In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
```

In [2]:
```
```

In [3]:
```
df_train=pd.read_csv("C2_train.gender_submission.csv")
df_test=pd.read_csv("C2_test.gender_submission.csv")
```

Out[3]:

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

891 rows × 12 columns

In [4]:

Out[4]:

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN
...	...	...	...	...	...	...	...	...	...	...
413	1305	3	Spector, Mr. Woolf	male	NaN	0	0	A.5. 3236	8.0500	NaN
414	1306	1	Oliva y Ocana, Dona. Fermina	female	39.0	0	0	PC 17758	108.9000	C105
415	1307	3	Saether, Mr. Simon Sivertsen	male	38.5	0	0	SOTON/O.Q. 3101262	7.2500	NaN
416	1308	3	Ware, Mr. Frederick	male	NaN	0	0	359309	8.0500	NaN
417	1309	3	Peter, Master. Michael J	male	NaN	1	1	2668	22.3583	NaN

418 rows × 11 columns

In [5]:

In [6]:

In [7]:
```
feature_matrix=df1[['PassengerId', 'Survived','Age', 'SibSp','Parch','Fare']]
```

In [8]:

Out[8]: (183, 6)

In [9]:

Out[9]: (183, 1)

In [10]:

In [11]:

In [12]:
```
logr=LogisticRegression()
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\utils\validation.py:63: Da
taConversionWarning: A column-vector y was passed when a 1d array was expecte
d. Please change the shape of y to (n_samples, ), for example using ravel().
  return f(*args, **kwargs)
```

Out[12]: LogisticRegression()

In [13]:

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 87 entries, 12 to 414
Data columns (total 11 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  87 non-null     int64
 1   Pclass       87 non-null     int64
 2   Name         87 non-null     object
 3   Sex          87 non-null     object
 4   Age          87 non-null     float64
 5   SibSp        87 non-null     int64
 6   Parch        87 non-null     int64
 7   Ticket       87 non-null     object
 8   Fare         87 non-null     float64
 9   Cabin        87 non-null     object
 10  Embarked     87 non-null     object
dtypes: float64(2), int64(4), object(5)
memory usage: 8.2+ KB
```

In [14]:

Out[14]:
```
Index(['PassengerId', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp', 'Parch',
       'Ticket', 'Fare', 'Cabin', 'Embarked'],
      dtype='object')
```

In [15]:
```
observation=df2[['PassengerId', 'Pclass','Age', 'SibSp', 'Parch',
```

In [16]:
```
prediction=logr.predict(observation)
```

```
['S' 'S' 'C' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'C' 'S' 'C' 'C' 'S' 'C' 'S' 'C'
 'C' 'S' 'C' 'S' 'S' 'S' 'S' 'S' 'C' 'S' 'S' 'S' 'S' 'C' 'S' 'C' 'S' 'S'
 'S' 'S' 'S' 'S' 'S' 'C' 'S' 'C' 'S' 'S' 'C' 'S' 'S' 'S' 'S' 'S' 'S' 'S'
 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'C' 'S' 'S' 'S'
 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S' 'S']
```

In [17]:

Out[17]: array(['C', 'Q', 'S'], dtype=object)

In [18]:

Out[18]: 2.8313345570546527e-29

In [19]:

Out[19]: 1.8438585166969872e-115

In [ ]: