Tugas Materi 2

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 [5]: data = pd.DataFrame(titanic_data, columns=['Name', 'Sex', 'Age', 'Pclass', 'Fare'])
        data
```

Code diatas untuk mengambil kolom tertentu dari file csv yg dibaca

Hasil No 3

Out[5]:

	Name	Sex	Age	Pclass	Fare
0	Braund, Mr. Owen Harris	male	22.0	3	7.2500
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	71.2833
2	Heikkinen, Miss. Laina	female	26.0	3	7.9250
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	53.1000
4	Allen, Mr. William Henry	male	35.0	3	8.0500
...	...	...	...	...	...
886	Montvila, Rev. Juozas	male	27.0	2	13.0000
887	Graham, Miss. Margaret Edith	female	19.0	1	30.0000
888	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	3	23.4500
889	Behr, Mr. Karl Howell	male	26.0	1	30.0000
890	Dooley, Mr. Patrick	male	32.0	3	7.7500

891 rows × 5 columns

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 [7]:
```
sibsp = titanic_data.loc[:, 'SibSp']
parch = titanic_data.loc[:, 'Parch']
relatives = sibsp + parch
relatives
```

Out[7]:
```
0      1
1      1
2      0
3      1
4      0
      ..
886    0
887    0
888    3
889    0
890    0
Length: 891, dtype: int64
```

Code diatas untuk menampung kolom sibsp dan parch, kemudian kedua kolom ditambahkan menjadi relatives

Hasil No 5

In [8]:
```
titanic_data['Relatives'] = pd.Series(relatives, index=titanic_data.index)
titanic_data
```

Out[8]:

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

891 rows × 13 columns

Code diatas digunakan untuk menambahkan kolom ‘Relatives’ pada file csv yang dibaca

Code No 5

```
In [9]: temp = titanic_data
        temp['Orang'] = pd.Series(1, index=temp.index)
        temp
```

Code diatas untuk membuat kolom baru ‘Orang’ yang mana setiap baris merupakan 1 Orang

```
]: pclass1 = pd.DataFrame(temp, columns=['Pclass', 'Orang'])
   pclass1
```

Code diatas untuk menampilkan kolom Pclass dan Orang

```
In [19]: pclass2 = pclass1.groupby('Pclass').sum()
         pclass2
```

Code diatas untuk membuat grup baru berdasarkan Pclass,
Kemudian selain kolom Pclass yaitu Orang ditambahkan valuenya berdasarkan Pclass

Hasil No 5

Out[9]:

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

891 rows × 14 columns

Hasil No 5

Out[18]:

	Pclass	Orang
0	3	1
1	1	1
2	3	1
3	1	1
4	3	1
...	...	...
886	2	1
887	1	1
888	3	1
889	1	1
890	3	1

891 rows × 2 columns

Out[19]:

Pclass	Orang
1	216
2	184
3	491

Code No 6

```
In [12]: sex1 = pd.DataFrame(temp, columns=['Sex', 'Orang'])
         sex1
```

Code diatas untuk menampilkan kolom Sex dan Orang

```
In [20]: sex2 = sex1.groupby('Sex').sum()
         sex2
```

Code diatas untuk membuat grup baru berdasarkan Sex,
Kemudian selain kolom Sex yaitu Orang ditambahkan valuenya berdasarkan Sex

Hasil No 6

Out[12]:

	Sex	Orang
0	male	1
1	female	1
2	female	1
3	female	1
4	male	1
...	...	...
886	male	1
887	female	1
888	female	1
889	male	1
890	male	1

891 rows × 2 columns

Out[20]:

Sex	Orang
female	314
male	577

Code No 7

```
In [28]: survive = titanic_data[titanic_data["Survived"] > 0]
         survive
```

Code diatas digunakan untuk mengambil data dengan kondisi hidup

```
In [38]: survive1 =pd.DataFrame(survive, columns=['Pclass', 'Orang'])
         survive2 = survive1.groupby('Pclass').sum()
         survive2
```

Code diatas digunakan untuk mengambil data dengan kolom Pclass dan Orang,
Kemudian di grupkan berdasarkan Pclass, kemudian selain kolom Pclass yaitu Orang
Dijumlahkan berdasarkan Pclass

Hasil No 7

3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S	1	1
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	NaN	S	2	1
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	NaN	C	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
875	876	1	3	Najib, Miss. Adele Kiamie "Jane"	female	15.0	0	0	2667	7.2250	NaN	C	0	1
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.1583	C50	C	1	1
880	881	1	2	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	1	230433	26.0000	NaN	S	1	1
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S	0	1
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C	0	1

342 rows × 14 columns

Out[38]:

	Orang
Pclass	
1	136
2	87
3	119

Code No 8

