| **Nama:**  **Adrian Halim**  **NIM: 064002200043** | C:\Users\RPL-SI 02\Pictures\288px-Trisakti_Logo.svg.png | **MODUL 6**  **Nama Dosen:**  **Dedy Sugiarto** |
| --- | --- | --- |
| **Hari/Tanggal:**  **Senin, 31 JULI 2023** | **Praktikum Statistika** | **Nama Asisten Labratorium:**  **1. Elen Fadilla Estri**  **064002000008**  **2. Rukhy Zaifa Aduhalim**  **064002000041** |

**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.

|  |
| --- |

1. **Alat dan Bahan**

Hardware : Laptop/PC

Software : R Studio

1. **Elemen Kompetensi**

Data yang digunakan: Iris.csv

* 1. Latihan pertama – Boxplot

1. Script

| summary(iris) |
| --- |

Output:

|  |
| --- |

2. Script

| head(iris) |
| --- |

Output:

|  |
| --- |

3. Script

| by(iris$Petal.Length, iris$Species, summary) |
| --- |

Output:

|  |
| --- |

4. Script

| mean(iris$Sepal.Length) |
| --- |

Output:

|  |
| --- |

5. Script

| var(iris$Sepal.Length) |
| --- |

Output:

|  |
| --- |

6. Script

| boxplot(iris$Petal.Length) |
| --- |

Output:

|  |
| --- |

7. Script

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

Output:

|  |
| --- |

* 1. 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:

|  |
| --- |

* 1. Latihan Ketiga – Scatter Plot

1. Script

| qplot(data\_namapraktikan$AmountSpent, data\_namapraktikan$Salary) |
| --- |

Output:

|  |
| --- |

2. Script

| plot(data\_namapraktikan$AmountSpent, data\_namapraktikan$Salary) |
| --- |

Output:

|  |
| --- |

* 1. 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

| data\_AdrianH=read.csv("C:/Users/ADRIAN HALIM/Downloads/DirectMarketing.csv",sep = ";")  hist(data\_AdrianH$AmountSpent)  d = density(data\_AdrianH$AmountSpent)  plot(d) |
| --- |

Output

|  |
| --- |

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

Script

| library(ggplot2)  data\_AdrianH=read.csv("C:/Users/ADRIAN HALIM/Downloads/houseprice.csv",sep = ";")  qplot(data\_AdrianH$Price, data\_AdrianH$Bedrooms)  qplot(data\_AdrianH$Price, data\_AdrianH$Bathrooms) |
| --- |

Output

|  |
| --- |

1. **File Praktikum**

Github Repository:

| <https://github.com/ADRIANHHALIM/ExplorasiData> |
| --- |

1. **Soal Latihan**

Soal:

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

Jawaban:  
1. Grafik boxplot menyajikan ringkasan statistik, histogram menampilkan distribusi diskrit dalam bentuk bar, dan density plot menampilkan distribusi dalam bentuk kurva kontinu  
  
2. Scatter plot adalah grafik untuk memvisualisasikan hubungan antara dua variabel. Digunakan untuk mengidentifikasi korelasi, pola sebaran data, dan pencilan. Berguna dalam analisis data dan eksplorasi visual.

1. **Kesimpulan**
   1. Dalam pengerjaan praktikum Statistika, Explorasi data adalah proses analisis awal untuk memahami karakteristik dan pola data sebelum langkah selanjutnya. Ini melibatkan visualisasi, statistik deskriptif, dan pemeriksaan keterhubungan. Hasilnya membantu menyusun hipotesis dan pemodelan lebih lanjut.
   2. Kita juga dapat mengetahui wawasan awal tentang dataset
2. **Cek List (✔)**

| **No** | **Elemen Kompetensi** | **Penyelesaian** | |
| --- | --- | --- | --- |
| **Selesai** | **Tidak Selesai** |
| **1.** | Latihan Pertama | **✅** |  |
| **2.** | Latihan Kedua | **✅** |  |
| **3.** | Latihan Ketiga | **✅** |  |
| **4.** | Latihan Keempat | **✅** |  |

1. **Formulir Umpan Balik**

| **No** | **Elemen Kompetensi** | **Waktu Pengerjaan** | **Kriteria** |
| --- | --- | --- | --- |
| **1.** | Latihan Pertama | 3 Menit | Menarik |
| **2.** | Latihan Kedua | 4 Menit | Menarik |
| **3.** | Latihan Ketiga | 10 Menit | Menarik |
| **4.** | Latihan Keempat | 15 Menit | Menarik |

Keterangan:

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