Tugas 2: Praktikum & Praktikum Mandiri 2

**Pandu Linggar Kumara - 0110221277,   
Link GitHub - https://github.com/PanduLgg/M\_Learning.git**

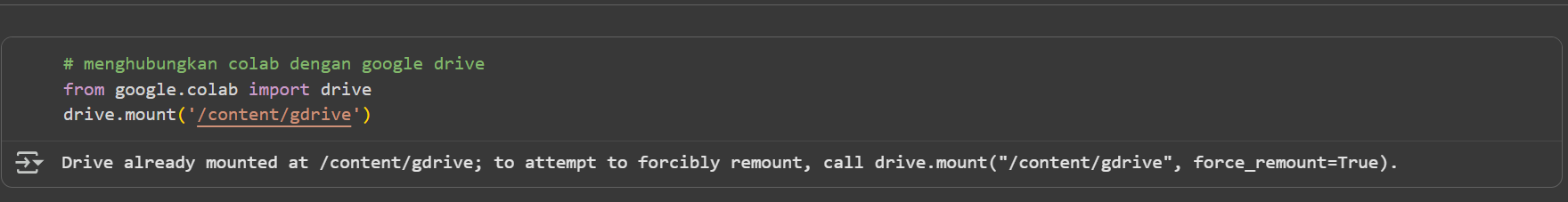
1 Teknik Informatika, STT Terpadu Nurul Fikri, Depok

\*E-mail: [pandulinggar1@gmail.com](mailto:pandulinggar1@gmail.com)

**Abstract.** Kegiatan praktikum mencakup pembacaan dataset, perhitungan nilai-nilai statistik dasar (seperti mean, median, modus, variansi, standar deviasi, dan kuartil), serta analisis korelasi antarvariabel. Selain itu, dilakukan pula visualisasi data menggunakan boxplot, histogram, dan scatter plot untuk memahami distribusi serta hubungan antar variabel.

1. Connecting Google Colab & Drive

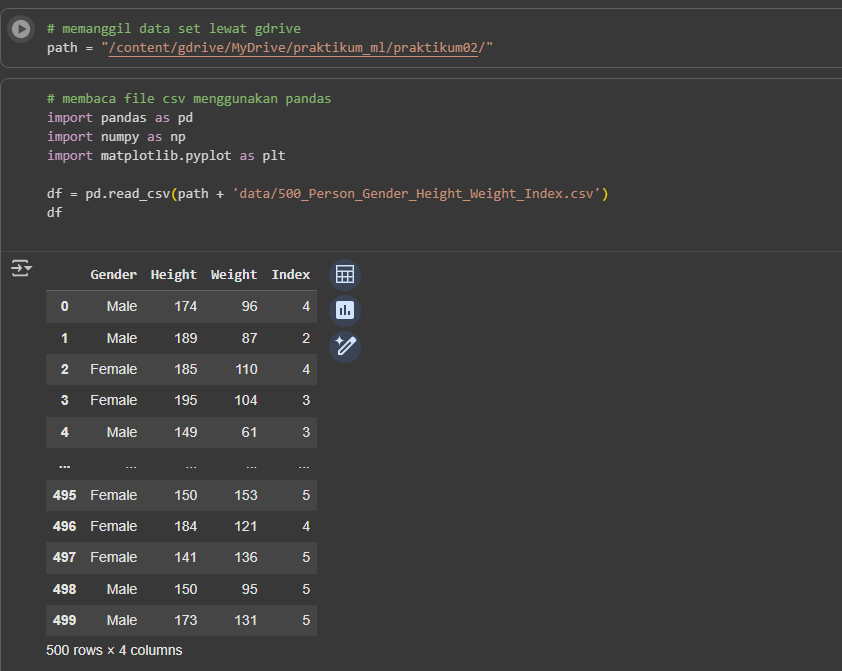
1.1 Menghubungkan lingkungan Google Colab dengan akun Google Drive

Sel ini berfungsi untuk menghubungkan lingkungan Google Colab dengan akun Google Drive

**Gambar 1.1.** Proses ini hanya perlu dilakukan satu kali per sesi.

1.2 Memanggil Data set dari Gdrive dan Membaca file .CSV menggunakan Pandas

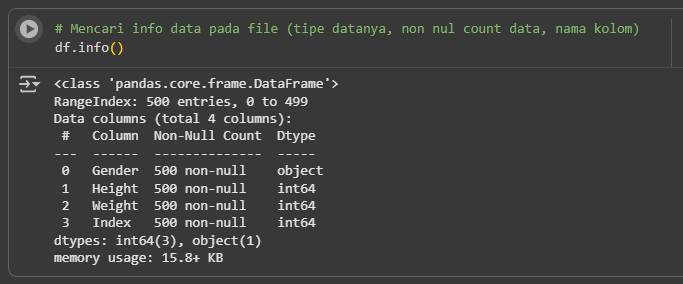
Sel ini menggunakan library Pandas untuk membaca file data, yang diinginkan yaitu ‘data/500\_Person\_Gender\_Height\_Weight\_Index.csv’.



