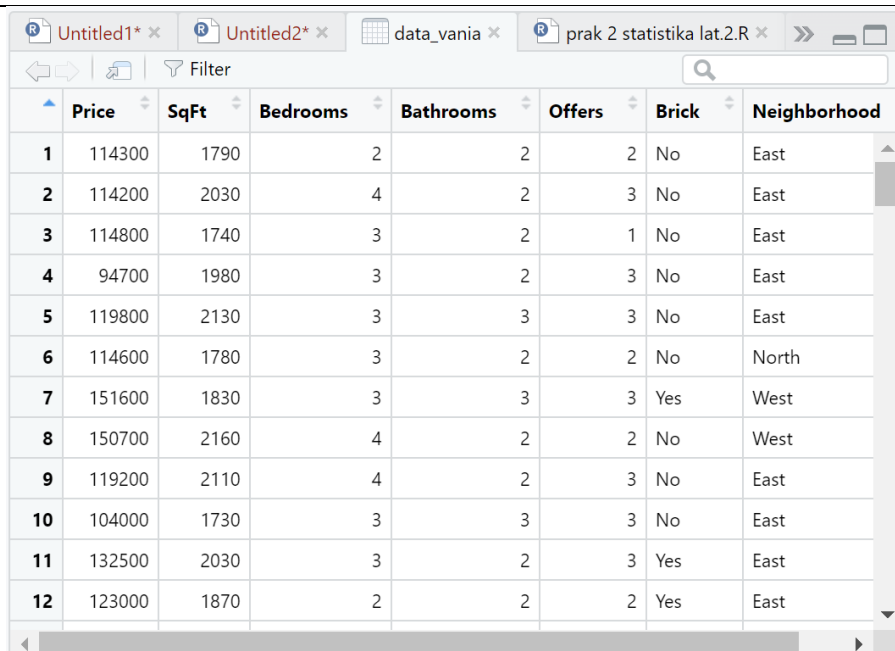
### a. Latihan pertama – Merge Data

1. Jalankan source 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:





	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
7	151600	1830	3	3	3	Yes	West
8	150700	2160	4	2	2	No	West
9	119200	2110	4	2	3	No	East
10	104000	1730	3	3	3	No	East
11	132500	2030	3	2	3	Yes	East
12	123000	1870	2	2	2	Yes	East

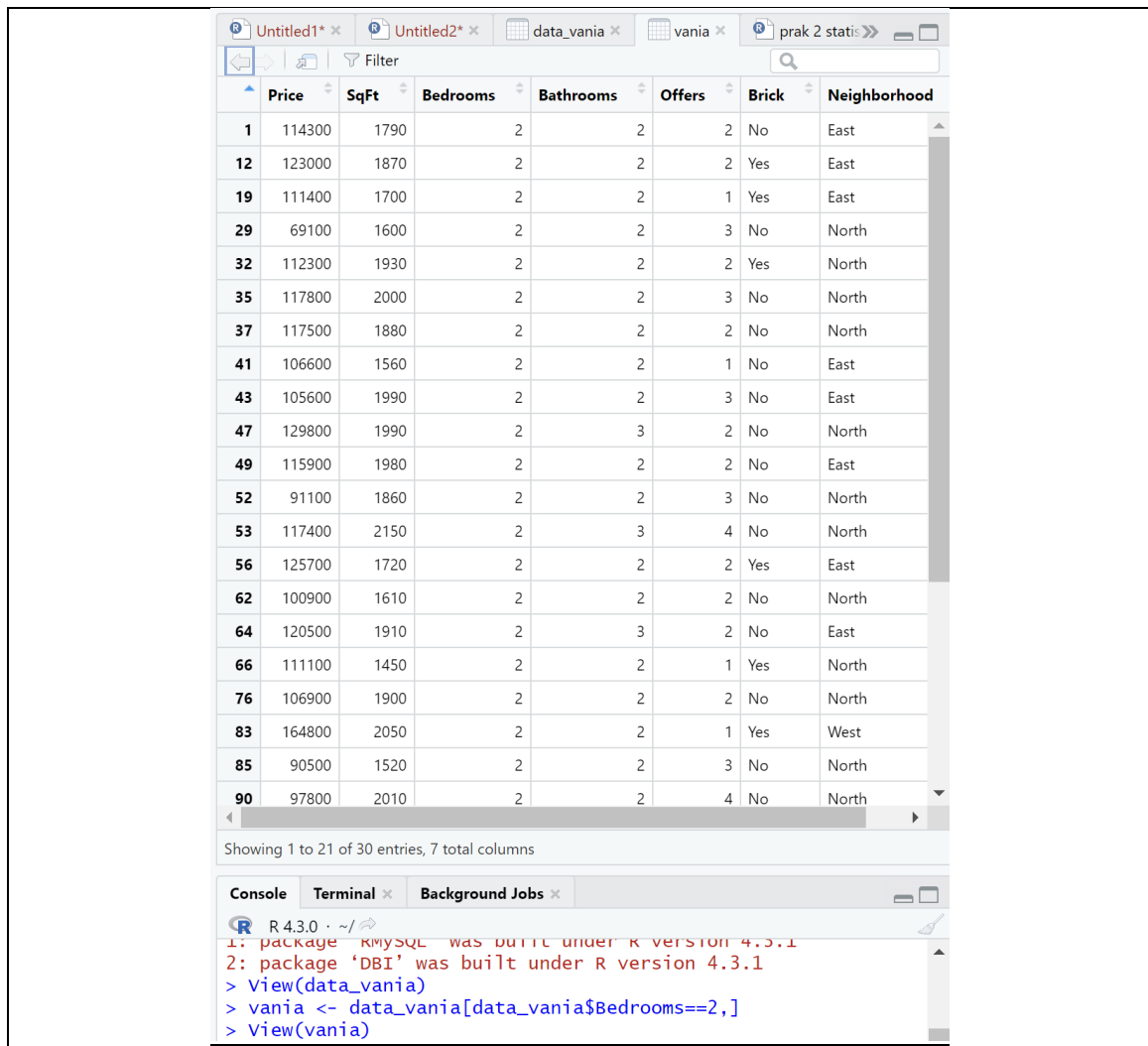
```
> data_vania = read.delim("clipboard")
Warning messages:
1: package 'RMySQL' was built under R version 4.3.1
2: package 'DBI' was built under R version 4.3.1
> View(data_vania)
```

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:





The screenshot shows an RStudio window with a data table and a console window. The data table has 7 columns: Price, SqFt, Bedrooms, Bathrooms, Offers, Brick, and Neighborhood. The console window shows the following R code and output:

```
R 4.3.0 ~ /
1: package 'rmysql' was built under R version 4.3.1
2: package 'DBI' was built under R version 4.3.1
> View(data_vania)
> vania <- data_vania[data_vania$Bedrooms==2,]
> View(vania)
```

