

ASSIGNMENT 13

Create a dataframe having following columns: ISBN (unique number for each book) • Title • Author • Publication • Year Published • Price • Copies sold of first edition • Copies sold in second edition • Copies sold in third edition You can use any of the studied approach to create dataframe. Add minimum 10 records to the dataframe. Perform following operations on dataframe using appropriate functions: 1. Display the number of rows and columns in dataframe 2. Give the descriptive statistics of the created dataset. 3. Display distinct publication names for the dataset. 4. Display the book title and author names published by "Metro Publication" 5. Rename columns "Copies sold in first edition", "Copies sold in second edition" and "Copies sold in third edition" to FE, SE and TE respectively 6. Add a column "Average sale" to your dataframe derived as (average of FE, SE and TE) * Price. 7. Display the details of books grouped by author. 8. For each group obtained in above query display maximum value of Average sale. 9. Reshape your dataframe such that rows will show 'Author', column will show "Publication" and values will be 'Title' of book.

```
[ ] import pandas as pd
import numpy as np
import random
```

```
[ ] ISBN=[random.randint(0,1000) for x in range(0,10)]
Title=['Networks','DSA','C Programming','Java','OS','Python','Physics','Chemistry','EM-1','Computer Science']
Author=['Faurozan','Thareja','Kulkarni','Rane','Russel','Billings','Sharma','Agarwal','Mehta','Patil']
Publication=['Metro','Oxford','MM','Metro','Techneo','Techknowledge','Metro','Oxford','MM','Pragati']
Year=[random.randint(2010,2020) for x in range(0,10)]
Price=[random.randrange(1000,5000,10) for x in range(0,10)]
FE=[random.randint(100,1000) for x in range(0,10)]
SE=[random.randint(100,1000) for x in range(0,10)]
TE=[random.randint(100,1000) for x in range(0,10)]
```

```
[ ] df=pd.DataFrame({'ISBN':ISBN,
                    'Title':Title,
                    'Author':Author,
                    'Publication':Publication,
                    'Year Published':Year,
                    'Price':Price,
                    'Copies sold of First edition':FE,
                    'Copies sold of Second edition':SE,
                    'Copies sold in third edition':TE})
```

```
[ ] df
```

	ISBN	Title	Author	Publication	Year Published	Price	Copies sold of First edition	Copies sold of Second edition	Copies sold in third edition
0	532	Networks	Faurozan	Metro	2020	4110	784	450	682
1	161	DSA	Thareja	Oxford	2019	3750	769	608	733
2	209	C Programming	Kulkarni	MM	2016	2940	398	844	575
3	951	Java	Rane	Metro	2015	1520	716	452	744
4	897	OS	Russel	Techneo	2019	3090	807	130	223
5	74	Python	Billings	Techknowledge	2010	3430	185	866	503
6	466	Physics	Sharma	Metro	2016	3740	641	262	511
7	779	Chemistry	Agarwal	Oxford	2014	2010	336	728	186
8	755	EM-1	Mehta	MM	2014	3210	521	640	196
9	720	Computer Science	Patil	Pragati	2012	1890	666	373	982

```
[ ] df.shape
```

```
(10, 9)
```

```
[ ] df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 9 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                ---
0   ISBN                                  10 non-null    int64
1   Title                                10 non-null    object
2   Author                              10 non-null    object
3   Publication                          10 non-null    object
4   Year Published                      10 non-null    int64
5   Price                               10 non-null    int64
6   Copies sold of First edition         10 non-null    int64
7   Copies sold of Second edition        10 non-null    int64
8   Copies sold in third edition         10 non-null    int64
dtypes: int64(6), object(3)
memory usage: 848.0+ bytes
```

```
[ ] df['Publication'].unique()
```

```
array(['Metro', 'Oxford', 'MM', 'Techneo', 'Techknowledge', 'Pragati'],  
      dtype=object)
```

```
[ ] df[df['Publication']=='Metro'][['Title','Author']]
```

	Title	Author
0	Networks	Faurozan
3	Java	Rane
6	Physics	Sharma

```
[ ] df.rename(columns={'Copies sold of First edition':'FE','Copies sold of Second edition':'SE','Copies sold in third edition':'TE'})
```

	ISBN	Title	Author	Publication	Year Published	Price	FE	SE	TE
0	532	Networks	Faurozan	Metro	2020	4110	784	450	682
1	161	DSA	Thareja	Oxford	2019	3750	769	608	733
2	209	C Programming	Kulkarni	MM	2016	2940	398	844	575
3	951	Java	Rane	Metro	2015	1520	716	452	744
4	897	OS	Russel	Techneo	2019	3090	807	130	223
5	74	Python	Billings	Techknowledge	2010	3430	185	866	503
6	466	Physics	Sharma	Metro	2016	3740	641	262	511
7	779	Chemistry	Agarwal	Oxford	2014	2010	336	728	186
8	755	EM-1	Mehta	MM	2014	3210	521	640	196
9	720	Computer Science	Patil	Pragati	2012	1890	666	373	982