**Gambar 1.2.** Proses ini hanya perlu dilakukan satu kali per sesi.

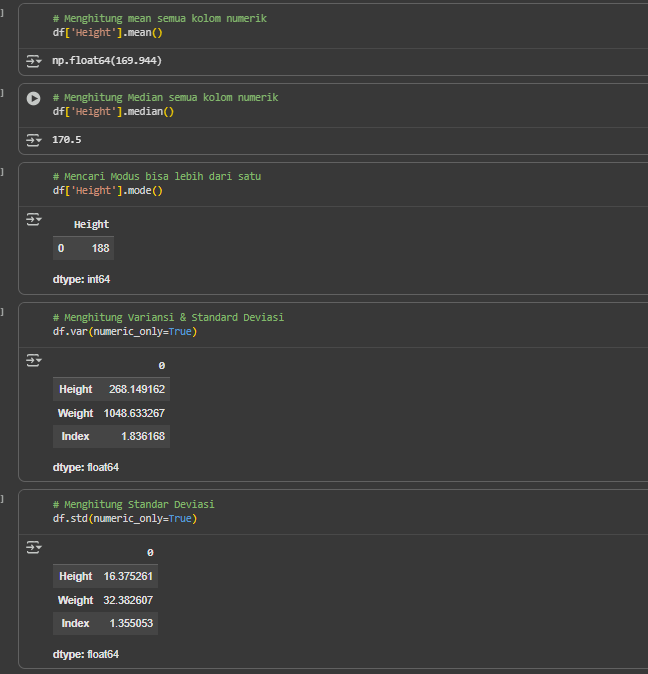
1.3 Mencari informasi data yang ada pada file

Sel ini menampilkan informasi yang ada di dalam file dari mulai tipe data nama kolom, dsb.



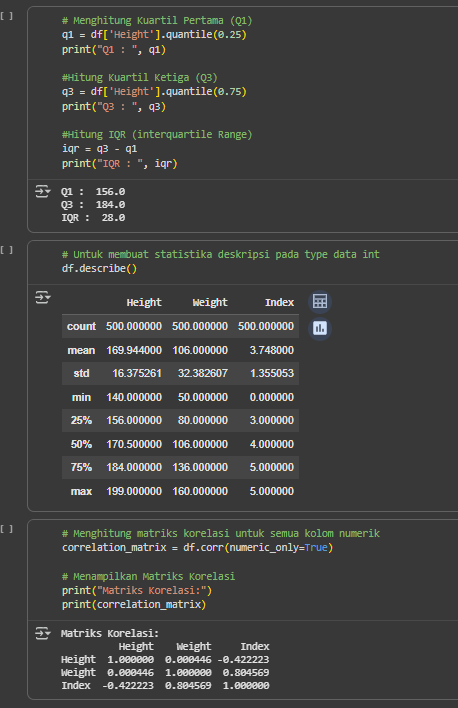
**Gambar 1.3.** Mencari info data pada file

1.4 Menghitung Ukuran Tendensi Sentral dan Ukuran Penyebaran Data   
Sel ini menghitung mean, median, modus, variansi, dan standard deviasi pada file data



**Gambar 1.4.** Menghitung Ukuran Tendensi Sentral dan Ukuran Penyebaran Data

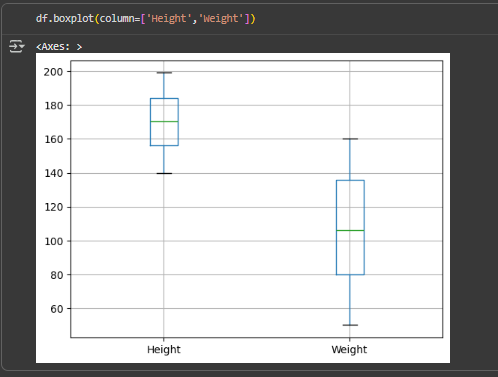
1.5 Menghitung Qartil dan Membuat Deskripsi pada Data   
Sel ini menghitung q1, q2, dan IQR, lalu membuat deskripsi pada type data Integer, lalu menghitung matriks korelasi dalam semua kolom numerik



