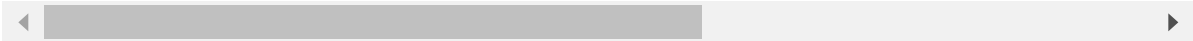


```
In [3]: 1 import pandas as pd
        2 df = pd.read_csv(r"C:\Users\ASUS\Dropbox\My PC (LAPTOP-4FONMOAM)\Document
        3 df
```

Out[3]:

	Date	DeviceID	Carrier	State
0	15-05-2018	d2f4c0b787a56848c22af641ec698b66d4cd9867	Bharti Airtel	KARNATAKA
1	15-05-2018	fa12d5960402c5e69ec0badda1254a982098d3f1	Bharti Airtel	TELANGANA
2	15-05-2018	cc14614074672b32e018a88c3490da785ccdf6e4	Jio	ANDHRA PRADESH
3	15-05-2018	537@f747a8@21e678899d15a@400463@63@93600	Jio	MAHARASHTRA
4	15-04-2018	11c35@46c1@1ece8453a3d5973a9@6@66@8d3@5e	Bharti Airtel	KARNATAKA
...
249994	15-05-2018	dc347836bb27a3e9f649863c4e96dc98244c7014	Vodafone Essar	GUJARAT
249995	15-05-2018	be0044cc76ef33baf8dea60e7dc4ce2f4dc51b24	Jio	MAHARASHTRA
249996	15-04-2018	1f4a2d863c157369c56342@3f7@293508cdd34@0	Idea	UTTAR PRADESH
249997	15-05-2018	9f@fda6ffdf4@65@32291aa3afc@6@c48f@8c56d	Jio	TELANGANA
249998	15-05-2018	e8f043ee7a41cd043c002471e960479ed0fd4210	Jio	TELANGANA

249999 rows × 9 columns




In [2]: 

```
1 df.describe()
2
```


Out[2]:

Mobile_internet_usage_in_min	
count	249999.000000
mean	119.933280
std	69.589373
min	0.000000
25%	60.000000
50%	120.000000
75%	180.000000
max	240.000000

In [4]: 

```
1 df['Date'].max()
```

Out[4]: '15-05-2018'

In [5]: 

```
1 df['Date'].min()
```

Out[5]: '15-04-2018'

In []: 

```
1
```

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

Out[7]:


	Date	DeviceID	Carrier	State
0	15-05-2018	d2f4c0b787a56848c22af641ec698b66d4cd9867	Bharti Airtel	KARNATAKA
1	15-05-2018	fa12d5960402c5e69ec0badda1254a982098d3f1	Bharti Airtel	TELANGANA
2	15-05-2018	cc14614074672b32e018a88c3490da785ccdf6e4	Jio	ANDHRA PRADESH
3	15-05-2018	537@f747a8@21e678899d15a@400463@63@93600	Jio	MAHARASHTRA
4	15-04-2018	11c35@46c1@1ece8453a3d5973a9@6@66@8d3@5e	Bharti Airtel	KARNATAKA
...
249994	15-05-2018	dc347836bb27a3e9f649863c4e96dc98244c7014	Vodafone Essar	GUJARAT
249995	15-05-2018	be0044cc76ef33baf8dea60e7dc4ce2f4dc51b24	Jio	MAHARASHTRA
249996	15-04-2018	1f4a2d863c157369c56342@3f7@293508cdd34@0	Idea	UTTAR PRADESH
249997	15-05-2018	9f@fda6ffdf4@65@32291aa3afc@6@c48f@8c56d	Jio	TELANGANA
249998	15-05-2018	e8f043ee7a41cd043c002471e960479ed0fd4210	Jio	TELANGANA

249999 rows × 9 columns




```
In [8]: 1 df.isnull().values.any().sum()
        2 # df.isnull().sum()
```

Out[8]: 0

In [9]:  1 pip install sqlalchemy

Requirement already satisfied: sqlalchemy in c:\users\asus\anaconda3\lib\site-packages (1.3.18)

Note: you may need to restart the kernel to use updated packages.

In [10]:  1 from sqlalchemy import create_engine
2 engine = create_engine('sqlite://', echo = False)
3
4 df.to_sql("inMobi", con = engine)
5 print(engine.execute("SELECT * FROM inMobi").fetchall())

IOPub data rate exceeded.

The notebook server will temporarily stop sending output to the client in order to avoid crashing it.


To change this limit, set the config variable

`--NotebookApp.iopub_data_rate_limit`.

Current values:

NotebookApp.iopub_data_rate_limit=1000000.0 (bytes/sec)

NotebookApp.rate_limit_window=3.0 (secs)

In [11]:  1 x = engine.execute("Select Carrier, count('Device ID') from inMobi group
2
3 x
4

Out[11]: [('Bharti Airtel', 55507),
('Idea', 27547),
('Jio', 131009),
('Vodafone Essar', 35936)]

Dates

In [12]:  1 engine.execute("Select distinct(Date) from inMobi").fetchall()

Out[12]: [('15-05-2018',), ('15-04-2018',)]

List of States

```
In [13]: 1 s = engine.execute("Select distinct(State) from inMobi").fetchall()
2 # for i in range()
3 s
```

```
Out[13]: [('KARNATAKA',),
('TELANGANA',),
('ANDHRA PRADESH',),
('MAHARASHTRA',),
('DELHI',),
('TAMIL NADU',),
('UTTARAKHAND',),
('UTTAR PRADESH',),
('GUJARAT',),
('WEST BENGAL',),
('JAMMU AND KASHMIR',),
('RAJASTHAN',),
('GOA',),
('KERALA',),
('BIHAR',),
('ASSAM',),
('HIMACHAL PRADESH',),
('MADHYA PRADESH',),
('ODISHA',),
('JHARKHAND',),
('CHHATTISGARH',),
('PUDUCHERRY',),
('HARYANA',),
('PUNJAB',),
('MANIPUR',),
('SIKKIM',),
('MIZORAM',),
('MEGHALAYA',),
('DAMAN AND DIU',),
('CHANDIGARH',),
('TRIPURA',),
('ARUNACHAL PRADESH',),
('NAGALAND',),
('DADRA AND NAGAR HAVELI',),
('ANDAMAN AND NICOBAR ISLANDS',)]
```

```
In [121]: ▶ 1 State = []  
2 for i in range(len(s)):  
3     State.append(s[i][0])  
4 State
```

```
Out[121]: ['KARNATAKA',  
'TELANGANA',  
'ANDHRA PRADESH',  
'MAHARASHTRA',  
'DELHI',  
'TAMIL NADU',  
'UTTARAKHAND',  
'UTTAR PRADESH',  
'GUJARAT',  
'WEST BENGAL',  
'JAMMU AND KASHMIR',  
'RAJASTHAN',  
'GOA',  
'KERALA',  
'BIHAR',  
'ASSAM',  
'HIMACHAL PRADESH',  
'MADHYA PRADESH',  
'ODISHA',  
'JHARKHAND',  
'CHHATTISGARH',  
'PUDUCHERRY',  
'HARYANA',  
'PUNJAB',  
'MANIPUR',  
'SIKKIM',  
'MIZORAM',  
'MEGHALAYA',  
'DAMAN AND DIU',  
'CHANDIGARH',  
'TRIPURA',  
'ARUNACHAL PRADESH',  
'NAGALAND',  
'DADRA AND NAGAR HAVELI',  
'ANDAMAN AND NICOBAR ISLANDS']
```

```
In [ ]: ▶ 1
```

```
Out[14]: [('LUNKARANSAR', 0),  
          ('PATUR', 0),  
          ('PENUGANCHIPROLU', 0),  
          ('BAJALI', 1),  
          ('KUTIIYANA', 1),  
          ('SALEKASA', 1),  
          ('ARDHAVEEDU', 2),  
          ('DHOLI (MORAU)', 2),  
          ('KALUAHI', 2),  
          ('SARADA', 2),  
          ('THELKOLOI', 2),  
          ('KOLABIRA', 3),  
          ('LUNDRA', 4),  
          ('MAJHIAON', 4),  
          ('MOHANGARH', 4),  
          ('RAHA', 4),  
          ('TISRI', 4),  
          ('LUNI', 5),  
          ('NARHARPUR', 5),  
          ('NUK KALADALI', 6)]
```

