

```
In [1]: import pandas as pd

url = 'https://raw.githubusercontent.com/datasciencedojo/datasets/master/titanic.csv'
df = pd.read_csv(url)

df.head()
```

```
Out[1]:
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...)	female	38.0	1	0	PC 17599	71.2833	C8
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C12
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN

```
In [2]: # Summary statistics by Pclass
summary_stats = df.groupby('Pclass')[['Age', 'Fare']].agg(['mean', 'median', 'min', 'max', 'std'])
print(summary_stats)
```

Pclass	Age					Fare			
	mean	median	min	max	std	mean	median	min	
1	38.233441	37.0	0.92	80.0	14.802856	84.154687	60.2875	0.0	
2	29.877630	29.0	0.67	70.0	14.001077	20.662183	14.2500	0.0	
3	25.140620	24.0	0.42	74.0	12.495398	13.675550	8.0500	0.0	

Pclass	max		std	
	Age	Fare	Age	Fare
1	512.3292	512.3292	78.380373	78.380373
2	73.5000	73.5000	13.417399	13.417399
3	69.5500	69.5500	11.778142	11.778142

```
In [8]:
```

Pclass	Age					Fare			
	mean	median	min	max	std	mean	median	min	
1	38.233441	37.0	0.92	80.0	14.802856	84.154687	60.2875	0.0	
2	29.877630	29.0	0.67	70.0	14.001077	20.662183	14.2500	0.0	
3	25.140620	24.0	0.42	74.0	12.495398	13.675550	8.0500	0.0	

	max	std
Pclass		
1	512.3292	78.380373
2	73.5000	13.417399
3	69.5500	11.778142

In [9]:

```
age_lists_by_class = {cls: df[df['Pclass'] == cls]['Age'].dropna().tolist() for cls in df['Pclass'].unique()}
print(age_lists_by_class)
```