**Gambar 1.5.** Menghitung Qartil dan Membuat Deskripsi pada Data

2. Visualisasi Data  
Memvisualisasikan Data menggunakan matplotlib, dengan metode Box Plot, Histogram, Scatter Plot

2.1 Visualisasi Data menggunakan Box Plot   
Sel ini menggunakan fungsi df.boxplot() mengambil data Tinggi dan Berat Badan



**Gambar 2.1.** Menampilkan Box Plot

2.2 Visualisasi Data menggunakan Histogram

*Sel ini membuat histogram untuk kolom 'Height'. Histogram adalah grafik yang menunjukkan*

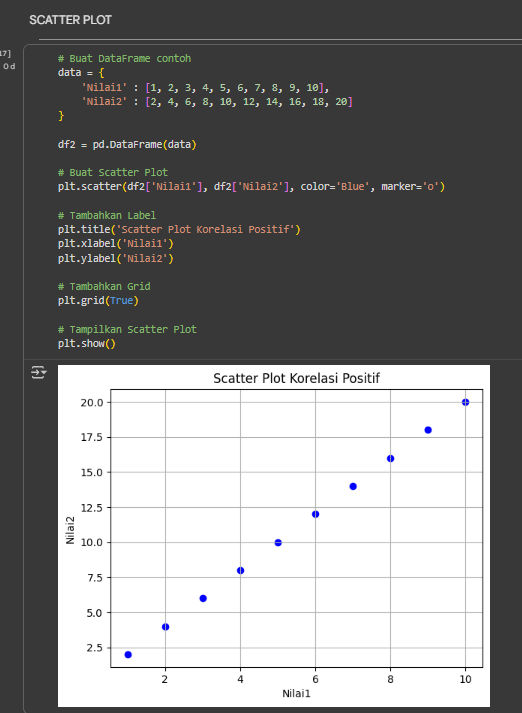
*distribusi frekuensi data*



**Gambar 2.2.** Menampilkan Histogram

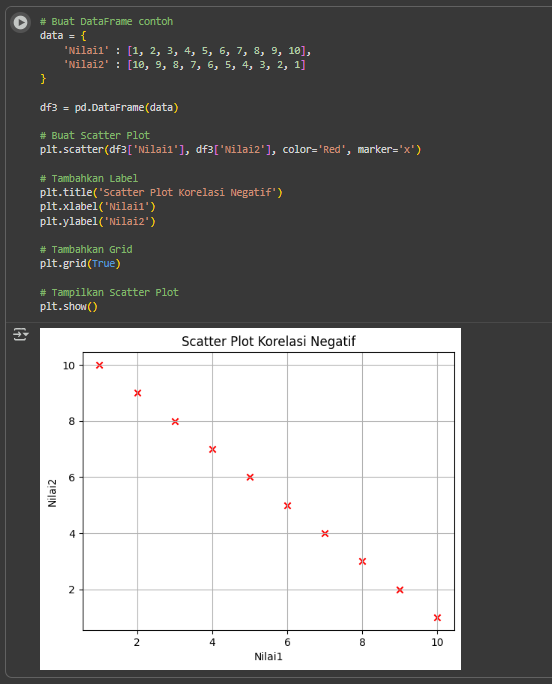
2.3 Visualisasi Data menggunakan Scatter Plot

*Sel ini membuat scatter plot atau diagram pencar untuk memvisualisasikan hubungan antara dua variabel numerik. Berisi dua blok kode untuk membuat dua scatter plot terpisah. Kode pertama menggunakan data Nilai1 dan Nilai2 yang memiliki korelasi positif*



**Gambar 2.3.** Menampilkan Scatter Plot Positif

sedangkan kode kedua memiliki korelasi negatif.



**Gambar 2.4.** Menampilkan Scatter Plot Negatif

Hasil:

1. Plot Korelasi Positif: Titik-titik data membentuk pola yang naik ke atas. Ini menunjukkan bahwa ketika Nilai1 meningkat, Nilai2 juga cenderung meningkat.
2. Plot Korelasi Negatif: Titik-titik data membentuk pola yang menurun. Ini menunjukkan bahwa ketika Nilai1 meningkat, Nilai2 justru cenderung menurun.