1

```
1 City_Usage = {}
2 City = []
3 Usage = []
4 for i in range(len(c_u)):
5     City.append(c_u[i][0])
6     Usage.append(c_u[i][1])
```

City_List

1	City
---	------

1	Usage
---	-------

City wise Usage


```
In [18]: 1 City_Usage = pd.DataFrame({'city':City,'usage':Usage})
          2 City_Usage.sort_values('usage', ascending = False).head(10)
```

Out[18]:

	city	usage
3920	BANGALORE	2430277
3919	MUMBAI (SUBURBAN)	1077442
3918	HYDERABAD	971254
3917	AHMEDABAD	774019
3916	CHENNAI	745330
3915	PUNE	620649
3914	LUCKNOW	609835
3913	KOLKATA	475327
3912	NORTH GUWAHATI	465731
3911	SURAT	373611

```
In [19]: 1 engine.execute("Select Carrier,sum(Mobile_internet_usage_in_min) from inMobi
          2")
```

Out[19]: [('Bharti Airtel', 6671042),
 ('Idea', 3293306),
 ('Jio', 15702290),
 ('Vodafone Essar', 4316562)]

```
In [20]: 1 M_u =engine.execute("Select Carrier,sum(Mobile_internet_usage_in_min) from inMobi
          2")
```

In []:

```
In [21]: 1 y = engine.execute("Select State,carrier,count('Device ID') from inMobi
          2 y")
```

Out[21]: [('MAHARASHTRA', 'Bharti Airtel', 5737),
 ('MAHARASHTRA', 'Idea', 5827),
 ('MAHARASHTRA', 'Jio', 18488),
 ('MAHARASHTRA', 'Vodafone Essar', 6245)]

```
In [22]: 1 y = engine.execute("Select State,carrier,count('Device ID') from inMobi
2 y
```

```
Out[22]: [('KARNATAKA', 'Bharti Airtel', 10450),
('KARNATAKA', 'Idea', 1532),
('KARNATAKA', 'Jio', 20852),
('KARNATAKA', 'Vodafone Essar', 2656)]
```

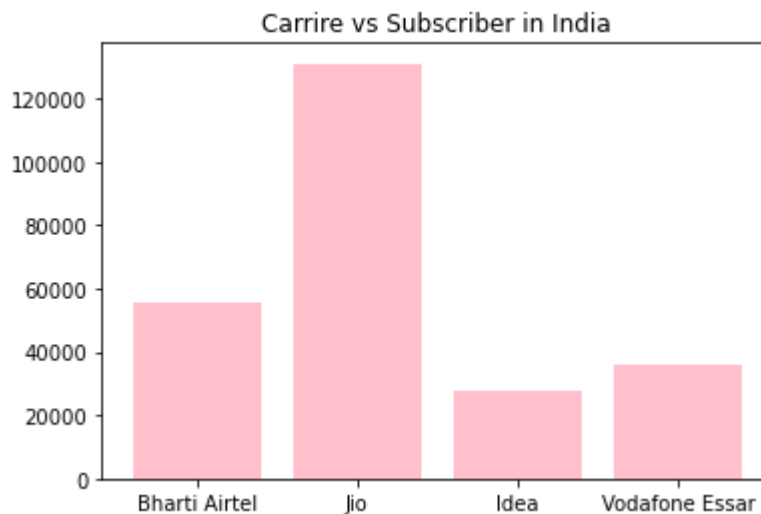
```
In [23]: 1 d = {'carrier': ['Bharti Airtel','Jio','Idea','Vodafone Essar'], 'Subscriber'
2
3 subsc = pd.DataFrame(d,index = [0,1,2,3])
4
5 subsc
```

```
Out[23]:
```

