

Tugas Materi 3

Zul Fauzi Oktaviansyah

2110181056

3 – D4 IT - B

Code No 1

```
In [2]: titanic_data = pd.read_csv('titanic.csv')  
titanic_data
```

--

Code diatas untuk membaca file csv dari library pandas

Hasil No 1

```
In [2]: titanic_data = pd.read_csv('titanic.csv')
titanic_data
```

Out[2]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

891 rows × 12 columns

Code No 2

```
In [3]: rows, cols = titanic_data.shape
```

```
In [4]: print('Jumlah Baris : ', rows)
        print('Jumlah Kolom : ', cols)
```

Code diatas untuk mengambil jumlah baris dan kolom pada data titanic yang telah dibaca

Hasil No 2

Jumlah Baris : 891
Jumlah Kolom : 12

Code No 3

```
In [113]: ageFare = pd.DataFrame(titanic_data, columns=['Age', 'Fare'])  
ageFare.head()
```

Code diatas untuk mengambil kolom tertentu dari file csv yg dibaca
Dan menampilkan 5 data pertama

Hasil No 3

Out[113]:

	Age	Fare
0	22.0	7.2500
1	38.0	71.2833
2	26.0	7.9250
3	35.0	53.1000
4	35.0	8.0500

Code No 4

```
In [6]: data = pd.DataFrame(titanic_data, columns=['Survived'])  
data
```

Code diatas untuk mengambil kolom tertentu dari file csv yg dibaca

Hasil No 4

Out[6]:

Survived	
0	0
1	1
2	1
3	1
4	0
...	...
886	0
887	1
888	0
889	1
890	0

891 rows × 1 columns

Code No 5

```
In [54]: miss_age = titanic_data.fillna(titanic_data.groupby('Orang').transform('mean'))
```

```
In [115]: data = miss_age[['Age', 'Fare']]  
data
```

Code diatas untuk mengisi NA menggunakan nilai rata-rata pada Orang, kemudian menampilkan kolom age, fare

Hasil No 5

Out[115]:

	Age	Fare
0	22.000000	7.2500
1	38.000000	71.2833
2	26.000000	7.9250
3	35.000000	53.1000
4	35.000000	8.0500
...
886	27.000000	13.0000
887	19.000000	30.0000
888	29.699118	23.4500
889	26.000000	30.0000
890	32.000000	7.7500

891 rows × 2 columns

Code No 6

```
In [115]: data = miss_age[['Age', 'Fare']]  
data
```

```
In [73]: def min_max_scaling(data):  
        data_norm = data.copy()  
  
        for column in data_norm.columns:  
            data_norm[column] = (data_norm[column] - data_norm[column].min()) / (data_norm[column].max() - data_norm[column].min())  
  
        return data_norm
```

```
In [76]: data_minMax = min_max_scaling(data)  
data_minMax.head()
```

Code diatas untuk menormalizekan nilai tiap kolom dengan menggunakan minmax metode, sebagai contoh pada kolom age dan fare

Hasil No 6

Output:

	Age	Fare
0	0.271174	0.014151
1	0.472229	0.139136
2	0.321438	0.015469
3	0.434531	0.103644
4	0.434531	0.015713

Code No 7

```
In [115]: data = miss_age[['Age', 'Fare']]  
data
```

```
In [77]: def z_score(data):  
        data_norm = data.copy()  
  
        for column in data_norm.columns:  
            data_norm[column] = (data_norm[column] - data_norm[column].mean()) / data_norm[column].std()  
  
        return data_norm
```

```
In [79]: data_zScore = z_score(data)  
data_zScore.head()
```

Code diatas untuk menormalizekan nilai tiap kolom dengan menggunakan z-score metode, sebagai contoh pada kolom age dan fare

Hasil No 7

Out[79]:

	Age	Fare
0	-0.592148	-0.502163
1	0.638430	0.786404
2	-0.284503	-0.488580
3	0.407697	0.420494
4	0.407697	-0.486064

Code No 8

```
In [115]: data = miss_age[['Age', 'Fare']]  
data
```

```
In [110]: import numpy as np
```

```
In [111]: def sigmoid(data):  
  
    data_norm = data.copy()  
  
    for column in data_norm.columns:  
        x = (data_norm[column] - data_norm[column].mean()) / data_norm[column].std()  
        data_norm[column] = (1 - np.exp(-x)) / (1 + np.exp(-x))  
  
    return data_norm
```

```
In [112]: data_sigmoid = sigmoid(data)  
data_sigmoid.head()
```

Code diatas untuk menormalizekan nilai tiap kolom dengan menggunakan sigmoid metode, sebagai contoh pada kolom age dan fare

Hasil No 8

Out[112]:

	Age	Fare
0	-0.287716	-0.245935
1	0.308797	0.374117
2	-0.141300	-0.239544
3	0.201071	0.207203
4	0.201071	-0.238358