```
{np.int64(3): [22.0, 26.0, 35.0, 2.0, 27.0, 4.0, 20.0, 39.0, 14.0, 2.0, 31.0, 15.0, 8.0, 38.0, 21.0, 18.0, 14.0, 40.0, 19.0, 18.0, 7.0, 21.0, 28.5, 11.0, 22.0, 4.0, 19.0, 17.0, 26.0, 16.0, 26.0, 32.0, 25.0, 30.0, 22.0, 29.0, 33.0, 16.0, 24.0, 29.0, 20.0, 26.0, 59.0, 28.0, 33.0, 37.0, 28.0, 21.0, 38.0, 14.5, 22.0, 20.0, 17.0, 21.0, 70.5, 2.0, 12.0, 24.0, 45.0, 33.0, 20.0, 47.0, 16.0, 22.0, 24.0, 19.0, 27.0, 9.0, 55.5, 40.5, 16.0, 30.0, 44.0, 26.0, 17.0, 1.0, 9.0, 45.0, 28.0, 4.0, 1.0, 21.0, 18.0, 36.0, 9.0, 4.0, 40.0, 36.0, 19.0, 42.0, 28.0, 34.0, 45.5, 18.0, 2.0, 32.0, 26.0, 16.0, 24.0, 22.0, 27.0, 16.0, 51.0, 22.0, 20.5, 29.0, 5.0, 22.0, 30.0, 25.0, 29.0, 30.0, 41.0, 29.0, 3.0, 16.0, 25.0, 25.0, 45.0, 7.0, 35.0, 65.0, 28.0, 16.0, 19.0, 33.0, 30.0, 22.0, 22.0, 24.0, 24.0, 23.5, 19.0, 28.0, 26.0, 22.0, 27.0, 61.0, 31.0, 16.0, 45.0, 3.0, 42.0, 23.0, 15.0, 25.0, 28.0, 40.0, 45.0, 35.0, 30.0, 18.0, 19.0, 3.0, 22.0, 20.0, 19.0, 1.0, 32.0, 1.0, 21.0, 28.0, 24.0, 22.0, 31.0, 39.0, 26.0, 21.0, 28.0, 20.0, 51.0, 21.0, 44.0, 10.0, 21.0, 29.0, 28.0, 18.0, 32.0, 17.0, 21.0, 20.0, 25.0, 5.0, 29.0, 34.0, 38.0, 0.75, 38.0, 22.0, 29.0, 22.0, 2.0, 9.0, 50.0, 63.0, 30.0, 9.0, 21.0, 21.0, 24.0, 17.0, 21.0, 37.0, 28.0, 26.0, 29.0, 24.0, 32.0, 22.0, 40.5, 39.0, 17.0, 30.0, 9.0, 11.0, 33.0, 22.0, 22.0, 36.0, 40.0, 24.0, 19.0, 29.0, 32.0, 16.0, 19.0, 32.0, 22.0, 35.0, 47.0, 36.0, 49.0, 44.0, 36.0, 30.0, 39.0, 35.0, 34.0, 26.0, 27.0, 20.0, 21.0, 21.0, 26.0, 51.0, 9.0, 32.0, 41.0, 20.0, 2.0, 0.75, 19.0, 23.0, 21.0, 18.0, 32.0, 40.0, 36.0, 20.0, 43.0, 18.0, 24.5, 18.0, 43.0, 20.0, 14.0, 14.0, 19.0, 18.0, 4.0, 25.0, 44.0, 42.0, 18.0, 25.0, 26.0, 29.0, 19.0, 33.0, 17.0, 20.0, 25.0, 11.0, 28.5, 48.0, 24.0, 31.0, 16.0, 31.0, 6.0, 33.0, 23.0, 28.0, 34.0, 41.0, 20.0, 16.0, 30.5, 32.0, 24.0, 48.0, 18.0, 5.0, 13.0, 25.0, 25.0, 18.0, 8.0, 1.0, 25.0, 31.0, 30.0, 30.0, 0.42, 27.0, 31.0, 18.0, 26.0, 39.0, 6.0, 30.5, 23.0, 43.0, 10.0, 27.0, 27.0, 2.0, 15.0, 23.0, 18.0, 21.0, 32.0, 20.0, 34.5, 17.0, 42.0, 35.0, 4.0, 74.0, 9.0, 18.0, 24.0, 41.0, 4.0, 26.0, 47.0, 15.0, 20.0, 19.0, 33.0, 22.0, 25.0, 39.0, 32.0], np.int64(1): [38.0, 35.0, 54.0, 58.0, 28.0, 19.0, 40.0, 28.0, 42.0, 49.0, 65.0, 38.0, 45.0, 28.0, 23.0, 46.0, 71.0, 23.0, 21.0, 47.0, 24.0, 54.0, 19.0, 37.0, 24.0, 22.0, 51.0, 61.0, 56.0, 50.0, 45.0, 44.0, 58.0, 40.0, 31.0, 32.0, 38.0, 35.0, 44.0, 37.0, 62.0, 30.0, 35.0, 52.0, 40.0, 58.0, 35.0, 37.0, 63.0, 26.0, 19.0, 2.0, 50.0, 0.92, 17.0, 30.0, 24.0, 18.0, 31.0, 40.0, 36.0, 16.0, 45.5, 38.0, 29.0, 41.0, 45.0, 24.0, 22.0, 60.0, 24.0, 25.0, 22.0, 27.0, 42.0, 35.0, 36.0, 23.0, 33.0, 28.0, 50.0, 14.0, 64.0, 4.0, 34.0, 52.0, 30.0, 49.0, 65.0, 48.0, 47.0, 56.0, 25.0, 35.0, 58.0, 55.0, 71.0, 54.0, 25.0, 16.0, 18.0, 36.0, 54.0, 47.0, 30.0, 44.0, 45.0, 30.0, 22.0, 36.0, 50.0, 64.0, 17.0, 62.0, 48.0, 39.0, 53.0, 36.0, 39.0, 39.0, 36.0, 18.0, 60.0, 52.0, 49.0, 35.0, 27.0, 40.0, 42.0, 61.0, 21.0, 80.0, 32.0, 24.0, 48.0, 56.0, 58.0, 50.0, 47.0, 31.0, 36.0, 27.0, 15.0, 31.0, 60.0, 49.0, 18.0, 35.0, 42.0, 22.0, 24.0, 48.0, 38.0, 27.0, 29.0, 35.0, 36.0, 21.0, 70.0, 19.0, 33.0, 36.0, 51.0, 43.0, 17.0, 29.0, 46.0, 49.0, 11.0, 39.0, 33.0, 52.0, 38.0, 62.0, 39.0, 30.0, 16.0, 45.0, 51.0, 48.0, 31.0, 47.0, 33.0, 56.0, 19.0, 26.0], np.int64(2): [14.0, 55.0, 35.0, 34.0, 66.0, 27.0, 3.0, 29.0, 21.0, 5.0, 29.0, 32.0, 21.0, 0.83, 17.0, 34.0, 34.0, 29.0, 21.0, 32.5, 32.5, 29.0, 25.0, 23.0, 18.0, 19.0, 36.5, 42.0, 51.0, 40.0, 30.0, 1.0, 32.0, 19.0, 3.0, 24.0, 35.0, 30.0, 42.0, 30.0, 27.0, 19.0, 18.0, 59.0, 24.0, 44.0, 8.0, 19.0, 33.0, 29.0, 24.0, 54.0, 50.0, 36.0, 41.0, 42.0, 36.0, 30.0, 26.0, 43.0, 24.0, 54.0, 30.0, 22.0, 36.0, 2.0, 28.0, 25.0, 36.0, 24.0, 40.0, 38.0, 29.0, 18.0, 36.0, 17.0, 46.0, 23.0, 28.0, 34.0, 3.0, 34.0, 18.0, 30.0, 28.0, 19.0, 42.0, 24.0, 31.0, 45.0, 28.0, 13.0, 36.0, 50.0, 48.0, 33.0, 23.0, 34.0, 33.0, 34.0, 36.0, 50.0, 23.0, 2.0, 7.0, 32.0, 19.0, 8.0, 27.0, 28.0, 62.0, 34.0, 25.0, 54.0, 47.0, 37.0, 24.0, 22.0, 24.0, 4.0, 26.0, 57.0, 28.0, 31.0, 18.0, 24.0, 23.0, 32.0, 25.0, 40.0, 70.0, 31.0, 60.0, 25.0, 52.0, 39.0, 45.0, 52.0, 27.0, 6.0, 34.0, 50.0, 30.0, 25.0, 23.0, 23.0, 30.0, 4.0, 48.0, 0.67, 18.0, 57.0, 54.0, 16.0, 3
```

```
9.0, 34.0, 31.0, 39.0, 35.0, 31.0, 1.0, 0.83, 16.0, 28.0, 44.0, 21.0, 24.0, 4
2.0, 27.0, 28.0, 25.0, 28.0, 27.0]}}
```

In [10]:

```
import numpy as np

for status in df['Survived'].unique():
    ages = df[df['Survived'] == status]['Age'].dropna()
    print(f"Survived = {status}")
    print("Mean:", round(ages.mean(), 2))
    print("Std Dev:", round(ages.std(), 2))
    print("25th percentile:", np.percentile(ages, 25))
    print("Median (50th):", np.percentile(ages, 50))
    print("75th percentile:", np.percentile(ages, 75))
    print("-----")
```

```
Survived = 0
Mean: 30.63
Std Dev: 14.17
25th percentile: 21.0
Median (50th): 28.0
75th percentile: 39.0
-----
Survived = 1
Mean: 28.34
Std Dev: 14.95
25th percentile: 19.0
Median (50th): 28.0
75th percentile: 36.0
-----
```

In []: