PRAKTIKUM 02

1. Membaca file day.csv
2. Mengambil 80% dari total dataset sebagai data training
3. Mengambil 20% dari total dataset sebagai data testing
4. Mengambil 10% dari data training sebagai data validation

```
import pandas as pd
from sklearn.model_selection import train_test_split

# Baca file day.csv
df2 = pd.read_csv(path + 'data/day.csv')

print("Jumlah total data:", len(df2))

# Bagi data menjadi Training (80%) dan Testing (20%)
train_df2, test_df2 = train_test_split(df2, test_size=0.2, random_state=42)

# Validation (10% dari Training)
train_df2, val_df2 = train_test_split(train_df2, test_size=0.1, random_state=42)

# Cek jumlah masing-masing
print("Training    :", len(train_df2))
print("Validation  :", len(val_df2))
print("Testing     :", len(test_df2))
```

```
Jumlah total data: 731
Training    : 525
Validation  : 59
Testing     : 147
```

5. Tampilkan jumlah data dan 5 baris data teratas untuk setiap set (Training, Validation, dan Testing)

Menggunakan Display() + markdown heading agara tampilan hasil menjadi rapih

```
from IPython.display import display, Markdown

display(Markdown("### 5 Baris Pertama Data Training"))
display(train_df.head())

display(Markdown("### 5 Baris Pertama Data Validation"))
display(val_df.head())

display(Markdown("### 5 Baris Pertama Data Testing"))
display(test_df.head())
```

a. Data Training

5 Baris Pertama Data Training

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
657	658	2012-10-19	4	1	10	0	5	1	2	0.563333	0.537896	0.815000	0.134954	753	4671	5424
163	164	2011-06-13	2	0	6	0	1	1	1	0.635000	0.601654	0.494583	0.305350	863	4157	5020
305	306	2011-11-02	4	0	11	0	3	1	1	0.377500	0.390133	0.718750	0.082092	370	3816	4186
111	112	2011-04-22	2	0	4	0	5	1	2	0.336667	0.321954	0.729583	0.219521	177	1506	1683
538	539	2012-06-22	3	1	6	0	5	1	1	0.777500	0.724121	0.573750	0.182842	964	4859	5823

b. Data Validation

5 Baris Pertama Data Validation

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
325	326	2011-11-22	4	0	11	0	2	1	3	0.416667	0.421696	0.962500	0.118792	69	1538	1607
410	411	2012-02-15	1	1	2	0	3	1	1	0.348333	0.351629	0.531250	0.181600	141	4028	4169
92	93	2011-04-03	2	0	4	0	0	0	1	0.378333	0.378767	0.480000	0.182213	1651	1598	3249
47	48	2011-02-17	1	0	2	0	4	1	1	0.435833	0.428658	0.505000	0.230104	259	2216	2475
508	509	2012-05-23	2	1	5	0	3	1	2	0.621667	0.584612	0.774583	0.102000	766	4494	5260

c. Data Testing

5 Baris Pertama Data Testing

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
703	704	2012-12-04	4	1	12	0	2	1	1	0.475833	0.469054	0.733750	0.174129	551	6055	6606
33	34	2011-02-03	1	0	2	0	4	1	1	0.186957	0.177878	0.437826	0.277752	61	1489	1550
300	301	2011-10-28	4	0	10	0	5	1	2	0.330833	0.318812	0.585833	0.229479	456	3291	3747
456	457	2012-04-01	2	1	4	0	0	0	2	0.425833	0.417287	0.676250	0.172267	2347	3694	6041
633	634	2012-09-25	4	1	9	0	2	1	1	0.550000	0.544179	0.570000	0.236321	845	6693	7538