

Tugas Praktikum: Eksplorasi Data

Deskripsi Tugas

Pada praktikum ini, mahasiswa akan melakukan eksplorasi data untuk memahami karakteristik dataset yang diberikan. Mahasiswa akan mengidentifikasi atribut, mendeteksi nilai yang hilang, menganalisis distribusi data, serta memvisualisasikan hubungan antar atribut.

⚠ Catatan Penting:

Terdapat **4 dataset berbeda** yang digunakan pada praktikum ini. Masing-masing kelompok Asisten Dosen (Asdos) akan menggunakan dataset yang berbeda sehingga soal eksplorasi dan visualisasi juga berbeda.

Tujuan Praktikum

1. Memahami langkah-langkah eksplorasi data.
2. Mengidentifikasi atribut dalam dataset (nominal, ordinal, numerik).
3. Menganalisis nilai yang hilang dan outlier.
4. Membuat visualisasi data untuk mendukung analisis.

Langkah-Langkah Praktikum

1. **Unduh Dataset**
Setiap kelompok akan mendapatkan link dataset yang berbeda sesuai arahan Asdos.
2. **Import dan Pemahaman Dataset**
 - Impor dataset menggunakan Python (pandas).
 - Tampilkan 5 baris pertama dataset untuk memahami struktur data.
 - Identifikasi tipe atribut: nominal, ordinal, numerik.
3. **Analisis Statistik**
 - Hitung statistik deskriptif (rata-rata, median, modus, standar deviasi) untuk atribut numerik.
 - Identifikasi jumlah nilai yang hilang pada setiap atribut.
4. **Deteksi Outlier**
 - Gunakan visualisasi seperti boxplot untuk mendeteksi outlier pada atribut numerik.
5. **Visualisasi Data (Disesuaikan dengan Dataset Asdos)**
 - a. Asdos Lucky – [Car Datasets](#)
 - Histogram: distribusi Total Speed (km/h).
 - Scatter Plot: hubungan antara HorsePower (Hp) dan Cars Prices.
 - Bar Chart: jumlah mobil berdasarkan Fuel Types.
 - b. Asdos Richard - [Spotify Dataset](#)
 - Histogram: distribusi songs_played_per_day.
 - Scatter Plot: hubungan antara age dan listening_time.
 - Bar Chart: jumlah user berdasarkan subscription_type.
 - c. Asdos Willy - [Global Coffee Health Dataset](#)
 - Histogram: distribusi heart_rate
 - Scatter Plot: hubungan antara coffee_intake dan sleep_hours
 - Bar Chart: jumlah user berdasarkan sleep_quality
 - d. Asdos Taib - [Flight Prices Dataset](#)
 - Histogram: distribusi durasi penerbangan (`duration`)

- Scatter Plot: hubungan antara h-keberangkatan (`days_left`) dengan harga tiket (`price`)
- Bar Chart: harga tiket rata-rata (`prices`) berdasarkan waktu keberangkatan (`departure_time`)

6. Laporan Analisis (Disesuaikan dengan Dataset Asdos)

Mahasiswa diminta menulis laporan singkat (1–2 halaman) berisi:

- Deskripsi Dataset: ringkasan atribut dan karakteristik dataset sesuai kelompok.
- Statistik Deskriptif: hasil perhitungan atribut numerik yang relevan.
- Penanganan Data: identifikasi nilai hilang/outlier dan langkah yang diambil.
- Visualisasi & Interpretasi: masukkan grafik yang sesuai dataset (lihat poin 5) dan tuliskan analisisnya.

Setiap kelompok hanya perlu melaporkan dataset sesuai Asdos pembimbingnya, tidak perlu mengerjakan dataset kelompok lain.

Hasil yang Diharapkan

1. Statistik deskriptif untuk atribut numerik.
2. Visualisasi data sesuai dataset masing-masing (histogram, scatter plot, diagram batang).
3. Penanganan nilai yang hilang (jika ada) dan identifikasi outlier.
4. Laporan analisis singkat (1–2 halaman).

Penilaian

1. Pemahaman Dataset (20%) – Mahasiswa mampu mengidentifikasi atribut dengan benar.
2. Analisis Statistik (30%) – Perhitungan statistik deskriptif dan identifikasi nilai hilang.
3. Visualisasi Data (30%) – Kualitas visualisasi sesuai dataset dan interpretasinya.
4. Laporan (20%) – Kerapihan, kelengkapan, kejelasan laporan.