

Ghibran Muhammad - 08

PROGRAM STUDI TEKNIK INFORMATIKA JURUSAN TEKNOLOGI INFORMASI



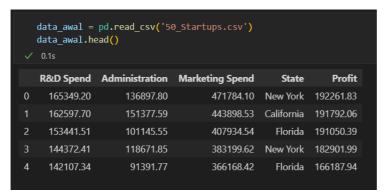
LAPORAN TUGAS MINGGU 3

1. Gambar di bawah ini merupakan pemanggilan library di python

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import OneHotEncoder
from sklearn.compose import ColumnTransformer
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

✓ 4.1s
```

2. library dari panda untuk membaca file "50_startups.csv"lalu ada kode program "data_awal.head()" digunakan untukmenampilkan 5 file teratas



3. dibawah ini merupakan proses dari OneHotEncoder

```
transformer_list = [
          ('encoded', OneHotEncoder(dtype='int'), ['State']),
          ('skip', 'passthrough', ["R&D Spend", "Administration", "Marketing Spend", "Profit"])
]
```



4. Setelah dilakukan proses OneHotEncoder maka dilakukan proses column transformer untuk mengubah isi dari column state

5. menggabungkan isian dari column state yang sudah dipecahmenadi satu dengan file "50 startups.csv



6. Proses selanjutnya proses Linear Regeresion, pada bagian tahap pertama dibuat variable X dan y, X digunakan untuk mengambil semua column dari table yang sudahdi lakukan proses Column Transformer



7. proses selanjutnya yaitu proses pemanggilan library LinearRegression, pada proses ini dibuat 4 variable dan dilakukan proses train_test_split yang isinya terdapat variable X, dan y

8. selanjutnya yaitu penggabungan dari Linear Regression

```
gabung = np.concatenate((ym_test, ym_pred), axis=1)
   gabung
array([[ 96712.8
                    , 84457.16115252],
      [ 90708.19
                   , 72543.44544734],
                   , 186070.97934498],
      [191792.06
                   , 67431.79651296],
      [ 81229.06
      [152211.77
                    , 150237.4771915 ],
                    , 113635.43439889],
      [126992.93
      [ 96479.51
                    , 88703.00224459],
                    , 124781.17620402],
      [134307.35
      [101004.64
                    , 101217.77951759],
      [108552.04
                     , 113863.68825244]])
```

9. Selanjutnya dilakukan prediksi menggunakan r2_score, nilai yang sempurna adalah nilai yang mendekati 1

```
r2_mlr = r2_score(ym_test, ym_pred)
print(r2_mlr)

✓ 0.5s
```





10.10.