##Bijen Maharjan 2431319

```python
import pandas as pd
data={
    'Name':['Alice','Bob','Charlie','David'],
    'Age':[24,19,22,25],
    'Score':[88,92,85,95]
}
df=pd.DataFrame(data)
print(df)
```

```
      Name  Age  Score
0    Alice   24     88
1      Bob   19     92
2  Charlie   22     85
3    David   25     95
```

```python
df_soted_by_age=df.sort_values(by='Age')
print(df_soted_by_age.head())
```

```
      Name  Age  Score
1      Bob   19     92
2  Charlie   22     85
0    Alice   24     88
3    David   25     95
```

```python
[7] sorted_by_index=df.sort_index()
print(sorted_by_index.head())
```

```
      Name  Age  Score
0    Alice   24     88
1      Bob   19     92
2  Charlie   22     85
3    David   25     95
```

```
[8] subset_iloc=df.iloc[1:3,0:2] # slicing row and column
    print("subset iloc:")
    print(subset_iloc)
```

```
subset iloc:
      Name  Age
1      Bob   19
2  Charlie   22
```

Loading...

```
    first_two_rows=df.iloc[1:2]
    print("\nfirst two rows\n",first_two_rows)
```

```
first two rows
   Name  Age  Score
1  Bob   19     92
```

```
[10] subset_loc=df.loc[df['Age']>20,['Name','Score']]
     print("\nSubset using loc:")
     print(subset_loc)
```

```
Subset using loc:
      Name  Score
0    Alice     88
2  Charlie     85
3    David     95
```

```
[12] name_column=df[['Name','Age']]
     print(name_column)
```

```
      Name  Age
0    Alice   24
1      Bob   19
2  Charlie   22
3    David   25
```

```
data={
    'Name':['Alice','Bob','Charlie','David'],
    'Department':['HR','IT','Finance','HR'],
    'Salary':[50000,60000,55000,62000]

}
df=pd.DataFrame(data)
print(df)
```

```
      Name Department  Salary
0    Alice         HR   50000
1      Bob         IT   60000
2  Charlie    Finance   55000
3    David         HR   62000
```

```
[15] filtered_single=df[df['Department']=='HR']
     print("\n filtered single column:",filtered_single)
```

```
 filtered single column:     Name Department  Salary
0  Alice         HR   50000
3  David         HR   62000
```

```
[17] filtered_multiple=df.loc[df['Department'].isin(['HR','IT'])]
     print("\n filtered multiple column:",filtered_multiple)
```

```
 filtered multiple column:     Name Department  Salary
0  Alice         HR   50000
1    Bob         IT   60000
3  David         HR   62000
```

```python
data={
    'Category':['A','B','A','B','A'],
    'Value':[10,20,30,40,50]
}
df=pd.DataFrame(data)
grouped=df.groupby('Category')['Value'].mean()
print(grouped)
```

```
Category
A    30.0
B    30.0
Name: Value, dtype: float64
```

```python
[20] df['Normalized']=df.groupby('Category')['Value'].transform(lambda x:x/x.sum())
     print(df)
```

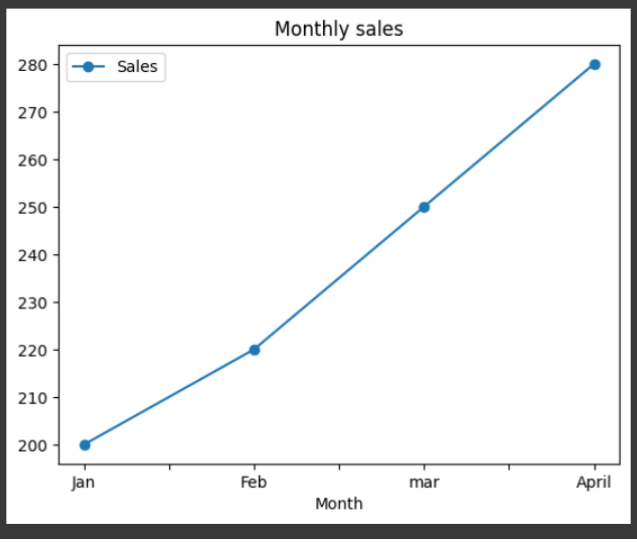
```
  Category  Value  Normalized
0        A     10    0.111111
1        B     20    0.333333
2        A     30    0.333333
3        B     40    0.666667
4        A     50    0.555556
```

```python
filtered=df.groupby('Category').filter(lambda x:x['Value'].sum()>60)
print(filtered)
```

```
  Category  Value  Normalized
0        A     10    0.111111
2        A     30    0.333333
4        A     50    0.555556
```

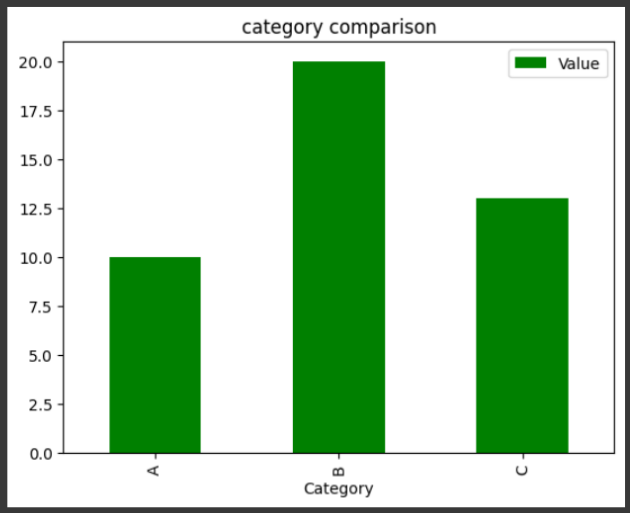
```python
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
data={'Month':['Jan','Feb','mar','April'],
      'Sales':[200,220,250,280]

      }
df=pd.DataFrame(data)
df.plot(x='Month',y='Sales',kind='line',marker='o',title='Monthly sales')
plt.show()
```



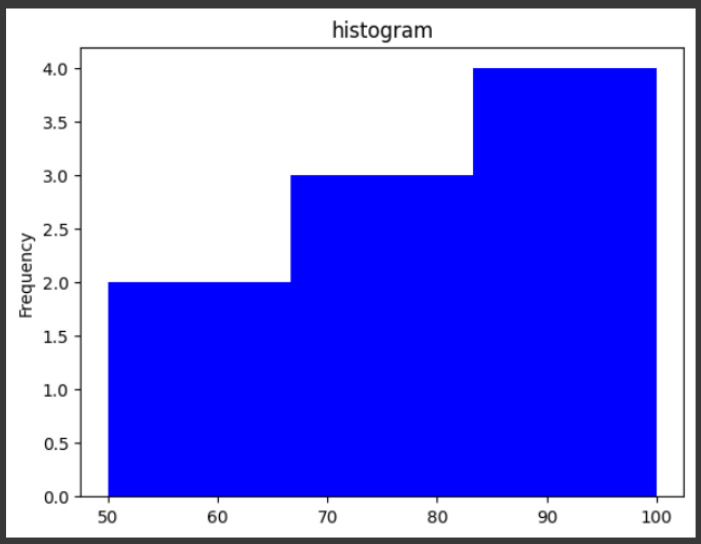

```python
data={'Category':['A','B','C'],
      'Value':[10,20,13]

      }
df=pd.DataFrame(data)
df.plot(x='Category',y='Value',kind='bar',title='category comparison',color='green')
plt.show()
```



```
data={
    'Score':[50,60,70,75,80,85,90,95,100]

}
df=pd.DataFrame(data)
df['Score'].plot(kind='hist',bins=3,title='histogram',color='blue')
plt.show()
```




```
import pandas as pd
df=pd.read_csv('/content/drive/MyDrive/Worksheet3/Titanic-Dataset.csv')
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  891 non-null    int64
 1   Survived     891 non-null    int64
 2   Pclass       891 non-null    int64
 3   Name         891 non-null    object
 4   Sex          891 non-null    object
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64
 7   Parch        891 non-null    int64
 8   Ticket       891 non-null    object
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object
 11  Embarked     889 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

```
fare=df['Fare']
print(fare.head())
```

```
0     7.2500
1    71.2833
2     7.9250
3    53.1000
4     8.0500
Name: Fare, dtype: float64
```

```
[30] class_age=df[['Pclass','Age']]
     print(class_age.head())
```

```
   Pclass   Age
0       3  22.0
1       1  38.0
2       3  26.0
3       1  35.0
4       3  35.0
```

```
survived_gender=df[['Survived','Sex']]
print(survived_gender.head())
```

```
   Survived     Sex
0         0    male
1         1  female
2         1  female
3         1  female
4         0    male
```

```
[32] passanger_Fare_gt=df[df['Fare']>100]
     print(passanger_Fare_gt)
```

```
passanger_Fare_gt=df[df['Fare']>100]
print(passanger_Fare_gt)
```

```
     PassengerId  Survived  Pclass  \
27            28         0       1
31            32         1       1
88            89         1       1
118          119         0       1
195          196         1       1
215          216         1       1
258          259         1       1
268          269         1       1
269          270         1       1
297          298         0       1
299          300         1       1
305          306         1       1
306          307         1       1
307          308         1       1
311          312         1       1
318          319         1       1
319          320         1       1
325          326         1       1
332          333         0       1
334          335         1       1
337          338         1       1
341          342         1       1
373          374         0       1
377          378         0       1
380          381         1       1
390          391         1       1
393          394         1       1
435          436         1       1
438          439         0       1
498          499         0       1
```

```
     PassengerId  Survived  Pclass  \
498          499         0       1
505          506         0       1
527          528         0       1
537          538         1       1
544          545         0       1
550          551         1       1
557          558         0       1
581          582         1       1
609          610         1       1
659          660         0       1
660          661         1       1
679          680         1       1
689          690         1       1
698          699         0       1
700          701         1       1
708          709         1       1
716          717         1       1
730          731         1       1
737          738         1       1
742          743         1       1
763          764         1       1
779          780         1       1
802          803         1       1
856          857         1       1

                                                Name     Sex     Age  SibSp  \
27                          Fortune, Mr. Charles Alexander    male   19.00      3
31          Spencer, Mrs. William Augustus (Marie Eugenie)  female     NaN      1
88                              Fortune, Miss. Mabel Helen  female   23.00      3
118                           Baxter, Mr. Quigg Edmond      male   24.00      0
195                               Lurette, Miss. Elise    female   58.00      0
215                            Newell, Miss. Madeleine    female   31.00      1
```

```
     498          499         0       1
     505          506         0       1
     527          528         0       1
     537          538         1       1
     544          545         0       1
     550          551         1       1
     557          558         0       1
     581          582         1       1
     609          610         1       1
     659          660         0       1
     660          661         1       1
     679          680         1       1
     689          690         1       1
     698          699         0       1
     700          701         1       1
     708          709         1       1
     716          717         1       1
     730          731         1       1
     737          738         1       1
     742          743         1       1
     763          764         1       1
     779          780         1       1
     802          803         1       1
     856          857         1       1

                                                  Name     Sex    Age  SibSp  \
     27                       Fortune, Mr. Charles Alexander    male  19.00      3
     31       Spencer, Mrs. William Augustus (Marie Eugenie)  female    NaN      1
     88                               Fortune, Miss. Mabel Helen  female  23.00      3
     118                              Baxter, Mr. Quigg Edmond    male  24.00      0
     195                                  Lurette, Miss. Elise  female  58.00      0
     215                              Newell, Miss. Madeleine  female  31.00      1
```

```
     527                                 Farthing, Mr. John    male    NaN      0
     537                               LeRoy, Miss. Bertha  female  30.00      0
     544                            Douglas, Mr. Walter Donald    male  50.00      1
     550                           Thayer, Mr. John Borland Jr    male  17.00      0
     557                                 Robbins, Mr. Victor    male    NaN      0
     581  Thayer, Mrs. John Borland (Marian Longstreth M...  female  39.00      1
     609                               Shutes, Miss. Elizabeth W  female  40.00      0
     659                             Newell, Mr. Arthur Webster    male  58.00      0
     660                          Frauenthal, Dr. Henry William    male  50.00      2
     679                      Cardeza, Mr. Thomas Drake Martinez    male  36.00      0
     689                       Madill, Miss. Georgette Alexandra  female  15.00      0
     698                               Thayer, Mr. John Borland    male  49.00      1
     700  Astor, Mrs. John Jacob (Madeleine Talmadge Force)  female  18.00      1
     708                                 Cleaver, Miss. Alice  female  22.00      0
     716                          Endres, Miss. Caroline Louise  female  38.00      0
     730                          Allen, Miss. Elisabeth Walton  female  29.00      0
     737                                 Lesurer, Mr. Gustave J    male  35.00      0
     742                 Ryerson, Miss. Susan Parker "Suzette"  female  21.00      2
     763              Carter, Mrs. William Ernest (Lucile Polk)  female  36.00      1
     779  Robert, Mrs. Edward Scott (Elisabeth Walton Mc...  female  43.00      0
     802                    Carter, Master. William Thornton II    male  11.00      1
     856            Wick, Mrs. George Dennick (Mary Hitchcock)  female  45.00      1