```
In [39]: died = titanic_data[titanic_data["Survived"] < 1]
         died
```

Code diatas digunakan untuk mengambil data dengan kondisi mati

```
In [40]: died1 =pd.DataFrame(died, columns=['Pclass', 'Orang'])
         died2 = died1.groupby('Pclass').sum()
         died2
```

Code diatas digunakan untuk mengambil data dengan kolom Pclass dan Orang,
Kemudian di grupkan berdasarkan Pclass, kemudian selain kolom Pclass yaitu Orang
Dijumlahkan berdasarkan Pclass

Hasil No 8

Out[39]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Relatives	Orang
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S	1	1
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S	0	1
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583	NaN	Q	0	1
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E46	S	0	1
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	NaN	S	4	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
884	885	0	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	SOTON/OQ 392076	7.0500	NaN	S	0	1
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.1250	NaN	Q	5	1
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S	0	1
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	S	3	1
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q	0	1

549 rows × 14 columns

Out[40]:

	Orang
Pclass	
1	80
2	97
3	372

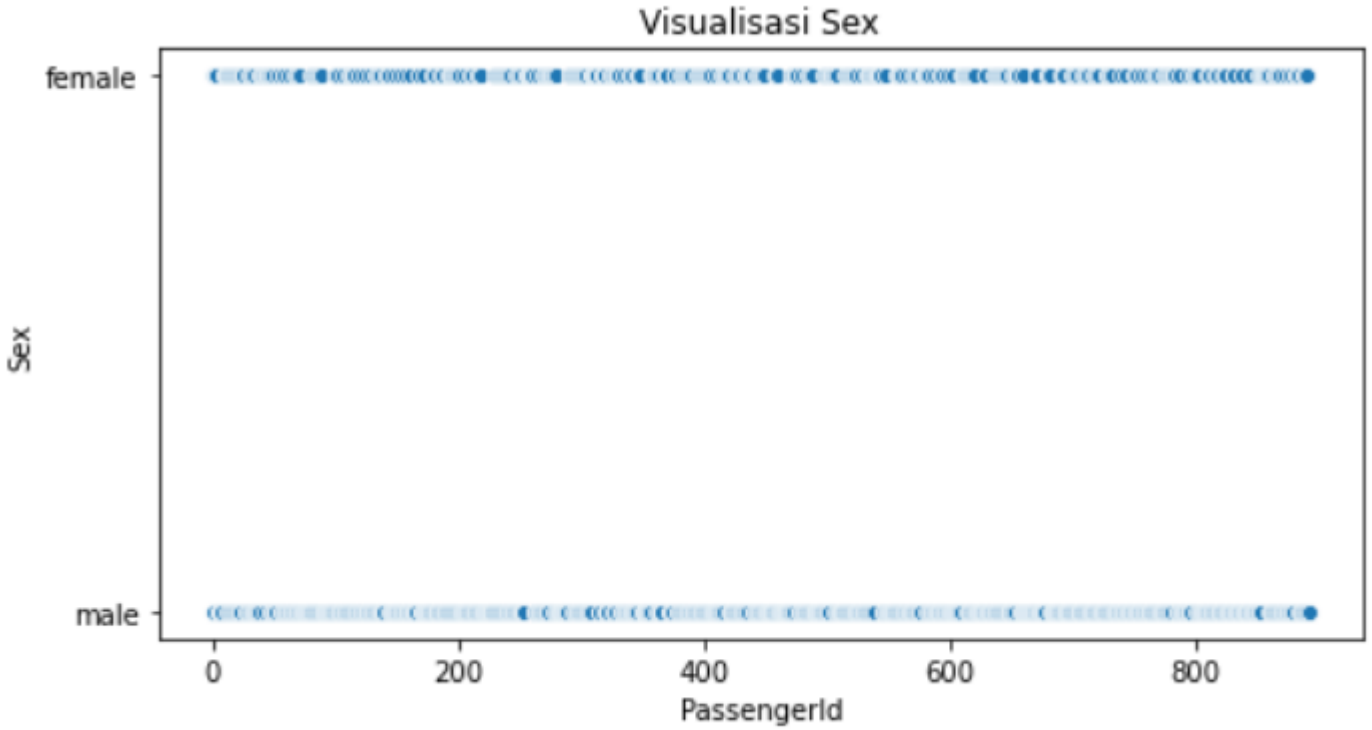
Code No 9

In [72]:
```
plt.figure(figsize=(8,4))
sns.scatterplot(temp['PassengerId'], temp['Sex'],
                palette=sns.color_palette('hls', 2))

plt.title('Visualisasi Sex')
plt.show()
```

Code diatas, plt untuk membuat plot dari library matplotlib.pyplot kemudian Sns dari library seaborn untuk menkonfigurasi.

Hasil No 9


Visualisasi Sex
female
male
Sex
0
200
400
600
800
PassengerId

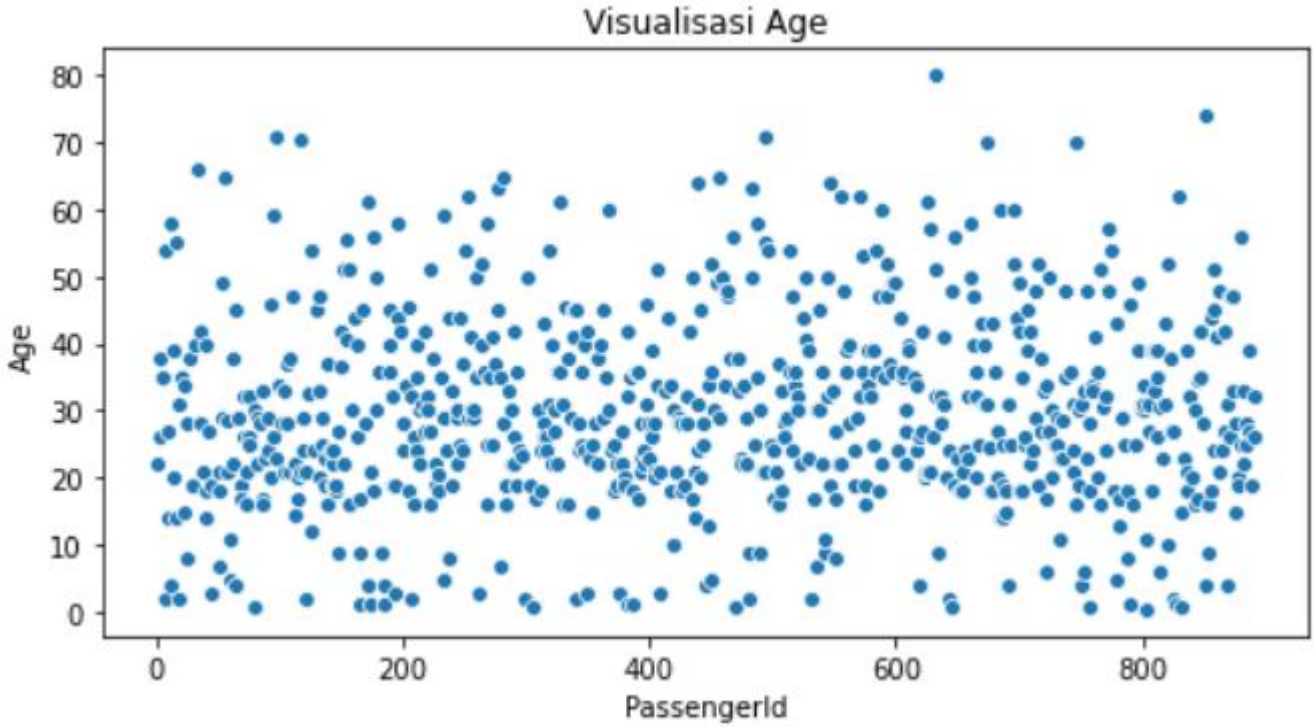
Code No 10

In [69]:
```
plt.figure(figsize=(8,4))
sns.scatterplot(temp['PassengerId'], temp['Age'],
                palette=sns.color_palette('hls', 2))

plt.title('Visualisasi Age')
plt.show()
```

Code diatas, plt untuk membuat plot dari library matplotlib.pyplot kemudian Sns dari library seaborn untuk menkonfigurasi.

Hasil No 10


Visualisasi Age
Age
PassengerId
0
10
20
30
40
50
60
70
80
0
200
400
600
800