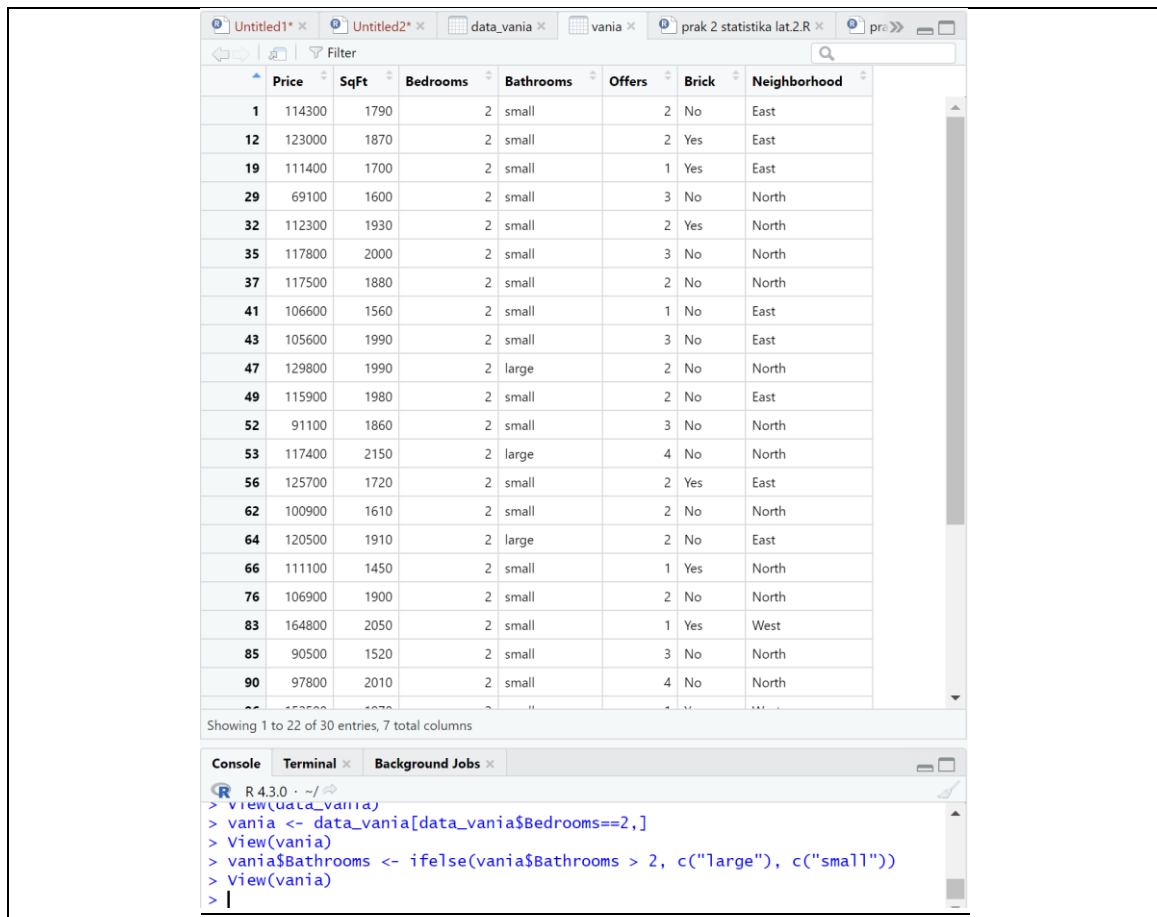
	Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood
1	114300	1790	2	2	2	No	East
12	123000	1870	2	2	2	Yes	East
19	111400	1700	2	2	1	Yes	East
29	69100	1600	2	2	3	No	North
32	112300	1930	2	2	2	Yes	North
35	117800	2000	2	2	3	No	North
37	117500	1880	2	2	2	No	North
41	106600	1560	2	2	1	No	East
43	105600	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
53	117400	2150	2	3	4	No	North
56	125700	1720	2	2	2	Yes	East
62	100900	1610	2	2	2	No	North
64	120500	1910	2	3	2	No	East
66	111100	1450	2	2	1	Yes	North
76	106900	1900	2	2	2	No	North
83	164800	2050	2	2	1	Yes	West
85	90500	1520	2	2	3	No	North
90	97800	2010	2	2	4	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
53	117400	2150	2	large	4	No	North
56	125700	1720	2	small	2	Yes	East
62	100900	1610	2	small	2	No	North
64	120500	1910	2	large	2	No	East
66	111100	1450	2	small	1	Yes	North
76	106900	1900	2	small	2	No	North
83	164800	2050	2	small	1	Yes	West
85	90500	1520	2	small	3	No	North
90	97800	2010	2	small	4	No	North

```

R 4.3.0 ~ /
> view(data_vania)
> vania <- data_vania[data_vania$Bedrooms==2,]
> View(vania)
> vania$Bathrooms <- ifelse(vania$Bathrooms > 2, c("large"), c("small"))
> View(vania)
> |

```

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
53	117400	2150	2	large	4	No	North	large
56	125700	1720	2	small	2	Yes	East	small
62	100900	1610	2	small	2	No	North	small
64	120500	1910	2	large	2	No	East	small
66	111100	1450	2	small	1	Yes	North	small
76	106900	1900	2	small	2	No	North	small
83	164800	2050	2	small	1	Yes	West	small
85	90500	1520	2	small	3	No	North	large
90	97800	2010	2	small	4	No	North	large

```

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

```

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

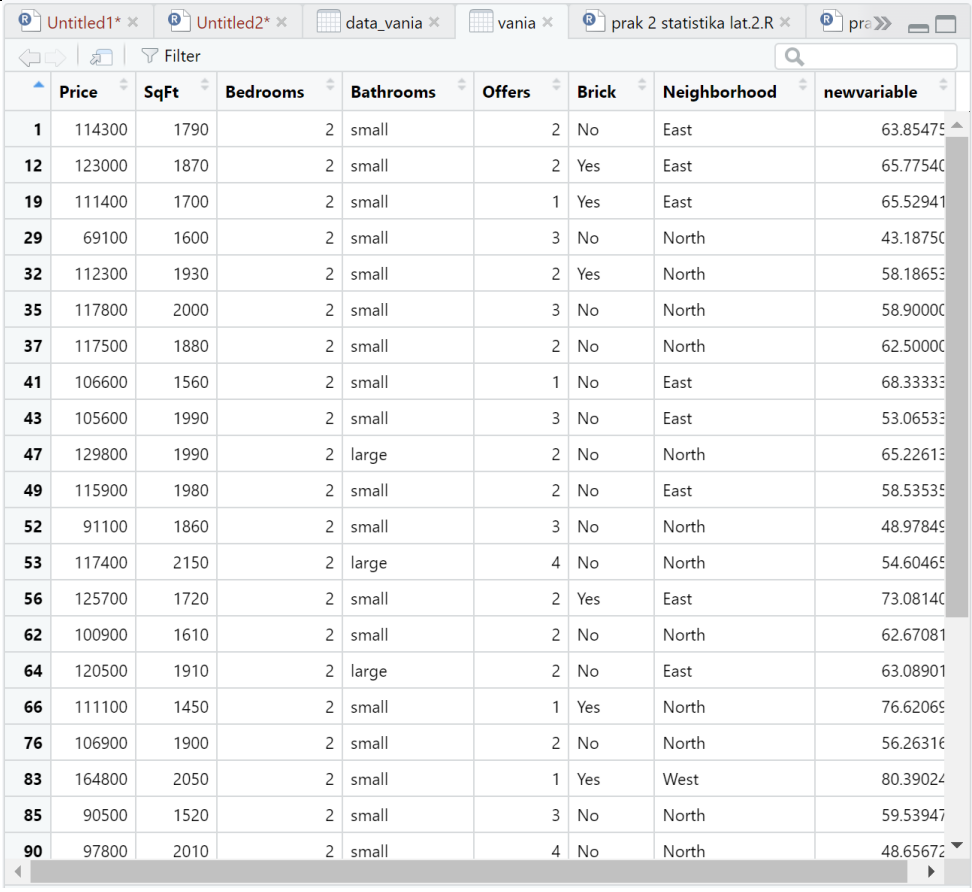
```

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
53	117400	2150	2	large	4	No	North	54.60465
56	125700	1720	2	small	2	Yes	East	73.08140
62	100900	1610	2	small	2	No	North	62.67081
64	120500	1910	2	large	2	No	East	63.08901
66	111100	1450	2	small	1	Yes	North	76.62069
76	106900	1900	2	small	2	No	North	56.26316
83	164800	2050	2	small	1	Yes	West	80.39024
85	90500	1520	2	small	3	No	North	59.53947
90	97800	2010	2	small	4	No	North	48.65672

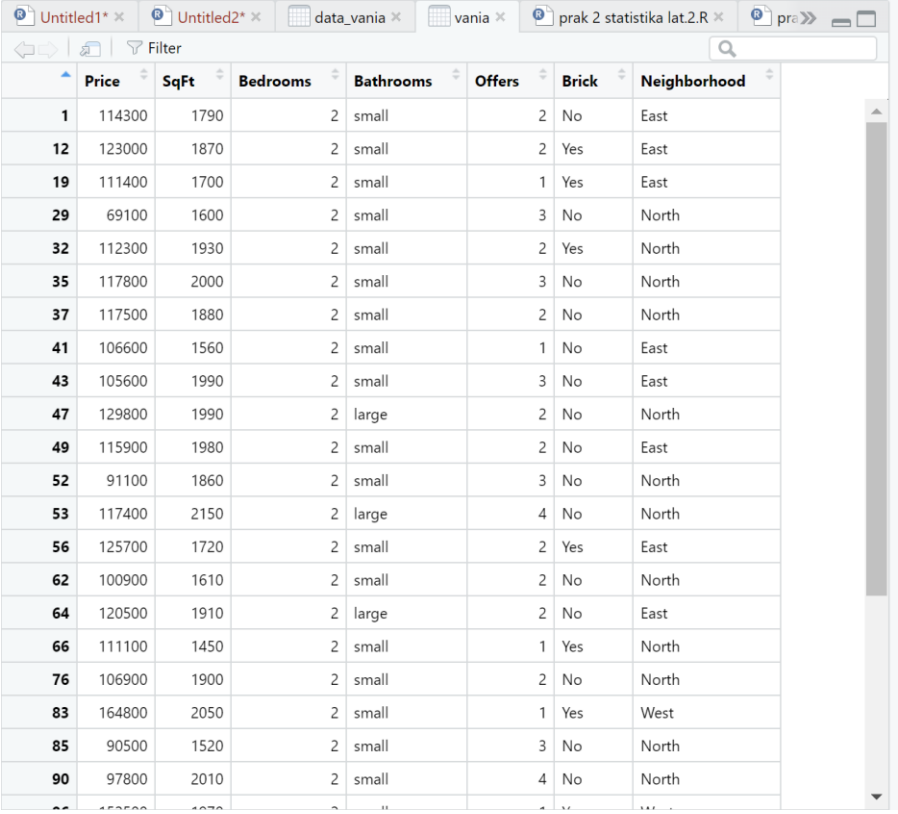
```
> vania$newvariable <- vania$Price / vania$SqFt
> View(vania)
```

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:





	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
53	117400	2150	2	large	4	No	North
56	125700	1720	2	small	2	Yes	East
62	100900	1610	2	small	2	No	North
64	120500	1910	2	large	2	No	East
66	111100	1450	2	small	1	Yes	North
76	106900	1900	2	small	2	No	North
83	164800	2050	2	small	1	Yes	West
85	90500	1520	2	small	3	No	North
90	97800	2010	2	small	4	No	North

```
> vania$newvariable <- NULL
> View(vania)
```

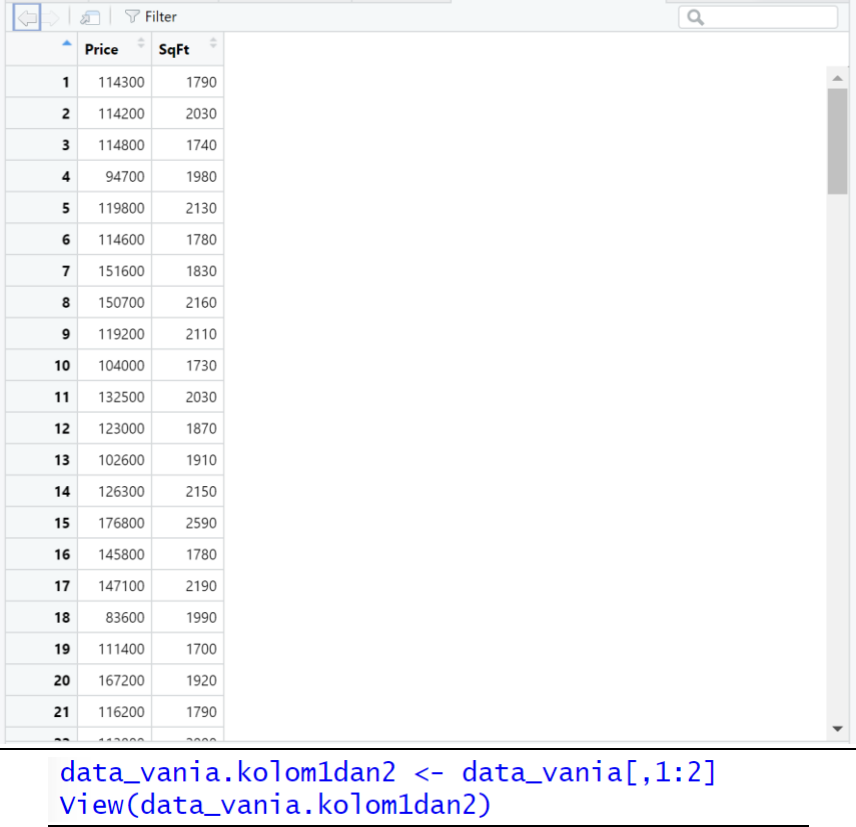
6. Merge Data Frame. Tahap 1 □ Merge kolom Artinya memisahkan data sesuai kolom yang diinginkan lalu kemudian digabungkan.

```
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
13	102600	1910
14	126300	2150
15	176800	2590
16	145800	1780
17	147100	2190
18	83600	1990
19	111400	1700
20	167200	1920
21	116200	1790

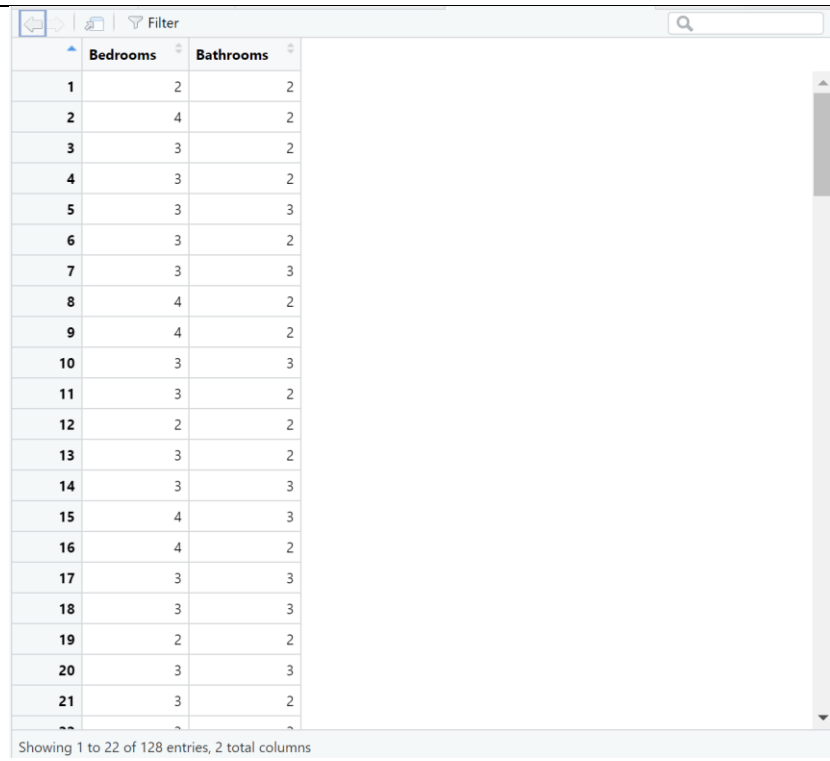
```
data_vania.kolom1dan2 <- data_vania[,1:2]
view(data_vania.kolom1dan2)
```

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:





	Bedrooms	Bathrooms
1	2	2
2	4	2
3	3	2
4	3	2
5	3	3
6	3	2
7	3	3
8	4	2
9	4	2
10	3	3
11	3	2
12	2	2
13	3	2
14	3	3
15	4	3
16	4	2
17	3	3
18	3	3
19	2	2
20	3	3
21	3	2

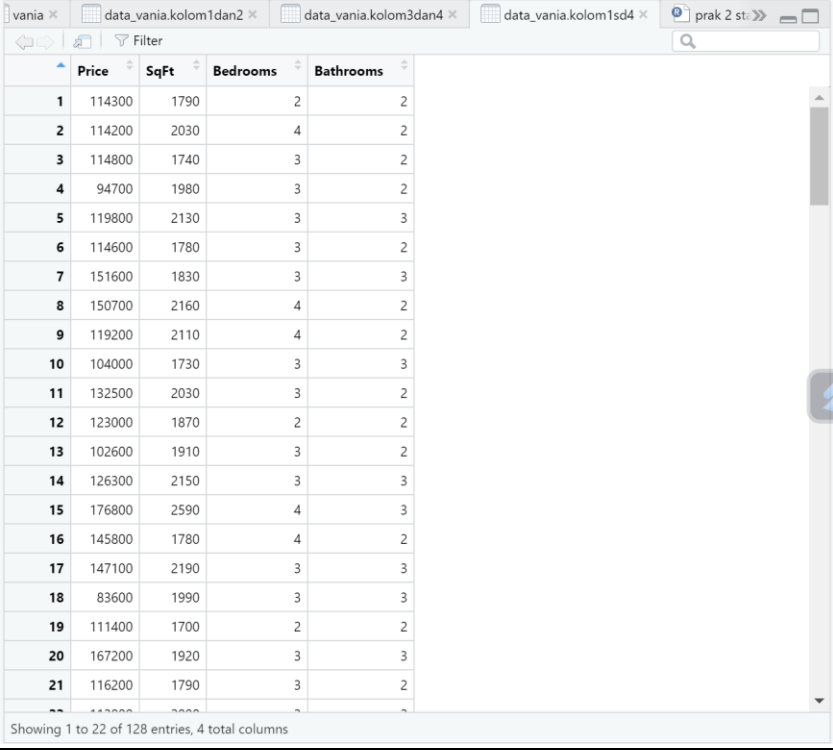
```
> data_vania.kolom3dan4 <- data_vania[,3:4]  
> View(data_vania.kolom3dan4)
```



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:



	Price	SqFt	Bedrooms	Bathrooms
1	114300	1790	2	2
2	114200	2030	4	2
3	114800	1740	3	2
4	94700	1980	3	2
5	119800	2130	3	3
6	114600	1780	3	2
7	151600	1830	3	3
8	150700	2160	4	2
9	119200	2110	4	2
10	104000	1730	3	3
11	132500	2030	3	2
12	123000	1870	2	2
13	102600	1910	3	2
14	126300	2150	3	3
15	176800	2590	4	3
16	145800	1780	4	2
17	147100	2190	3	3
18	83600	1990	3	3
19	111400	1700	2	2
20	167200	1920	3	3
21	116200	1790	3	2

```
> data_vania.kolom1sd4 <- cbind(data_vania.kolom1dan2, data_vania.kolom3dan4)
> view(data_vania.kolom1sd4)
```

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

```

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

```

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:



	Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood
29	69100	1600	2	2	3	No	North
55	81300	1650	3	2	3	No	North
105	82300	1910	3	2	4	No	East
18	83600	1990	3	3	4	No	North
48	90300	2050	3	2	6	No	North
85	90500	1520	2	2	3	No	North
52	91100	1860	2	2	3	No	North
23	91700	1690	3	2	3	No	North
69	93600	2140	3	2	4	No	North
4	94700	1980	3	2	3	No	East
90	97800	2010	2	2	4	No	North
28	99300	1700	3	2	2	No	East
62	100900	1610	2	2	2	No	North
116	102500	1900	3	3	3	No	North
13	102600	1910	3	2	4	No	North
87	102700	1900	4	2	4	No	North

```

> data_vania.sort<-data_vania[order(data_vania$Price),]
> view
Error: object 'view' not found
> View(data_vania.sort)

```

b. Latihan Kedua – Tugas

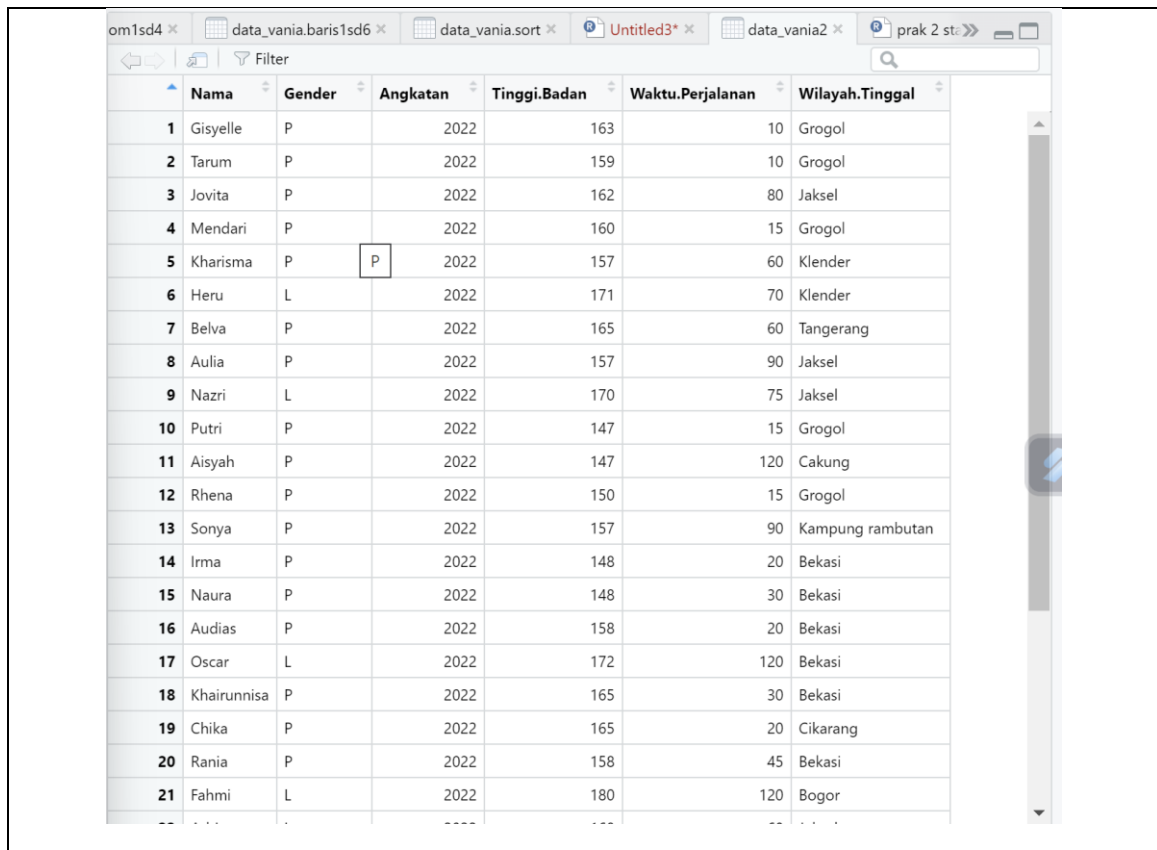
Gunakan dataset pada tugas 1 yang telah ditambah lagi datanya dengan 10 mhs TIF/SI. Sehingga total baris data pada file tersebut berjumlah 30. Ulangi kembali perintah-perintah di atas dan sesuaikan dengan data anda. Dan lampirkan Screen Capture untuk tiap poin yang ada.

1. Read CSV

```
data_vania2 = read.delim("clipboard")
View(data_vania2)
```

Output:



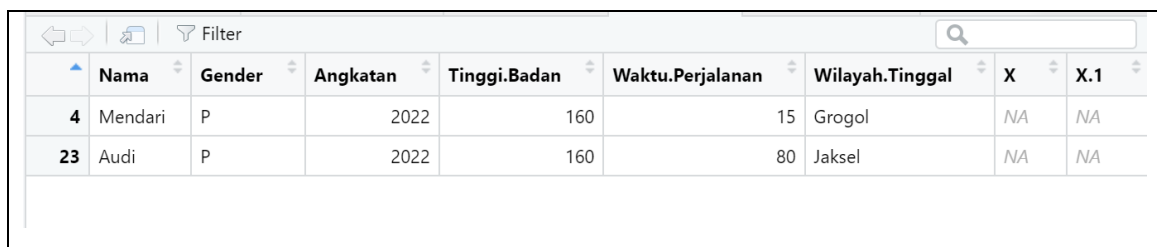


	Nama	Gender	Angkatan	Tinggi.Badan	Waktu.Perjalanan	Wilayah.Tinggal
1	Gisyelle	P	2022	163	10	Grogol
2	Tarum	P	2022	159	10	Grogol
3	Jovita	P	2022	162	80	Jaksel
4	Mendari	P	2022	160	15	Grogol
5	Kharisma	P	P	157	60	Klender
6	Heru	L	2022	171	70	Klender
7	Belva	P	2022	165	60	Tangerang
8	Aulia	P	2022	157	90	Jaksel
9	Nazri	L	2022	170	75	Jaksel
10	Putri	P	2022	147	15	Grogol
11	Aisyah	P	2022	147	120	Cakung
12	Rhena	P	2022	150	15	Grogol
13	Sonya	P	2022	157	90	Kampung rambutan
14	Irma	P	2022	148	20	Bekasi
15	Naura	P	2022	148	30	Bekasi
16	Audias	P	2022	158	20	Bekasi
17	Oscar	L	2022	172	120	Bekasi
18	Khairunnisa	P	2022	165	30	Bekasi
19	Chika	P	2022	165	20	Cikarang
20	Rania	P	2022	158	45	Bekasi
21	Fahmi	L	2022	180	120	Bogor

2. Subset Data □ Gunakan Kolom Tinggi Badan untuk bagian ini

```
vania2 <- data_vania2[data_vania2$Tinggi.Badan==160,]
View(vania2)
```

Output:



	Nama	Gender	Angkatan	Tinggi.Badan	Waktu.Perjalanan	Wilayah.Tinggal	X	X.1
4	Mendari	P	2022	160	15	Grogol	NA	NA
23	Audi	P	2022	160	80	Jaksel	NA	NA

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



```
data_vania2$Tinggi.Badan <- ifelse(data_vania2$Tinggi.Badan > 160, c("tinggi"),
c("pendek"))
View(data_vania2)
```

Output:

	Nama	Gender	Angkatan	Tinggi.Badan	Waktu.Perjalanan	Wilayah.Tinggal
1	Gisyelle	P	2022	tinggi	10	Grogol
2	Tarum	P	2022	pendek	10	Grogol
3	Jovita	P	2022	tinggi	80	Jaksel
4	Mendari	P	2022	pendek	15	Grogol
5	Kharisma	P	2022	pendek	60	Klender
6	Heru	L	2022	tinggi	70	Klender
7	Belva	P	2022	tinggi	60	Tangerang
8	Aulia	P	2022	pendek	90	Jaksel
9	Nazri	L	2022	tinggi	75	Jaksel
10	Putri	P	2022	pendek	15	Grogol
11	Aisyah	P	2022	pendek	120	Cakung
12	Rhena	P	2022	pendek	15	Grogol
13	Sonya	P	2022	pendek	90	Kampung rambutan
14	Irma	P	2022	pendek	20	Bekasi
15	Naura	P	2022	pendek	30	Bekasi
16	Audias	P	2022	pendek	20	Bekasi
17	Oscar	L	2022	tinggi	120	Bekasi
18	Khairunnisa	P	2022	tinggi	30	Bekasi
19	Chika	P	2022	tinggi	20	Cikarang
20	Rania	P	2022	pendek	45	Bekasi
21	Fahmi	L	2022	tinggi	120	Bogor

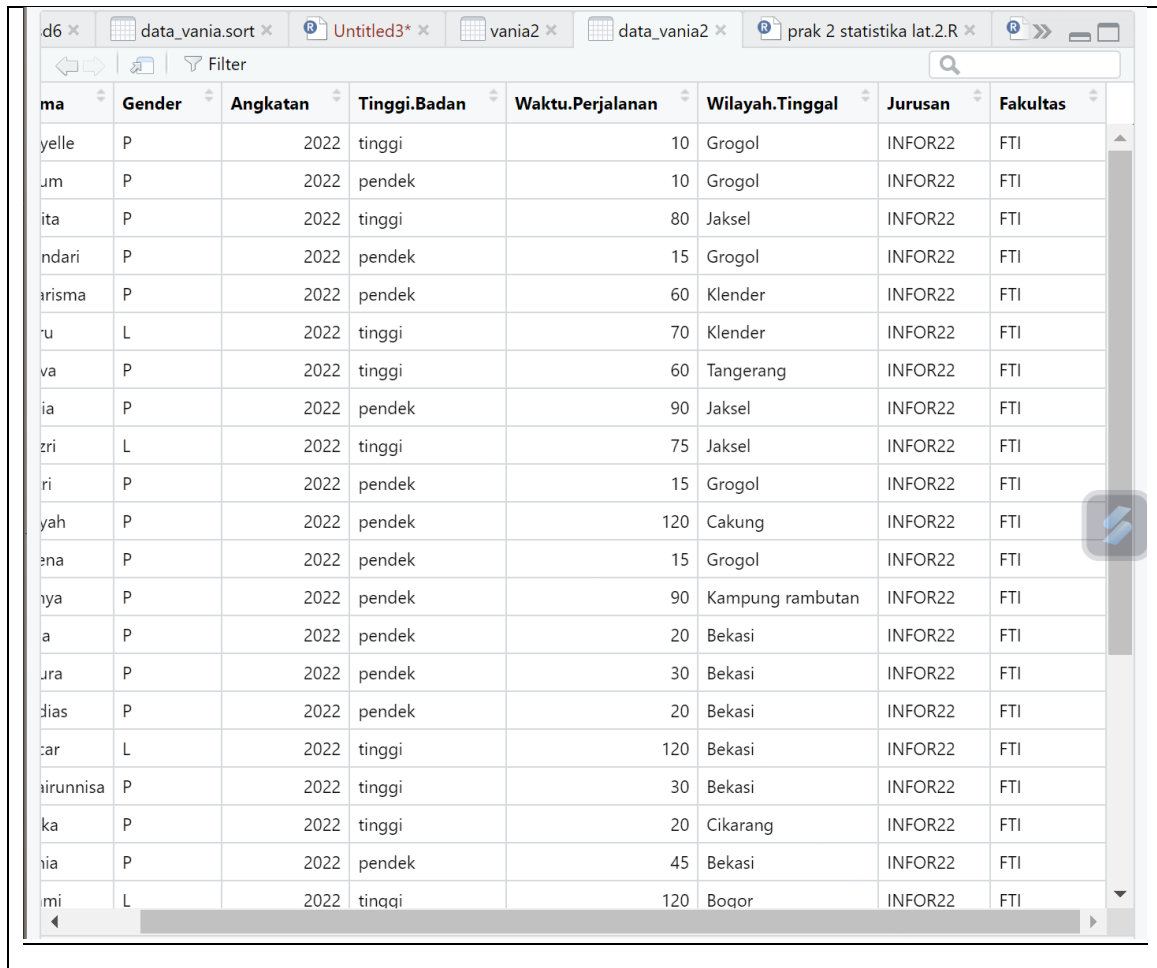
#### 4. Menambah Variabel

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

```
data_vania2$Jurusan <- ("INFOR22")
View(data_vania2)
data_vania2$Fakultas <- ("FTI")
View(data_vania2)
```



Output:



ma	Gender	Angkatan	Tinggi.Badan	Waktu.Perjalanan	Wilayah.Tinggal	Jurusan	Fakultas
yelle	P	2022	tinggi	10	Grogol	INFOR22	FTI
um	P	2022	pendek	10	Grogol	INFOR22	FTI
ita	P	2022	tinggi	80	Jaksel	INFOR22	FTI
ndari	P	2022	pendek	15	Grogol	INFOR22	FTI
arisma	P	2022	pendek	60	Klender	INFOR22	FTI
u	L	2022	tinggi	70	Klender	INFOR22	FTI
va	P	2022	tinggi	60	Tangerang	INFOR22	FTI
ia	P	2022	pendek	90	Jaksel	INFOR22	FTI
zri	L	2022	tinggi	75	Jaksel	INFOR22	FTI
ri	P	2022	pendek	15	Grogol	INFOR22	FTI
yah	P	2022	pendek	120	Cakung	INFOR22	FTI
ena	P	2022	pendek	15	Grogol	INFOR22	FTI
ya	P	2022	pendek	90	Kampung rambutan	INFOR22	FTI
a	P	2022	pendek	20	Bekasi	INFOR22	FTI
ura	P	2022	pendek	30	Bekasi	INFOR22	FTI
dias	P	2022	pendek	20	Bekasi	INFOR22	FTI
car	L	2022	tinggi	120	Bekasi	INFOR22	FTI
airunnisa	P	2022	tinggi	30	Bekasi	INFOR22	FTI
ka	P	2022	tinggi	20	Cikarang	INFOR22	FTI
ia	P	2022	pendek	45	Bekasi	INFOR22	FTI
mi	L	2022	tinggi	120	Boqor	INFOR22	FTI

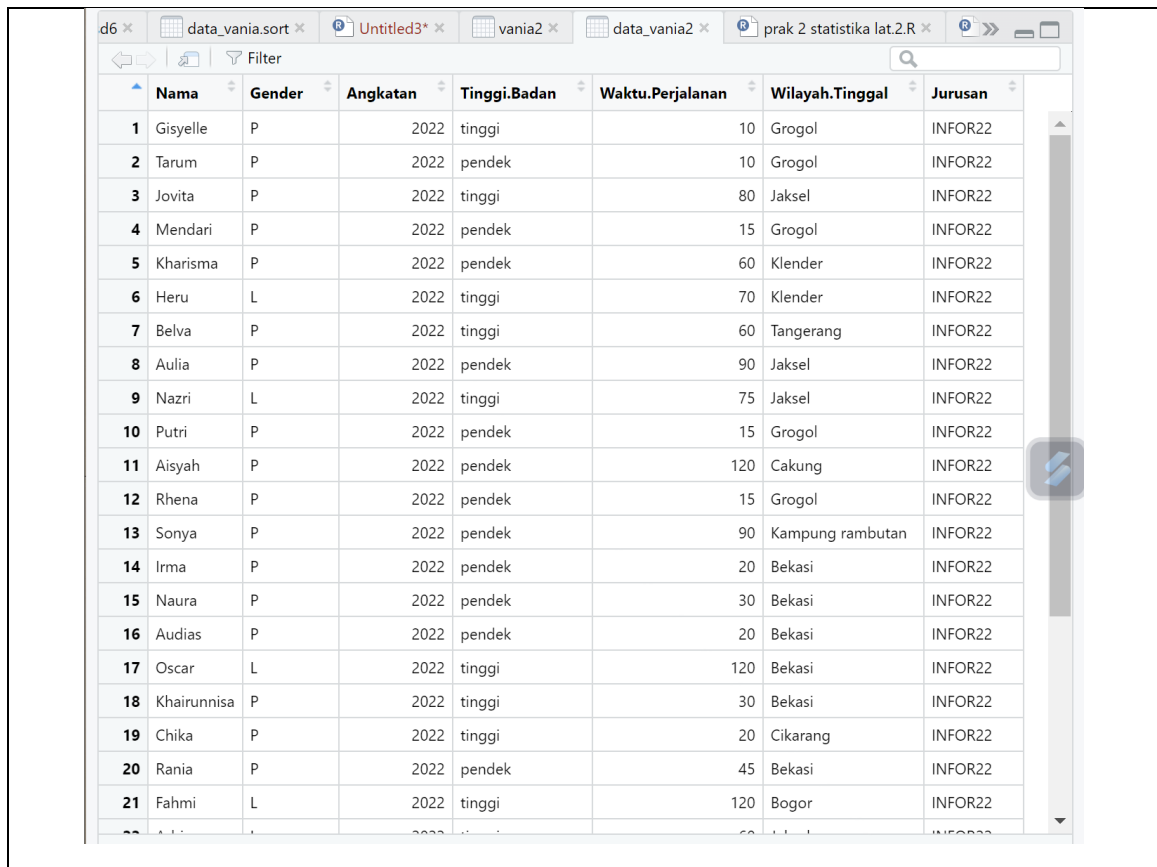
5. Delete Variabel □ Hapus kolom Fakultas

```
data_vania2$Fakultas <- NULL
```

Output:







	Nama	Gender	Angkatan	Tinggi.Badan	Waktu.Perjalanan	Wilayah.Tinggal	Jurusan
1	Gisyelle	P	2022	tinggi	10	Grogol	INFOR22
2	Tarum	P	2022	pendek	10	Grogol	INFOR22
3	Jovita	P	2022	tinggi	80	Jaksel	INFOR22
4	Mendari	P	2022	pendek	15	Grogol	INFOR22
5	Kharisma	P	2022	pendek	60	Klender	INFOR22
6	Heru	L	2022	tinggi	70	Klender	INFOR22
7	Belva	P	2022	tinggi	60	Tangerang	INFOR22
8	Aulia	P	2022	pendek	90	Jaksel	INFOR22
9	Nazri	L	2022	tinggi	75	Jaksel	INFOR22
10	Putri	P	2022	pendek	15	Grogol	INFOR22
11	Aisyah	P	2022	pendek	120	Cakung	INFOR22
12	Rhena	P	2022	pendek	15	Grogol	INFOR22
13	Sonya	P	2022	pendek	90	Kampung rambutan	INFOR22
14	Irma	P	2022	pendek	20	Bekasi	INFOR22
15	Naura	P	2022	pendek	30	Bekasi	INFOR22
16	Audias	P	2022	pendek	20	Bekasi	INFOR22
17	Oscar	L	2022	tinggi	120	Bekasi	INFOR22
18	Khairunnisa	P	2022	tinggi	30	Bekasi	INFOR22
19	Chika	P	2022	tinggi	20	Cikarang	INFOR22
20	Rania	P	2022	pendek	45	Bekasi	INFOR22
21	Fahmi	L	2022	tinggi	120	Bogor	INFOR22

## 6. Merge Data Frame

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

```
data_vania2.kolom1dan2 <- data_vania2[,1:2]
View(data_vania2.kolom1dan2)
data_vania2.kolom3dan4 <- data_vania2[,3:4]
View(data_vania2.kolom3dan4)
data_vania2.kolom1sd4 <- cbind(data_vania2.kolom1dan2, data_vania2.kolom3dan4)
View(data_vania2.kolom1sd4)
```

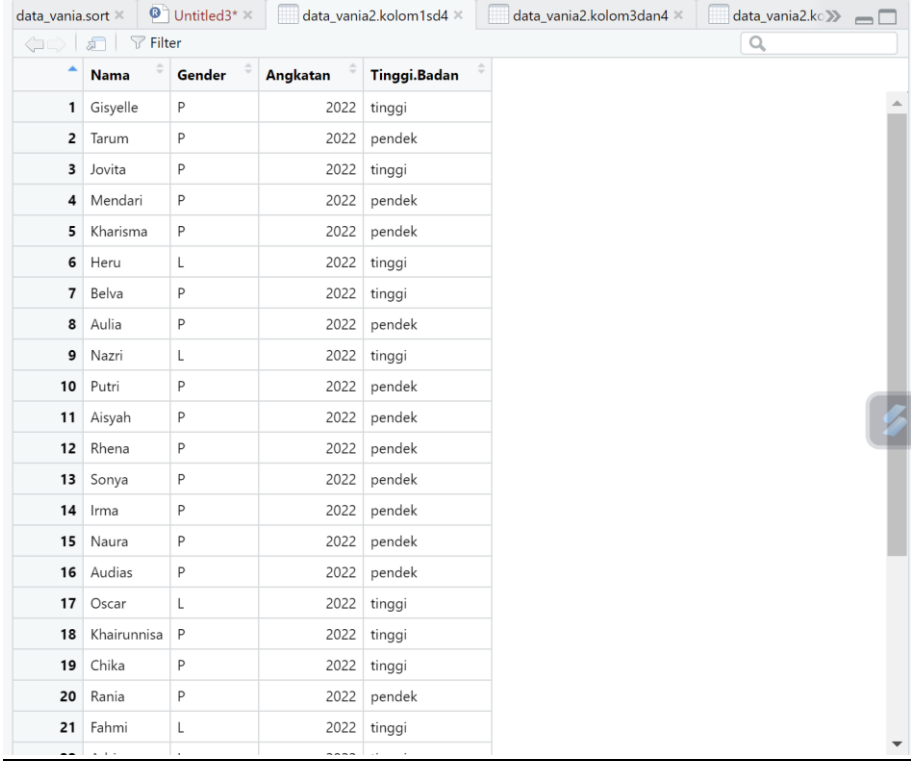
Output:



	Nama	Gender
1	Gisyelle	P
2	Tarum	P
3	Jovita	P
4	Mendari	P
5	Kharisma	P
6	Heru	L
7	Belva	P
8	Aulia	P
9	Nazri	L
10	Putri	P
11	Aisyah	P
12	Rhena	P
13	Sonya	P
14	Irma	P
15	Naura	P
16	Audias	P
17	Oscar	L
18	Khairunnisa	P
19	Chika	P
20	Rania	P
21	Fahmi	L

	Angkatan	Tinggi.Badan
1	2022	tinggi
2	2022	pendek
3	2022	tinggi
4	2022	pendek
5	2022	pendek
6	2022	tinggi
7	2022	tinggi
8	2022	pendek
9	2022	tinggi
10	2022	pendek
11	2022	pendek
12	2022	pendek
13	2022	pendek
14	2022	pendek
15	2022	pendek
16	2022	pendek
17	2022	tinggi
18	2022	tinggi
19	2022	tinggi
20	2022	pendek
21	2022	tinggi





	Nama	Gender	Angkatan	Tinggi.Badan
1	Gisyelle	P	2022	tinggi
2	Tarum	P	2022	pendek
3	Jovita	P	2022	tinggi
4	Mendari	P	2022	pendek
5	Kharisma	P	2022	pendek
6	Heru	L	2022	tinggi
7	Belva	P	2022	tinggi
8	Aulia	P	2022	pendek
9	Nazri	L	2022	tinggi
10	Putri	P	2022	pendek
11	Aisyah	P	2022	pendek
12	Rhena	P	2022	pendek
13	Sonya	P	2022	pendek
14	Irma	P	2022	pendek
15	Naura	P	2022	pendek
16	Audias	P	2022	pendek
17	Oscar	L	2022	tinggi
18	Khairunnisa	P	2022	tinggi
19	Chika	P	2022	tinggi
20	Rania	P	2022	pendek
21	Fahmi	L	2022	tinggi

## 7. Merge Baris

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

```
data_vania2.baris1sd5 <- data_vania2.kolom1sd4[1:5,]
data_vania2.baris25sd30 <- data_vania2.kolom1sd4[25:30,]
data_vania2.baris1sd30 <- rbind(data_vania2.baris1sd5, data_vania2.baris25sd30)
View(data_vania2.baris1sd30)
```

Output:



	Nama	Gender	Angkatan	Tinggi.Badan
1	Gisyelle	P	2022	tinggi
2	Tarum	P	2022	pendek
3	Jovita	P	2022	tinggi
4	Mendari	P	2022	pendek
5	Kharisma	P	2022	pendek
25	Kayla	P	2022	tinggi
26	Azalea	P	2022	pendek
27	Mona	P	2022	tinggi
28	Laras	P	2022	tinggi
29	Dara	P	2022	pendek
30	Brigitta	P	2022	pendek

8. Sort Data Frame □ Lakukan sort berdasarkan waktu perjalanan

```
data_vania2.sort<-data_vania2[order(data_vania2$Waktu.Perjalanan),]
View(data_vania2.sort)
```

Output:



	Nama	Gender	Angkatan	Tinggi.Badan	Waktu.Perjalanan	Wilayah.Tinggal	Jurusan
1	Gisyelle	P	2022	tinggi	10	Grogol	INFOR22
2	Tarum	P	2022	pendek	10	Grogol	INFOR22
27	Mona	P	2022	tinggi	10	Medan	INFOR22
30	Brigitta	P	2022	pendek	10	Cikarang	INFOR22
4	Mendari	P	2022	pendek	15	Grogol	INFOR22
10	Putri	P	2022	pendek	15	Grogol	INFOR22
12	Rhena	P	2022	pendek	15	Grogol	INFOR22
26	Azalea	P	2022	pendek	15	Semarang	INFOR22
14	Irma	P	2022	pendek	20	Bekasi	INFOR22
16	Audias	P	2022	pendek	20	Bekasi	INFOR22
19	Chika	P	2022	tinggi	20	Cikarang	INFOR22
24	Davina	P	2022	tinggi	20	Jogja	INFOR22
28	Laras	P	2022	tinggi	20	Semarang	INFOR22
15	Naura	P	2022	pendek	30	Bekasi	INFOR22
18	Khairunnisa	P	2022	tinggi	30	Bekasi	INFOR22
29	Dara	P	2022	pendek	30	Surabaya	INFOR22
20	Rania	P	2022	pendek	45	Bekasi	INFOR22
5	Kharisma	P	2022	pendek	60	Klender	INFOR22

#### 4. File Praktikum

Github Repository:

[https://github.com/10velys/praktikumStatistika\\_VaniaRD](https://github.com/10velys/praktikumStatistika_VaniaRD)

#### 5. Soal Latihan

Soal:

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

Jawaban:

1. Kegunaan pengelolaan data pada data frame, yaitu :
  - Penyimpanan dan representasi data
  - Pembersihan data
  - Penggabungan data
  - Seleksi dan subsetting data
  - Transformasi data



2. Kasus data perlu dihapus :

- Data tidak relevan
- Data duplikat
- Data tidak valid

## 6. Kesimpulan

- Dalam pengerjaan praktikum Statistika, saya dapat memahami bagaimana cara melakukan pengelolaan data pada R studio dengan cara merge data, subset data frame, mengubah dan menambah variabel, serta menghapus variabel.
- Kita juga dapat mengetahui bagaimana cara mengubah nilai suatu variabel dengan cara mengubah isi kolom variabel tersebut (dengan menggunakan condition if else).

## 7. Cek List (✓)

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

## 8. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	30 Menit	Menarik
2.	Latihan Kedua	45 Menit	Menarik

Keterangan:

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