

MACHINE LEARNING | PAGI

Nama : Sintia Sari
NIM : 0110222135
Prodi : Teknik Informatika
Rombel : TI08 - 2022

Tugas Latihan Mandiri:

1. Memuat, membaca dan menampilkan 5 baris data dalam dataset

Kode :

```
# Import library
import pandas as pd

# Memanggil file di dalam Google Drive
from google.colab import drive
drive.mount('/content/gdrive')
import os

# Inisiasi directory tempat data akan dibaca dan disimpan
path = "gdrive/MyDrive/Machine Learning/Praktikum/Praktikum01/Data/"

# Menampilkan 5 baris data dalam dataset
df_h = pd.read_csv(path + 'hour.csv')
df_h.head()
```

Hasil :

	instant	dteday	season	yr	mnth	hr	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
0	1	2011-01-01	1	0	1	0	0	6	0	1	0.24	0.2879	0.81	0.0	3	13	16
1	2	2011-01-01	1	0	1	1	0	6	0	1	0.22	0.2727	0.80	0.0	8	32	40
2	3	2011-01-01	1	0	1	2	0	6	0	1	0.22	0.2727	0.80	0.0	5	27	32
3	4	2011-01-01	1	0	1	3	0	6	0	1	0.24	0.2879	0.75	0.0	3	10	13
4	5	2011-01-01	1	0	1	4	0	6	0	1	0.24	0.2879	0.75	0.0	0	1	1

Penjelasan Kode :

- import pandas as pd
Mengimpor library pandas untuk mengolah data dalam bentuk tabel.
- from google.colab import drive & drive.mount('/content/gdrive')
Menghubungkan Google Colab dengan Google Drive supaya file dataset bisa diakses langsung dari Drive.
- import os
Digunakan untuk mengatur direktori/file di sistem (opsional, untuk navigasi folder).
- path = "gdrive/MyDrive/Machine Learning/Praktikum/Praktikum01/Data/"

Menyimpan alamat folder tempat file dataset berada.

- `df_h = pd.read_csv(path + 'hour.csv')`

Membaca file `hour.csv` yang ada di folder tersebut dan menyimpannya ke dalam variabel `df` (dataframe), `df_h` merupakan dataframe yang digunakan untuk menyimpan dataset `hour.csv` karena `df` sebelumnya sudah dipakai untuk menyimpan dataset `day.csv` (dikerjakan dalam satu file google collab yang sama).

- `df_h.head()`

Menampilkan 5 baris pertama dari dataset supaya kita bisa melihat isi datanya.

Penjelasan Output :

Tabel yang muncul menampilkan 5 baris pertama dari dataset Bike Sharing dengan beberapa kolom, antara kolom `instant`, `dteday`, `season`, `yr`, `mnth`, `hr`, `holiday`, `weekday`, `workingday`, `weathersit`, `temp`, `atemp`, `hum`, `windspeed`, `casual`, `registered`, dan `cnt`.

2. Buat dataframe baru (misal: `df1`) dari dataframe sebelumnya, dengan data yang diambil adalah `instant`, `dteday`, `season`, `temp`, `hum`, `windspeed`

Kode :

```
# Buat dataframe baru dengan kolom yang dipilih
df1 = df_h[['instant', 'dteday', 'season', 'temp', 'hum', 'windspeed']]
```

Penjelasan Kode :

- `df1 = df_h[['instant', 'dteday', 'season', 'temp', 'hum', 'windspeed']]`

Dari dataset utama `df_h`, dibuat dataframe baru bernama `df1` yang hanya berisi beberapa kolom saja, yaitu:

Nama Kolom	Keterangan
<code>instant</code>	nomor urut data
<code>dteday</code>	tanggal
<code>season</code>	musim
<code>temp</code>	suhu
<code>hum</code>	kelembaban
<code>windspeed</code>	kecepatan angin

3. Cetak dataframe baru kelayar

Kode :

```
# Tampilkan 5 data teratas  
df1.head()
```

Hasil :

	instant	dteday	season	temp	hum	windspeed
0	1	2011-01-01	1	0.24	0.81	0.0
1	2	2011-01-01	1	0.22	0.80	0.0
2	3	2011-01-01	1	0.22	0.80	0.0
3	4	2011-01-01	1	0.24	0.75	0.0
4	5	2011-01-01	1	0.24	0.75	0.0

Penjelasan Kode :

- `df1.head()`

Menampilkan 5 baris pertama dari dataframe `df1` untuk memastikan kolom yang dipilih sudah benar.

Penjelasan Hasil :

Tabel hasil menampilkan data dengan 6 kolom sesuai arahan point 2.

4. Simpan dataframe ke file csv baru ke direktori data dengan nama latihan01.csv

Kode :

```
# Menentukan direktori tujuan  
save_path = "/content/gdrive/MyDrive/Machine Learning/Praktikum/Praktikum01/Data/latihan01.csv"  
  
# Menyimpan dataframe ke CSV  
df1.to_csv(save_path, index=False)  
  
print(f"Dataframe berhasil disimpan di: {save_path}")
```

Hasil :

Dataframe berhasil disimpan di: /content/gdrive/MyDrive/Machine Learning/Praktikum/Praktikum01/Data/latihan01.csv

...

> Praktikum01

> Data

▼

✓

☰

ⓘ

Type ▼

People ▼

Modified ▼

Source ▼

Name	Owner	Last m... ▼	↓	File size	⋮
<div>📄</div> <div>latihan01.csv</div>	<div>👤</div> <div>me</div>	11:07		587 KB	⋮
<div>📖</div> <div>Readme.txt</div>	<div>👤</div> <div>me</div>	18 Sept 2025		5 KB	⋮

Penjelasan Kode :

- `save_path = "/content/gdrive/MyDrive/Machine Learning/Praktikum/Praktikum01/Data/latihan01.csv"`
Menentukan alamat folder dan nama file tempat dataframe akan disimpan. Di sini file akan disimpan dengan nama latihan01.csv di Google Drive.
- `df1.to_csv(save_path, index=False)`
Menyimpan dataframe df1 ke dalam file CSV. `index=False` artinya indeks baris (0,1,2,...) tidak ikut disimpan ke file, jadi datanya lebih bersih.
- `print(f'Dataframe berhasil disimpan di: {save_path}')`
Menampilkan pesan konfirmasi jika file sudah berhasil disimpan, lengkap dengan lokasi penyimpanan.

Penjelasan Output :

file latihan01.csv sudah tersimpan di folder yang ditentukan di Google Drive.

Kesimpulan :

Dari hasil implementasi di atas dapat disimpulkan bahwa proses menampilkan data sudah berjalan dengan baik. Dataset berhasil dibaca langsung dari Google Drive agar lebih praktis dan aman, karena file tidak perlu diunggah berulang kali setiap kali membuka Google Colab. Setelah itu, data ditampilkan untuk memastikan isi file sesuai dan bisa digunakan. Selanjutnya, beberapa kolom penting dipilih untuk membuat dataframe baru yang lebih ringkas sehingga lebih mudah dianalisis. Dataframe tersebut kemudian berhasil disimpan kembali ke dalam file CSV baru dengan nama latihan01.csv, sehingga dapat digunakan kembali tanpa perlu mengolah dari awal.