importing required library

```
In [5]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

Load the Titanic dataset

```
data = sns.load_dataset("titanic")
In [9]:
         data.head()
Out[9]:
             survived pclass
                                sex
                                    age sibsp parch
                                                          fare embarked class
                                                                                  who adult_n
          0
                                     22.0
                                                        7.2500
                                                                          Third
                               male
                                                                                   man
                                                                                             F
          1
                   1
                           1 female 38.0
                                              1
                                                     0 71.2833
                                                                       С
                                                                           First woman
          2
                                                                                             F
                             female 26.0
                                              0
                                                        7.9250
                                                                       S
                                                                          Third woman
                                                     0
          3
                                                                       S
                                                                                             F
                   1
                             female 35.0
                                              1
                                                     0 53.1000
                                                                           First woman
                   0
                           3
                               male 35.0
                                              0
                                                     0
                                                        8.0500
                                                                       S
                                                                          Third
                                                                                   man
```

In [10]: data.describe()

Out[10]:

	survived	pclass	age	sibsp	parch	fare
count	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

1 of 3 03-03-2024, 10:17 pm

```
In [11]: data.info()
         <class 'pandas.core.frame.DataFrame'>
         RangeIndex: 891 entries, 0 to 890
         Data columns (total 15 columns):
         survived
                         891 non-null int64
                         891 non-null int64
         pclass
                         891 non-null object
         sex
                         714 non-null float64
         age
                         891 non-null int64
         sibsp
                         891 non-null int64
         parch
         fare
                         891 non-null float64
                         889 non-null object
         embarked
         class
                         891 non-null category
         who
                         891 non-null object
         adult_male
                         891 non-null bool
         deck
                         203 non-null category
         embark_town
                         889 non-null object
         alive
                         891 non-null object
                         891 non-null bool
         alone
         dtypes: bool(2), category(2), float64(2), int64(4), object(5)
         memory usage: 80.6+ KB
In [12]:
         data.isnull().sum()
Out[12]: survived
                           0
         pclass
                           0
                           0
         sex
         age
                         177
         sibsp
                           0
                           0
         parch
         fare
                           0
         embarked
                           2
         class
                           0
                           0
         who
         adult_male
                           0
         deck
                         688
                           2
         embark_town
         alive
                           0
         alone
                           0
         dtype: int64
In [18]:
         data['age'] = data['age'].fillna(np.mean(data['age']))
         data['deck'] = data['deck'].fillna(data['deck'].mode()[0])
         data['embarked'] = data['embarked'].fillna(data['embarked'].mode()[0])
```

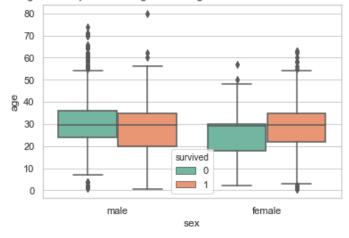
2 of 3 03-03-2024, 10:17 pm

```
In [19]: data.isnull().sum()
Out[19]: survived
                          0
          pclass
                          0
                          0
          sex
          age
                          0
                          0
          sibsp
          parch
                          0
          fare
                          0
          embarked
                          0
          class
                          0
                          0
          who
          adult_male
                          0
          deck
                          0
                          2
          embark_town
          alive
                          0
          alone
                          0
          dtype: int64
```

Vidualize

In [22]: sns.boxplot(data['sex'], data["age"], data["survived"], palette = 'Set
2').set_title('Plot for distribution of age with respect to each gender a
long with the information about whether they survived or not')
plt.show()

Plot for distribution of age with respect to each gender along with the information about whether they survived or not



3 of 3 03-03-2024, 10:17 pm