

[Week9_homework_성혜원]

1. Challenge 1

The screenshot shows a Jupyter Notebook interface with the following code and output:

```
In [51]: df.loc[df['Sex']=='male', 'Sex'] = 0
df.loc[df['Sex']=='female', 'Sex'] = 1
df.head()
```

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	FamilySize	
0	1	0	3	Braund, Mr. Owen Harris	0	2.0	1	0	A/5 21171	7.2500	NaN	S	1
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	1	3.0	1	0	PC 17569	71.2833	C65	C	1
2	3	1	3	Heikinen, Miss. Laina	1	2.0	0	0	STON/O2, 3101282	7.9250	NaN	S	0
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	1	3.0	1	0	113803	53.1000	C123	S	1
4	5	0	3	Allen, Mr. William Henry	0	3.0	0	0	373450	8.0500	NaN	S	0

```
In [52]: df_test.loc[df_test['Sex']=='male', 'Sex'] = 0
df_test.loc[df_test['Sex']=='female', 'Sex'] = 1
df_test.head()
```

PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	FamilySize	
0	892	3	Kelly, Mr. James	0	3.0	0	0	330911	7.6292	NaN	Q	0
1	893	3	Wilkes, Mrs. James (Ellen Needs)	1	4.0	1	0	363272	7.0000	NaN	S	1
2	894	2	Myles, Mr. Thomas Francis	0	5.0	0	0	240276	9.6875	NaN	Q	0
3	895	3	Wirz, Mr. Albert	0	2.0	0	0	315154	8.6625	NaN	S	0
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	1	2.0	1	1	3101298	12.2875	NaN	S	2

The screenshot shows a Jupyter Notebook interface with the following code and output:

```
In [55]: df.loc[df['Embarked']=='S', 'Embarked'] = 0
df.loc[df['Embarked']=='C', 'Embarked'] = 1
df.loc[df['Embarked']=='Q', 'Embarked'] = 2
df_test.loc[df_test['Embarked']=='S', 'Embarked'] = 0
df_test.loc[df_test['Embarked']=='C', 'Embarked'] = 1
df_test.loc[df_test['Embarked']=='Q', 'Embarked'] = 2
```

```
In [58]: train = df[['Survived', 'Sex', 'Age', 'FamilySize', 'Fare', 'Embarked']]
test = df_test[['Sex', 'Age', 'FamilySize', 'Fare', 'Embarked']]
train.head()
```

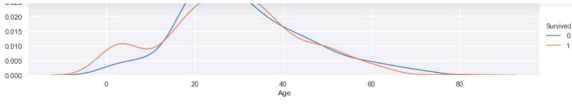
Survived	Sex	Age	FamilySize	Fare	Embarked	
0	0	0	2.0	1	7.2500	0
1	1	1	3.0	1	71.2833	1
2	1	1	2.0	0	7.9250	0
3	1	1	3.0	1	53.1000	0
4	0	0	3.0	0	8.0500	0

2. Challenge 2

jupyter challenge2 (unsaved changes) Python

File Edit View Insert Cell Kernel Widgets Help Trusted Python

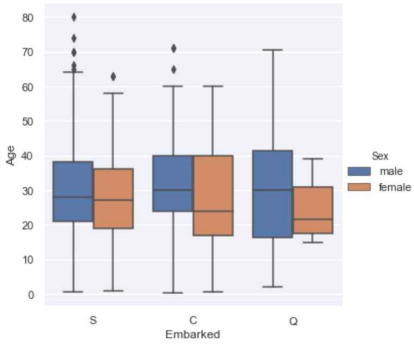
Code



A line plot showing the relationship between Age (x-axis, 0 to 80) and Survival (y-axis, 0.000 to 0.020). Two lines are plotted: a blue line for Survival = 0 and an orange line for Survival = 1. The orange line shows a peak in survival probability around age 20, while the blue line shows a peak around age 40.

```
In [12]: sns.catplot(data=train, x='Embarked', y='Age', hue='Sex', kind='box')
```

Out[12]: <seaborn.axisgrid.FacetGrid at 0x1f620e4ac08>



A box plot showing the distribution of Age (y-axis, 0 to 80) across three embarkment categories (S, C, Q) on the x-axis. The plot is faceted by Sex (male, female). The legend indicates that blue represents 'male' and orange represents 'female'. The plot shows that for each embarkment category, the age distribution is generally similar for both sexes, with some outliers present.

Embarked	Sex	Approx. Median Age
S	male	28
S	female	25
C	male	30
C	female	28
Q	male	30
Q	female	20