3. Praktikum Mandiri  
Buat program untuk membagi dataset day.csv menjadi tiga bagian, yaitu:

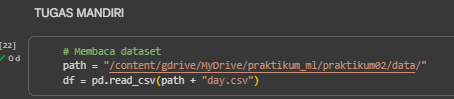
(a) Data Training: 80% dari total dataset

(b) Data Validation: 10% dari data training

(c) Data Testing: 20% dari total dataset

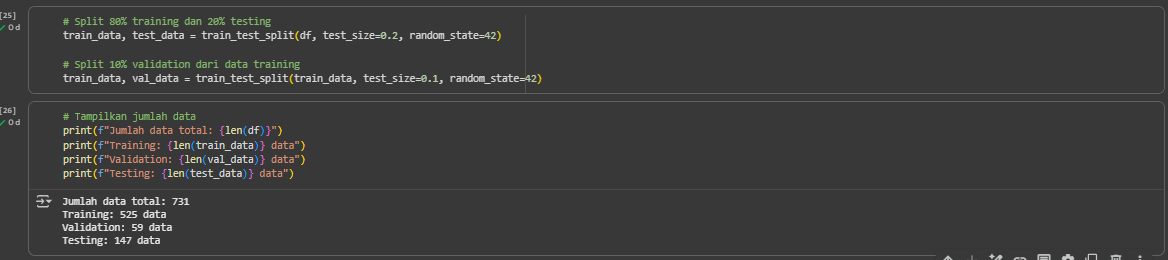
3.1 Membaca Dataset File

*Sel ini membaca dataset yang diinginkan yaitu day.csv*

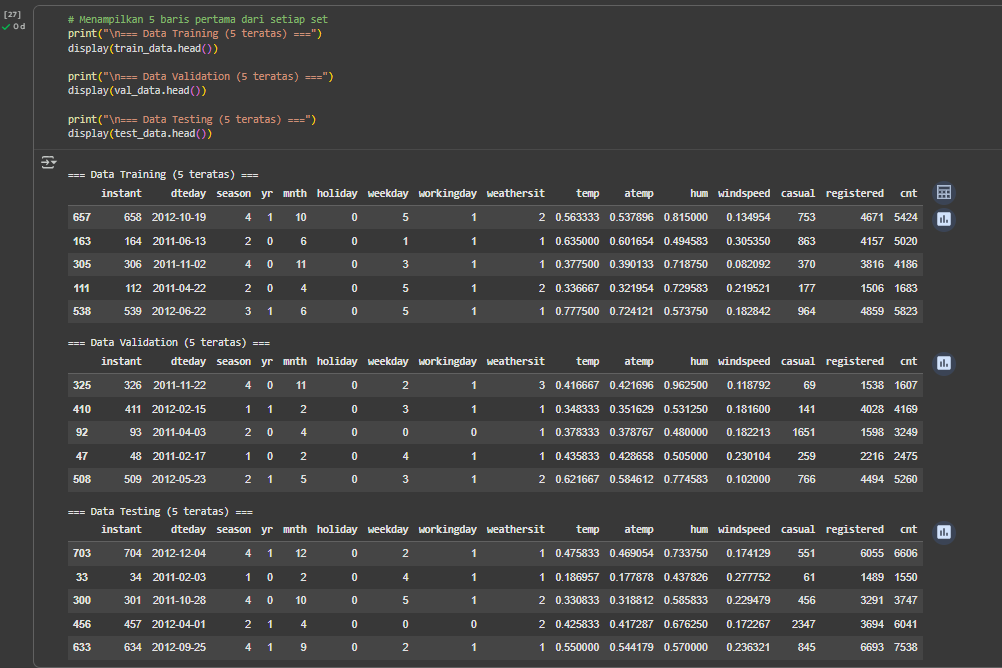


**Gambar 3.1.** Proses ini hanya perlu dilakukan satu kali per sesi.

3.2 Membaca Dataset File   
Sel ini membagi data Training 80%, Data Testing 20%, dan Data validation 10%, lalu menampilkan jumlah datanya

  
**Gambar 3.2.** Menampilkan Jumlah Data hasil pembagian persentase

3.3 Menampilkan 5 baris perta pada setiap Set   
Sel ini menampilkan hasil dari 5 data teratas dari Data Training, Data Validation, dan Data Testing



**Gambar 3.3.** Menampilkan 5 Baris Data Teratas setiap Set

Referensi:

Munir, S., Seminar, K. B., Sudradjat, Sukoco, H., & Buono, A. (2022). The Use of Random Forest Regression for Estimating Leaf Nitrogen Content of Oil Palm Based on Sentinel 1-A Imagery. *Information*, *14*(1), 10. https://doi.org/10.3390/info14010010

Seminar, K. B., Imantho, H., Sudradjat, Yahya, S., Munir, S., Kaliana, I., Mei Haryadi, F., Noor Baroroh, A., Supriyanto, Handoyo, G. C., Kurnia Wijayanto, A., Ijang Wahyudin, C., Liyantono, Budiman, R., Bakir Pasaman, A., Rusiawan, D., & Sulastri. (2024). PreciPalm: An Intelligent System for Calculating Macronutrient Status and Fertilizer Recommendations for Oil Palm on Mineral Soils Based on a Precision Agriculture Approach. *Scientific World Journal*, *2024*(1). https://doi.org/10.1155/2024/1788726