

# Praktikum Statistika

## Latihan 1:

Pada praktikum kali ini, kita akan mempelajari Statistika Deskriptif, yaitu bagian dari statistika yang digunakan untuk mengolah, menyajikan, dan menggambarkan data agar lebih mudah dipahami. Fokus utama praktikum mencakup analisis distribusi frekuensi, ukuran gejala pusat, ukuran letak, serta ukuran simpangan dan dispersi. Melalui kegiatan ini, mahasiswa diharapkan mampu memahami bagaimana data dikumpulkan, ditampilkan dalam bentuk tabel maupun grafik, serta dianalisis menggunakan ukuran-ukuran statistik yang sesuai. Dengan pemahaman tersebut, mahasiswa dapat menarik kesimpulan sederhana dari data, sekaligus melatih keterampilan interpretasi dalam kasus nyata

## Langkah-Langkah Analisis Statistika Deskriptif

### 0. Upload Dataset

```
from google.colab import drive
drive.mount('/content/drive')

import pandas as pd

path = "/content/drive/MyDrive/STATISTIKA/dataset.csv"
df = pd.read_csv(path)

df.head()
```

#### dataset :

[https://drive.google.com/file/d/1SGMM\\_QVFKbfQ\\_Hbl7wJaXoB2SCOLWTOh/view?usp=sharing](https://drive.google.com/file/d/1SGMM_QVFKbfQ_Hbl7wJaXoB2SCOLWTOh/view?usp=sharing)

### 1. Eksplorasi Data Awal

```
#Eksplorasi Data Awal
df.head()
df.info()
df.describe()
```

Tujuan: mengetahui bentuk data, jumlah observasi, tipe variabel (numerik/kategorikal), dan ringkasan awal.

### 2. Distribusi Frekuensi & Grafik

```
#Distribusi Frekuensi & Grafik
df['final_score'].hist(bins=10)
plt.title("Histogram Nilai Akhir")
plt.xlabel("Nilai Akhir"); plt.ylabel("Frekuensi")
plt.show()
```

Tujuan: melihat pola distribusi nilai (apakah merata, condong tinggi, condong rendah).

### 3. Ukuran Gejala Pusat

```
#Ukuran Gejala Pusat
mean_val = df['final_score'].mean()
median_val = df['final_score'].median()
mode_val = df['final_score'].mode()[0]

print("Mean:", mean_val)
print("Median:", median_val)
print("Mode:", mode_val)
```

Tujuan: mengetahui nilai rata-rata, tengah, dan yang sering muncul.

### 4. Ukuran Letak Data

```
#Ukuran Letak Data
Q1 = df['final_score'].quantile(0.25)
Q2 = df['final_score'].quantile(0.5) # median
Q3 = df['final_score'].quantile(0.75)
IQR = Q3 - Q1

print("Q1:", Q1, "Q2:", Q2, "Q3:", Q3, "IQR:", IQR)
```

Tujuan: memahami penyebaran data di kuartil bawah, tengah, dan atas.

### 5. Ukuran Simpangan & Dispersi

```
#Ukuran Simpangan & Dispersi
var_val = df['final_score'].var()
std_val = df['final_score'].std()
range_val = df['final_score'].max() - df['final_score'].min()

print("Variansi:", var_val)
print("Standar Deviasi:", std_val)
print("Rentang:", range_val)
```

Tujuan: mengetahui seberapa besar keragaman/penyebaran nilai mahasiswa.

### 6. Bentuk Distribusi

```
#Bentuk Distribusi
print("Skewness :", df['final_score'].skew())
print("Kurtosis :", df['final_score'].kurt())
```

Tujuan: melihat apakah distribusi miring ke kiri/kanan (skewness) atau runcing/datar (kurtosis).

### 7. Visualisasi Tambahan

*Boxplot per kelas → melihat median dan sebaran per kelompok.*

```
#Visualisasi Tambahan
sns.boxplot(x="class", y="final_score", data=df)
plt.title("Boxplot Nilai Akhir per Kelas")
```

```
plt.show()
```

*Scatter plot → melihat hubungan jam belajar dengan nilai.*

```
sns.regplot(x="hours_study_per_week", y="final_score", data=df)
plt.title("Jam Belajar vs Nilai Akhir")
plt.show()
```

## Kesimpulan dari Analisis Sederhana

1. Distribusi nilai akhir menunjukkan mayoritas mahasiswa mendapat nilai pada kisaran menengah hingga tinggi, dengan sedikit nilai rendah. Histogram memperlihatkan distribusi cenderung normal dengan puncak di sekitar 75–85.
2. Nilai rata-rata (mean) berada di sekitar 78, median mendekati 80, dan modus juga di kisaran tersebut, menunjukkan data cukup simetris.
3. Rentang nilai cukup lebar (misalnya 35–100), tetapi interquartile range (IQR) lebih sempit, artinya sebagian besar mahasiswa berkumpul di tengah.
4. Standar deviasi relatif moderat, yang berarti variasi antar mahasiswa ada, tetapi tidak terlalu ekstrem.
5. Boxplot per kelas menunjukkan perbedaan sebaran kecil antar kelas, dengan median yang hampir sama, menandakan performa kelas cukup merata.
6. Scatter plot jam belajar vs nilai akhir memperlihatkan kecenderungan positif: semakin banyak jam belajar, semakin tinggi nilai akhir, meskipun ada beberapa outlier.

### Kesimpulan umum:

Mahasiswa dalam dataset ini cenderung memiliki nilai akhir yang merata dengan penyebaran moderat. Tidak ada perbedaan besar antar kelas, tetapi faktor jam belajar tetap berpengaruh terhadap pencapaian nilai.

## Latihan 2:

Dataset: stock\_dataset.csv (berisi data 10 perusahaan, 60 hari perdagangan terakhir).

[https://drive.google.com/file/d/1LR8lFHheqH74O4L3hpt\\_0Pi9m5FUR6la/view?usp=sharing](https://drive.google.com/file/d/1LR8lFHheqH74O4L3hpt_0Pi9m5FUR6la/view?usp=sharing)

Lakukan analisis statistik deskriptif dari dataset diatas, dan beri kesimpulan berdasarkan analisis yang telah dilakukan !

Di kumpulkan di class room, hasil run dan analisis kesimpulannya, berupa file **pdf** dan **.ipynb**