data manipulation

In [5]: ▶|
```
1  #Name : Achal Gajanan Ghorad
2  #Roll no. 39
3  #Section :3A
4  #Date:24/08/2024
```

In [23]: ▶|
```
1  # Aim: to perform data manipulation
```

In [4]: ▶|
```
1  import os
```

In [5]: ▶|
```
1  import pandas as pd
```

In [6]: ▶|
```
1  os.getcwd()
```

Out[6]: 'C:\\Users\\ACHAL'

In [7]: ▶|
```
1  os.chdir("C:\\Users\\ACHAL\\OneDrive\\Desktop")
```

In [8]: ▶|
```
1  data=pd.read_csv("archive (5).zip")
```

In [9]: ▶| 1 data

Out[9]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket
0	892	0	3	Kelly, Mr. James	male	34.5	0	0	330911
1	893	1	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272
2	894	0	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276
3	895	0	3	Wirz, Mr. Albert	male	27.0	0	0	315154
4	896	1	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298
...	...	...	...	...	...	...	...	...	...
413	1305	0	3	Spector, Mr. Woolf	male	NaN	0	0	A.5. 3236
414	1306	1	1	Oliva y Ocana, Dona. Fermina	female	39.0	0	0	PC 17758
415	1307	0	3	Saether, Mr. Simon Sivertsen	male	38.5	0	0	SOTON/O.Q. 3101262
416	1308	0	3	Ware, Mr. Frederick	male	NaN	0	0	359309
417	1309	0	3	Peter, Master. Michael J	male	NaN	1	1	2668

418 rows × 12 columns

In [10]: ▶ | 1 `data.head()`

Out[10]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fa
0	892	0	3	Kelly, Mr. James	male	34.5	0	0	330911	7.82
1	893	1	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.00
2	894	0	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.68
3	895	0	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.66
4	896	1	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.28

In [11]: ▶ | 1 `data.tail()`

Out[11]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket
413	1305	0	3	Spector, Mr. Woolf	male	NaN	0	0	A.5. 3236
414	1306	1	1	Oliva y Ocana, Dona. Fermina	female	39.0	0	0	PC 17758
415	1307	0	3	Saether, Mr. Simon Sivertsen	male	38.5	0	0	SOTON/O.Q. 3101262
416	1308	0	3	Ware, Mr. Frederick	male	NaN	0	0	359309
417	1309	0	3	Peter, Master. Michael J	male	NaN	1	1	2668

In [12]: ▶| 1 data.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 418 entries, 0 to 417
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  418 non-null    int64
 1   Survived     418 non-null    int64
 2   Pclass       418 non-null    int64
 3   Name         418 non-null    object
 4   Sex          418 non-null    object
 5   Age          332 non-null    float64
 6   SibSp        418 non-null    int64
 7   Parch        418 non-null    int64
 8   Ticket       418 non-null    object
 9   Fare         417 non-null    float64
 10  Cabin        91 non-null     object
 11  Embarked     418 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 39.3+ KB
```

In [13]: ▶| 1 data.describe()

Out[13]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	
count	418.000000	418.000000	418.000000	332.000000	418.000000	418.000000	417.00
mean	1100.500000	0.363636	2.265550	30.272590	0.447368	0.392344	35.62
std	120.810458	0.481622	0.841838	14.181209	0.896760	0.981429	55.90
min	892.000000	0.000000	1.000000	0.170000	0.000000	0.000000	0.00
25%	996.250000	0.000000	1.000000	21.000000	0.000000	0.000000	7.89
50%	1100.500000	0.000000	3.000000	27.000000	0.000000	0.000000	14.45
75%	1204.750000	1.000000	3.000000	39.000000	1.000000	0.000000	31.50
max	1309.000000	1.000000	3.000000	76.000000	8.000000	9.000000	512.32

In [14]: ▶| 1 data.shape
          2

Out[14]: (418, 12)

In [15]: ▶| 1 data.size

Out[15]: 5016

In [16]: ▶| 1 data.ndim

Out[16]: 2

In [17]: ▶| 1 data.columns

Out[17]: Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'Sib
         Sp',
                'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
               dtype='object')

         1

In [18]: ▶| 1 data.drop(labels="Age",axis=1)

Out[18]:

	PassengerId	Survived	Pclass	Name	Sex	SibSp	Parch	Ticket	
0	892	0	3	Kelly, Mr. James	male	0	0	330911	7.8
1	893	1	3	Wilkes, Mrs. James (Ellen Needs)	female	1	0	363272	7.0
2	894	0	2	Myles, Mr. Thomas Francis	male	0	0	240276	9.0
3	895	0	3	Wirz, Mr. Albert	male	0	0	315154	8.0
4	896	1	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	1	1	3101298	12.2
...	...	...	...	...	...	...	...	...	
413	1305	0	3	Spector, Mr. Woolf	male	0	0	A.5. 3236	8.0
414	1306	1	1	Oliva y Ocana, Dona. Fermina	female	0	0	PC 17758	108.9
415	1307	0	3	Saether, Mr. Simon Sivertsen	male	0	0	SOTON/O.Q. 3101262	7.2
416	1308	0	3	Ware, Mr. Frederick	male	0	0	359309	8.0
417	1309	0	3	Peter, Master. Michael J	male	1	1	2668	22.3

418 rows × 11 columns

In [19]: ▶|
```
1  data.drop(labels=2,axis=0)
```

Out[19]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket
0	892	0	3	Kelly, Mr. James	male	34.5	0	0	330911
1	893	1	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272
3	895	0	3	Wirz, Mr. Albert	male	27.0	0	0	315154
4	896	1	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298
5	897	0	3	Svensson, Mr. Johan Cervin	male	14.0	0	0	7538
...	...	...	...	...	...	...	...	...	...
413	1305	0	3	Spector, Mr. Woolf	male	NaN	0	0	A.5. 3236
414	1306	1	1	Oliva y Ocana, Dona. Fermina	female	39.0	0	0	PC 17758
415	1307	0	3	Saether, Mr. Simon Sivertsen	male	38.5	0	0	SOTON/O.Q. 3101262
416	1308	0	3	Ware, Mr. Frederick	male	NaN	0	0	359309
417	1309	0	3	Peter, Master. Michael J	male	NaN	1	1	2668

417 rows × 12 columns

◄   ►

In [20]: ▶|
```
1
```

In [21]: ▶|
```
1
```

In [22]: ▶|
```
1
```

In [ ]: ▶|
```
1
```