          Parch     Ticket      Fare            Cabin Embarked
     27       2      19950  263.0000      C23 C25 C27        S
     31       0  PC 17569  146.5208              B78        C
     88       2      19950  263.0000      C23 C25 C27        S
     118      1  PC 17558  247.5208          B58 B60        C
     195      0  PC 17569  146.5208              B80        C
     215      0      35273  113.2750              D36        C
     258      0  PC 17755  512.3292              NaN        C
     268      1  PC 17582  153.4625             C125        S
```

```
filtered_Pclass=df[df['Pclass']==1]
print(filtered_Pclass)
```

```
     PassengerId  Survived  Pclass  \
1              2         1       1
3              4         1       1
6              7         0       1
11            12         1       1
23            24         1       1
..           ...       ...     ...
871         872         1       1
872         873         0       1
879         880         1       1
887         888         1       1
889         890         1       1

                                                  Name     Sex   Age  SibSp  \
1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
3         Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
6                              McCarthy, Mr. Timothy J    male  54.0      0
11                          Bonnell, Miss. Elizabeth  female  58.0      0
23                       Sloper, Mr. William Thompson    male  28.0      0
..                                                ...     ...   ...    ...
871   Beckwith, Mrs. Richard Leonard (Sallie Monypeny)  female  47.0      1
872                           Carlsson, Mr. Frans Olof    male  33.0      0
879      Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)  female  56.0      0
887                        Graham, Miss. Margaret Edith  female  19.0      0
889                              Behr, Mr. Karl Howell    male  26.0      0

     Parch    Ticket     Fare        Cabin Embarked
1        0  PC 17599  71.2833         C85        C
3        0    113803  53.1000        C123        S
6        0     17463  51.8625         E46        S
11       0    113783  26.5500        C103        S
23       0    113788  35.5000          A6        S
..     ...       ...      ...         ...      ...
871      1     11751  52.5542         D35        S
872      0       695   5.0000  B51 B53 B55        S
879      1     11767  83.1583         C50        C
887      0    112053  30.0000         B42        S
889      0    111369  30.0000        C148        C

[216 rows x 12 columns]
```

```
[34] passangerAge=df[(df['Age']<18) & (df['Sex']=='female')]
     print(passangerAge.head())
```

```
     PassengerId  Survived  Pclass                                     Name  \
9             10         1       2        Nasser, Mrs. Nicholas (Adele Achem)
10            11         1       3              Sandstrom, Miss. Marguerite Rut
14            15         0       3       Vestrom, Miss. Hulda Amanda Adolfina
22            23         1       3                   McGowan, Miss. Anna "Annie"
24            25         0       3                  Palsson, Miss. Torborg Danira

        Sex   Age  SibSp  Parch   Ticket     Fare Cabin Embarked
9    female  14.0      1      0   237736  30.0708   NaN        C
10   female   4.0      1      1  PP 9549  16.7000    G6        S
14   female  14.0      0      0   350406   7.8542   NaN        S
22   female  15.0      0      0   330923   8.0292   NaN        Q
24   female   8.0      3      1   349909  21.0750   NaN        S
```

```
embarked_c_or_s=df[(df['Embarked']=='C') | (df['Embarked']=='S')]
print(embarked_c_or_s.head())
```

```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3
3            4         1       1
4            5         0       3

                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2                             Heikkinen, Miss. Laina  female  26.0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                           Allen, Mr. William Henry    male  35.0      0

   Parch            Ticket     Fare Cabin Embarked
0      0         A/5 21171   7.2500   NaN        S
1      0          PC 17599  71.2833   C85        C
2      0  STON/O2. 3101282   7.9250   NaN        S
3      0            113803  53.1000  C123        S
4      0            373450   8.0500   NaN        S
```

```
filtered_Pclass=df[(df['Pclass'].isin([1,2]))]
print(filtered_Pclass)
```

```
     PassengerId  Survived  Pclass  \
1              2         1       1
3              4         1       1
6              7         0       1
9             10         1       2
11            12         1       1
..           ...       ...     ...
880           881         1       2
883           884         0       2
886           887         0       2
887           888         1       1
889           890         1       1

                                                 Name     Sex   Age  SibSp  \
1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
3         Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
6                              McCarthy, Mr. Timothy J    male  54.0      0
9                  Nasser, Mrs. Nicholas (Adele Achem)  female  14.0      1
11                             Bonnell, Miss. Elizabeth  female  58.0      0
..                                                  ...     ...   ...    ...
880       Shelley, Mrs. William (Imanita Parrish Hall)  female  25.0      0
883                       Banfield, Mr. Frederick James    male  28.0      0
886                              Montvila, Rev. Juozas    male  27.0      0
887                       Graham, Miss. Margaret Edith  female  19.0      0
889                              Behr, Mr. Karl Howell    male  26.0      0

     Parch            Ticket     Fare Cabin Embarked
1        0          PC 17599  71.2833   C85        C
3        0            113803  53.1000  C123        S
6        0             17463  51.8625   E46        S
9        0            237736  30.0708   NaN        C
11       0            113783  26.5500  C103        S
..     ...               ...      ...   ...      ...
880      1            230433  26.0000   NaN        S
883      0  C.A./SOTON 34068  10.5000   NaN        S
886      0            211536  13.0000   NaN        S
887      0            112053  30.0000   B42        S
889      0            111369  30.0000  C148        C

[400 rows x 12 columns]
```

```
[37] df['Fare_per_year']=df['Fare']/df['Age']
     print(df.head())
```

```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3
3            4         1       1
4            5         0       3

                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2                             Heikkinen, Miss. Laina  female  26.0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                           Allen, Mr. William Henry    male  35.0      0

   Parch            Ticket     Fare Cabin Embarked  Fare_per_year
0      0         A/5 21171   7.2500   NaN        S       0.329545
1      0          PC 17599  71.2833   C85        C       1.875876
2      0  STON/O2. 3101282   7.9250   NaN        S       0.304808
3      0            113803  53.1000  C123        S       1.517143
4      0            373450   8.0500   NaN        S       0.230000
```

```
high=df[df['Fare_per_year']>5]
print(high)
```

```
     PassengerId  Survived  Pclass                                     Name  \
7              8         0       3              Palsson, Master. Gosta Leonard
16            17         0       3                      Rice, Master. Eugene
27            28         0       1              Fortune, Mr. Charles Alexander
43            44         1       2  Laroche, Miss. Simonne Marie Anne Andree
50            51         0       3                  Panula, Master. Juha Niilo
..           ...       ...     ...                                        ...
813          814         0       3          Andersson, Miss. Ebba Iris Alfrida
824          825         0       3                Panula, Master. Urho Abraham
827          828         1       2                       Mallet, Master. Andre
831          832         1       2              Richards, Master. George Sibley
850          851         0       3   Andersson, Master. Sigvard Harald Elias

         Sex    Age  SibSp  Parch           Ticket      Fare        Cabin  \
7       male   2.00      3      1           349909   21.0750          NaN
16      male   2.00      4      1           382652   29.1250          NaN
27      male  19.00      3      2            19950  263.0000  C23 C25 C27
43    female   3.00      1      2    SC/Paris 2123   41.5792          NaN
50      male   7.00      4      1          3101295   39.6875          NaN
..       ...    ...    ...    ...              ...       ...          ...
813   female   6.00      4      2           347082   31.2750          NaN
824     male   2.00      4      1          3101295   39.6875          NaN
827     male   1.00      0      2  S.C./PARIS 2079   37.0042          NaN
831     male   0.83      1      1            29106   18.7500          NaN
850     male   4.00      4      2           347082   31.2750          NaN

     Embarked  Fare_per_year
7           S      10.537500
16          Q      14.562500
27          S      13.842105
43          C      13.859733
50          S       5.669643
..        ...            ...
813         S       5.212500
824         S      19.843750
827         C      37.004200
831         S      22.590361
850         S       7.818750

[68 rows x 13 columns]
```

```
[39] high_fare_age_srt=df.sort_values(by='Fare_per_year',ascending=False)
     print(high_fare_age_srt.head())
```

```
     PassengerId  Survived  Pclass                                Name     Sex  \
305          306         1       1   Allison, Master. Hudson Trevor    male
297          298         0       1     Allison, Miss. Helen Loraine  female
386          387         0       3  Goodwin, Master. Sidney Leonard    male
164          165         0       3     Panula, Master. Eino Viljami    male
183          184         1       2        Becker, Master. Richard F    male

      Age  SibSp  Parch   Ticket      Fare    Cabin Embarked  Fare_per_year
305  0.92      1      2   113781  151.5500  C22 C26        S      164.728261
297  2.00      1      2   113781  151.5500  C22 C26        S       75.775000
386  1.00      5      2  CA 2144   46.9000      NaN        S       46.900000
164  1.00      4      1  3101295   39.6875      NaN        S       39.687500
183  1.00      2      1   230136   39.0000       F4        S       39.000000
```

```
result=high_fare_age_srt[['Fare_per_year','Name']]
print(result)
```

```
     Fare_per_year                                     Name
305     164.728261           Allison, Master. Hudson Trevor
297      75.775000             Allison, Miss. Helen Loraine
386      46.900000          Goodwin, Master. Sidney Leonard
164      39.687500             Panula, Master. Eino Viljami
183      39.000000                Becker, Master. Richard F
..             ...                                      ...
859            NaN                         Razi, Mr. Raihed
863            NaN  Sage, Miss. Dorothy Edith "Dolly"
868            NaN               van Melkebeke, Mr. Philemon
878            NaN                           Laleff, Mr. Kristo
888            NaN  Johnston, Miss. Catherine Helen "Carrie"

[891 rows x 2 columns]
```

```
[41] total=df['Fare'].sum()
     print("Total Fare:",total)
```

```
Total Fare: 28693.9493
```

```
[42] firstClass=df[df['Pclass']==1]['Fare']
     totalf=firstClass.sum()
     print(totalf)
```

```
18177.4125
```

```
[43] secondClass=df[df['Pclass']==2]['Fare']
     totalS=secondClass.sum()
     print("Total of second class:",totalS)
     thirdClass=df[df['Pclass']==3]['Fare']
     totalT=thirdClass.sum()
     print("Total of third class:",totalT)
```

```
Total of second class: 3801.8417
Total of third class: 6714.6951
```

```
     lst=[total,totalS,totalT]
     print(lst)
```

```
[28693.9493, 3801.8417, 6714.6951]
```

+ Code + Text

```
[45] proptionfirst=totalf/total
     proptionsecond=totalS/total
     proptionthird=totalT/total
     print("First:",proptionfirst)
     print("Second:",proptionsecond)
     print("Third:",proptionthird)
```

```
First: 0.6334928771899656
Second: 0.1324962855496507
Third: 0.23401083726038366
```

```
     df['ageGroup']=df['Age'].apply(lambda Age:'child' if Age<18 else 'adult' if 18<= Age<=65 else 'senior' if Age>65 else 'invalid')
     print(df.head())
```

```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3
3            4         1       1
4            5         0       3

                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2                             Heikkinen, Miss. Laina  female  26.0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                           Allen, Mr. William Henry    male  35.0      0

   Parch            Ticket     Fare Cabin Embarked  Fare_per_year ageGroup
0      0         A/5 21171   7.2500   NaN        S       0.329545    adult
1      0          PC 17599  71.2833   C85        C       1.875876    adult
2      0  STON/O2. 3101282   7.9250   NaN        S       0.304808    adult
3      0            113803  53.1000  C123        S       1.517143    adult
4      0            373450   8.0500   NaN        S       0.230000    adult
```

```
[47] total_passenger=len(df)
     print(total_passenger)
```

```
891
```

```
[48] numberChild=len(df[df['ageGroup']=='child'])
     numberAdult=len(df[df['ageGroup']=='adult'])
     numberSenior=len(df[df['ageGroup']=='senior'])
     print("Number of children:",numberChild)
     print("Number of adult:",numberAdult)
     print("Number of senior:",numberSenior)
```

```
Number of children: 113
Number of adult: 593
Number of senior: 8
```

```
[49] countChildprop= numberChild/total_passenger
     countAdultprop=numberAdult/total_passenger
     countSeniorprop=numberSenior/total_passenger
     print("Proportion of children:",countChildprop)
     print("Proportion of adult:",countAdultprop)
     print("Proportion of senior:",countSeniorprop)
```

```
Proportion of children: 0.12682379349046016
Proportion of adult: 0.6655443322109988
Proportion of senior: 0.008978675645342313
```

```
print("Proportion of children:",countChildprop,"%")
print("Proportion of adult:",countAdultprop,"%")
print("Proportion of senior:",countSeniorprop,"%")
```

```
Proportion of children: 0.12682379349046016 %
Proportion of adult: 0.6655443322109988 %
Proportion of senior: 0.008978675645342313 %
```