	carrier	Subscriber
0	Bharti Airtel	55507
1	Jio	131009
2	Idea	27547
3	Vodafone Essar	35936

```
In [24]: 1 import matplotlib.pyplot as plt
2 plt.bar(subsc['carrier'],subsc['Subscriber'],color = 'Pink')
3 plt.title('Carrire vs Subscriber in India')
```

```
Out[24]: Text(0.5, 1.0, 'Carrire vs Subscriber in India')
```



```
In [25]: 1 l1 = []
2 l2 = []
3 for i in range(4):
4     l1.append(y[i][1])
5     l2.append(y[i][2])
```

Subscribers of different Carriers in

Maharashtra

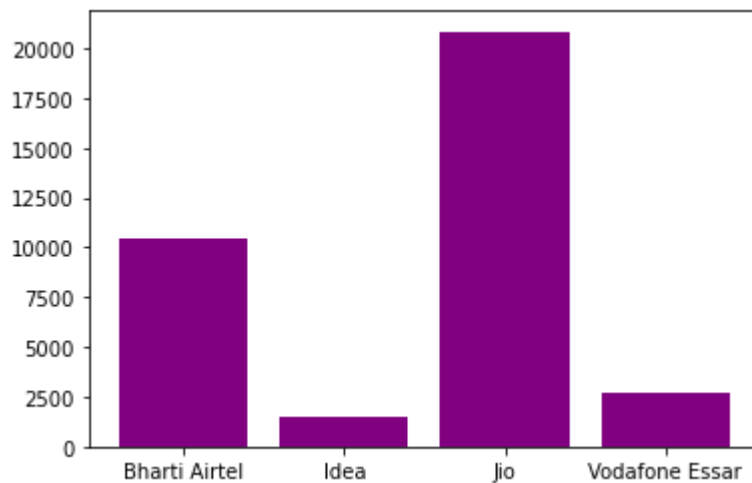
```
In [26]: 1 d_MH = pd.DataFrame({'Carrier':  
2             11, 'Subscriber':12}, index = [0,1,2,3])  
3 d_MH
```

Out[26]:

	Carrier	Subscriber
0	Bharti Airtel	10450
1	Idea	1532
2	Jio	20852
3	Vodafone Essar	2656

```
In [27]: 1 plt.bar(d_MH['Carrier'],d_MH['Subscriber'],color = 'purple')
```

Out[27]: <BarContainer object of 4 artists>



```
Out[28]: [('₹ 649',),
 ('₹ 500',),
 ('₹ 199',),
 ('₹ 500',),
 ('₹ 199',),
 ('₹ 649',),
 ('₹ 649',),
 ('₹ 649',),
 ('₹ 500',),
 ('₹ 649',),
 ('₹ 99',),
 ('₹ 649',),
 ('₹ 99',),
 ('₹ 500',),
 ('₹ 199',),
 ('₹ 649',),
 ('₹ 199',),
 ('₹ 649',),
 ('₹ 500',),
 ('₹ 649',)]
```

