

**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**

**EXPERIMENT :- 09**

**Aim : Import italy-covid-daywise.csv and do analysis of given dataset using matplotlib**

**Source code :**

```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
df = pd.read_csv('italy-covid-daywise.csv')
df.head()
df.info()

#DatetimeIndex functio
df['year']=pd.DatetimeIndex(df.date).year
df['month']=pd.DatetimeIndex(df.date).month
df['weekdays']=pd.DatetimeIndex(df.date).weekday
covid_may=df[df['month']==5]
covid_may

#new cases in year 2022
ys=covid_tt.groupby('year')['new_cases'].sum()
ys

# Grouping the data by 'month' and calculating the sum for each variable
nc=covid_tt.groupby('month')['new_cases'].sum()
nc.head()
nt=covid_tt.groupby('month')['new_tests'].sum()
nt.head()
nd=covid_tt.groupby('month')['new_deaths'].sum()
nd.head()

# Extracting the month values for the x-axis
# Creating a plot for new_cases in red, new_deaths in green, and new_tests in blue
x=nc.index
plt.plot(nc.index, nc.values, color='r', label='New Cases')
plt.plot(x, nt, color='g', label='New Tests')
plt.plot(x, nd, color='b', label='New Deaths')
plt.legend()
plt.show()

#only plotting new cases graph by day
```

**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**

```
ncd=covid_tt.groupby('date')['new_cases'].sum()
ntd=covid_tt.groupby('date')['new_tests'].sum()
nnd=covid_tt.groupby('date')['new_deaths'].sum()
x=ncd.index
plt.plot(ncd)
plt.legend()
plt.show()

#only plotting for new test
x=ntd.index
plt.plot(ntd)
plt.legend()
plt.show()

#only plotting for new death
x=nnd.index
plt.plot(nnd)
plt.show()

#plotting all of them together
x=ncd.index
plt.plot(ncd.index, ncd.values, color='r', label='New Cases')
plt.plot(x, ntd, color='g', label='New Tests')
plt.plot(x, nnd, color='b', label='New Deaths')
plt.legend()
plt.show()

#more clear graph
import matplotlib.dates as d
import matplotlib.ticker as t
plt.figure(figsize=(10, 4))
plt.plot(ncd.index, ncd.values)
plt.xlabel('Date')
plt.ylabel('new_cases')
plt.legend()
plt.show()
```

**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**

**Output :**

	date	new_cases	new_deaths	new_tests
0	2019-12-31	0.0	0.0	NaN
1	2020-01-01	0.0	0.0	NaN
2	2020-01-02	0.0	0.0	NaN
3	2020-01-03	0.0	0.0	NaN
4	2020-01-04	0.0	0.0	NaN

