

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
df_aerofit = pd.read_csv("https://d2beiqkhq929f0.cloudfront.net/public_assets/asset
```

```
df_aerofit.head()
```

	Product	Age	Gender	Education	MaritalStatus	Usage	Fitness	Income	Miles
0	KP281	18	Male	14	Single	3	4	29562	11
1	KP281	19	Male	15	Single	2	3	31836	7
2	KP281	19	Female	14	Partnered	4	3	30699	6
3	KP281	19	Male	12	Single	3	3	32973	8
4	KP281	20	Male	13	Partnered	4	2	35247	4

```
df_aerofit.shape
```

```
(180, 9)
```

```
df_aerofit.isnull().sum()
```

```
Product      0
Age           0
Gender        0
Education     0
MaritalStatus 0
Usage         0
Fitness       0
Income        0
Miles         0
dtype: int64
```

```
df_aerofit.isna().sum()
```

```
Product      0
Age           0
Gender        0
Education     0
MaritalStatus 0
Usage         0
Fitness       0
Income        0
Miles         0
dtype: int64
```

```
df_aerofit.describe()
```



	Age	Education	Usage	Fitness	Income	Miles
count	180.000000	180.000000	180.000000	180.000000	180.000000	180.000000
mean	28.788889	15.572222	3.455556	3.311111	53719.577778	103.194444
std	6.943498	1.617055	1.084797	0.958869	16506.684226	51.863605
min	18.000000	12.000000	2.000000	1.000000	29562.000000	21.000000
25%	24.000000	14.000000	3.000000	3.000000	44058.750000	66.000000
50%	26.000000	16.000000	3.000000	3.000000	50596.500000	94.000000
75%	33.000000	16.000000	4.000000	4.000000	58668.000000	114.750000
max	50.000000	21.000000	7.000000	5.000000	104581.000000	360.000000

```
df_aerofit.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 180 entries, 0 to 179
Data columns (total 9 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Product         180 non-null   object
1   Age             180 non-null   int64
2   Gender          180 non-null   object
3   Education       180 non-null   int64
4   MaritalStatus   180 non-null   object
5   Usage           180 non-null   int64
6   Fitness         180 non-null   int64
7   Income          180 non-null   int64
8   Miles           180 non-null   int64
dtypes: int64(6), object(3)
memory usage: 12.8+ KB
```

```
df_aerofit.nunique()
```

```
Product      3
Age          32
Gender        2
Education     8
MaritalStatus 2
Usage         6
Fitness       5
Income       62
Miles        37
dtype: int64
```

```
df_aerofit.columns
```

```
Index(['Product', 'Age', 'Gender', 'Education', 'MaritalStatus', 'Usage',
       'Fitness', 'Income', 'Miles'],
      dtype='object')
```

```
df_aerofit["Product"].value_counts()
```

```

KP281    80
KP481    60
KP781    40
Name: Product, dtype: int64

```

```
df_aerofit["Usage"].value_counts()
```

```

3    69
4    52
2    33
5    17
6     7
7     2
Name: Usage, dtype: int64

```

```
df_aerofit["Fitness"].value_counts()
```

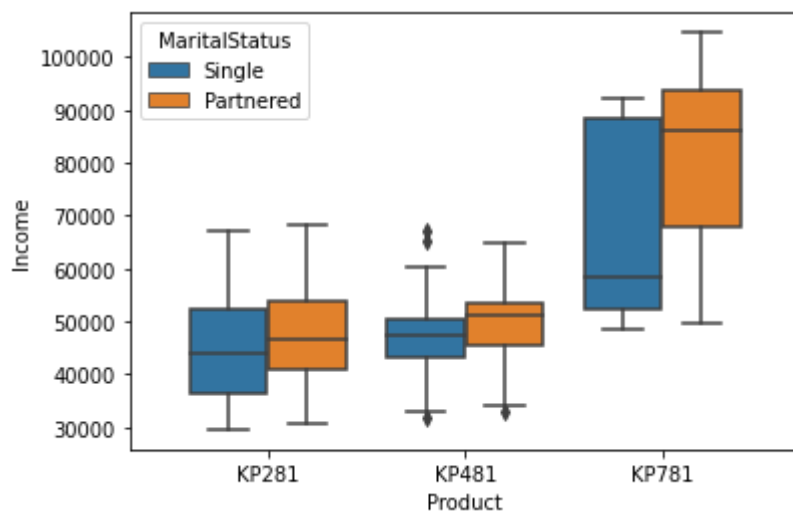
```

3    97
5    31
2    26
4    24
1     2
Name: Fitness, dtype: int64

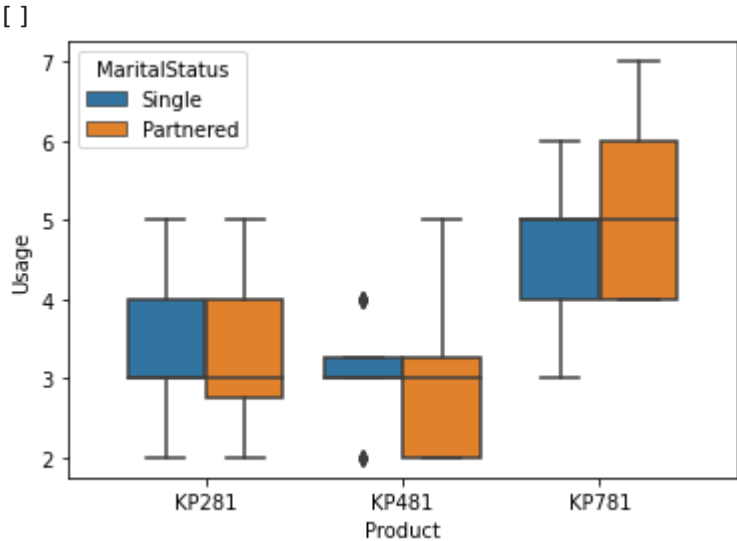
```

```
sns.boxplot(x=df_aerofit["Product"],y=df_aerofit["Income"],hue=df_aerofit["MaritalStatus"],plt.plot())
```

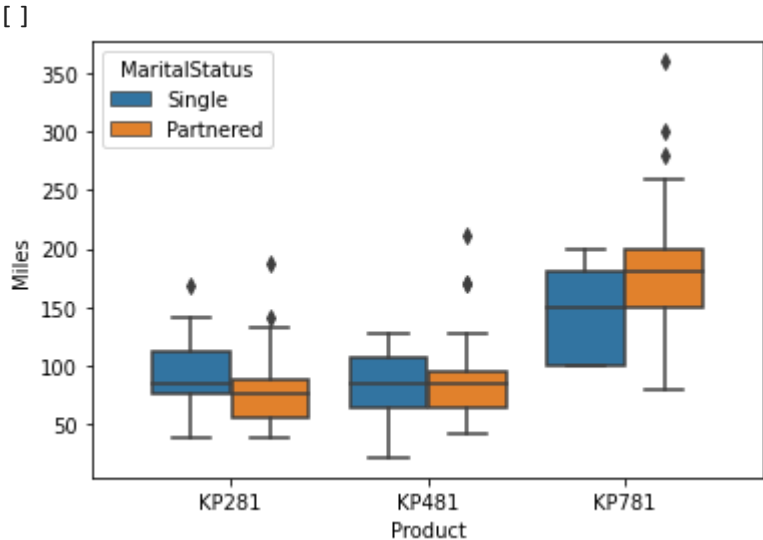
```
[ ]
```



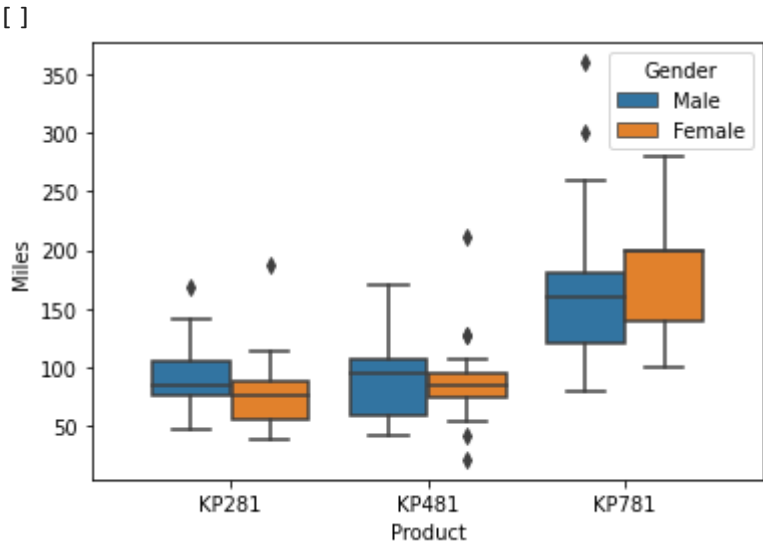
```
sns.boxplot(x=df_aerofit["Product"],y=df_aerofit["Usage"],hue=df_aerofit["MaritalStatus"],plt.plot())
```



```
sns.boxplot(x=df_aerofit["Product"],y=df_aerofit["Miles"],hue=df_aerofit["MaritalSt
plt.plot()
```

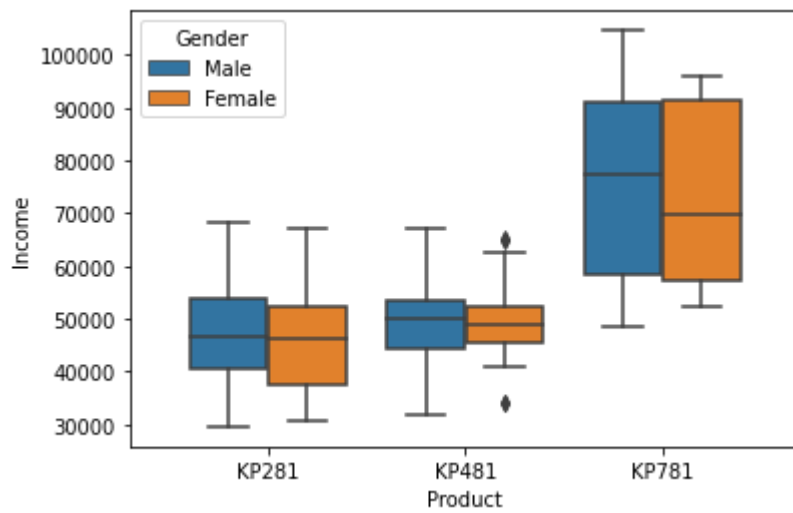


```
sns.boxplot(x=df_aerofit["Product"],y=df_aerofit["Miles"],hue=df_aerofit["Gender"],
plt.plot()
```



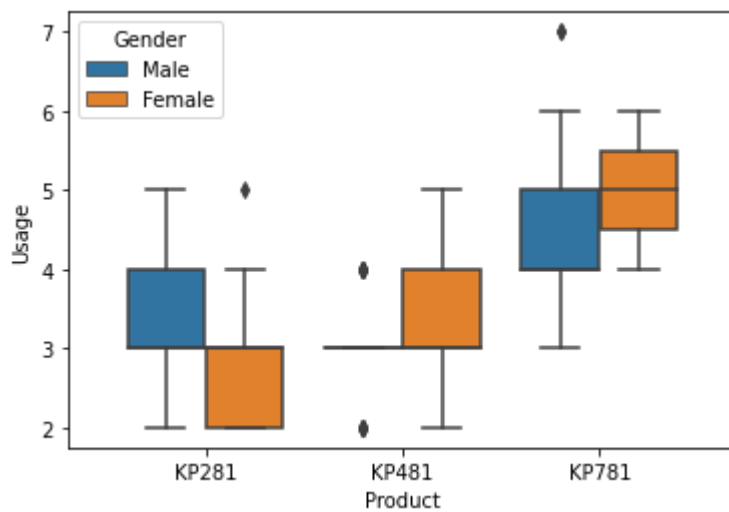
```
sns.boxplot(x=df_aerofit["Product"],y=df_aerofit["Income"],hue=df_aerofit["Gender"])
plt.plot()
```

[]

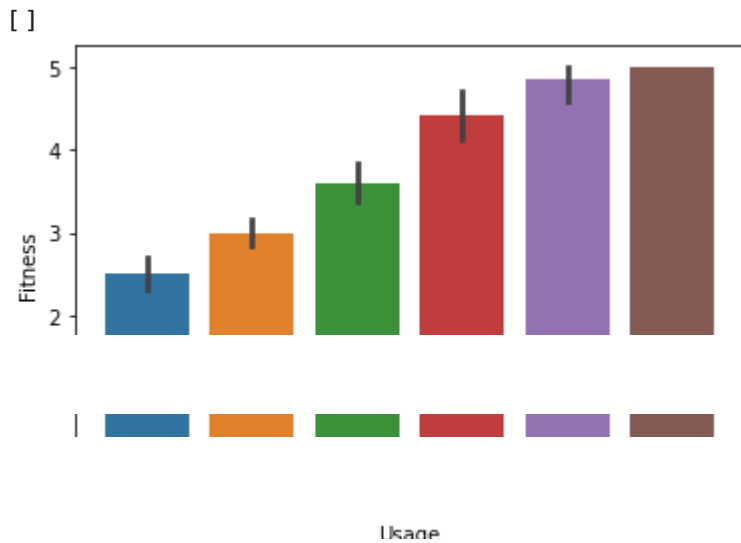


```
sns.boxplot(x=df_aerofit["Product"],y=df_aerofit["Usage"],hue=df_aerofit["Gender"])
plt.plot()
```

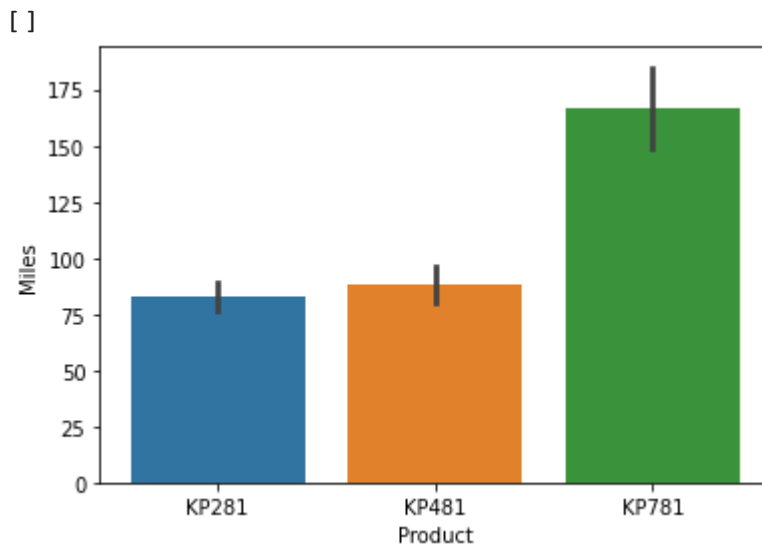
[]



```
sns.barplot(x=df_aerofit["Usage"],y=df_aerofit["Fitness"])
plt.plot()
```



```
sns.barplot(x=df_aerofit["Product"],y=df_aerofit["Miles"])
plt.plot()
```



```
df_aerofit.corr()
```

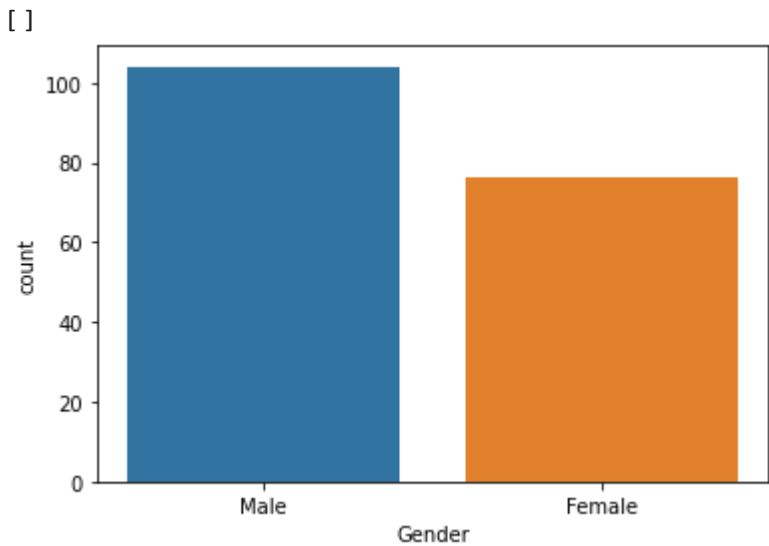
	Age	Education	Usage	Fitness	Income	Miles	
Age	1.000000	0.280496	0.015064	0.061105	0.513414	0.036618	
Education	0.280496	1.000000	0.395155	0.410581	0.625827	0.307284	
Usage	0.015064	0.395155	1.000000	0.668606	0.519537	0.759130	
Fitness	0.061105	0.410581	0.668606	1.000000	0.535005	0.785702	
Income	0.513414	0.625827	0.519537	0.535005	1.000000	0.543473	
Miles	0.036618	0.307284	0.759130	0.785702	0.543473	1.000000	



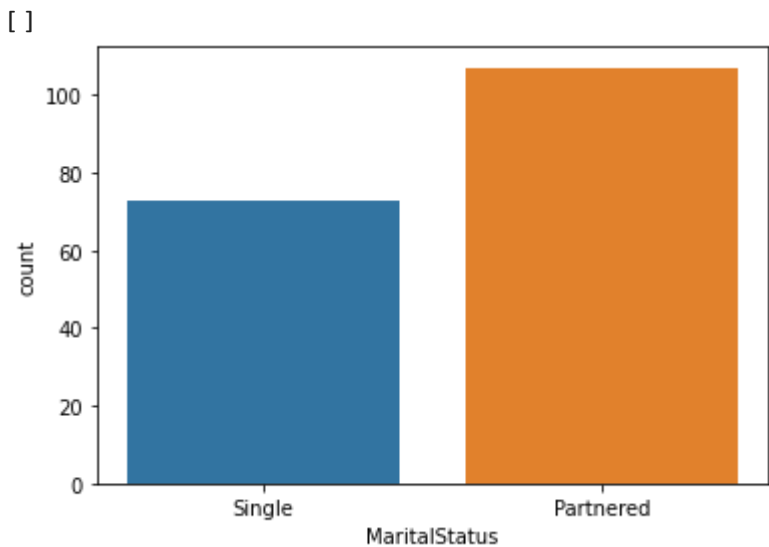
```
df_aerofit.head()
```

	Product	Age	Gender	Education	MaritalStatus	Usage	Fitness	Income	Mileage
0	KP281	18	Male	14	Single	3	4	29562	11000
1	KP281	19	Male	15	Single	2	3	31836	7000
2	KP281	19	Female	14	Partnered	4	3	30699	6000
3	KP281	19	Male	12	Single	3	3	32973	8000
4	KP281	20	Male	13	Partnered	4	2	35247	4000

```
sns.countplot(x="Gender",data=df_aerofit)
plt.plot()
```

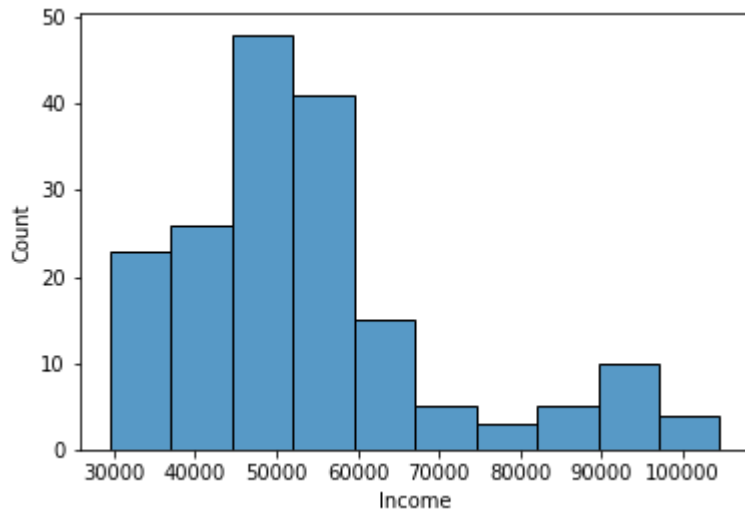


```
sns.countplot(x="MaritalStatus",data=df_aerofit)
plt.plot()
```



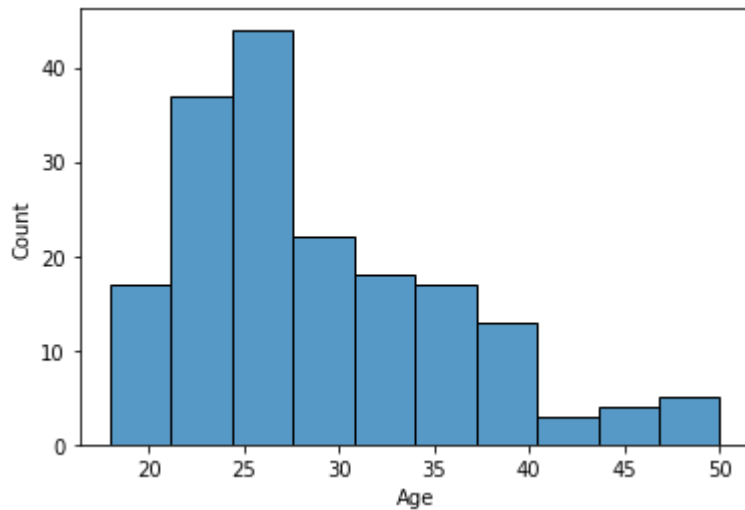
```
sns.histplot(data=df_aerofit["Income"],bins=10)
plt.plot()
```

[]



```
sns.histplot(data=df_aerofit["Age"],bins=10)
plt.plot()
```


[]




```
pd.crosstab(index=df_aerofit["Product"],columns=df_aerofit["Usage"],margins=True)
```

Usage	2	3	4	5	6	7	All	
Product								
KP281	19	37	22	2	0	0	80	
KP481	14	31	12	3	0	0	60	
KP781	0	1	18	12	7	2	40	
All	33	69	52	17	7	2	180	



```
pd.crosstab(index=df_aerofit["Product"],columns=df_aerofit["Fitness"],margins=True)
```

Fitness	1	2	3	4	5	All	
Product							
KP281	1	14	54	9	2	80	
KP481	1	12	39	8	0	60	
KP781	0	0	4	7	29	40	
All	2	26	97	24	31	180	

```
pd.crosstab(index=df_aerofit["Product"],columns=df_aerofit["Gender"],margins=True)
```

Gender	Female	Male	All	
Product				
KP281	40	40	80	
KP481	29	31	60	
KP781	7	33	40	
All	76	104	180	

```
pd.crosstab(index=df_aerofit["Product"],columns=df_aerofit["MaritalStatus"],margins
```

MaritalStatus	Partnered	Single	All	
Product				
KP281	48	32	80	
KP481	36	24	60	
KP781	23	17	40	
All	107	73	180	

```
df_aerofit[df_aerofit["Product"] == "KP281"].shape[0]/df_aerofit.shape[0]
```

```
0.4444444444444444
```

```
df_aerofit[df_aerofit["Product"] == "KP481"].shape[0]/df_aerofit.shape[0]
```

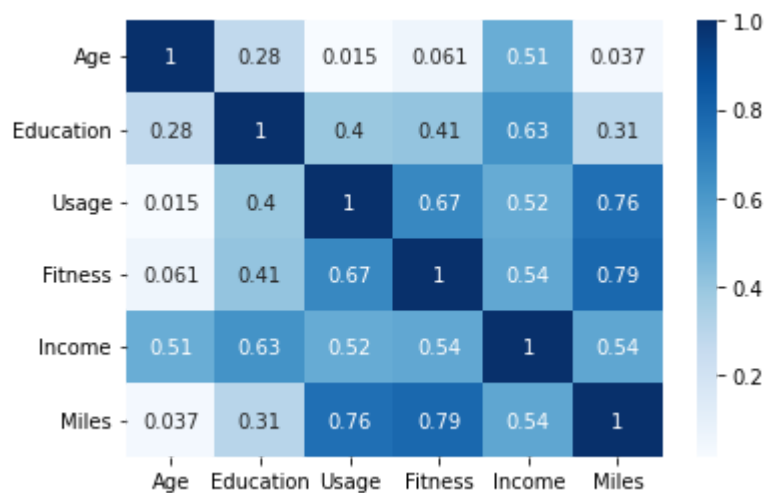
```
0.3333333333333333
```

```
df_aerofit[df_aerofit["Product"] == "KP781"].shape[0]/df_aerofit.shape[0]
```

```
0.2222222222222222
```

```
sns.heatmap(data=df_aerofit.corr(),annot=True,cmap="Blues")
plt.plot()
```

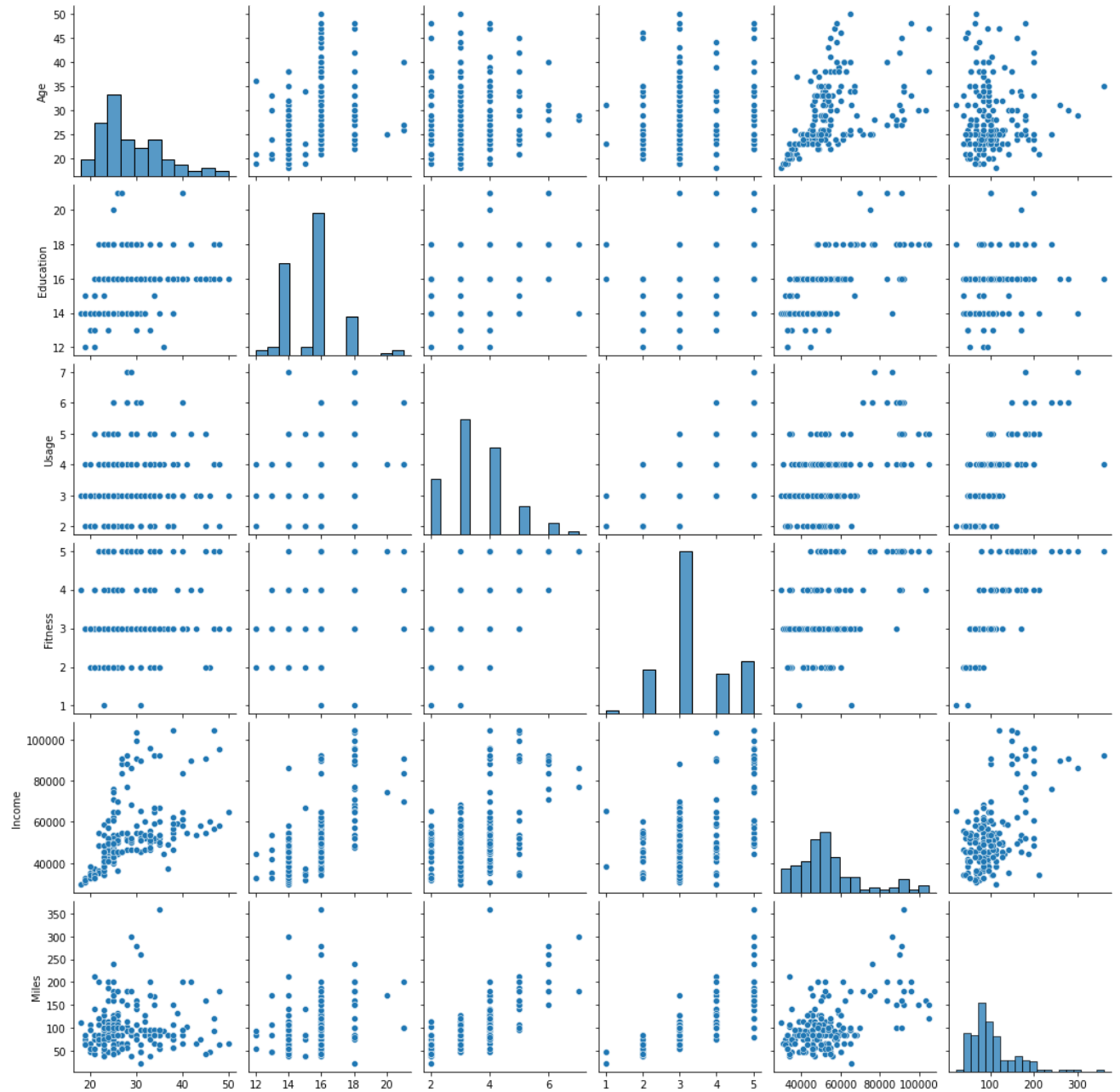
```
[ ]
```



```
sns.pairplot(data=df_aerofit)
plt.plot()
```



[]



```
df_aerofit[df_aerofit["Gender"] == "Male"].shape[0]/df_aerofit.shape[0]
```

```
0.5777777777777777
```

```
1 - df_aerofit[df_aerofit["Gender"] == "Male"].shape[0]/df_aerofit.shape[0]
```

```
0.42222222222222223
```

[Colab paid products](#) - [Cancel contracts here](#)

✓ 0s completed at 9:05 PM

