|  |  |  |
| --- | --- | --- |
| **Nama :** Ibnu Fajar Setiawan  **NIM :** 065002000006 | c9824-logouniversitastrisakti  **Praktikum Data Analitik** | **Modul 2**  **Nama Dosen:** **Syandra Sari, S.Kom, M.Kom** |
| **Hari/Tanggal** :  Kamis, 29 Oktober 2022 | **Nama Aslab :**   1. **Ida Jubaidah**   **(06500190037)**   1. **Azzahra Nuranisa (065001900044)** |

**Praktikum 2 – Data Analitik**  
**HUBUNGAN ANTAR DUA VARIABEL KATEGORIK DAN**

**ANTAR DUA VARIABEL NUMERIK**

**DESKRIPSI MODUL** : Melihat hubungan antara variabel Katagorik dengan Variabel Numerik.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Elemen Kompetensi | Indikator Kinerja | Jml  Jam | hlm |
| 1 | Mampu memahami cara menganalisis hubungan antar dua variable kategorik dan dua variable numerik | Dapat menganalisis hubungan antar dua variable kategorik dan dua variable numerik | 2 |  |

#### 

#### TEORI SINGKAT

## Dalam praktikum ini akan dipelajari dan dipraktekkan bagaimana mengetahui adanya hubungan antar variabel dalam sebuah data multivariat atau data yang berisi lebih dari dua variabel. Variabel sendiri dapat diartikan sebagai sebuah konsep yang memiliki variasi nilai. Hubungan (relationship) antar variable dapat terjadi antara sebuah variabel kategorik dengan sebuah variabel kontinu, antar dua variabel kategorik serta antar dua variabel kontinu.

#### LAB SETUP

Untuk dapat menjalankan praktikum ini maka yang harus disiapkan adalah :

1. Aplikasi RStudio
2. Xampp

**ELEMEN KOMPETENSI I**

**Deskripsi :** Dapat menganalisis hubungan antar dua variabel

**Kompetensi Dasar :** Mampu memahami cara menganalisis hubungan antar dua variable

**PRAKTIKUM**

Dalam praktikum ini akan dipelajari dan dipraktekkan bagaimana mengetahui adanya hubungan antar dua variabel kategorik serta antar dua variabel numerik. Hubungan antar dua variabel kategori dapat diketahui menggunakan teknik tabulasi silang atau cross tabulation serta uji statistiknya menggunakan uji khi-kuadrat atau uji kebebasan. Sedangkan hubungan antar dua variabel numerik dapat dilihat menggunakan Teknik scatter plot serta koefisien korelasi. Nilai koefisien korelasi berkisar dari -1 sd 1.

**Hubungan antar dua variabel kategorik**

|  |
| --- |
| > data()  > View(Titanic)  > cross <- xtabs(Freq ~ Class+Sex, data=Titanic)  > cross  > prop.table(cross, 1)  > chisq.test(cross)  **Output:** |

**Hubungan antar dua variabel numerik**

|  |
| --- |
| > head(iris)    > cor(iris$sepal.length, iris$sepal.width)  > cor(iris$petal.length, iris$petal.width)  > library(ggplot2)  > ggplot(iris, aes(x=petal.length, y=petal.width)) + geom\_point()  **Output:** |

**SCATTER PLOT**

**ELEMEN KOMPETENSI II**

**Tugas 1:**

**Deskripsi :** Dapat menganalisis hubungan antar dua variable kategorik dan dua variable numerik

**Kompetensi Dasar :** Mampu memahami cara menganalisis hubungan antar dua variable kategorik dan dua variable numerik dari data Houseprices dengan menggunakan crostab-uji chi square, koefisien korelasi dan scatter plot

**Data: House Price Data**

Tujuan utama analisis dalam tugas ini adalah melihat adanya hubungan antar dua variabel kategorik (misalnya bahan dinding dengan lokasi) serta hubungan antar dua variabel numerik (misalnya harga rumah dengan luas bangunan). Jawaban atas pertanyaan-pertanyaan ini bisa membantu  
calon pembeli memutuskan berapa besar tawaran untuk rumah.

Berikan penjelasan untuk setiap output yang dihasilkan serta tulis R code yang digunakan. Tampilan grafik yang berwarna tentunya lebih baik.

Catatan : Kalimat yang dipergunakan untuk membuat penjelasan setiap output antar praktikan tidak boleh sama

|  |
| --- |
| > library(RMySQL)  > con = dbConnect(MySQL(), user = 'root', password = '', dbname = 'houseprices', host = 'localhost')  > myQuery <- "select \* from tabel1;"  > df\_nama<- dbGetQuery(con, myQuery)  > View(df\_nama) |

**Output:**

Table

Description automatically generated

**Hubungan antar dua variabel kategorik**

|  |
| --- |
| > cross <- xtabs(Price ~ Brick+Neighborhood, data=df\_nama)  > cross  > prop.table(cross, 1)  > chisq.test(cross) |

**Output:**

**Text

Description automatically generatedA picture containing table

Description automatically generated**

**Hubungan antar dua variabel numerik**

|  |
| --- |
| > head(df\_nama)  > cor(df\_nama$Price, df\_nama$SqFt)  > cor(df\_nama$Price, df\_nama$Bedrooms)  > cor(df\_nama$Price, df\_nama$Bathrooms)  library(ggplot2)  ggplot(df\_nama, aes(x=SqFt, y=Price, color=Neighborhood)) + geom\_point() |

