

gaa0hx8cb

December 13, 2024

1 Data Manipulation

```
[1]: #Name : Devesh J Arbat  
#Roll no. : 06  
#Section : A
```

```
[2]: #Aim : To perform Opreation on Data Manipulation
```

```
[3]: import pandas as pd
```

```
[4]: import os
```

```
[5]: os.getcwd()
```

```
[5]: 'C:\\Users\\91876'
```

```
[6]: os.chdir("C:\\Users\\91876\\OneDrive\\Desktop\\Data Science")
```

```
[7]: data=pd.read_csv("titanic.csv")
```

```
[8]: data
```

```
[8]:
```

	pclass	survived	name	\
0	1.0	1.0	Allen, Miss. Elisabeth Walton	
1	1.0	1.0	Allison, Master. Hudson Trevor	
2	1.0	0.0	Allison, Miss. Helen Loraine	
3	1.0	0.0	Allison, Mr. Hudson Joshua Creighton	
4	1.0	0.0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	
...	
1305	3.0	0.0	Zabour, Miss. Thamine	
1306	3.0	0.0	Zakarian, Mr. Mapriededer	
1307	3.0	0.0	Zakarian, Mr. Ortin	
1308	3.0	0.0	Zimmerman, Mr. Leo	
1309	NaN	NaN	NaN	

	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	\
0	female	29.0000	0.0	0.0	24160	211.3375	B5	S	2	

1	male	0.9167	1.0	2.0	113781	151.5500	C22 C26		S	11
2	female	2.0000	1.0	2.0	113781	151.5500	C22 C26		S	NaN
3	male	30.0000	1.0	2.0	113781	151.5500	C22 C26		S	NaN
4	female	25.0000	1.0	2.0	113781	151.5500	C22 C26		S	NaN
...		
1305	female	NaN	1.0	0.0	2665	14.4542		NaN	C	NaN
1306	male	26.5000	0.0	0.0	2656	7.2250		NaN	C	NaN
1307	male	27.0000	0.0	0.0	2670	7.2250		NaN	C	NaN
1308	male	29.0000	0.0	0.0	315082	7.8750		NaN	S	NaN
1309	NaN	NaN	NaN	NaN	NaN	NaN		NaN	NaN	NaN

	body	home.dest
0	NaN	St Louis, MO
1	NaN	Montreal, PQ / Chesterville, ON
2	NaN	Montreal, PQ / Chesterville, ON
3	135.0	Montreal, PQ / Chesterville, ON
4	NaN	Montreal, PQ / Chesterville, ON
...
1305	NaN	NaN
1306	304.0	NaN
1307	NaN	NaN
1308	NaN	NaN
1309	NaN	NaN

[1310 rows x 14 columns]

```
[9]: data.head()
```

```
[9]:
```

	pclass	survived	name	sex	\
0	1.0	1.0	Allen, Miss. Elisabeth Walton	female	
1	1.0	1.0	Allison, Master. Hudson Trevor	male	
2	1.0	0.0	Allison, Miss. Helen Loraine	female	
3	1.0	0.0	Allison, Mr. Hudson Joshua Creighton	male	
4	1.0	0.0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	

	age	sibsp	parch	ticket	fare	cabin	embarked	boat	body	\
0	29.0000	0.0	0.0	24160	211.3375	B5	S	2	NaN	
1	0.9167	1.0	2.0	113781	151.5500	C22 C26	S	11	NaN	
2	2.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN	NaN	
3	30.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN	135.0	
4	25.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN	NaN	

	home.dest
0	St Louis, MO
1	Montreal, PQ / Chesterville, ON
2	Montreal, PQ / Chesterville, ON
3	Montreal, PQ / Chesterville, ON

4 Montreal, PQ / Chesterville, ON

```
[11]: data.tail()
```

```
[11]:
```

	pclass	survived	name	sex	age	sibsp	parch	\
1305	3.0	0.0	Zabour, Miss. Thamine	female	NaN	1.0	0.0	
1306	3.0	0.0	Zakarian, Mr. Mapriededer	male	26.5	0.0	0.0	
1307	3.0	0.0	Zakarian, Mr. Ortin	male	27.0	0.0	0.0	
1308	3.0	0.0	Zimmerman, Mr. Leo	male	29.0	0.0	0.0	
1309	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

	ticket	fare	cabin	embarked	boat	body	home.dest
1305	2665	14.4542	NaN	C	NaN	NaN	NaN
1306	2656	7.2250	NaN	C	NaN	304.0	NaN
1307	2670	7.2250	NaN	C	NaN	NaN	NaN
1308	315082	7.8750	NaN	S	NaN	NaN	NaN
1309	NaN	NaN	NaN	NaN	NaN	NaN	NaN

```
[12]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1310 entries, 0 to 1309
Data columns (total 14 columns):
#   Column      Non-Null Count  Dtype
---  -
0   pclass      1309 non-null   float64
1   survived    1309 non-null   float64
2   name        1309 non-null   object
3   sex         1309 non-null   object
4   age         1046 non-null   float64
5   sibsp       1309 non-null   float64
6   parch       1309 non-null   float64
7   ticket      1309 non-null   object
8   fare        1308 non-null   float64
9   cabin       295 non-null    object
10  embarked    1307 non-null   object
11  boat        486 non-null    object
12  body        121 non-null    float64
13  home.dest    745 non-null    object
dtypes: float64(7), object(7)
memory usage: 143.4+ KB
```

```
[13]: data.describe()
```

```
[13]:
```

	pclass	survived	age	sibsp	parch	\
count	1309.000000	1309.000000	1046.000000	1309.000000	1309.000000	
mean	2.294882	0.381971	29.881135	0.498854	0.385027	

std	0.837836	0.486055	14.413500	1.041658	0.865560
min	1.000000	0.000000	0.166700	0.000000	0.000000
25%	2.000000	0.000000	21.000000	0.000000	0.000000
50%	3.000000	0.000000	28.000000	0.000000	0.000000
75%	3.000000	1.000000	39.000000	1.000000	0.000000
max	3.000000	1.000000	80.000000	8.000000	9.000000

	fare	body
count	1308.000000	121.000000
mean	33.295479	160.809917
std	51.758668	97.696922
min	0.000000	1.000000
25%	7.895800	72.000000
50%	14.454200	155.000000
75%	31.275000	256.000000
max	512.329200	328.000000

```
[14]: data.shape
```

```
[14]: (1310, 14)
```

```
[15]: data.size
```

```
[15]: 18340
```

```
[16]: data.ndim
```

```
[16]: 2
```

```
[17]: data.columns
```

```
[17]: Index(['pclass', 'survived', 'name', 'sex', 'age', 'sibsp', 'parch', 'ticket',
         'fare', 'cabin', 'embarked', 'boat', 'body', 'home.dest'],
         dtype='object')
```

```
[18]: data.drop(labels="age",axis=1)
```

```
[18]:
```

	pclass	survived	name \
0	1.0	1.0	Allen, Miss. Elisabeth Walton
1	1.0	1.0	Allison, Master. Hudson Trevor
2	1.0	0.0	Allison, Miss. Helen Loraine
3	1.0	0.0	Allison, Mr. Hudson Joshua Creighton
4	1.0	0.0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)
...
1305	3.0	0.0	Zabour, Miss. Thamine
1306	3.0	0.0	Zakarian, Mr. Mapriededer
1307	3.0	0.0	Zakarian, Mr. Ortin

1308	3.0	0.0					Zimmerman, Mr. Leo		
1309	NaN	NaN					NaN		

	sex	sibsp	parch	ticket	fare	cabin	embarked	boat	body \
0	female	0.0	0.0	24160	211.3375	B5	S	2	NaN
1	male	1.0	2.0	113781	151.5500	C22 C26	S	11	NaN
2	female	1.0	2.0	113781	151.5500	C22 C26	S	NaN	NaN
3	male	1.0	2.0	113781	151.5500	C22 C26	S	NaN	135.0
4	female	1.0	2.0	113781	151.5500	C22 C26	S	NaN	NaN
...
1305	female	1.0	0.0	2665	14.4542	NaN	C	NaN	NaN
1306	male	0.0	0.0	2656	7.2250	NaN	C	NaN	304.0
1307	male	0.0	0.0	2670	7.2250	NaN	C	NaN	NaN
1308	male	0.0	0.0	315082	7.8750	NaN	S	NaN	NaN
1309	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

	home.dest
0	St Louis, MO
1	Montreal, PQ / Chesterville, ON
2	Montreal, PQ / Chesterville, ON
3	Montreal, PQ / Chesterville, ON
4	Montreal, PQ / Chesterville, ON
...	...
1305	NaN
1306	NaN
1307	NaN
1308	NaN
1309	NaN

[1310 rows x 13 columns]

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

```
[19]:
```

	pclass	survived	name \
0	1.0	1.0	Allen, Miss. Elisabeth Walton
1	1.0	1.0	Allison, Master. Hudson Trevor
3	1.0	0.0	Allison, Mr. Hudson Joshua Creighton
4	1.0	0.0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)
5	1.0	1.0	Anderson, Mr. Harry
...
1305	3.0	0.0	Zabour, Miss. Thamine
1306	3.0	0.0	Zakarian, Mr. Mapriededer
1307	3.0	0.0	Zakarian, Mr. Ortin
1308	3.0	0.0	Zimmerman, Mr. Leo
1309	NaN	NaN	NaN

	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	\
--	-----	-----	-------	-------	--------	------	-------	----------	------	---

0	female	29.0000	0.0	0.0	24160	211.3375	B5	S	2
1	male	0.9167	1.0	2.0	113781	151.5500	C22 C26	S	11
3	male	30.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN
4	female	25.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN
5	male	48.0000	0.0	0.0	19952	26.5500	E12	S	3
...		
1305	female	NaN	1.0	0.0	2665	14.4542	NaN	C	NaN
1306	male	26.5000	0.0	0.0	2656	7.2250	NaN	C	NaN
1307	male	27.0000	0.0	0.0	2670	7.2250	NaN	C	NaN
1308	male	29.0000	0.0	0.0	315082	7.8750	NaN	S	NaN
1309	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

	body	home.dest
0	NaN	St Louis, MO
1	NaN	Montreal, PQ / Chesterville, ON
3	135.0	Montreal, PQ / Chesterville, ON
4	NaN	Montreal, PQ / Chesterville, ON
5	NaN	New York, NY
...
1305	NaN	NaN
1306	304.0	NaN
1307	NaN	NaN
1308	NaN	NaN
1309	NaN	NaN

[1309 rows x 14 columns]