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 248 entries, 0 to 247  
Data columns (total 4 columns):  
#   Column      Non-Null Count  Dtype  
---  ---  
0   date        248 non-null   object  
1   new_cases   248 non-null   float64  
2   new_deaths  248 non-null   float64  
3   new_tests   135 non-null   float64  
dtypes: float64(3), object(1)  
memory usage: 7.9+ KB
```

**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**

	date	new_cases	new_deaths	new_tests	year	month	weekdays
122	2020-05-01	1872.0	285.0	43732.0	2020	5	4
123	2020-05-02	1965.0	269.0	31231.0	2020	5	5
124	2020-05-03	1900.0	474.0	27047.0	2020	5	6
125	2020-05-04	1389.0	174.0	22999.0	2020	5	0
126	2020-05-05	1221.0	195.0	32211.0	2020	5	1
127	2020-05-06	1075.0	236.0	37771.0	2020	5	2
128	2020-05-07	1444.0	369.0	13665.0	2020	5	3
129	2020-05-08	1401.0	274.0	45428.0	2020	5	4
130	2020-05-09	1327.0	243.0	36091.0	2020	5	5
131	2020-05-10	1083.0	194.0	31384.0	2020	5	6
132	2020-05-11	802.0	165.0	25823.0	2020	5	0
133	2020-05-12	744.0	179.0	39620.0	2020	5	1
134	2020-05-13	1402.0	172.0	37049.0	2020	5	2
135	2020-05-14	888.0	195.0	41131.0	2020	5	3
136	2020-05-15	992.0	262.0	39027.0	2020	5	4
137	2020-05-16	789.0	242.0	40657.0	2020	5	5