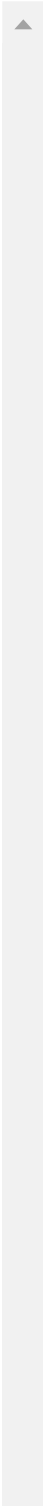
```
Requirement already satisfied: pandas-profiling[notebook] in c:\users\asus\anaconda3\lib\site-packages (2.11.0)
Requirement already satisfied: requests>=2.24.0 in c:\users\asus\anaconda3\lib\site-packages (from pandas-profiling[notebook]) (2.24.0)
Requirement already satisfied: scipy>=1.4.1 in c:\users\asus\anaconda3\lib\site-packages (from pandas-profiling[notebook]) (1.5.0)
Requirement already satisfied: visions[type_image_path]==0.6.0 in c:\users\asus\anaconda3\lib\site-packages (from pandas-profiling[notebook]) (0.6.0)
Requirement already satisfied: htmlmin>=0.1.12 in c:\users\asus\anaconda3\lib\site-packages (from pandas-profiling[notebook]) (0.1.12)
Requirement already satisfied: numpy>=1.16.0 in c:\users\asus\anaconda3\lib\site-packages (from pandas-profiling[notebook]) (1.18.5)
Requirement already satisfied: tqdm>=4.48.2 in c:\users\asus\anaconda3\lib\site-packages (from pandas-profiling[notebook]) (4.57.0)
Requirement already satisfied: matplotlib>=3.2.0 in c:\users\asus\anaconda3\lib\site-packages (from pandas-profiling[notebook]) (3.2.2)
Requirement already satisfied: pandas!=1.0.0,!=1.0.1,!=1.0.2,!=1.1.0,>=0.25.3 in c:\users\asus\anaconda3\lib\site-packages (from pandas-profiling[notebook]) (1.0.5)
```

12/21

In [33]: ▶

1 pandas_profiling.ProfileReport(df)

Summarize dataset:	22/22 [00:17<00:00, 2.77it/s,
100%	Completed]
Generate report structure:	1/1 [00:04<00:00,
100%	4.06s/it]
Render HTML: 100%	1/1 [00:00<00:00, 2.29it/s]



Overview

Dataset statistics

Number of variables	9
Number of observations	249999
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	17.2 MiB
Average record size in memory	72.0 B

Variable types

Categorical	8
Numeric	1

Warnings

DeviceID has a high cardinality: 245699 distinct values	High cardinality
City has a high cardinality: 3921 distinct values	High cardinality
DeviceID is uniformly distributed	Uniform

Reproduction

Out[33]:

In [34]: 

```
1 c_jio = engine.execute("select Carrier,Date,count(DeviceID) from inMobi v
2 c_jio
```

Out[34]: [(('Jio', '15-04-2018', 65147), ('Jio', '15-05-2018', 65862))]

```
In [35]: 1 l1 = []
2 l2 = []
3 for i in range(len(c_jio)):
4     l1.append(c_jio[i][1])
5     l2.append(c_jio[i][2])
6
```

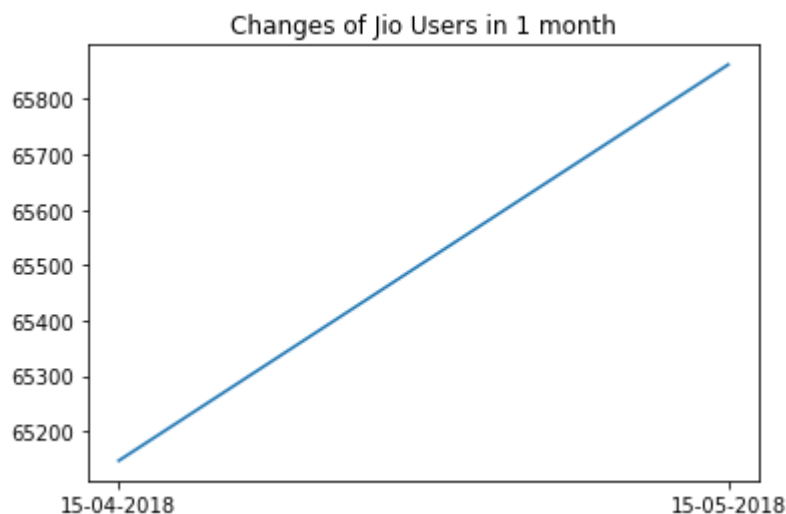
```
In [36]: 1 Change_Jio = pd.DataFrame({'Date':l1,'Users':l2})
2 Change_Jio
```

Out[36]:

	Date	Users
0	15-04-2018	65147
1	15-05-2018	65862

```
In [37]: 1 plt.plot(Change_Jio['Date'],Change_Jio['Users'])
2 plt.title('Changes of Jio Users in 1 month')
```

Out[37]: Text(0.5, 1.0, 'Changes of Jio Users in 1 month')



```
In [38]: 1 c_airtel = engine.execute("select Carrier,Date,count(DeviceID) from inMol")
2 c_airtel
```

Out[38]: [('Bharti Airtel', '15-04-2018', 27524),
('Bharti Airtel', '15-05-2018', 27983)]

```
In [39]: 1 l1 = []
2 l2 = []
3 for i in range(len(c_jio)):
4     l1.append(c_airtel[i][1])
5     l2.append(c_airtel[i][2])
6
7
```

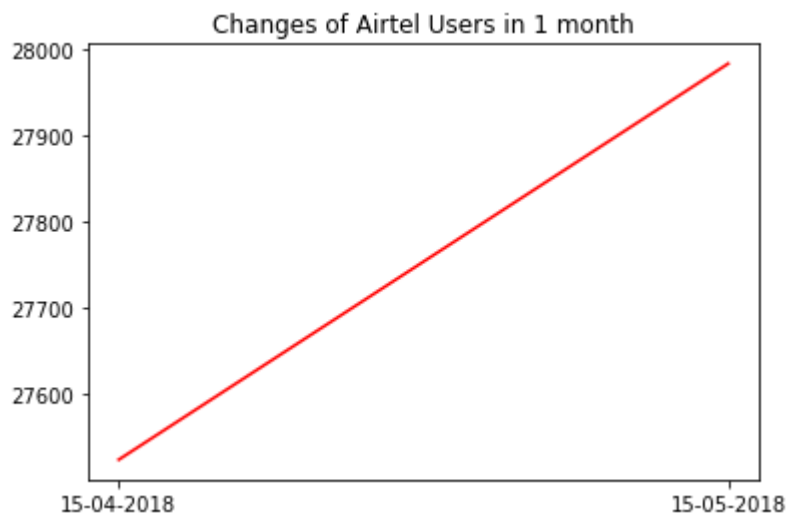
```
In [40]: 1 Change_Airtel = pd.DataFrame({'Date':l1, 'Users':l2})
          2 Change_Airtel
```

Out[40]:

	Date	Users
0	15-04-2018	27524
1	15-05-2018	27983

```
In [41]: 1 plt.plot(Change_Airtel['Date'],Change_Airtel['Users'], color = 'r')
          2 plt.title('Changes of Airtel Users in 1 month')
```

Out[41]: Text(0.5, 1.0, 'Changes of Airtel Users in 1 month')



```
In [42]: 1 c_idea = engine.execute("select Carrier,Date,count(DeviceID) from inMobi
          2 c_idea
```

Out[42]: [('Idea', '15-04-2018', 13826), ('Idea', '15-05-2018', 13721)]

```
In [ ]: 1
```

```
In [43]: 1 l1 = []
          2 l2 = []
          3 for i in range(len(c_jio)):
          4     l1.append(c_idea[i][1])
          5     l2.append(c_idea[i][2])
          6
          7
```



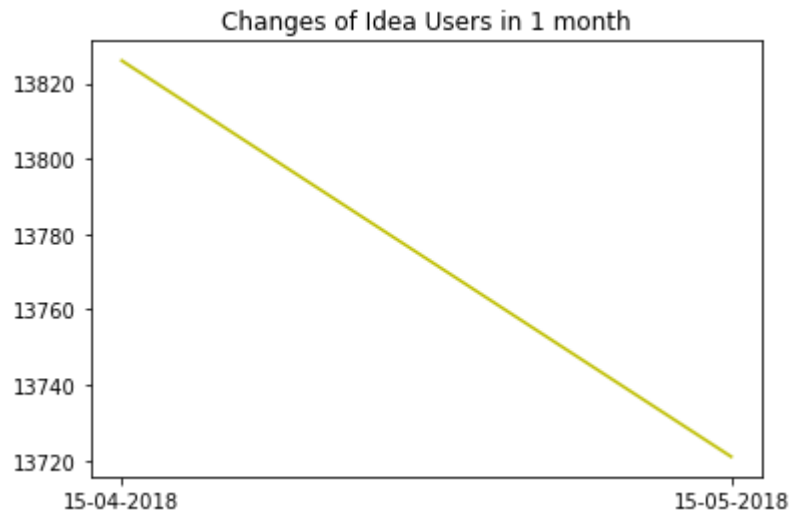
```
In [44]: 1 Change_idea = pd.DataFrame({'Date':l1, 'Users':l2})
          2 Change_idea
```

Out[44]:

	Date	Users
0	15-04-2018	13826
1	15-05-2018	13721

```
In [45]: 1 plt.plot(Change_idea['Date'],Change_idea['Users'], color = 'y')
          2 plt.title('Changes of Idea Users in 1 month')
```

Out[45]: Text(0.5, 1.0, 'Changes of Idea Users in 1 month')



```
In [46]: 1 c_vodafone = engine.execute("select Carrier,Date,count(DeviceID) from inf
          2 c_vodafone
```

Out[46]: [('Vodafone Essar', '15-04-2018', 17978),
 ('Vodafone Essar', '15-05-2018', 17958)]

```
In [47]: 1 l1 = []
          2 l2 = []
          3 for i in range(len(c_jio)):
          4     l1.append(c_vodafone[i][1])
          5     l2.append(c_vodafone[i][2])
          6
          7
```

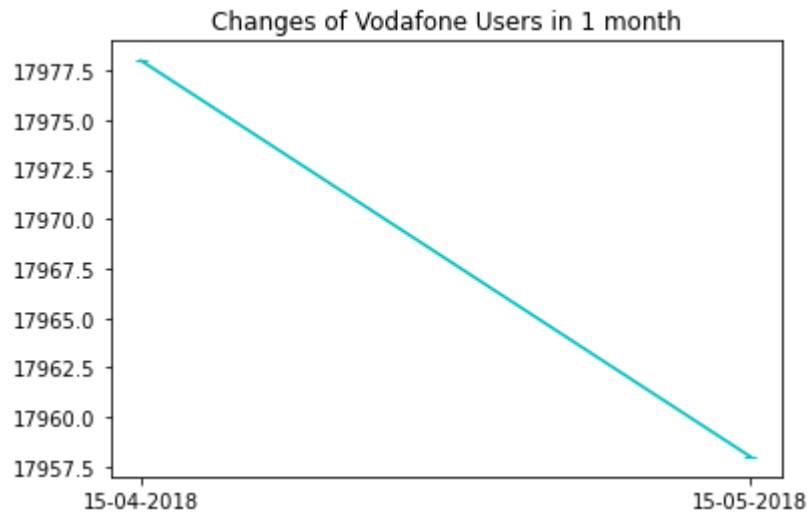
```
In [48]: 1 Change_vodafone = pd.DataFrame({'Date':l1, 'Users':l2})
          2 Change_vodafone
```

Out[48]:

	Date	Users
0	15-04-2018	17978
1	15-05-2018	17958

```
In [49]: 1 plt.plot(Change_vodafone['Date'],Change_vodafone['Users'], color = 'c',
2 plt.title('Changes of Vodafone Users in 1 month'))
```

Out[49]: Text(0.5, 1.0, 'Changes of Vodafone Users in 1 month')



```
In [50]: 1 plt.scatter(Change_vodafone['Date'],Change_vodafone['Users'])
```

Out[50]: <matplotlib.collections.PathCollection at 0x1b3ab5e3a30>



```
In [ ]: 1
```

```
In [ ]: 1
```

```
In [ ]: 1
```

```
In [51]: 1 b = engine.execute('Select * from inMobi where Carrier = "Bharti Airtel"')
```

In [52]: `1 pd.DataFrame(b,columns= ['id','Date','DeviceID','Carrier','State','City']`

