

# Mata Kuliah - Penggalian Data

**Nama Kelompok** : [dikosongi]

**Anggota** :

[202110370311306-Priya Yudha Prakasa]

[202110370311456-Aditya Afrian Aunur Chandra]

[202110370311442 -Ahmad Reva Elmiawan]

**Nilai Total: 120 poin**

## **Tahap 0 (poin: 25): Business Objective**

Menganalisis Kinerja Penjualan Maven Roasters Coffee Shop untuk Memahami Tren Penjualan, Hari Paling Sibuk, dan Produk yang Mendatangkan Pendapatan Terbesar.

## **Tahap 1 (poin: 25): Original Data**

- Urgensi dan Kasus yang Dipilih:
- Topik: Analisis Penjualan Maven Roasters Coffee Shop.
- Urgensi: Memahami tren penjualan, hari-hari paling sibuk, dan produk yang berkontribusi pada pendapatan.

Data yang Digunakan:

Dataset ini mencakup detail transaksi Maven Roasters Coffee Shop, termasuk tanggal, waktu, lokasi, dan detail produk.

Atribut Data:

### 1. Transaction Information:

transaction\_id: ID unik transaksi.

transaction\_date: Tanggal transaksi (MM/DD/YY).

transaction\_time: Waktu transaksi (HH:MM:SS).

transaction\_qty: Jumlah item yang terjual.

### 2. Store Information:

store\_id: ID unik toko kopi.

store\_location: Lokasi toko kopi.

### 3. Product Details:

product\_id: ID unik produk.

unit\_price: Harga eceran produk.

product\_category: Kategori produk.

product\_type: Tipe produk.

product\_detail: Detail produk.

- Sumber Data: <https://www.kaggle.com/datasets/ahmedabbas757/coffee-sales/data>

## **Tahap 2 (poin: 10): Target Data (Optional)**

Fitur yang Digunakan untuk Analisis:

- Transaction Information:

Tren Penjualan dari Waktu ke Waktu: Melalui analisis tanggal dan waktu transaksi, kita dapat mengidentifikasi tren penjualan harian, mingguan, bulanan, dan tahunan. Hal ini akan memberikan pemahaman yang lebih baik tentang pola pembelian pelanggan dari waktu ke waktu.

- Store Information:

Hari Paling Sibuk: Dengan memperhatikan jumlah transaksi per hari, kita dapat mengidentifikasi hari-hari dalam seminggu yang paling sibuk. Informasi ini penting untuk perencanaan operasional, manajemen stok, dan alokasi sumber daya.

- Product Details:

Produk Terlaris dan Tidak Laris: Analisis produk yang paling sering terjual dan jarang terjual akan membantu kita memahami preferensi pelanggan dan potensi peningkatan penjualan.

Produk dengan Pendapatan Tertinggi: Identifikasi produk yang memberikan kontribusi terbesar pada pendapatan bisnis akan membantu dalam mengoptimalkan strategi pemasaran dan penjualan.

## **Tahap 3-4 (poin: 25): Data Pre-processing & Transformation**

Beberapa teknik yang bisa digunakan yaitu (tentu sesuai kondisi dan kebutuhan):

- data cleaning (and/or data correcting)

- data integration
- data transformation
- data reduction
- data discretization
- data normalization
- feature selection

### **Tahap 5 (poin: 25): Data Mining**

- Algoritma data mining yang digunakan (sesuai data mining task).
- Skenario eksperimen sederhana.

### **Tahap 6 (poin: 20): Knowledge Interpretation**

- Pola-pola *useful* yang telah ditemukan.

### **Tahap 7 (poin: 15): Reporting**

- Simple academic Poster.
- Jupiter Notebook (Python)