**JOBSHEET 7**

**EXPLORATORY DATA ANALYSIS**

1. **TUJUAN**
2. Mahasiswa dapat menentukan pertanyaan bisnis
3. Mahasiswa dapat melakukan eksplorasi data untuk menjawab pertanyaan bisnis
4. Mahasiswa dapat menerapkan teknik pivot table untuk eksplorasi data
5. **PETUNJUK**

Awali setiap aktivitas dengan doa, semoga berkah dan mendapat kemudahan. Pahami tujuan, dasar teori, dan latihan-latihan praktikum dengan baik dan benar. Kerjakan tugas-tugas praktikum dengan baik, sabar, dan jujur. Tanyakan kepada asisten/dosen apabila ada hal-hal yang kurang jelas.

1. **DASAR TEORI**
2. **LATIHAN**

Setelah memahami semua teori tentang proses EDA, kini saatnya kita menerapkannya dalam sebuah proyek analisis data sederhana. Pada materi ini, kita akan melanjutkan proyek analisis data Adinda Collection**.**

Pertama, kita akan melanjutkan proyek analisis data Adinda Collection dengan menjalankan proses EDA. Proses ini memiliki tujuan antara lain

1. Membuat pertanyaan analisis atau bisnis yang ingin dicari jawabannya; dan
2. melakukan eksplorasi terhadap setiap data untuk mencari insight menarik guna menjawab pertanyaan bisnis tersebut.

**Alur Latihan**

Berikut tahapan dalam latihan ini.

* Tahap persiapan.
* Tahap penentuan pertanyaan bisnis untuk explorasi data.
* Tahap eksplorasi data.

**Persiapan**

Sebelum mulai mengerjakan latihan ini, terdapat beberapa hal yang harus disiapkan terlebih dahulu.

* Pertama, siapkanlah Google Colab atau Jupyter Notebook yang sebelumnya digunakan pada materi Latihan Data Wrangling.
* Pastikan Anda menjalankan kembali seluruh kode yang ada pada notebook tersebut. File Jupyter Notebook pada Materi Data Wrangling dapat diunduh [disini](https://drive.google.com/file/d/1bh1WJ45lpxeHmND2p4-HcwxrO6IiSMo2/view?usp=sharing).
* Kerjakan bagian Exploratory Data Analysis.

1. **TUGAS PRAKTIKUM**

Pada jobsheet kalo ini, anda harus melakukan analisis menggunakan salah satu dari beberapa dataset berikut.

* [Bike Sharing Dataset](https://drive.google.com/file/d/1RaBmV6Q6FYWU4HWZs80Suqd7KQC34diQ/view?usp=sharing) ([Sumber](https://www.kaggle.com/datasets/lakshmi25npathi/bike-sharing-dataset))
* [E-Commerce Public Dataset](https://drive.google.com/file/d/1MsAjPM7oKtVfJL_wRp1qmCajtSG1mdcK/view?usp=sharing) ([Sumber](https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce))
* [Air Quality Dataset](https://drive.google.com/file/d/1RhU3gJlkteaAQfyn9XOVAz7a5o1-etgr/view?usp=share_link) ([Sumber](https://github.com/marceloreis/HTI/tree/master))

Mirip seperti berbagai materi latihan yang telah dibahas sebelumnya, Anda harus melakukan seluruh proses analisis data mulai dari mendefinisikan pertanyaan hingga membuat kesimpulan dari hasil analisis. Selain itu, proyek analisis data yang Anda buat harus memenuhi ketentuan berikut.

* Minimal terdapat 2 buah pertanyaan bisnis (pertanyaan analisis) yang ingin dijawab melalui proses analisis data. Pertanyaan tersebut haruslah efektif sesuai dengan materi Mendefinisikan Pertanyaan Untuk Explorasi Data.
* Minimal terdapat 2 buah visualisasi data untuk menjawab pertanyaan bisnis yang telah dibuat.

Kerjakan proyek analisis data menggunakan template proyek yang telah disediakan. Template yang dimaksud dapat diakses pada tautan berikut: [template notebook](https://drive.google.com/file/d/1sXfepCPHhfW1KbwxIEspSN8vFsG0waDd/view?usp=sharing).