Out[52]:

	id	Date	DeviceID	Carrier	State	
0	0	15-05-2018	d2f4c0b787a56848c22af641ec698b66d4cd9867	Bharti Airtel	KARNATAKA	YE
1	1	15-05-2018	fa12d5960402c5e69ec0badda1254a982098d3f1	Bharti Airtel	TELANGANA	YE
2	4	15-04-2018	11c35@46c1@1ece8453a3d5973a9@6@66@8d3@5e	Bharti Airtel	KARNATAKA	
3	6	15-05-2018	11d2dac@355773d49061ff86f239973@a4704a3d	Bharti Airtel	TELANGANA	WAI
4	7	15-05-2018	82302ddacf5479df0d66c605@8e9@d78a5c88342	Bharti Airtel	TELANGANA	WAI
5	9	15-04-2018	c03dcc2bc4da480bee0dd43cacf4ba3e0d334f	Bharti Airtel	DELHI	VIVE
6	11	15-04-2018	94e0508e6@5c893e6@fdc1d4fa95@6a1f123d856	Bharti Airtel	ANDHRA PRADESH	VISAKHA
7	12	15-04-2018	caa996f397c7bc19aaf3f1602dc399f7c1dbde5b	Bharti Airtel	ANDHRA PRADESH	VISAKHA
8	17	15-05-2018	3a@c72c@3cca7e903f826c@fdf3ee22a41115d18	Bharti Airtel	ANDHRA PRADESH	VISAKHA
9	18	15-05-2018	4048cd8e87d@87474f@8d724cf7e3@9791add8@a	Bharti Airtel	ANDHRA PRADESH	VISAKHA

In [53]: `1 engine.execute('Select distinct(Mtariff) from inMobi where Carrier = "Bha`

Out[53]: `[('₹ 649',), ('₹ 500',), ('₹ 199',), ('₹ 99',)]`

Airtel Users on the basis of Monthly Tarrif on 15-04-2018

```
In [54]: 1 a_t1 = engine.execute('Select Date,Mtarrrif,count(Mtarrrif) from inMobi where
2 a_t1
```

```
Out[54]: [('15-04-2018', '₹ 199', 6903),
          ('15-04-2018', '₹ 500', 6854),
          ('15-04-2018', '₹ 649', 6903),
          ('15-04-2018', '₹ 99', 6864)]
```

Airtel Users on the basis of Monthly Tarrif on 15-05-2018

```
In [55]: 1 a_t2 = engine.execute('Select Date,Mtarrrif,count(Mtarrrif) from inMobi where
2 a_t2
```

```
Out[55]: [('15-05-2018', '₹ 199', 7088),
          ('15-05-2018', '₹ 500', 6906),
          ('15-05-2018', '₹ 649', 7074),
          ('15-05-2018', '₹ 99', 6915)]
```

```
In [56]: 1 Airtel_Users = pd.DataFrame({'Mtariff':[199,500,649,99], 'Users_inApril':
2 Airtel_Users
```

Out[56]:

	Mtariff	Users_inApril	Users_inMarch
0	199	6903	7088
1	500	6854	6906
2	649	6903	7074
3	99	6864	6915

```
In [57]: 1 a_ = engine.execute('Select Age,Date,count(Age) from inMobi where Carrier
2 a_
```

```
Out[57]: [('18-25', '15-05-2018', 8230),
          ('26-35', '15-05-2018', 8219),
          ('36-50', '15-05-2018', 7130),
          ('51+', '15-05-2018', 4404)]
```

```
In [58]: 1 a_ = engine.execute('Select Age,Date,count(Age) from inMobi where Carrier
2 a_
```

```
Out[58]: [('18-25', '15-04-2018', 8104),
          ('26-35', '15-04-2018', 8140),
          ('36-50', '15-04-2018', 6859),
          ('51+', '15-04-2018', 4421)]
```

```
In [59]: 1 Airtel_Users_a = pd.DataFrame({'Age': ['18-25', '26-35', '36-50', '51+'], 'Us  
2 Airtel_Users_a
```

Out[59]:

	Age	Users_inApril	Users_inMarch
0	18-25	8104	8230
1	26-35	8140	8219
2	36-50	6859	7130
3	51+	4420	4404