


<p>Nama: Putra Mario Santoso</p> <p>NIM: 065002200037</p>	 <p>Praktikum Statistika</p>	<p>MODUL 6</p> <p>Nama Dosen: Dedy Sugiarto</p>
<p>Hari/Tanggal: Senin, 31 Juli 2023</p>		<p>Nama Asisten Labratorium: 1. Elen Fadilla Estri 064002000008 2. Rukhy Zaifa Aduhalim 064002000041</p>

Explorasi Data

1. Teori Singkat

Boxplot

Box plot atau boxplot (juga dikenal sebagai diagram box-and-whisker) merupakan suatu teknik grafikal dalam statistik deskriptif untuk menggambarkan secara grafik dari data numerik melalui lima ukuran sebagai berikut:

- Nilai observasi terkecil (minimum)
- Kuartil terendah atau kuartil pertama (Q1), yang memotong 25 % dari data terendah
- Median (Q2) atau nilai pertengahan
- Kuartil tertinggi atau kuartil ketiga (Q3), yang memotong 25 % dari data tertinggi
- Nilai observasi terbesar (maksimum)

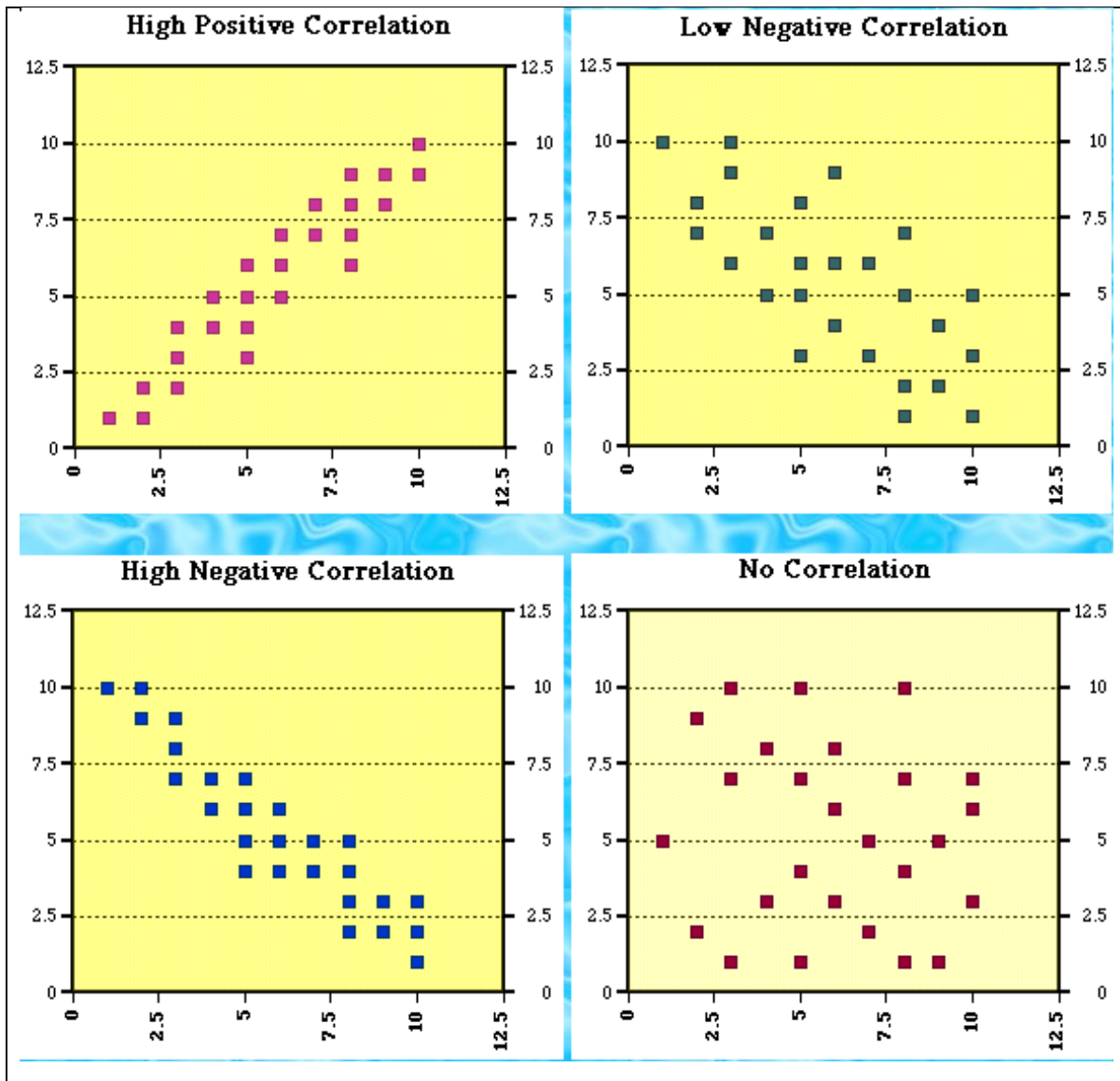
Histogram

Histogram adalah salah satu grafik statistik untuk mengevaluasi bentuk dan sebaran dari data numerik. Histogram membagi nilai sampel menjadi beberapa interval yang disebut bins. Batangnya menggambarkan jumlah pengamatan (frekuensi) yang jatuh di dalam bin.

Scatter Plot

Scatter plot dapat digunakan untuk melihat hubungan antara dua variabel numeric berdasarkan dari pola tebaran titiknya.





