Name - Varunkumar Dadaji Sonawane ¶

Batch - T4

Roll no. - 73

TE IT

Title - Perform the following operations using Python on the Facebook metrics data sets a. Create data subsets b. Merge Data c. Sort Data d. Transposing Data e. Shape and reshape Data

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

In [2]:
```
df =  pd.read_csv('facebook.csv')
df.head()
```

Out[2]:

	userid	age	dob_day	dob_year	dob_month	gender	tenure	friend_count	friendships_initiated	likes
0	2094382	14	19	1999	11	male	266.0	0	0	0
1	1192601	14	2	1999	11	female	6.0	0	0	0
2	2083884	14	16	1999	11	male	13.0	0	0	0
3	1203168	14	25	1999	12	female	93.0	0	0	0
4	1733186	14	4	1999	12	male	82.0	0	0	0

Creating Data Subset

In [3]:
```
d_f = df[0:10]
d_f
```

Out[3]:

	userid	age	dob_day	dob_year	dob_month	gender	tenure	friend_count	friendships_initiated	likes
0	2094382	14	19	1999	11	male	266.0	0	0	0
1	1192601	14	2	1999	11	female	6.0	0	0	0
2	2083884	14	16	1999	11	male	13.0	0	0	0
3	1203168	14	25	1999	12	female	93.0	0	0	0
4	1733186	14	4	1999	12	male	82.0	0	0	0
5	1524765	14	1	1999	12	male	15.0	0	0	0
6	1136133	13	14	2000	1	male	12.0	0	0	0
7	1680361	13	4	2000	1	female	0.0	0	0	0
8	1365174	13	1	2000	1	male	81.0	0	0	0
9	1712567	13	2	2000	2	male	171.0	0	0	0

Sorting Data

In [4]:
```
Sorted_by_Gender = df.sort_values(by='gender')
Sorted_by_Gender.head()
```

Out[4]:

	userid	age	dob_day	dob_year	dob_month	gender	tenure	friend_count	friendships_initiated	li
49502	1099429	25	30	1988	5	female	429.0	84	25	
54719	2131675	16	21	1997	9	female	271.0	100	81	
54723	2184314	63	25	1950	2	female	2103.0	100	36	
54724	1323947	16	17	1997	3	female	90.0	99	85	
54725	1052390	37	2	1976	12	female	400.0	100	43	

In [5]:
```
Sorted_by_Gender.tail()
```

Out[5]:

	userid	age	dob_day	dob_year	dob_month	gender	tenure	friend_count	friendships_initiated	li
94778	1748557	68	27	1945	4	NaN	1862.0	790	435	
95317	2173780	50	11	1963	9	NaN	2129.0	863	389	
95479	1442490	108	1	1905	7	NaN	1332.0	967	537	
97530	2119521	74	26	1939	9	NaN	1998.0	1609	348	
98216	1966857	102	6	1911	12	NaN	2389.0	2548	1130	

Transposing Data

In [6]:
```
1 df.transpose()
```

Out[6]:

	0	1	2	3	4	5	6	7	8
userid	2094382	1192601	2083884	1203168	1733186	1524765	1136133	1680361	1365174
age	14	14	14	14	14	14	13	13	13
dob_day	19	2	16	25	4	1	14	4	1
dob_year	1999	1999	1999	1999	1999	1999	2000	2000	2000
dob_month	11	11	11	12	12	12	1	1	1
gender	male	female	male	female	male	male	male	female	male
tenure	266.0	6.0	13.0	93.0	82.0	15.0	12.0	0.0	81.0
friend_count	0	0	0	0	0	0	0	0	0
friendships_initiated	0	0	0	0	0	0	0	0	0
likes	0	0	0	0	0	0	0	0	0
likes_received	0	0	0	0	0	0	0	0	0
mobile_likes	0	0	0	0	0	0	0	0	0
mobile_likes_received	0	0	0	0	0	0	0	0	0
www_likes	0	0	0	0	0	0	0	0	0
www_likes_received	0	0	0	0	0	0	0	0	0

15 rows × 99003 columns

Shape of Data

In [7]:
```
1 df.shape
```

Out[7]:
```
(99003, 15)
```

Reshape Data

In [8]:
```
1 x = np.array([[1,2,3], [7,8,9]])
2 x
```

Out[8]:
```
array([[1, 2, 3],
       [7, 8, 9]])
```

In [9]:
```
1 np.reshape(x, (3, 2))
```

Out[9]:
```
array([[1, 2],
       [3, 7],
       [8, 9]])
```

Merge Data

In [10]:
```
Data1 = pd.DataFrame({'students': ['Varun', 'Rohit', 'Vipul', 'Rushikesh'],
                      'subject': ['Maths', 'Bio', 'Marathi', 'Hindi']})
Data2 = pd.DataFrame({'students': ['Varun', 'Rohit', 'Vipul', 'Rushikesh'],
                      'birth_year': [2001, 2002, 2003, 2004]})
display(Data1, Data2)
```

	students	subject
0	Varun	Maths
1	Rohit	Bio
2	Vipul	Marathi
3	Rushikesh	Hindi

	students	birth_year
0	Varun	2001
1	Rohit	2002
2	Vipul	2003
3	Rushikesh	2004

In [11]:
```
Merged_Data = pd.merge(Data1, Data2)
```

In [12]:
```
Merged_Data
```

Out[12]:

	students	subject	birth_year
0	Varun	Maths	2001
1	Rohit	Bio	2002
2	Vipul	Marathi	2003
3	Rushikesh	Hindi	2004

In [13]:
```
1 df
```

Out[13]:

	userid	age	dob_day	dob_year	dob_month	gender	tenure	friend_count	friendships_initiated
0	2094382	14	19	1999	11	male	266.0	0	0
1	1192601	14	2	1999	11	female	6.0	0	0
2	2083884	14	16	1999	11	male	13.0	0	0
3	1203168	14	25	1999	12	female	93.0	0	0
4	1733186	14	4	1999	12	male	82.0	0	0
...	...	...	...	...	...	...	...	...	...
98998	1268299	68	4	1945	4	female	541.0	2118	341
98999	1256153	18	12	1995	3	female	21.0	1968	1720
99000	1195943	15	10	1998	5	female	111.0	2002	1524
99001	1468023	23	11	1990	4	female	416.0	2560	185
99002	1397896	39	15	1974	5	female	397.0	2049	768

99003 rows × 15 columns