Name: Samvid Shrestha

Worksheet3 Submission

```python
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

df = pd.read_csv('/content/drive/MyDrive/Concept And Technology Of AI/Titanic-Dataset.csv')

print(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
None
```

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

```
      Fare
0   7.2500
1  71.2833
2   7.9250
3  53.1000
4   8.0500
```

```
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
```

```
fare_gt_100 = df[df['Fare'] > 100]
print(fare_gt_100)
```

```
     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
```

| | 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 | |
| 258 | Ward, Miss. Anna | female | 35.00 | 0 | |
| 268 | Graham, Mrs. William Thompson (Edith Junkins) | female | 58.00 | 0 | |
| 269 | Bissette, Miss. Amelia | female | 35.00 | 0 | |
| 297 | Allison, Miss. Helen Loraine | female | 2.00 | 1 | |
| 299 | Baxter, Mrs. James (Helene DeLaudeniere Chaput) | female | 50.00 | 0 | |
| 305 | Allison, Master. Hudson Trevor | male | 0.92 | 1 | |
| 306 | Fleming, Miss. Margaret | female | NaN | 0 | |
| 307 | Penasco y Castellana, Mrs. Victor de Satode (M... | female | 17.00 | 1 | |
| 311 | Ryerson, Miss. Emily Borie | female | 18.00 | 2 | |
| 318 | Wick, Miss. Mary Natalie | female | 31.00 | 0 | |
| 319 | Spedden, Mrs. Frederic Oakley (Margaretta Corn... | female | 40.00 | 1 | |
| 325 | Young, Miss. Marie Grice | female | 36.00 | 0 | |
| 332 | Graham, Mr. George Edward | male | 38.00 | 0 | |
| 334 | Frauenthal, Mrs. Henry William (Clara Heinshei... | female | NaN | 1 | |
| 337 | Burns, Miss. Elizabeth Margaret | female | 41.00 | 0 | |
| 341 | Fortune, Miss. Alice Elizabeth | female | 24.00 | 3 | |
| 373 | Ringhini, Mr. Sante | male | 22.00 | 0 | |

| | 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 |
| 269 | 0 | PC 17760 | 135.6333 | C99 | S |
| 297 | 2 | 113781 | 151.5500 | C22 C26 | S |
| 299 | 1 | PC 17558 | 247.5208 | B58 B60 | C |
| 305 | 2 | 113781 | 151.5500 | C22 C26 | S |
| 306 | 0 | 17421 | 110.8833 | NaN | C |
| 307 | 0 | PC 17758 | 108.9000 | C65 | C |
| 311 | 2 | PC 17608 | 262.3750 | B57 B59 B63 B66 | C |
| 318 | 2 | 36928 | 164.8667 | C7 | S |
| 319 | 1 | 16966 | 134.5000 | E34 | C |
| 325 | 0 | PC 17760 | 135.6333 | C32 | C |
| 332 | 1 | PC 17582 | 153.4625 | C91 | S |
| 334 | 0 | PC 17611 | 133.6500 | NaN | S |
| 337 | 0 | 16966 | 134.5000 | E40 | C |
| 341 | 2 | 19950 | 263.0000 | C23 C25 C27 | S |
| 373 | 0 | PC 17760 | 135.6333 | NaN | C |
| 377 | 2 | 113503 | 211.5000 | C82 | C |
| 380 | 0 | PC 17757 | 227.5250 | NaN | C |
| 390 | 2 | 113760 | 120.0000 | B96 B98 | S |
| 393 | 0 | 35273 | 113.2750 | D36 | C |
| 435 | 2 | 113760 | 120.0000 | B96 B98 | S |
| 438 | 4 | 19950 | 263.0000 | C23 C25 C27 | S |
| 498 | 2 | 113781 | 151.5500 | C22 C26 | S |
| 505 | 0 | PC 17758 | 108.9000 | C65 | C |
| 527 | 0 | PC 17483 | 221.7792 | C95 | S |
| 537 | 0 | PC 17761 | 106.4250 | NaN | C |
| 544 | 0 | PC 17761 | 106.4250 | C86 | C |
| 550 | 2 | 17421 | 110.8833 | C70 | C |

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

```
     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
```

```
female_under_18 = df[(df['Age'] < 18) & (df['Sex'] == 'female')]
print(female_under_18)
```

```
830           Yasbeck, Mrs. Antoni (Selini Alexander)  female  15.00      1
852                             Boulos, Miss. Nourelain  female   9.00      1
853                           Lines, Miss. Mary Conover  female  16.00      0
875                 Najib, Miss. Adele Kiamie "Jane"  female  15.00      0

     Parch         Ticket      Fare    Cabin Embarked
9        0         237736   30.0708      NaN        C
10       1        PP 9549   16.7000       G6        S
14       0         350406    7.8542      NaN        S
22       0         330923    8.0292      NaN        Q
24       1         349909   21.0750      NaN        S
39       0           2651   11.2417      NaN        C
43       2  SC/Paris 2123   41.5792      NaN        C
58       2     C.A. 34651   27.7500      NaN        S
68       2        3101281    7.9250      NaN        S
71       2        CA 2144   46.9000      NaN        S
84       0     SO/C 14885   10.5000      NaN        S
111      0           2665   14.4542      NaN        C
114      0           2627   14.4583      NaN        C
119      2         347082   31.2750      NaN        S
147      2     W./C. 6608   34.3750      NaN        S
156      0          35851    7.7333      NaN        Q
172      1         347742   11.1333      NaN        S
184      2         315153   22.0250      NaN        S
205      1         347054   10.4625       G6        S
208      0         367231    7.7500      NaN        Q
233      2         347077   31.3875      NaN        S
237      2     C.A. 31921   26.2500      NaN        S
297      2         113781  151.5500  C22 C26        S
307      0       PC 17758  108.9000      C65        C
329      1         111361   57.9792      B18        C
374      1         349909   21.0750      NaN        S
381      2           2653   15.7417      NaN        C
389      0        SC 1748   12.0000      NaN        C
419      2         345773   24.1500      NaN        S
```

```
embarked_c_or_s = df[df['Embarked'].isin(['C', 'S'])]
print(embarked_c_or_s)
```

```
     PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           ...       ...     ...
884          885         0       3
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1

                                                 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
..                                               ...     ...   ...    ...
884                           Sutehall, Mr. Henry Jr    male  25.0      0
886                            Montvila, Rev. Juozas    male  27.0      0
887                     Graham, Miss. Margaret Edith  female  19.0      0
888         Johnston, Miss. Catherine Helen "Carrie"  female   NaN      1
889                            Behr, Mr. Karl Howell    male  26.0      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
..     ...               ...      ...   ...      ...
884      0   SOTON/OQ 392076   7.0500   NaN        S
886      0            211536  13.0000   NaN        S
```

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

```
     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
```

```python
df['Age'] = df['Age'].fillna(df['Age'].median())
df['fare_per_year'] = df['Fare'] / df['Age']
high_fare_age = df[df['fare_per_year'] > 5]
high_fare_age_srt = high_fare_age.sort_values(by='fare_per_year', ascending=False)
result = high_fare_age_srt[['Name', 'fare_per_year']]
print(result)
```

```
                                          Name  fare_per_year
305                 Allison, Master. Hudson Trevor     164.728261
297                   Allison, Miss. Helen Loraine      75.775000
386                Goodwin, Master. Sidney Leonard      46.900000
164                   Panula, Master. Eino Viljami      39.687500
183                      Becker, Master. Richard F      39.000000
..                                             ...            ...
348          Coutts, Master. William Loch "William"       5.300000
31   Spencer, Mrs. William Augustus (Marie Eugenie)       5.232886
205                       Strom, Miss. Telma Matilda       5.231250
813              Andersson, Miss. Ebba Iris Alfrida       5.212500
480                  Goodwin, Master. Harold Victor       5.211111

[71 rows x 2 columns]
```

```python
df['fare_per_class'] = df['Fare'] / df['Pclass']
adult_males = df[(df['Sex'] == 'male') & (df['Age'] >= 18)]
adult_males_srt = adult_males.sort_values(by='fare_per_class', ascending=False)
result = adult_males_srt[['Name', 'Age', 'fare_per_class']]
print(result)
```

```
                                  Name   Age  fare_per_class
737             Lesurer, Mr. Gustave J  35.0        512.3292
679  Cardeza, Mr. Thomas Drake Martinez  36.0        512.3292
27       Fortune, Mr. Charles Alexander  19.0        263.0000
438                  Fortune, Mr. Mark  64.0        263.0000
118           Baxter, Mr. Quigg Edmond  24.0        247.5208
..                                 ...   ...             ...
179                Leonard, Mr. Lionel  36.0          0.0000
732               Knight, Mr. Robert J  28.0          0.0000
822     Reuchlin, Jonkheer. John George  38.0          0.0000
806              Andrews, Mr. Thomas Jr  39.0          0.0000
815                   Fry, Mr. Richard  28.0          0.0000

[519 rows x 3 columns]
```

```
total_fare = df['Fare'].sum()
first_class_fare = df[df['Pclass'] == 1]['Fare'].sum()
second_class_fare = df[df['Pclass'] == 2]['Fare'].sum()
third_class_fare = df[df['Pclass'] == 3]['Fare'].sum()
fare_totals = [first_class_fare, second_class_fare, third_class_fare]
proportions = [total / total_fare for total in fare_totals]
```

```
def age_group(age):
    if age < 18:
        return 'child'
    elif age <= 64:
        return 'adult'
    else:
        return 'senior'

df['age_group'] = df['Age'].apply(age_group)
total_passengers = df.shape[0]
age_counts = df['age_group'].value_counts()
age_percentages = (age_counts / total_passengers) * 100
print(age_percentages)
```

```
age_group
adult     86.083053
child     12.682379
senior     1.234568
Name: count, dtype: float64
```