2. Alat dan Bahan

Hardware : Laptop/PC

Software : R Studio

3. Elemen Kompetensi

Data yang digunakan: Iris.csv

a. Latihan pertama – Boxplot

1. Script

```
summary(iris)
```



Output:

```
> summary(iris)
      Sepal.Length      Sepal.Width      Petal.Length      Petal.Width
Min.      :4.300    Min.      :2.000    Min.      :1.000    Min.      :0.100
1st Qu.:5.100    1st Qu.:2.800    1st Qu.:1.600    1st Qu.:0.300
Median :5.800    Median :3.000    Median :4.350    Median :1.300
Mean   :5.843    Mean   :3.057    Mean   :3.758    Mean   :1.199
3rd Qu.:6.400    3rd Qu.:3.300    3rd Qu.:5.100    3rd Qu.:1.800
Max.   :7.900    Max.   :4.400    Max.   :6.900    Max.   :2.500

Species
setosa      :50
versicolor:50
virginica   :50
```

2. Script

```
head(iris)
```

Output:

The screenshot shows the RStudio interface. In the console, the command `head(iris)` has been executed, resulting in a data frame with 6 rows and 5 columns: `Sepal.Length`, `Sepal.Width`, `Petal.Length`, `Petal.Width`, and `Species`. The first row shows values 5.1, 3.5, 1.4, 0.2, and 'setosa'. The Environment pane on the right shows the `data` object with 19 observations and 6 variables.

```
> head(iris)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5         1.4         0.2   setosa
2          4.9         3.0         1.4         0.2   setosa
3          4.7         3.2         1.3         0.2   setosa
4          4.6         3.1         1.5         0.2   setosa
5          5.0         3.6         1.4         0.2   setosa
6          5.4         3.9         1.7         0.4   setosa
```

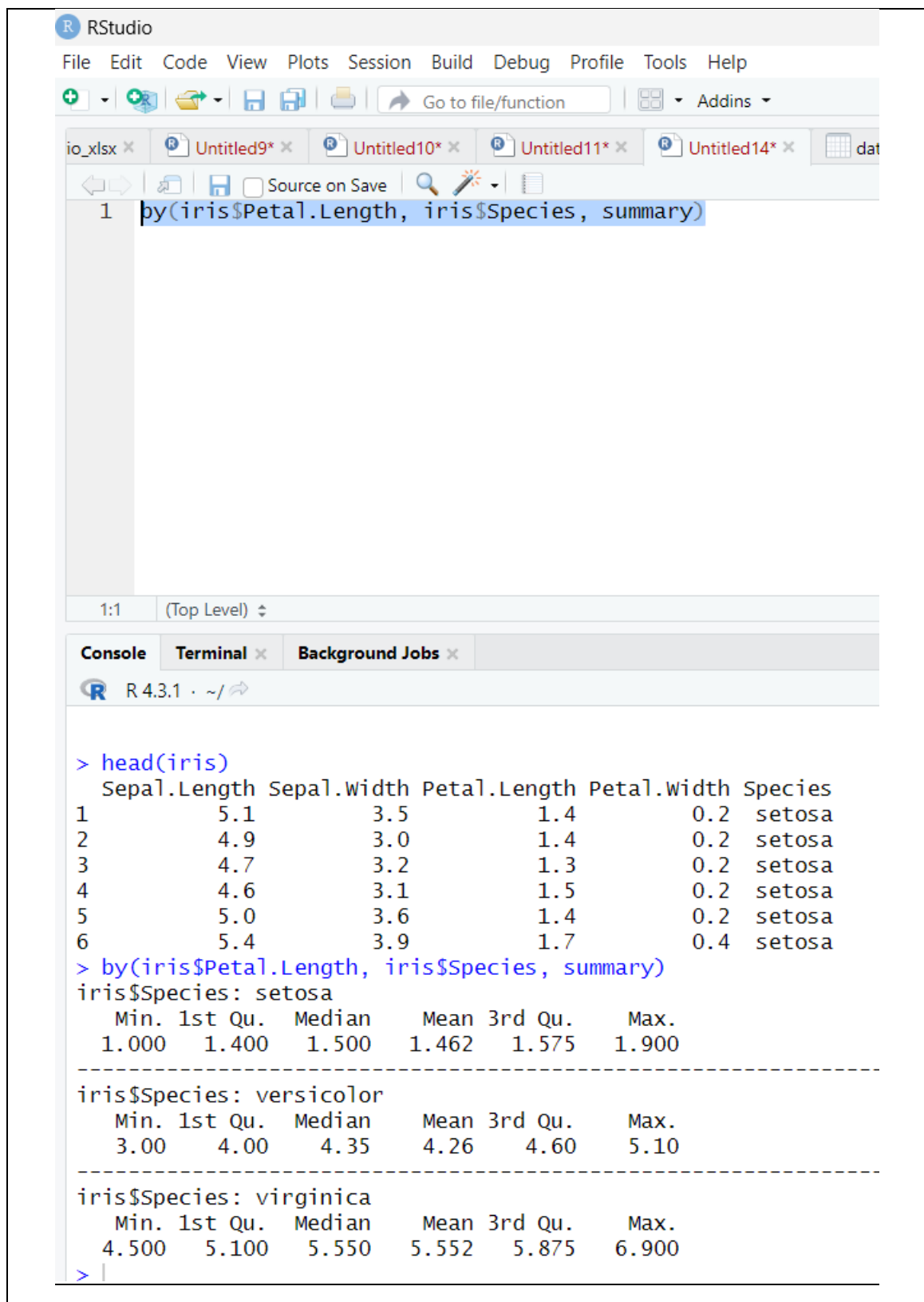


3. Script

```
by(iris$Petal.Length, iris$Species, summary)
```

Output:





The screenshot shows the RStudio environment. The script editor contains the following code:

```
1 by(iris$Petal.Length, iris$Species, summary)
```

The console shows the output of the code execution:

```
> head(iris)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1           3.5          1.4          0.2  setosa
2          4.9           3.0          1.4          0.2  setosa
3          4.7           3.2          1.3          0.2  setosa
4          4.6           3.1          1.5          0.2  setosa
5          5.0           3.6          1.4          0.2  setosa
6          5.4           3.9          1.7          0.4  setosa

> by(iris$Petal.Length, iris$Species, summary)
iris$Species: setosa
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
1.000  1.400   1.500   1.462  1.575   1.900

-----
iris$Species: versicolor
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 3.00   4.00   4.35   4.26   4.60   5.10

-----
iris$Species: virginica
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
4.500  5.100  5.550  5.552  5.875  6.900

>
```



4. Script

```
mean(iris$Sepal.Length)
```

Output:

```
> mean(iris$Sepal.Length)
[1] 5.843333
> |
```

5. Script

```
var(iris$Sepal.Length)
```

Output:

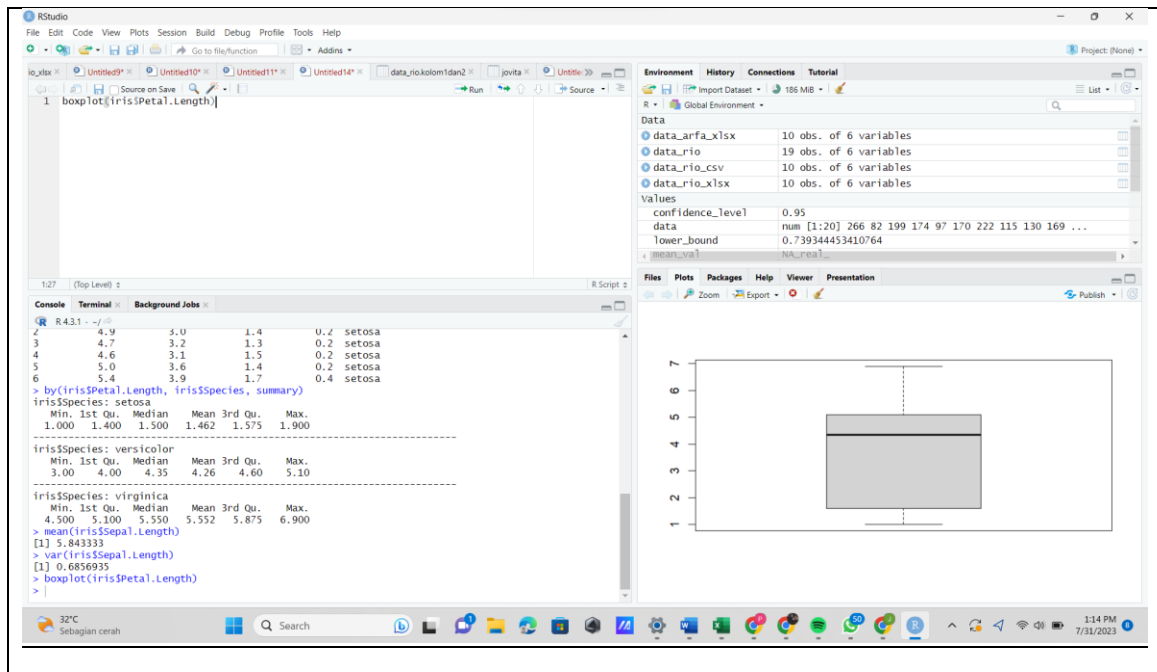
```
> var(iris$Sepal.Length)
[1] 0.6856935
> |
```

6. Script

```
boxplot(iris$Petal.Length)
```

Output:



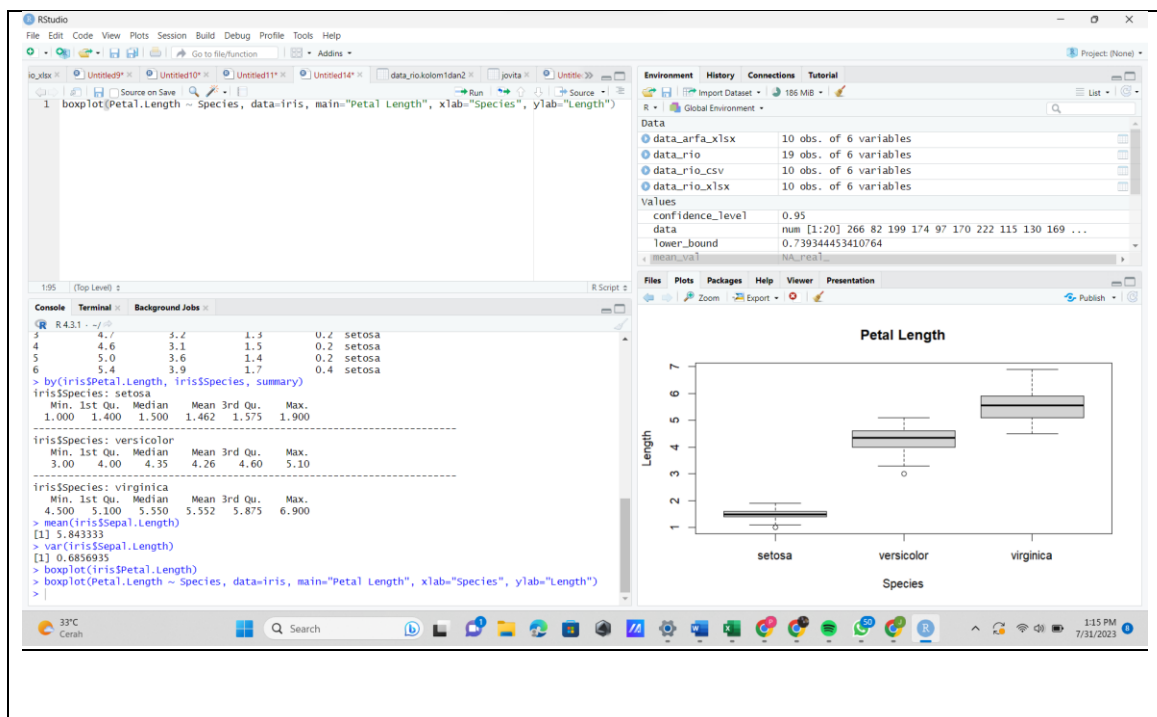


7. Script

```
boxplot(Petal.Length ~ Species, data=iris, main="Petal Length", xlab="Species",  
ylab="Length")
```

Output:





