Nama: **(Adrian Maulana Yusup)** **NIM:** **(065002100036)**	 **Praktikum Statistika**	**MODUL 3** **Nama Dosen:** **Dedy Sugiarto**
Hari/Tanggal: **Kamis, 29 September 2022**		**Nama Asisten Labratorium:** **1. Azhar Rizki Zulma** **065001900001** **2. Arfa Maulana** **064001900039**

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

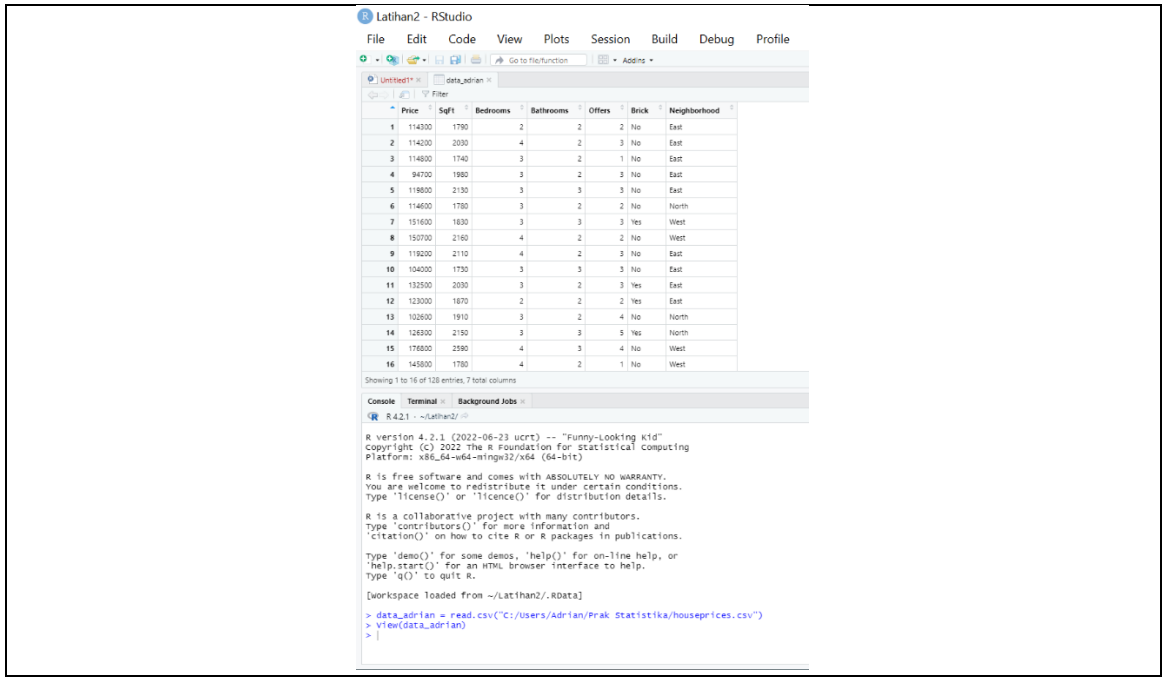
3. Elemen Kompetensi

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.csv("C:/Users/arfa/Bahan Aslab Statistika 2021/houseprices.csv")
View(data_nama)
```

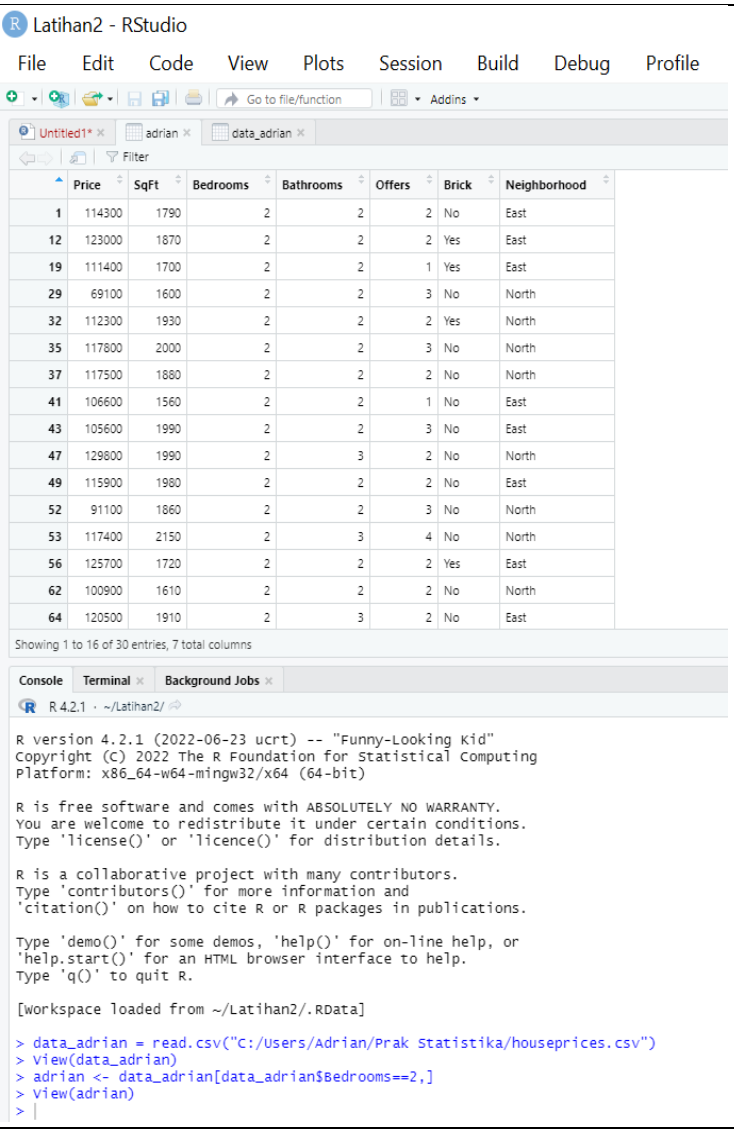
Output:



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)
```

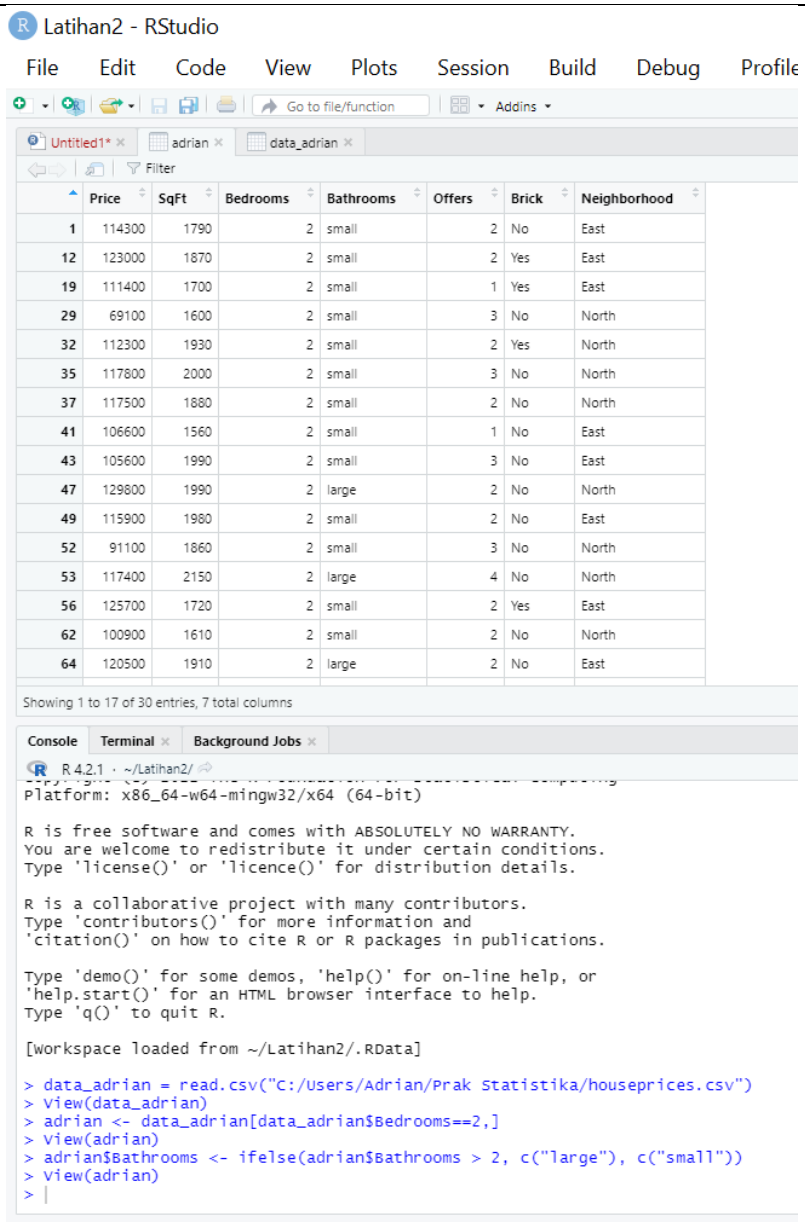
Output:



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)
```

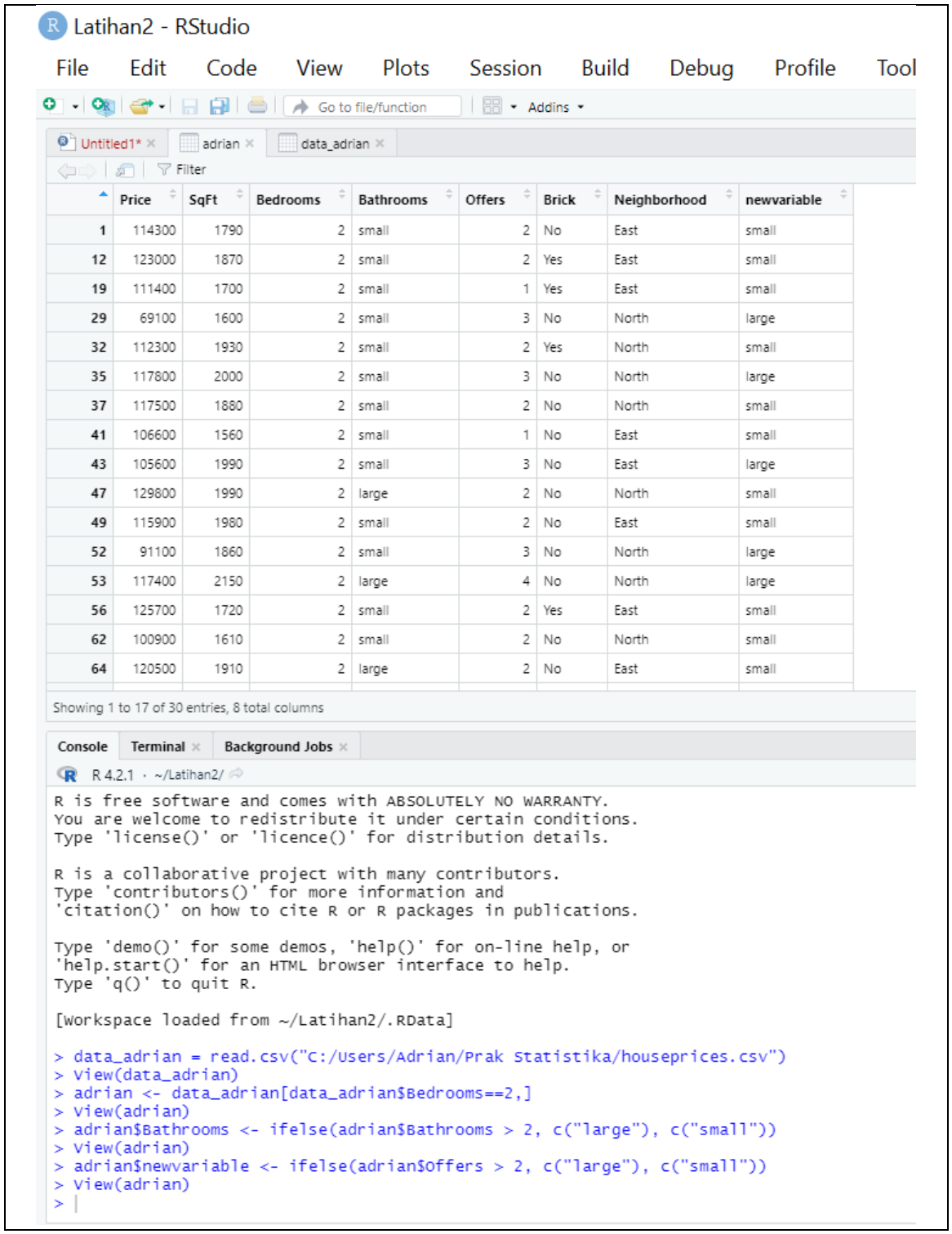
Output:



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

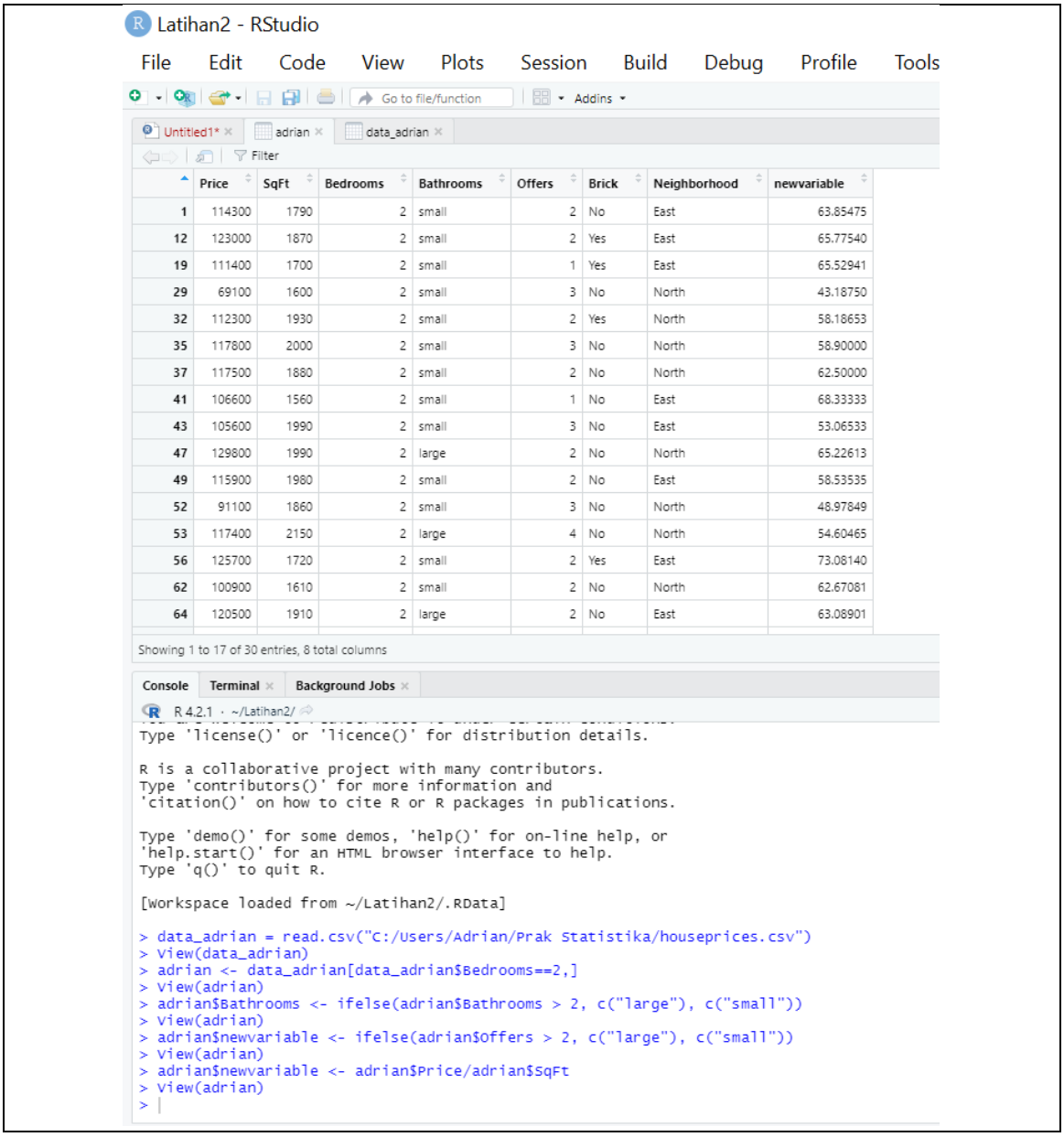


Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood	newvariable
1	114300	1790	2	small	2	No	East	small
12	123000	1870	2	small	2	Yes	East	small
19	111400	1700	2	small	1	Yes	East	small
29	69100	1600	2	small	3	No	North	large
32	112300	1930	2	small	2	Yes	North	small
35	117800	2000	2	small	3	No	North	large
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

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

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

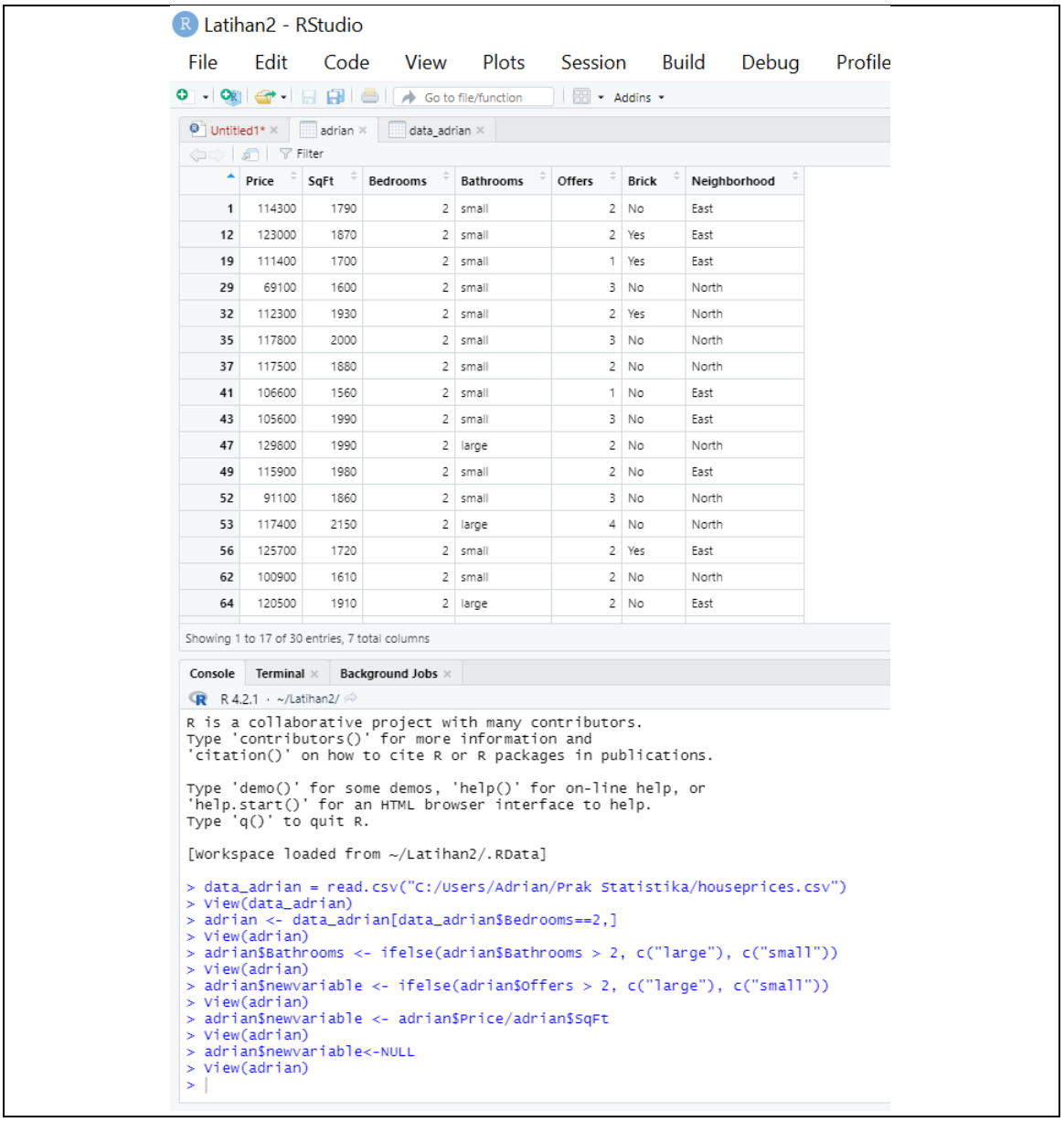
Output:



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:

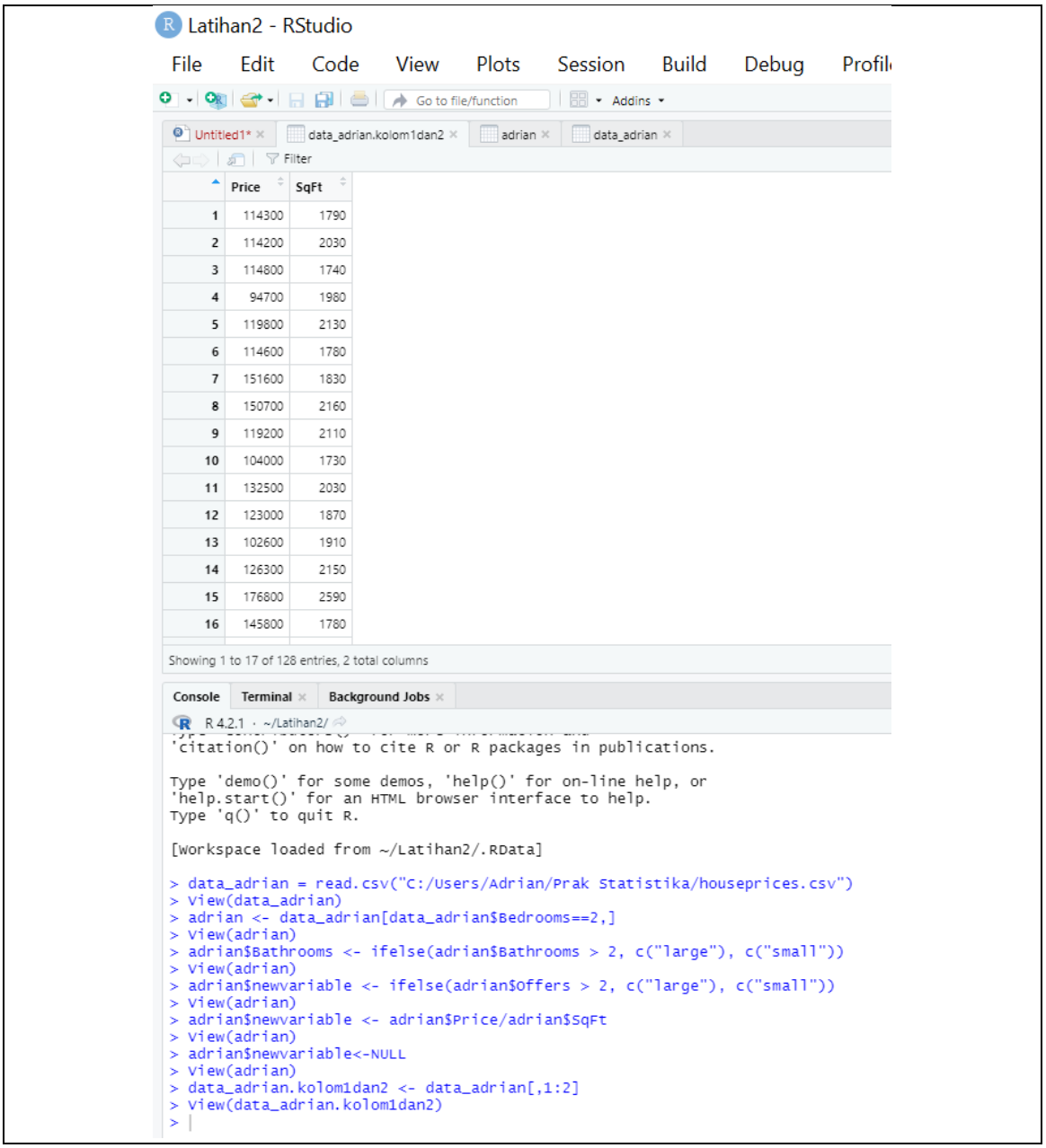


Price	SqFt	Bedrooms	Bathrooms	Offers	Brick	Neighborhood
1	114300	1790	2	small	2	No	East
12	123000	1870	2	small	2	Yes	East
19	111400	1700	2	small	1	Yes	East
29	69100	1600	2	small	3	No	North
32	112300	1930	2	small	2	Yes	North
35	117800	2000	2	small	3	No	North
37	117500	1880	2	small	2	No	North
41	106600	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

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)
```

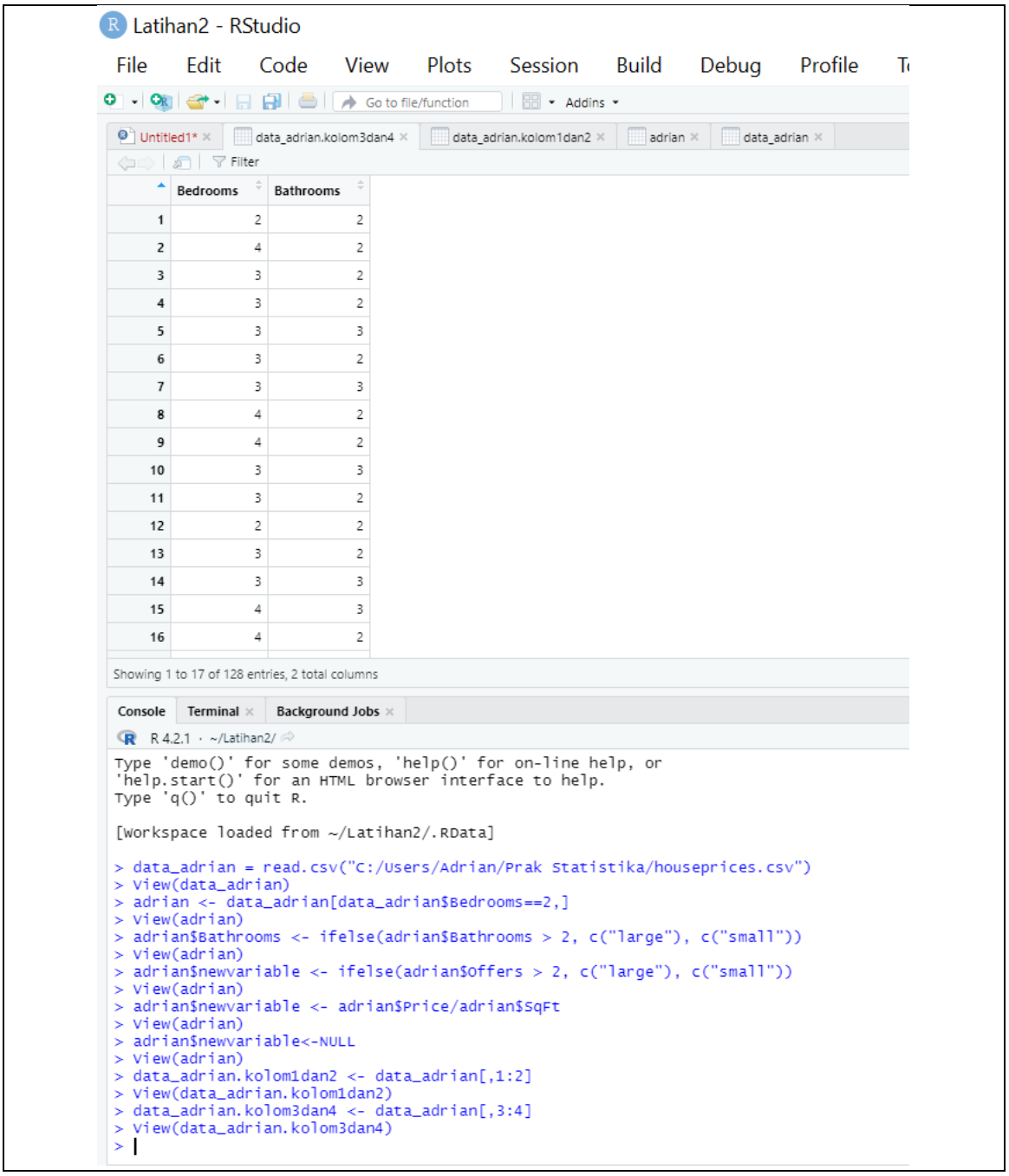
Output:



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

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

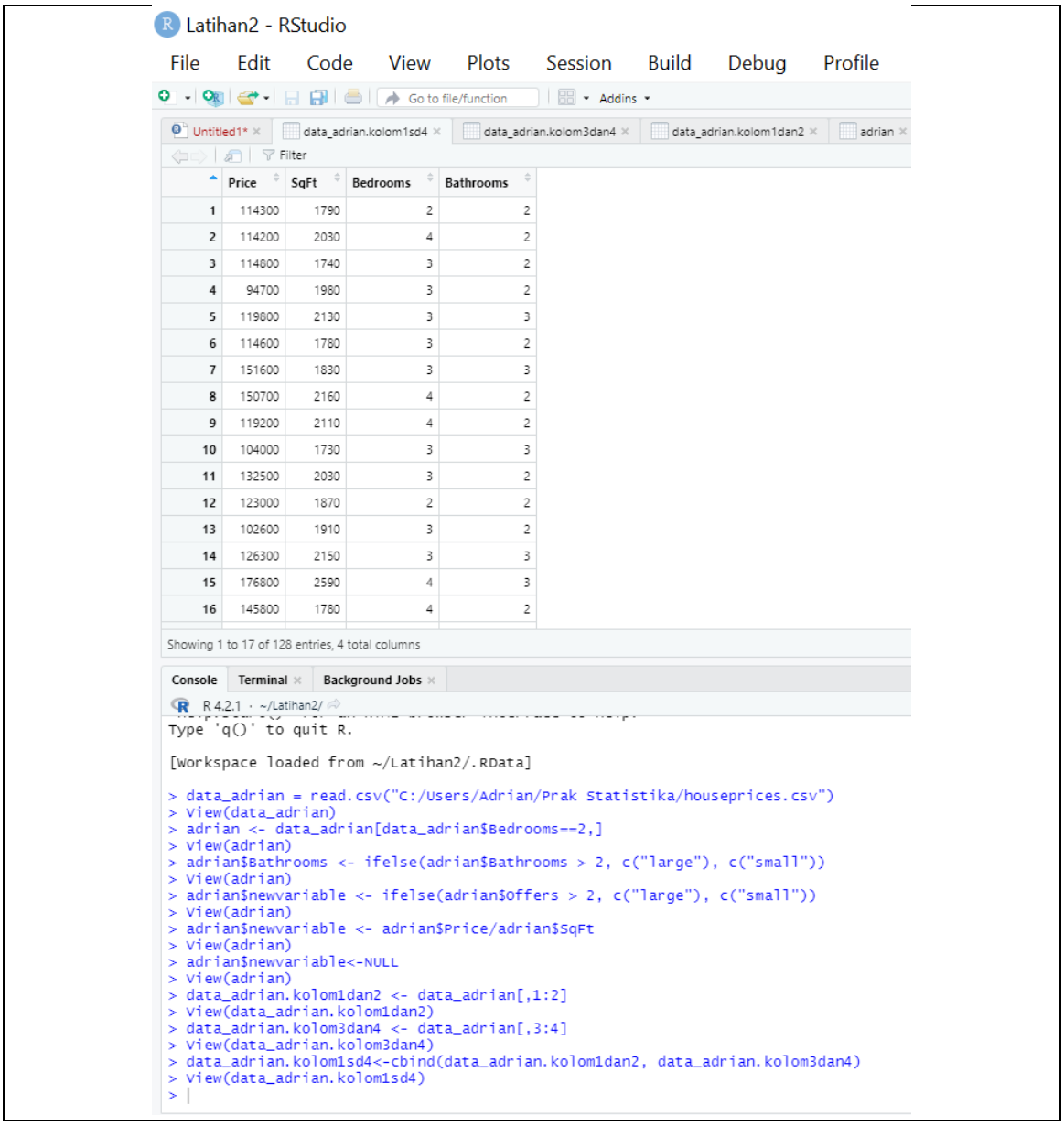
Output:



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)
```

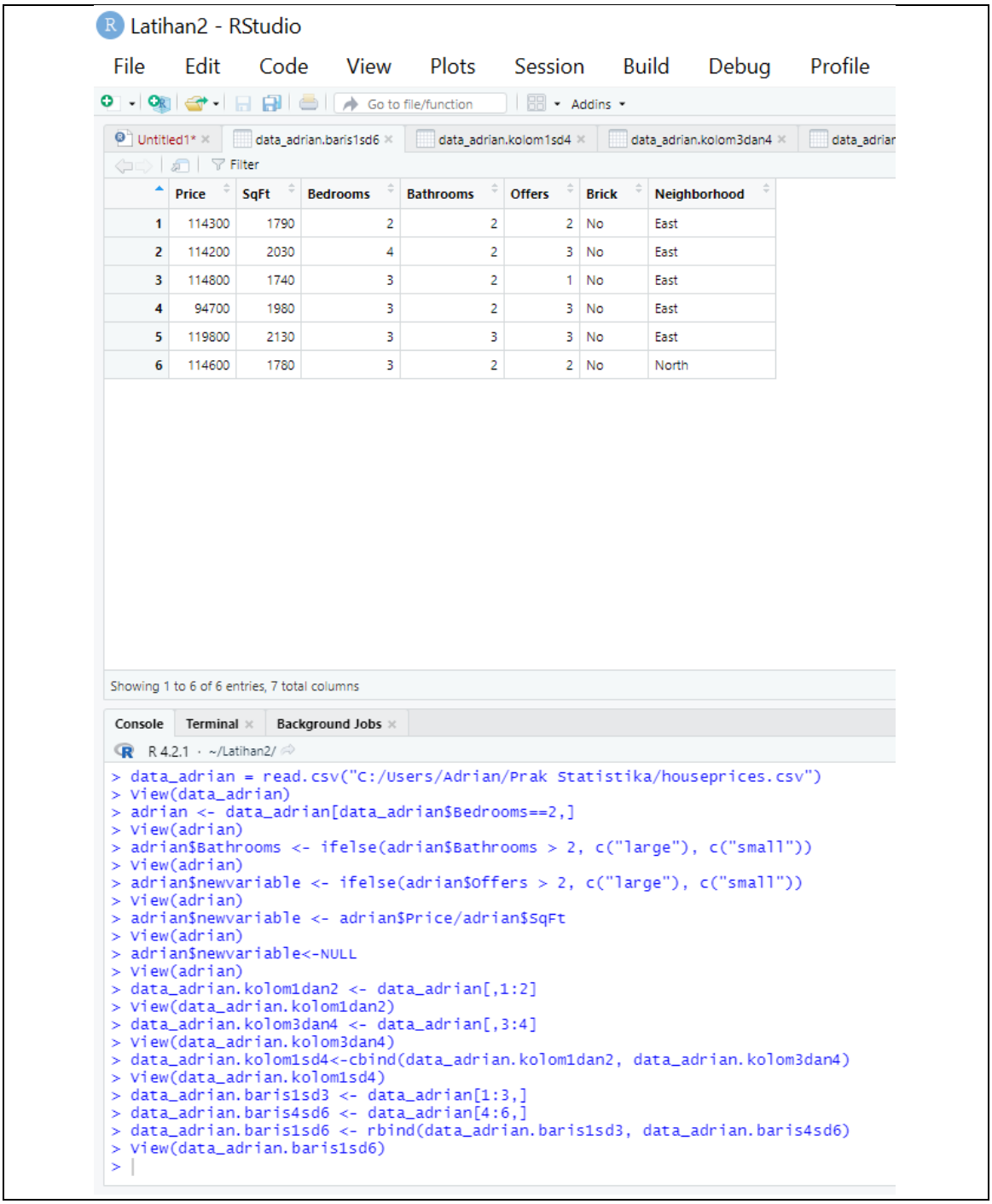
Output:



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)
```

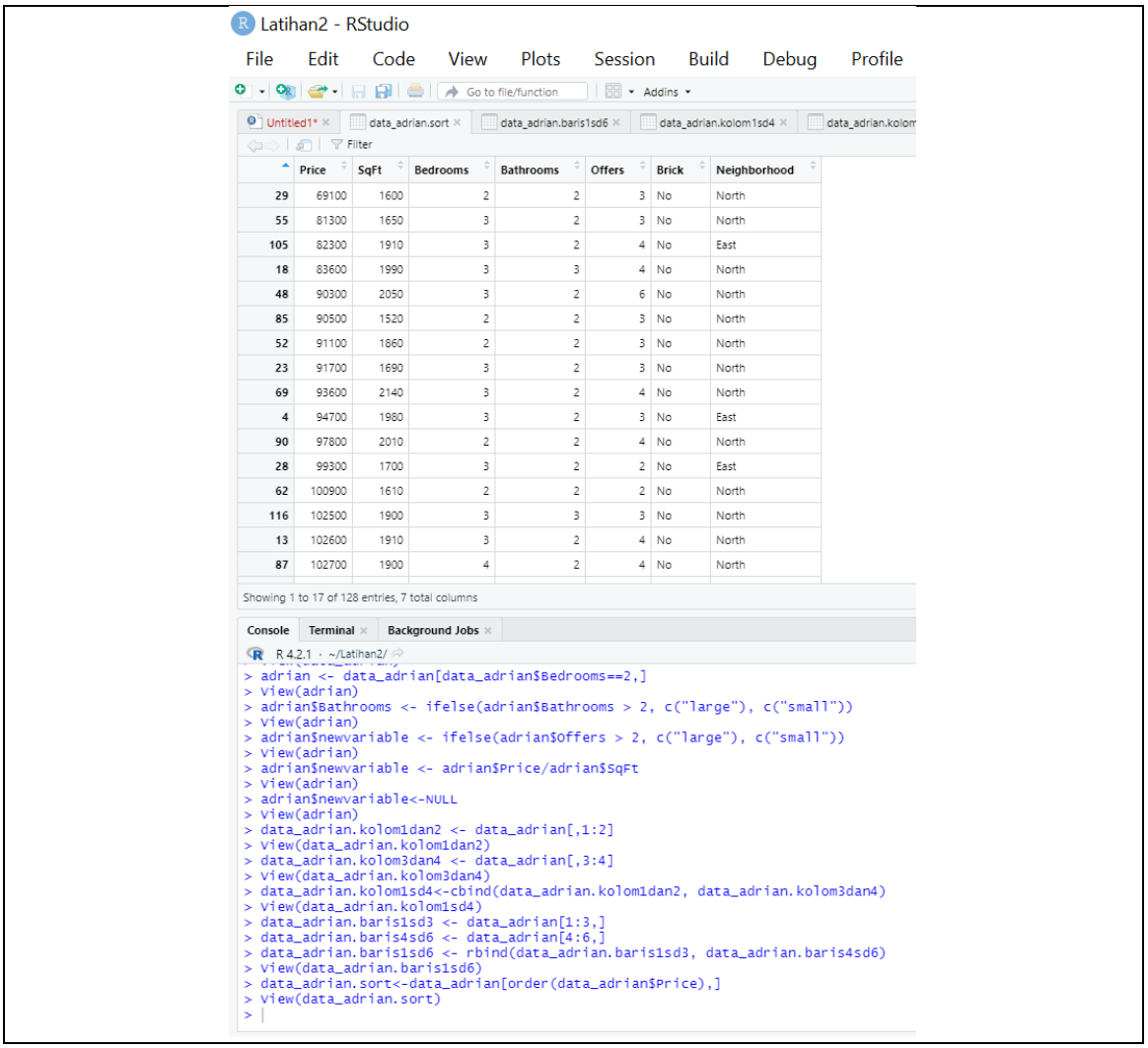
Output:



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:



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

Taruh baris source code disini

Output:

Taruh baris screenshot output disini

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

Taruh baris source code disini

Output:

Taruh baris screenshot output disini

3. Mengubah nilai suatu variabel → Ubah isi kolom tinggi badan untuk tinggi > 160 menjadi "Tinggi" dan jika bukan berarti "Pendek"

Taruh baris source code disini

Output:

Taruh baris screenshot output disini

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"

Taruh baris source code disini

Output:

Taruh baris screenshot output disini

5. Delete Variabel → Hapus kolom Fakultas

Taruh baris source code disini

Output:

Taruh baris screenshot output disini

6. Merge Data Frame

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

Taruh baris source code disini

Output:

Taruh baris screenshot output disini

7. Merge Baris

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

Taruh baris source code disini

Output:

Taruh baris screenshot output disini

8. Sort Data Frame → Lakukan sort berdasarkan waktu perjalanan

Taruh baris source code disini

Output:

Taruh baris screenshot output disini

4. File Praktikum

Github Repository:

5. Soal Latihan

Soal:

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

Jawaban:

1.
2.

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	**Selesai**	
2.	Latihan Kedua	**…**	

8. Formulir Umpan Balik

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

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