**Output:**

Graphical user interface, application, Word

Description automatically generated

Dari hasil grafik diatas dapat disimpulkan bahwa …

**Tugas 2:**

**Deskripsi :** Dapat menganalisis hubungan antar dua variable kategorik dan dua variable numerik

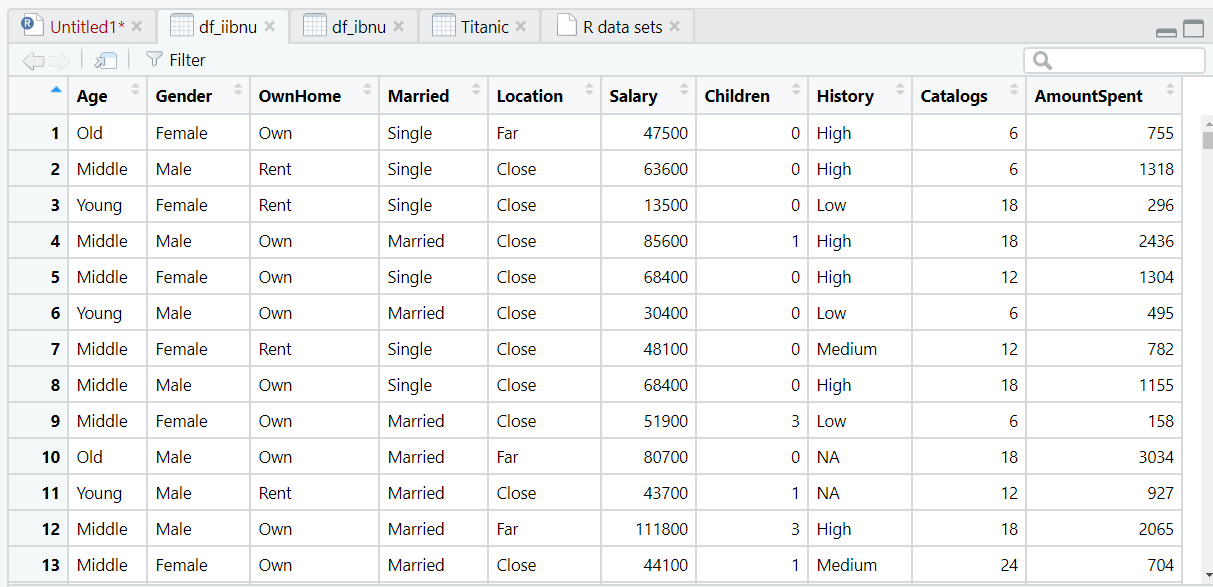
**Kompetensi Dasar :** Mampu memahami cara menganalisis hubungan antar dua variable kategorik dan dua variable numerik dari data Direct Marketing dengan menggunakan crostab-uji chi square, koefisien korelasi dan scatter plot

**Data : DirectMarketing.**

Lihat apakah di dalam datanya terdapat hubungan antar variable kategorik dan antar variable numerik. Buat tampilan grafiknya berwarna.

|  |
| --- |
| > library(RMySQL)  > con = dbConnect(MySQL(), user = 'root', password = '', dbname = 'directmarketing', host = 'localhost')  > myQuery <- "select \* from tabel1;"  > df\_nama <- dbGetQuery(con, myQuery)  > View(df\_nama) |

**Output:**



**Hubungan antar dua variabel kategorik**

|  |
| --- |
| > cross <- xtabs(Salary ~ OwnHome+Married, data=df\_nama)  > cross  > prop.table(cross, 1)  > chisq.test(cross) |

**Output:**

**Text, letter

Description automatically generated**

**Hubungan antar dua variabel numerik**

|  |
| --- |
| > head(df)  > cor(df\_nama$Salary, df\_nama$AmountSpent)  > cor(df\_nama$Salary, df\_nama$Catalogs)  > cor(df\_nama$Salary, df\_nama$Children)  > library(ggplot2)  > ggplot(df\_nama, aes(x=AmountSpent, y=Salary, color=Age)) + geom\_point() |

Dari hasil grafik diatas dapat disimpulkan bahwa age young banyak yg belum berpenghasilan/salary dan amount spend nya dominan 0-2000.

**Output:**

1. **Cek List**

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Elemen Kompetensi** | **Penyelesaian** | |
| **Selesai** | **Tidak** |
| **1** | **Elemen Kompetensi I**  **Memahami hubungan antar variabel katagorik dengan Numerik** |  |  |
| **2** | **Elemen Kompetensi II**  **TUGAS 1 :** Mampu memahami cara menganalisis hubungan antar dua variable kategorik dan dua variable numerik dari data Houseprices dengan menggunakan Scatter Plot |  |  |
| **3** | **Elemen Kompetensi II**  **TUGAS 2 :** Mampu memahami cara menganalisis hubungan antar dua variable kategorik dan dua variable numerik dari data Direct Marketing dengan menggunakan Scatter Plot |  |  |

1. **Form Umpan Balik**

|  |  |  |
| --- | --- | --- |
| **Elemen Kompetensi** | **Waktu Pengerjaan** | **Kriteria** |
| **Elemen Kompetensi I**  **Memahami hubungan antar variabel katagorik dengan Numerik** | 30 | 1 |
| **Elemen Kompetensi II**  **TUGAS 1 :** Mampu memahami cara menganalisis hubungan antar dua variable kategorik dan dua variable numerik dari data Houseprices dengan menggunakan Scatter Plot | 30 | 1 |
| **Elemen Kompetensi II**  **TUGAS 2 :** Mampu memahami cara menganalisis hubungan antar dua variable kategorik dan dua variable numerik dari data Direct Marketing dengan menggunakan Scatter Plot | 30 | 1 |

Kriteria

1.Sangat Menarik

2.Cukup Menarik

3.Kurang Menarik

4.Sangat Kurang Menarik