|  |  |  |
| --- | --- | --- |
| **Nama:**  **Dimas Dwi Saputtra**  **NIM: 064002200041** | C:\Users\RPL-SI 02\Pictures\288px-Trisakti_Logo.svg.png | **MODUL 6**  **Nama Dosen:**  **Dedy Sugiarto** |
| **Hari/Tanggal:**  **Kamis, 10, agustus 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

|  |
| --- |
| # Membuat histogram pengeluaran  ggplot(data\_dimasmarket, aes(x = AmountSpent)) +  geom\_histogram(binwidth = 100, fill = "blue", color = "black") +  labs(title = "Histogram Pengeluaran", x = "Jumlah Pengeluaran", y = "Frekuensi") |

Output

|  |
| --- |
| Histogram    Density |

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

Script

|  |
| --- |
| # Membaca data dari teks  > data\_text <- "Gender OwnHome Married Location Salary Children History Catalogs AmountSpent  + Female Own Single Far 47500 0 High 6 755  + Male Rent Single Close 63600 0 High 6 1318  + Female Rent Single Close 13500 0 Low 18 296  + Male Own Married Close 85600 1 High 18 2436  + Female Own Single Close 68400 0 High 12 1304  + Male Own Married Close 30400 0 Low 6 495  + Female Rent Single Close 48100 0 Medium 12 782  + Male Own Single Close 68400 0 High 18 1155  + Female Own Married Close 51900 3 Low 6 158  + Male Own Married Far 80700 0 NA 18 3034  + Male Rent Married Close 43700 1 NA 12 927  + Male Own Married Far 111800 3 High 18 2065"  >  > # Membaca data dari teks ke dalam dataframe  > data\_houseprices <- read.table(text = data\_text, header = TRUE)  >  > # Membuat scatter plot Salary vs. Location (asumsi Salary adalah variabel numerik)  > ggplot(data\_houseprices, aes(x = Salary, y = Location)) +  + geom\_point() +  + labs(title = "Scatter Plot Salary vs. Location", x = "Gaji", y = "Lokasi") |

Output

|  |
| --- |
|  |

1. **File Praktikum**

Github Repository:

|  |
| --- |
|  |

1. **Soal Latihan**

Soal:

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

Jawaban:  
1. Box plot atau diagram kotak adalah grafik yang menampilkan distribusi data numerik dalam bentuk ringkasan statistik, termasuk kuartil, median, dan pencilan (outliers).  
2. Scatter plot digunakan untuk menampilkan hubungan antara dua variabel numerik. Masing-masing titik pada plot mewakili pasangan nilai dari dua variabel tersebut.

Fungsi:

Pengenalan Hubungan: Scatter plot membantu melihat apakah ada hubungan antara dua variabel. Misalnya, apakah ada korelasi positif atau negatif.

Visualisasi Pencar: Scatter plot membantu mengidentifikasi pola atau cluster dalam data yang mungkin tidak terlihat dalam tabel angka.

1. **Kesimpulan**

Dalam analisis data dan visualisasi, kami telah melihat beberapa jenis plot dan bagaimana mereka membantu memahami dan menganalisis data.

1. **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 | … Menit | menarik |
| **2.** | Latihan Kedua | … Menit | menarik |
| **3.** | Latihan Ketiga | … Menit | menarik |
| **4.** | Latihan Keempat | … Menit | menarik |

Keterangan:

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