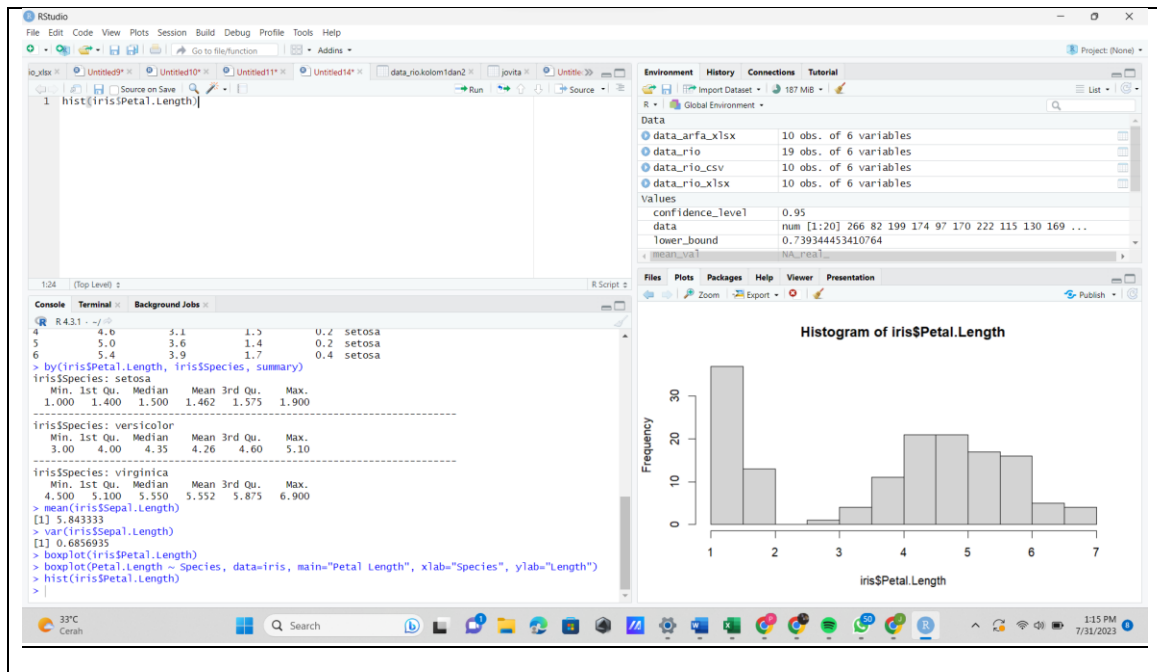
b. Latihan Kedua – Histogram & Destiny

1. Script

```
hist(iris$Petal.Length)
```

Output:

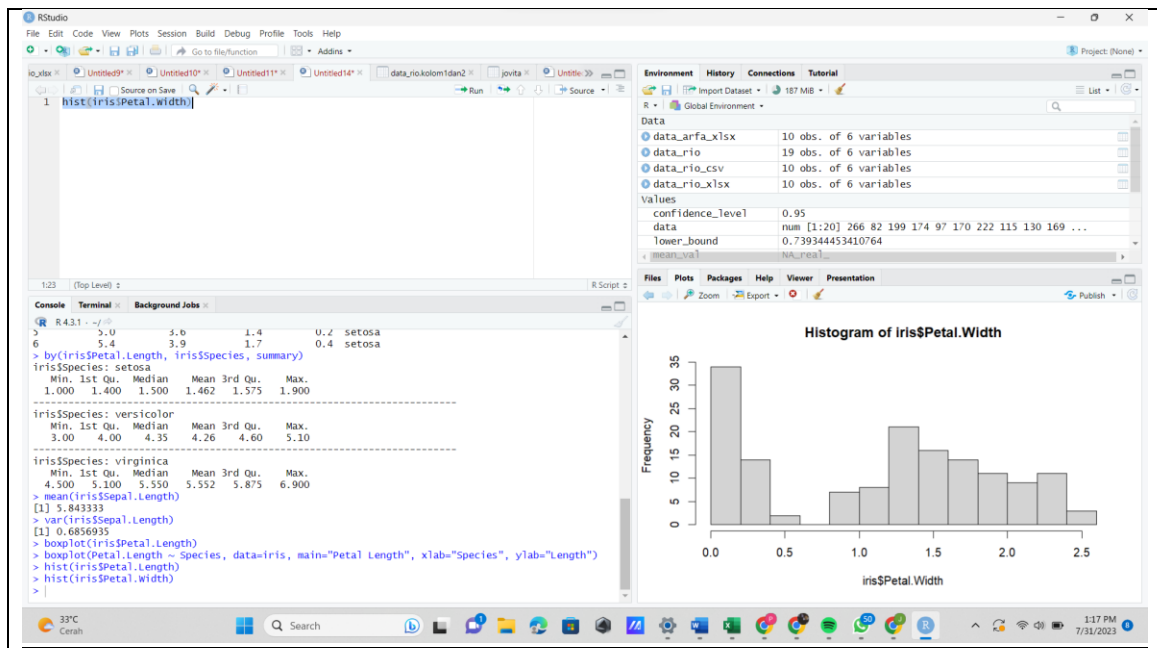




2. Script

```
hist(iris$Sepal.Width)
```

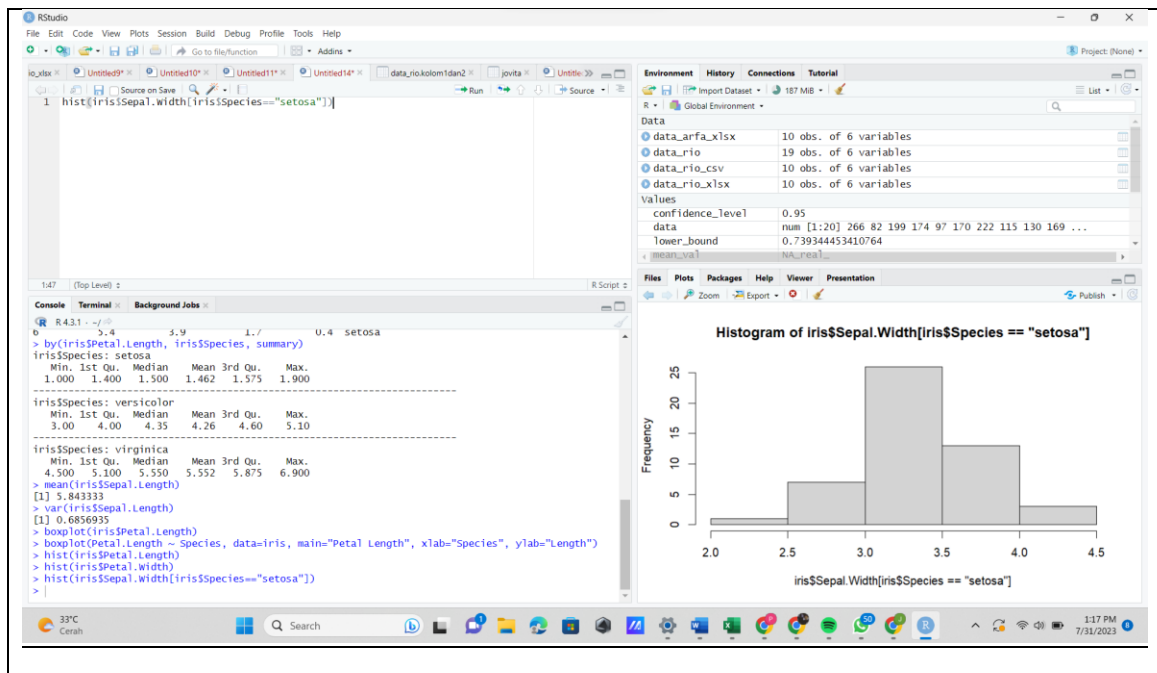
Output:



3. Script

```
hist(iris$Sepal.Width[iris$Species=="setosa"])
```

Output:

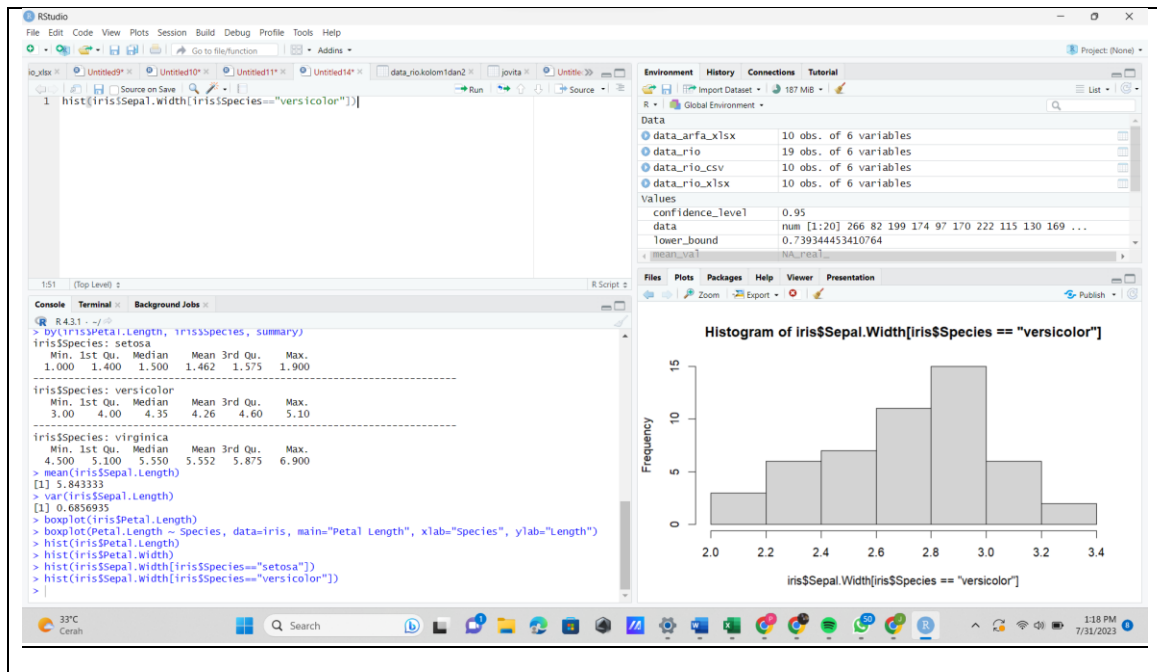


4. Script

```
hist(iris$Sepal.Width[iris$Species=="versicolor"])
```

Output:



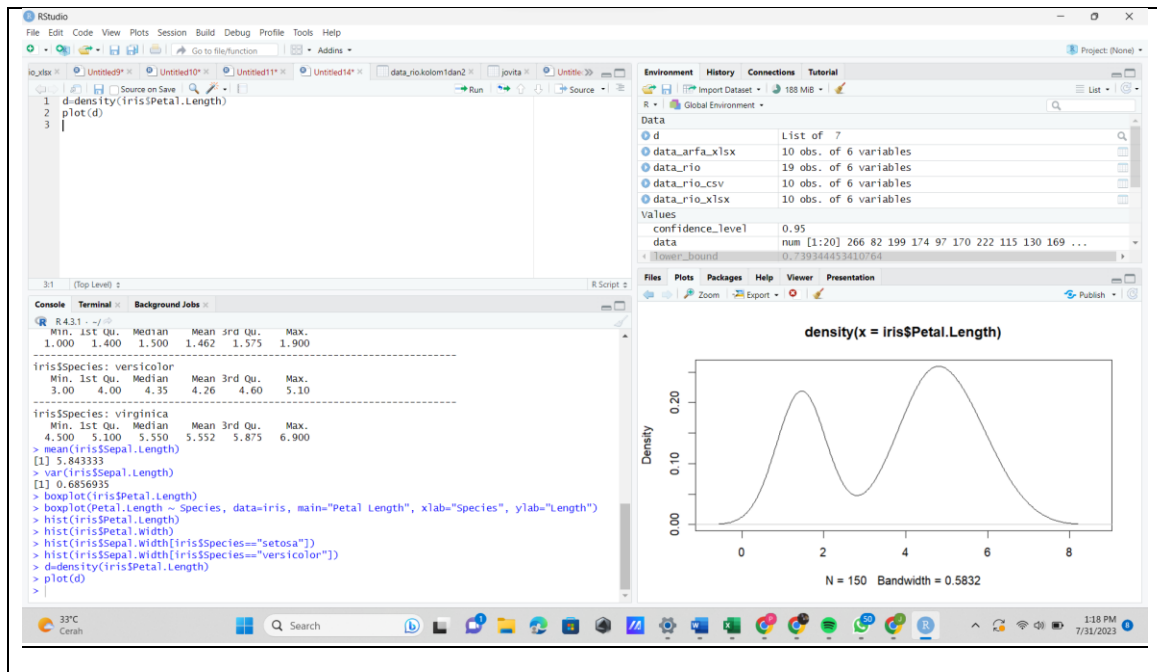


5. Script

```
d=density(iris$Petal.Length)
plot(d)
```

Output:



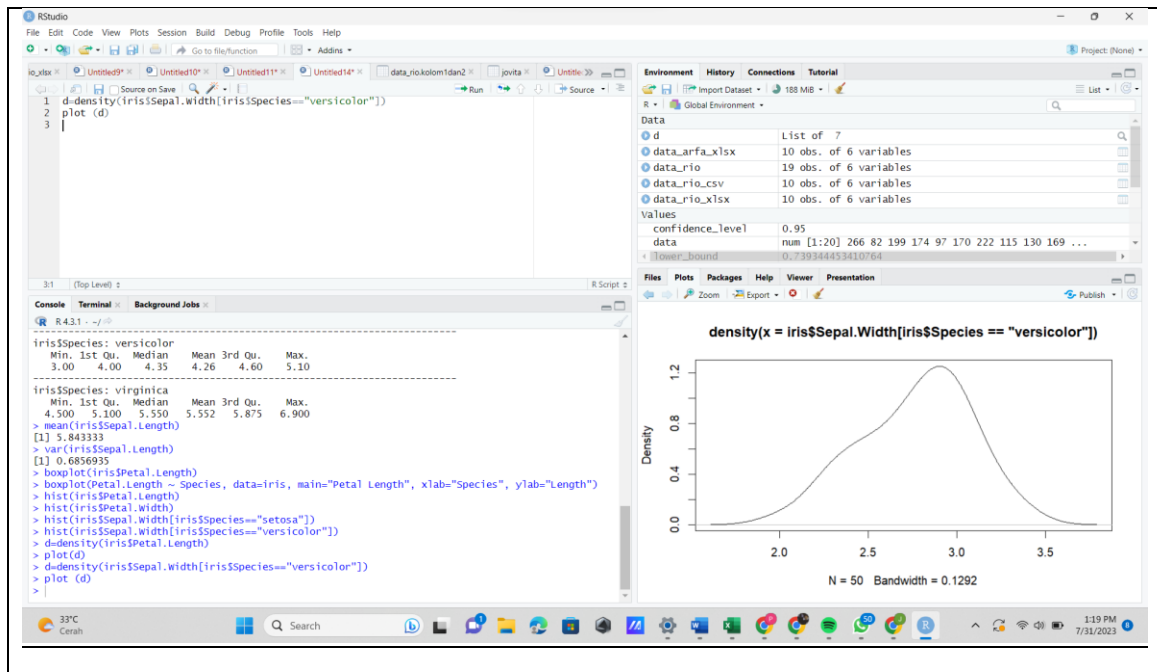


6. Script

```
d=density(iris$Sepal.Width[iris$Species=="versicolor"])
plot (d)
```

Output:





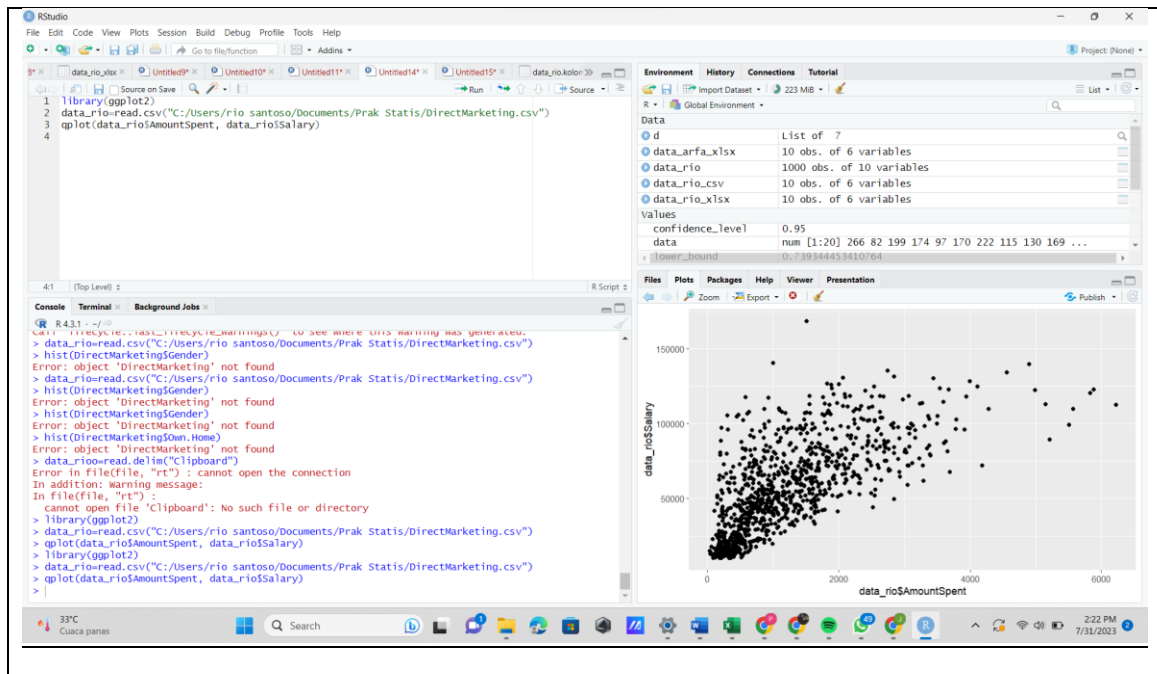
c. Latihan Ketiga – Scatter Plot

1. Script

```
qplot(data_namapraktikan$AmountSpent, data_namapraktikan$Salary)
```

Output:

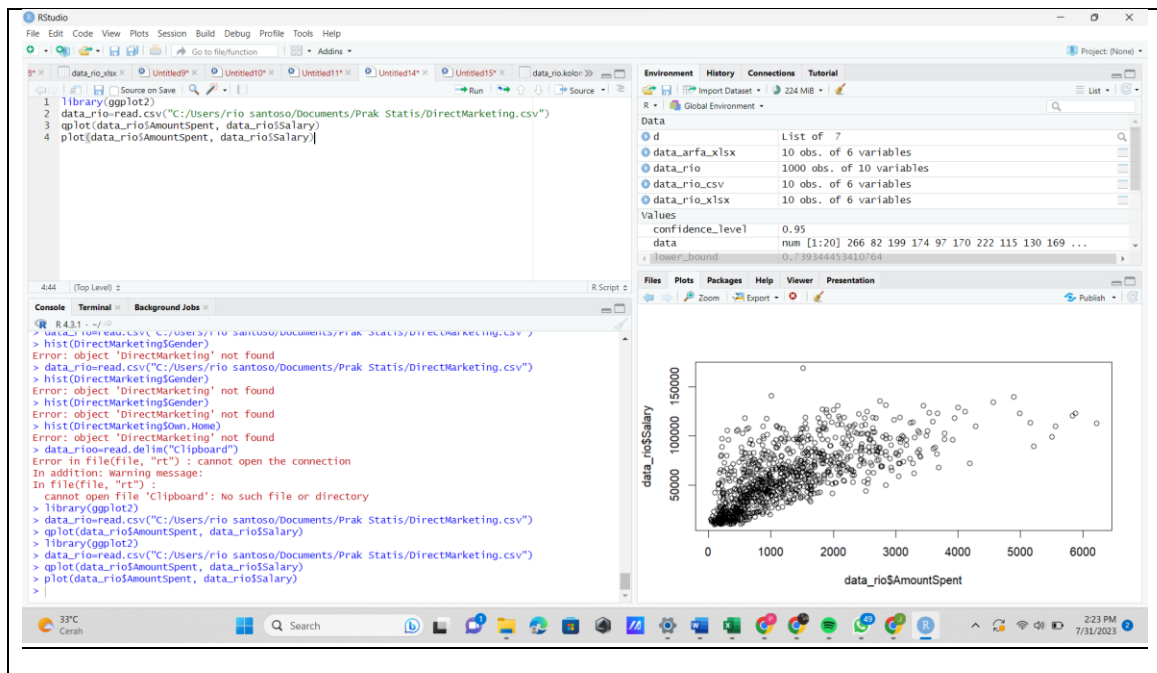




2. Script

```
plot(data_namapraktikan$AmountSpent, data_namapraktikan$Salary)
```

Output:



d. Latihan Keempat – Tugas

1. Gunakan data DirectMarketing.csv. Lakukan analisis dengan menggunakan histogram dan density plot untuk mengetahui faktor-faktor yang mempengaruhi pengeluaran (AmountSpent)

Script

```
Histogram

DirectMarketing = read.delim("clipboard")

hist(DirectMarketing$AmountSpent)

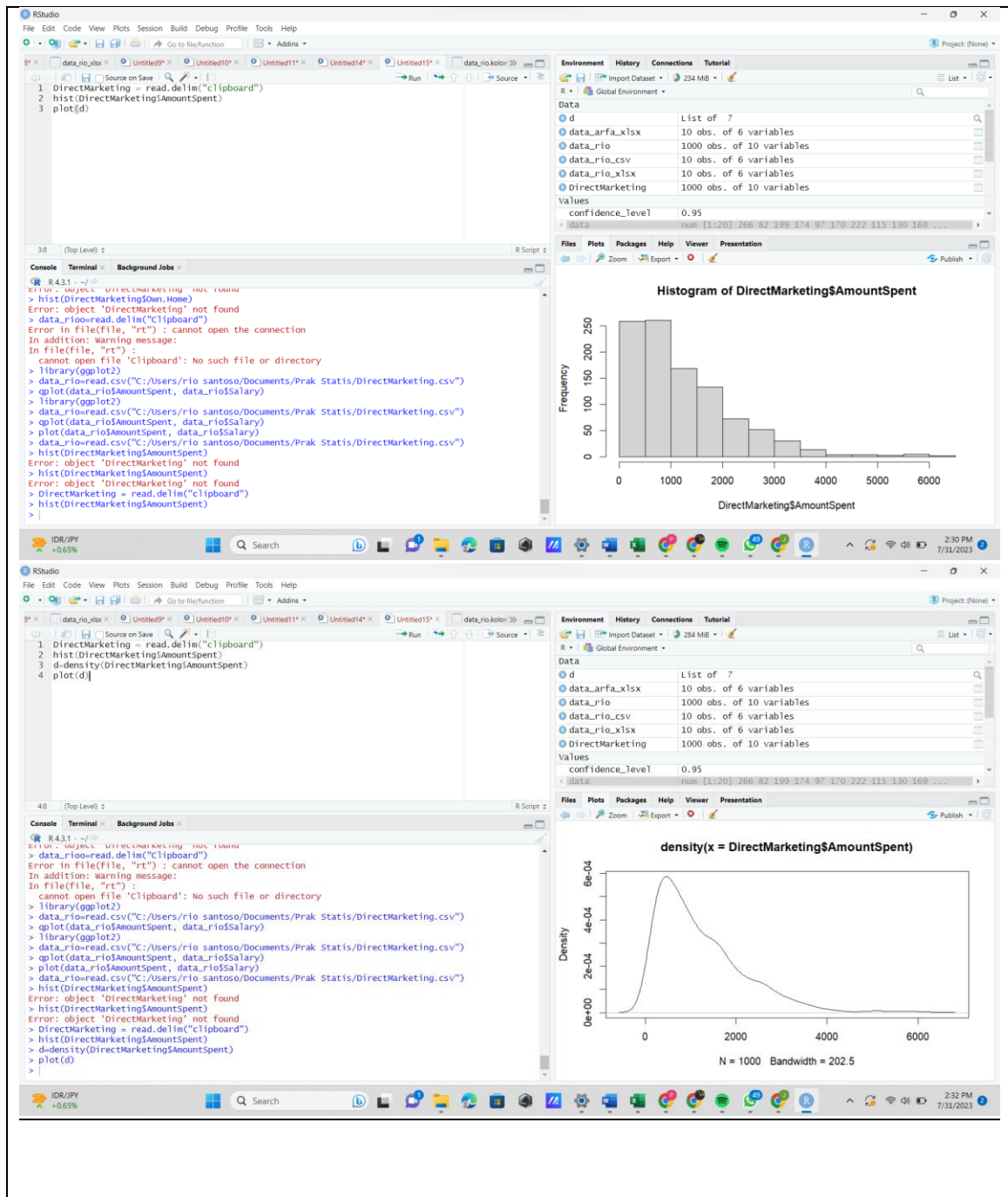
d=density(DirectMarketing$AmountSpent)

plot(d)
```

Output

Histogram





2. Gunakan data Houseprices.csv. Lakukan analisis dengan menggunakan scatter plot untuk mengetahui berbagai hubungan variabel numeric dengan harga rumah.

