# Installation

In [9]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
%matplotlib inline
```

In [47]:
```
titanic_train=pd.read_csv('train.csv')
titanic_test=pd.read_csv('test.csv')
```

In [12]:
```
titanic_train.head()
```

Out[12]:

| | PassengerId | Survived | Pclass | Name | Sex | Age | SibSp | Parch | Tic |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 0 | 3 | Braund, Mr. Owen Harris | male | 22.0 | 1 | 0 | 21 |
| **1** | 2 | 1 | 1 | Cumings, Mrs. John Bradley (Florence Briggs Th... | female | 38.0 | 1 | 0 | PC 17 |
| **2** | 3 | 1 | 3 | Heikkinen, Miss. Laina | female | 26.0 | 0 | 0 | STON/ 3101 |
| **3** | 4 | 1 | 1 | Futrelle, Mrs. Jacques Heath (Lily May Peel) | female | 35.0 | 1 | 0 | 113 |
| **4** | 5 | 0 | 3 | Allen, Mr. William Henry | male | 35.0 | 0 | 0 | 373 |

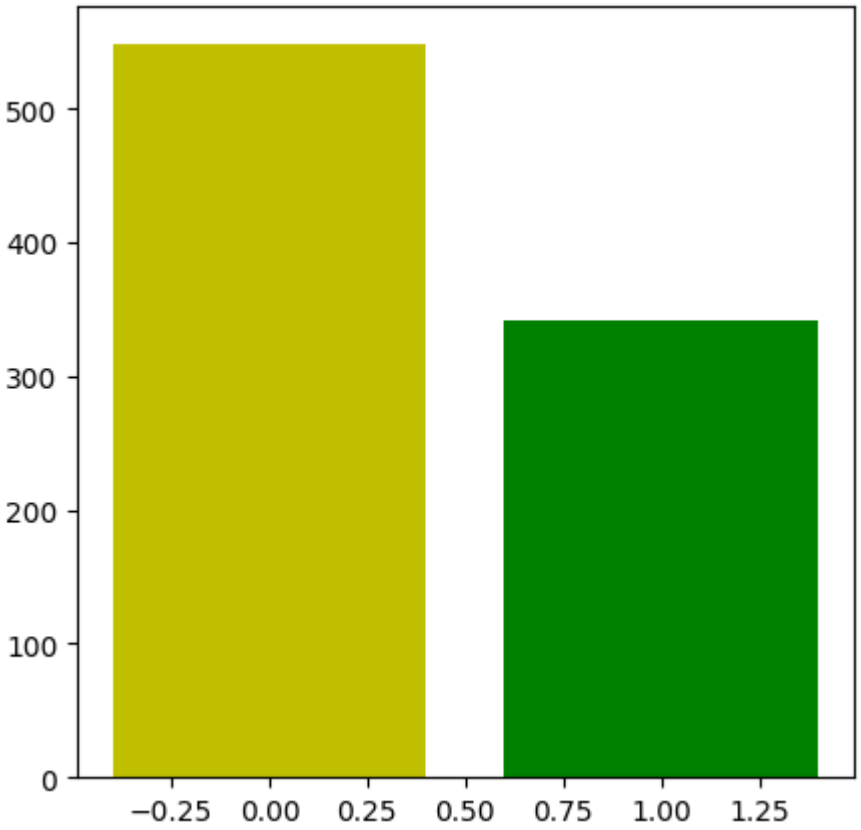
In [13]:
```
titanic_train.shape
```

Out[13]:
```
(891, 12)
```

In [14]:
```
titanic_train['Survived'].value_counts()
```

Out[14]:
```
Survived
0    549
1    342
Name: count, dtype: int64
```

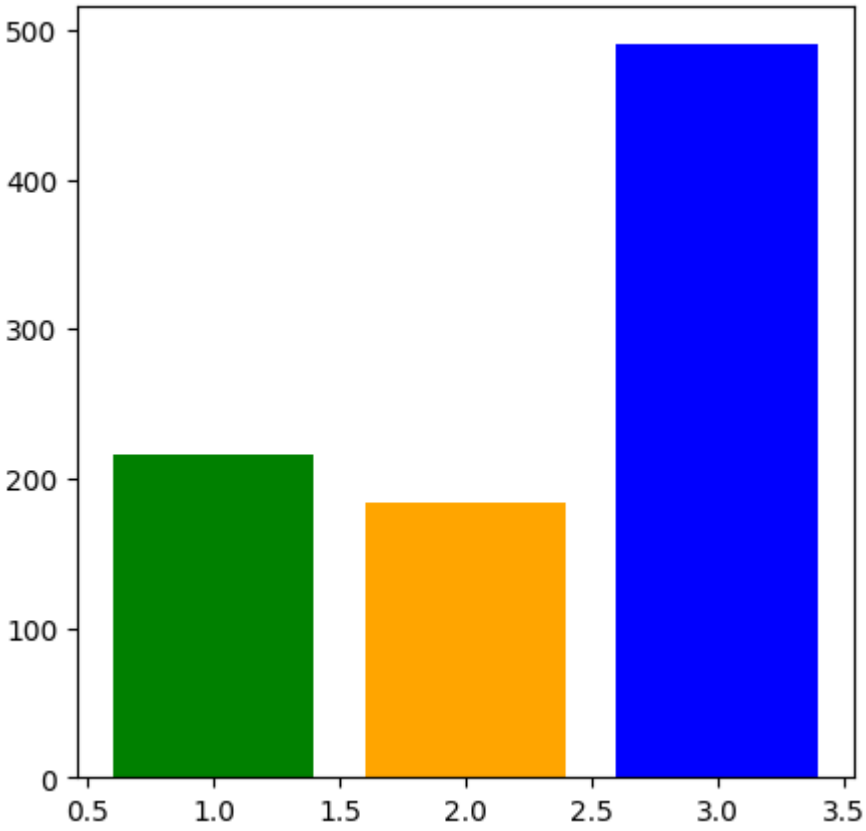
In [16]:
```
plt.figure(figsize=(5,5))
plt.bar(list(titanic_train['Survived'].value_counts().keys()),list(titanic_t
plt.show()
```




In [17]:
```
titanic_train['Pclass'].value_counts()
```

Out[17]:
```
Pclass
3    491
1    216
2    184
Name: count, dtype: int64
```

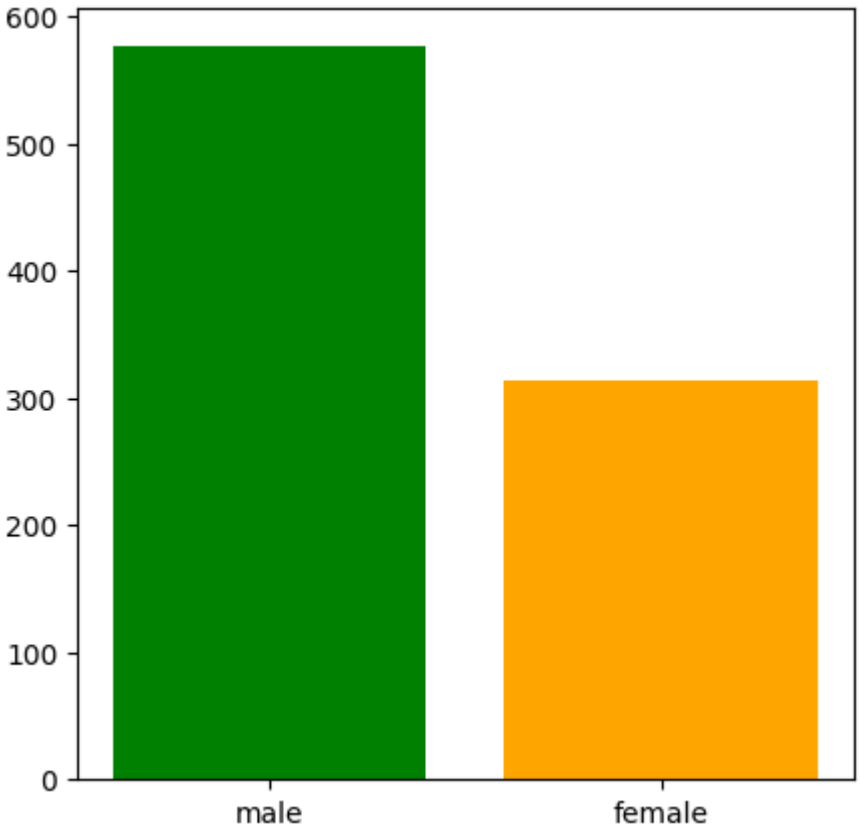
In [23]:
```
plt.figure(figsize=(5,5))
plt.bar(list(titanic_train['Pclass'].value_counts().keys()),list(titanic_tra
plt.show()
```



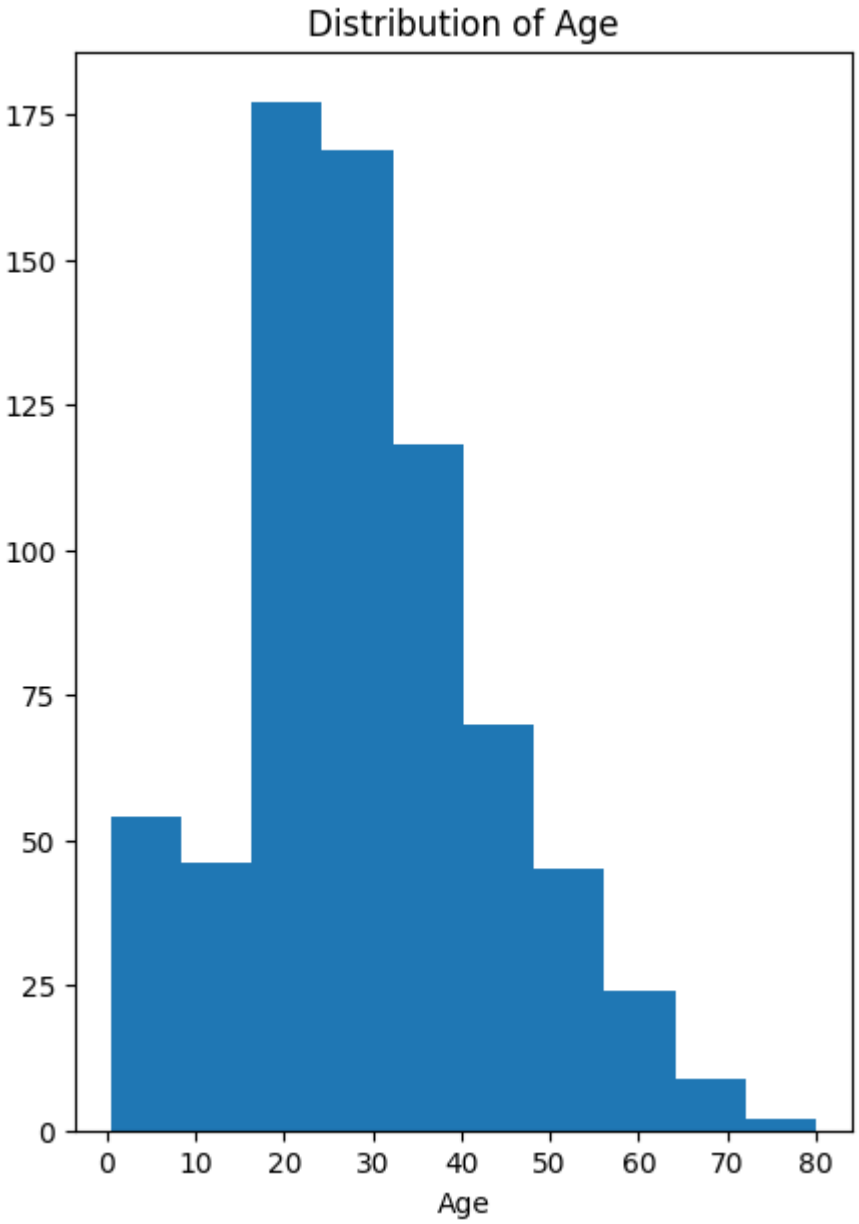

In [24]:
```
titanic_train['Sex'].value_counts()
```

Out[24]:
```
Sex
male      577
female    314
Name: count, dtype: int64
```

In [27]:
```
plt.figure(figsize=(5,5))
plt.bar(list(titanic_train['Sex'].value_counts().keys()),list(titanic_train[
plt.show()
```




In [28]:
```
plt.figure(figsize=(5,7))
plt.hist(titanic_train['Age'])
plt.title("Distribution of Age")
plt.xlabel("Age")
plt.show()
```




In [31]:
```
titanic_train['Survived'].isnull()
```

Out[31]:
```
0      False
1      False
2      False
3      False
4      False
       ...
886    False
887    False
888    False
889    False
890    False
Name: Survived, Length: 891, dtype: bool
```

In [32]:
```
sum(titanic_train['Survived'].isnull())
```

Out[32]: 0

In [33]:
```
sum(titanic_train['Age'].isnull())
```

Out[33]: 177

In [34]:
```
titanic_train=titanic_train.dropna()
```

## Building model

In [37]:
```
sum(titanic_train['Survived'].isnull())
```

Out[37]: 0

In [38]:
```
sum(titanic_train['Age'].isnull())
```

Out[38]: 0

In [39]:
```
x_train=titanic_train[['Age']]
y_train=titanic_train[['Survived']]
```

In [43]:
```
from sklearn.tree import DecisionTreeClassifier
```

In [44]:
```
dtc = DecisionTreeClassifier()
```

In [45]:
```
dtc.fit(x_train,y_train)
```

Out[45]:
```
▼ DecisionTreeClassifier
DecisionTreeClassifier()
```

## Predicting value

In [53]:
```
sum(titanic_test['Age'].isnull())
```

Out[53]: 86

In [54]:
```
titanic_test=titanic_test.dropna()
```

In [57]:
```
sum(titanic_test['Age'].isnull())
```

Out[57]: 0

In [55]:
```
x_test=titanic_test[['Age']]
```

In [56]:
```
y_pred=dtc.predict(x_test)
```

In [58]:
```
y_pred
```

Out[58]:
```
array([1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0,
       1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1,
       1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1],
      dtype=int64)
```

In [ ]: