data manipulation

```
M
                  #Name : Achal Gajanan Ghorad
 In [5]:
                  #Roll no. 39
               2
                  #Section :3A
                 #Date:24/08/2024
In [23]:
                 # Aim: to perform data manipulation
In [4]:
               1
                  import os
                  import pandas as pd
In [5]:
 In [6]:
                 os.getcwd()
    Out[6]: 'C:\\Users\\ACHAL'
                  os.chdir("C:\\Users\\ACHAL\\OneDrive\\Desktop")
 In [7]:
                  data=pd.read_csv("archive (5).zip")
 In [8]:
```

In [9]: № 1 data

Ticket	Parch	SibSp	Age	Sex	Name	Pclass	Survived	Passengerld		Out[9]:
330911	0	0	34.5	male	Kelly, Mr. James	3	0	892	0	
363272	0	1	47.0	female	Wilkes, Mrs. James (Ellen Needs)	3	1	893	1	
240276	0	0	62.0	male	Myles, Mr. Thomas Francis	2	0	894	2	
315154	0	0	27.0	male	Wirz, Mr. Albert	3	0	895	3	
3101298	1	1	22.0	female	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	3	1	896	4	
A.5. 3236	0	0	NaN	male	Spector, Mr. Woolf	3	0	1305	413	
PC 17758	0	0	39.0	female	Oliva y Ocana, Dona. Fermina	1	1	1306	414	
SOTON/O.Q. 3101262	0	0	38.5	male	Saether, Mr. Simon Sivertsen	3	0	1307	415	
359309	0	0	NaN	male	Ware, Mr. Frederick	3	0	1308	416	
2668	1	1	NaN	male	Peter, Master. Michael J	3	0	1309	417	

In [10]: ▶	1	data.head()								
Out[10]:	ı	Passengerld	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fi
	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
	4										•
In [11]: ▶	1	data.tail()								
In [11]: N Out[11]:	1	data.tail(d Pclas	s Name	e Se	k Age	s SibSp	o Parch	ı T	icket
	413	Passengerld	Survive	d Pclas	Spector, 3 Mr.	, . male	x Age e NaN				Ticket 3236
		PassengerId	Survive		Spector, 3 Mr.	, male f ' female		l () () A.5.	
	413	PassengerId	Survive	0	Spector, 3 Mr. Woolf Oliva y Ocana, Dona.	, male f , female i	e NaN) () () A.5.) PC 1	3236
	413	PassengerId 1305 1306	Survive	0	Spector, 3 Mr. Woolf Oliva y Cocana, Dona, Fermina Saether, Mr. Simon	male f ' female male male	e NaN) () (C) A.5.) PC 1) SOTON 310	3236 17758
	413 414 415	PassengerId 1305 1306 1307	Survive	0	Spector, 3 Mr. Woolf Oliva y 1 Ocana, Fermina Fermina Saether, 3 Mr. Simon Sivertsen Ware, 3 Mr.	male f f f f female male male	e NaNe 39.0) () () (A.5. PC 1 SOTON 310	3236 17758 1/O.Q. 01262
	413 414 415	PassengerId 1305 1306 1307	Survive	0	Spector, 3 Mr. Woolf Oliva y 1 Ocana, Fermina Saether, 3 Mr. Simon Sivertsen Ware, 4 Mr. Frederick Peter, Master, Michae	male f f f f female male male	e 39.0) () (A.5. PC 1 SOTON 310	3236 17758 1/O.Q. 01262

In [12]:

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
                                  418 non-null
                                                    int64
                    SibSp
               7
                    Parch
                                  418 non-null
                                                    int64
               8
                                                    object
                    Ticket
                                  418 non-null
               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]:
                   data.describe()
    Out[13]:
                      PassengerId
                                    Survived
                                                Pclass
                                                                                  Parch
                                                              Age
                                                                       SibSp
                                  418.000000 418.000000
                       418.000000
                                                        332.000000 418.000000 418.000000 417.00
               count
                      1100.500000
                                               2.265550
                                                                                0.392344
               mean
                                    0.363636
                                                         30.272590
                                                                     0.447368
                                                                                          35.62
                       120.810458
                                    0.481622
                                                                                          55.90
                                               0.841838
                                                         14.181209
                                                                     0.896760
                                                                                0.981429
                 std
                       892.000000
                                    0.000000
                                               1.000000
                                                          0.170000
                                                                     0.000000
                                                                                0.000000
                                                                                           0.00
                 min
                 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
                      1309.000000
                                    1.000000
                                               3.000000
                                                         76.000000
                                                                     8.000000
                                                                                9.000000
                                                                                        512.32
                 max
In [14]:
                1
                   data.shape
                2
    Out[14]: (418, 12)
In [15]:
                   data.size
    Out[15]: 5016
In [16]:
                   data.ndim
    Out[16]: 2
```

In [17]: data.columns Out[17]: Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'Sib Sp', 'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'], dtype='object') 1 data.drop(labels="Age",axis=1) In [18]: M Out[18]: Passengerld Survived Pclass Name Sex SibSp Parch **Ticket** Kelly, Mr. 0 0 0 892 male 0 330911 7.8 **James** Wilkes, Mrs. 1 893 1 3 James female 0 363272 7.1 1 (Ellen Needs) Myles, Mr. 2 894 0 2 0 0 240276 9.1 male Thomas Francis Wirz, Mr. 0 3 895 3 male 0 0 315154 8.1 Albert Hirvonen, Mrs. 4 896 1 Alexander female 1 3101298 12.: (Helga E Lindqvist) Spector, 1305 413 0 0 0 A.5. 3236 8.1 3 male Mr. Woolf Oliva y Ocana, 414 1306 1 female PC 17758 108.9 Dona. Fermina Saether, SOTON/O.Q. 415 1307 0 7.; 0 Mr. Simon male 0 3101262 Sivertsen Ware, Mr. 416 1308 0 0 359309 8.1 male Frederick Peter, 417 1309 0 3 Master. 1 2668 22.3 1 male Michael J

418 rows × 11 columns

In [19]: ► data.drop(labels=2,axis=0)

										Į.
Ticket	Parch	SibSp	Age	Sex	Name	Pclass	Survived	Passengerld		Out[19]:
330911	0	0	34.5	male	Kelly, Mr. James	3	0	892	0	
363272	0	1	47.0	female	Wilkes, Mrs. James (Ellen Needs)	3	1	893	1	
315154	0	0	27.0	male	Wirz, Mr. Albert	3	0	895	3	
3101298	1	1	22.0	female	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	3	1	896	4	
7538	0	0	14.0	male	Svensson, Mr. Johan Cervin	3	0	897	5	
A.5. 3236	0	0	NaN	male	Spector, Mr. Woolf	3	0	1305	413	
PC 17758	0	0	39.0	female	Oliva y Ocana, Dona. Fermina	1	1	1306	414	
SOTON/O.Q. 3101262	0	0	38.5	male	Saether, Mr. Simon Sivertsen	3	0	1307	415	
359309	0	0	NaN	male	Ware, Mr. Frederick	3	0	1308	416	
2668	1	1	NaN	male	Peter, Master. Michael J	3	0	1309	417	

417 rows × 12 columns