Script



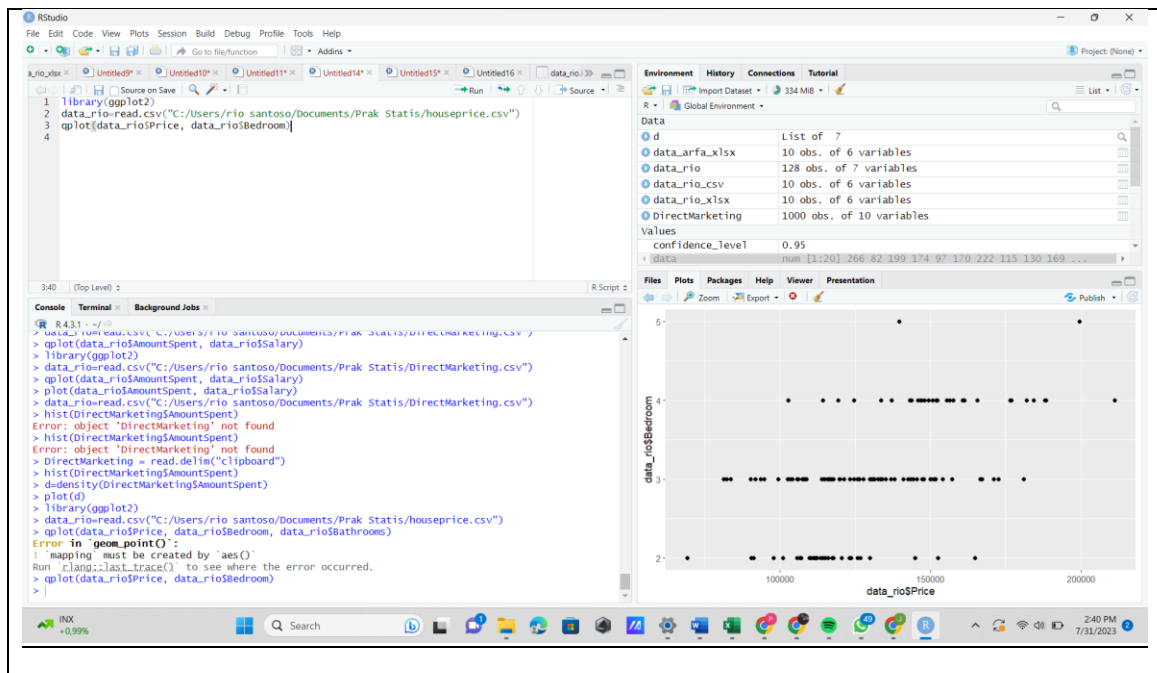

```
library(ggplot2)

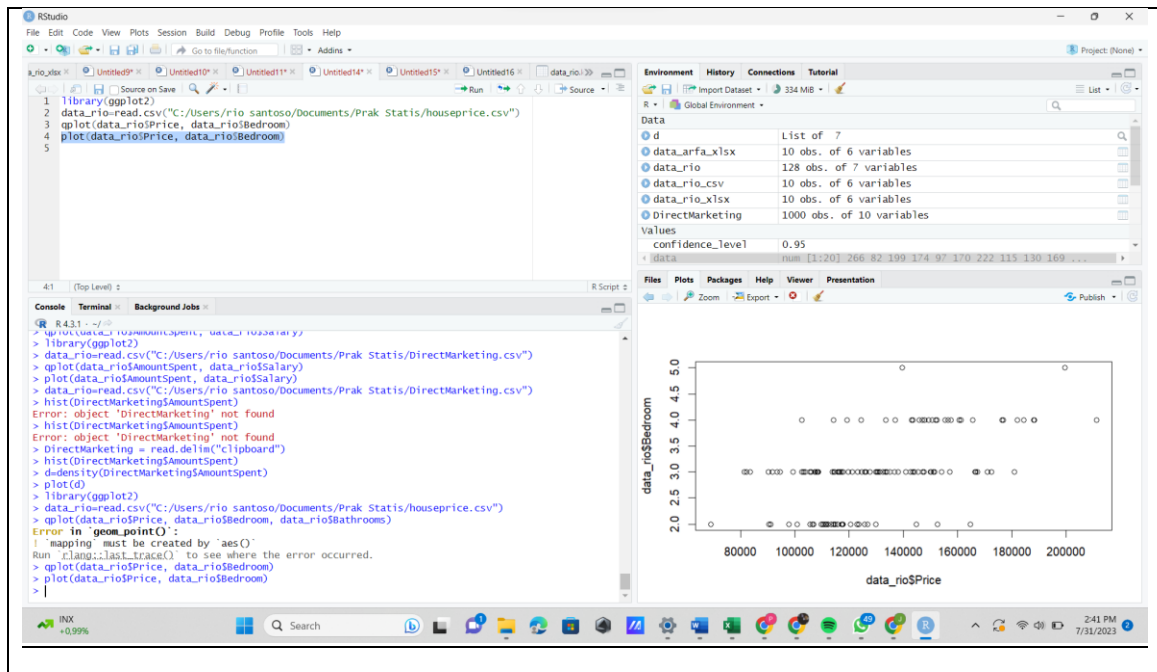
data_rio=read.csv("C:/Users/rio santoso/Documents/Prak Statis/houseprice.csv")

qplot(data_rio$Price, data_rio$Bedroom)

plot(data_rio$Price, data_rio$Bedroom)
```

Output





4. File Praktikum

Github Repository:

5. Soal Latihan

Soal:

1. Apa perbedaan utama pada grafik boxplot, histogram, dan density?
2. Jelaskan kegunaan dan fungsi dari Scatter Plot?

1. Jawaban:

1. Perbedaan utama antara ketiganya adalah metode representasi data dan informasi yang ingin disampaikan. Boxplot cocok untuk melihat perbandingan distribusi beberapa kelompok data, melihat keberadaan pencilan, dan memahami statistik ringkasan. Histogram berguna untuk melihat bentuk distribusi secara visual, sedangkan density plot memberikan informasi yang lebih halus tentang kepadatan probabilitas data
2. a. Memvisualisasikan hubungan antara variabel.

- b. Mendeteksi pola atau tren dalam data.
- c. Mengidentifikasi outliers atau data ekstrem.
- d. Klasifikasi data dengan membedakan warna atau simbol untuk kelas yang berbeda.



- e. Analisis regresi untuk memodelkan hubungan antara variabel independen dan dependen.
- f. Memvisualisasikan hubungan multivariabel dengan matriks scatter plot.

6. Kesimpulan

- a. Dalam pengerjaan praktikum Statistika, bagaimana cara mengolah suatu data lewat grafik boxpot, histogram
- b. Kita juga dapat mengetahui lalu kita dapat mengetahui bagaimana cara mengolah data juga dengan menggunakan grafik density dan scatter plot

7. Cek List (□)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	V	
2.	Latihan Kedua	V	
3.	Latihan Ketiga	v	
4.	Latihan Keempat	v	

8. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	20 Menit	Menarik
2.	Latihan Kedua	20 Menit	Menarik
3.	Latihan Ketiga	20 Menit	menarik



4.	Latihan Keempat	20 Menit	...
----	-----------------	----------	-----

Keterangan:

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