138	2020-05-17	875.0	153.0	33505.0	2020	5	6

year

2020 271515.0

Name: new\_cases, dtype: float64

month

1 3.0

2 885.0

3 100851.0

4 101852.0

5 29073.0

**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**

Name: new\_cases, dtype: float64

month

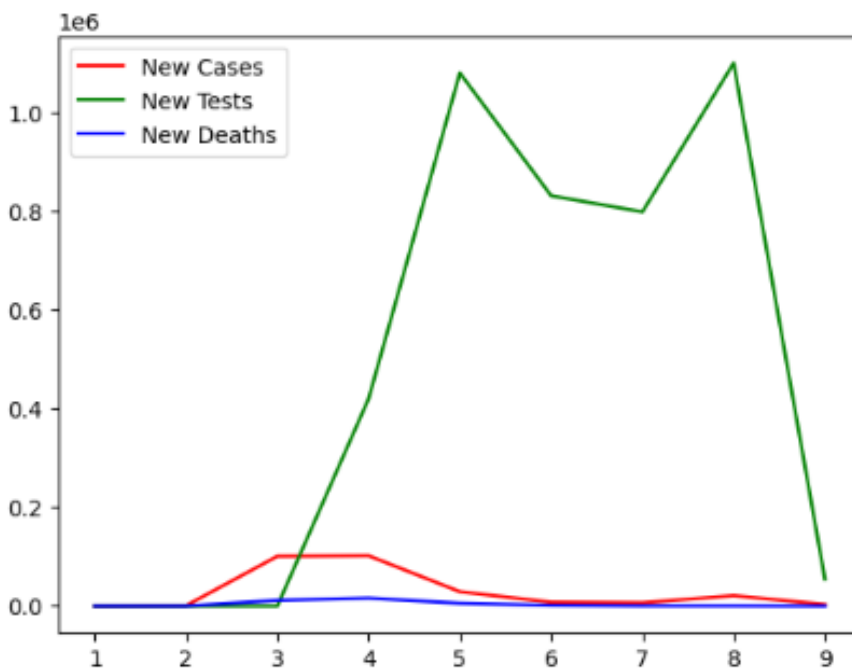
1	0.0
2	0.0
3	0.0
4	419591.0
5	1078720.0

Name: new\_tests, dtype: float64

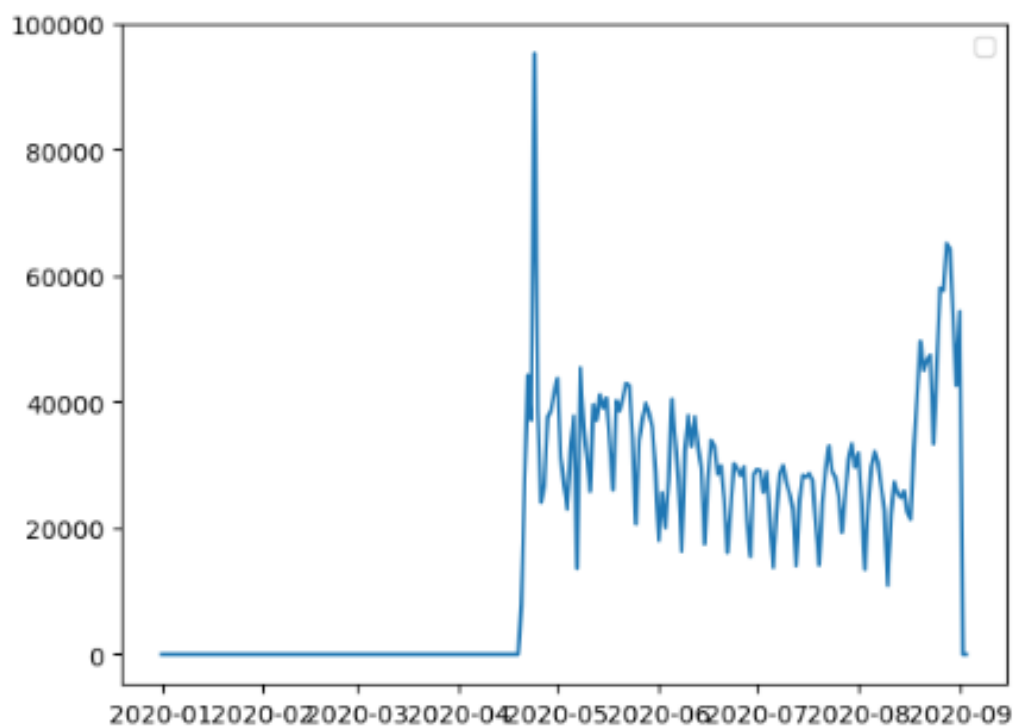
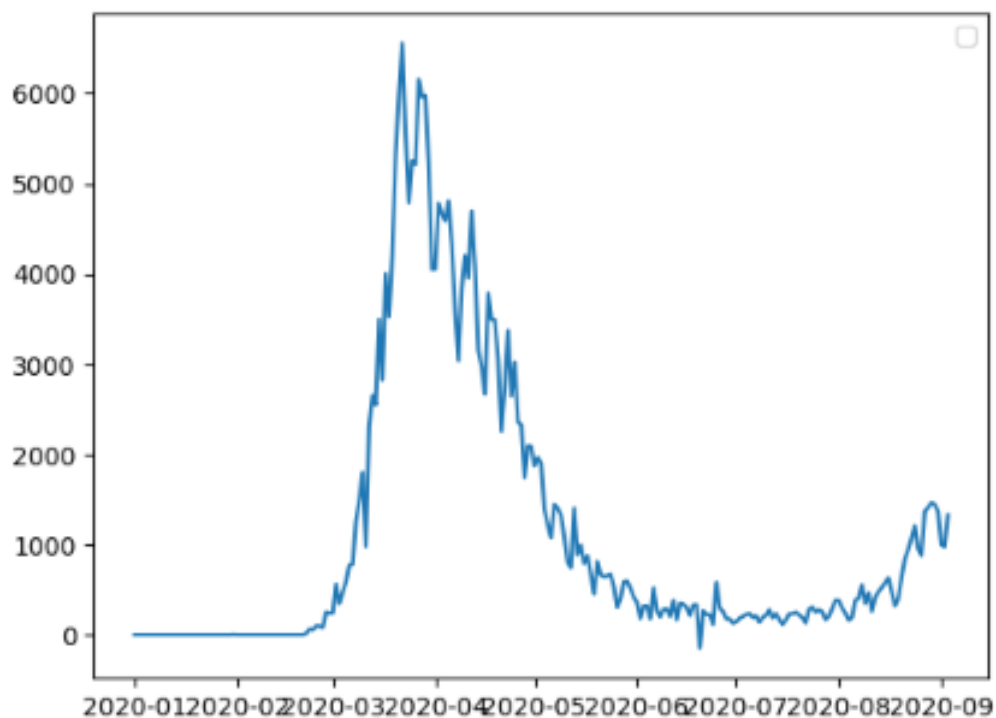
month

1	0.0
2	21.0
3	11570.0
4	16091.0
5	5658.0

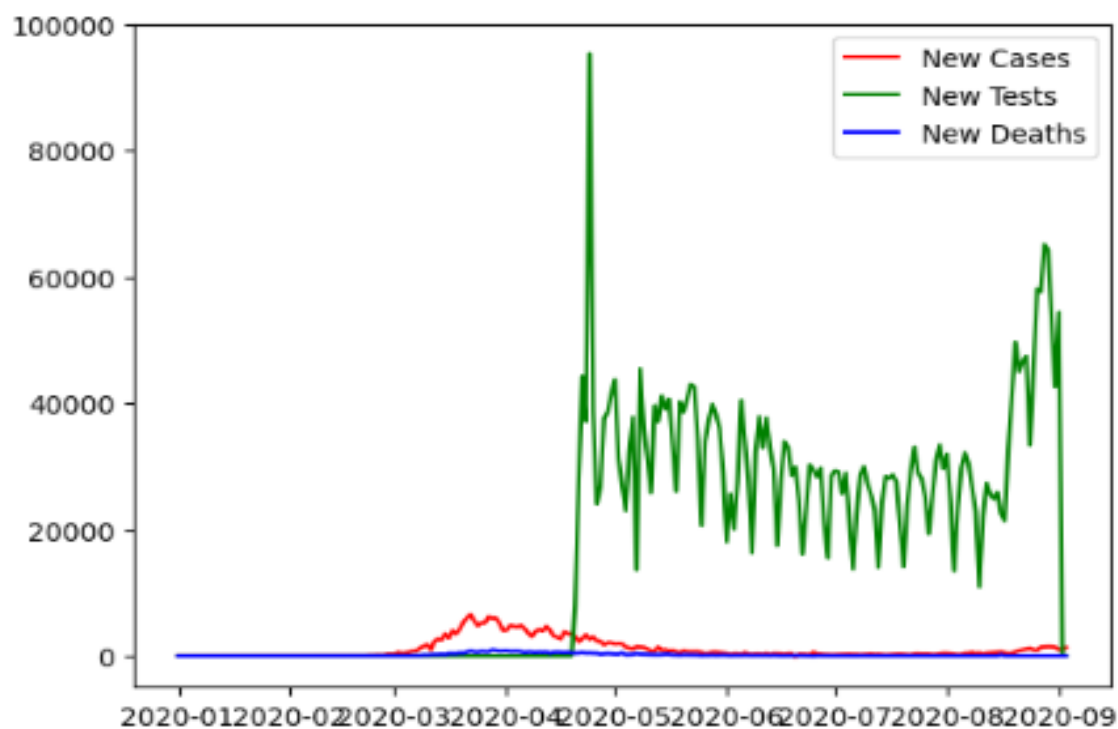
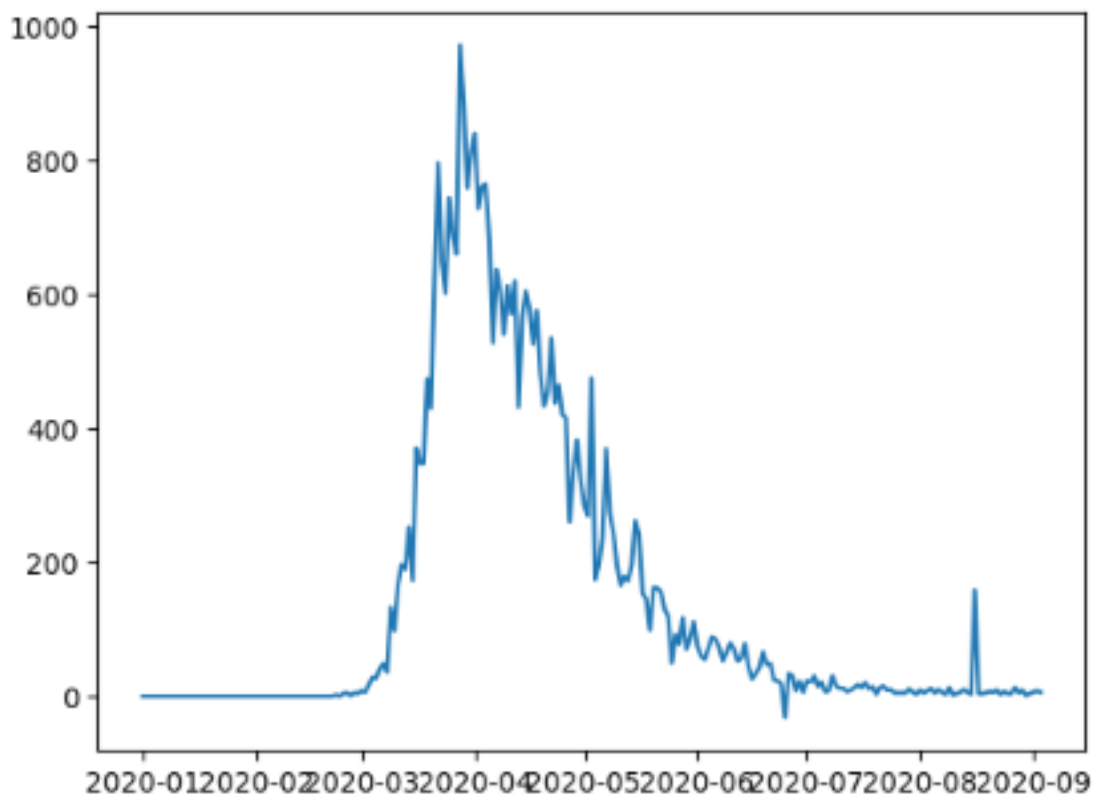
Name: new\_deaths, dtype: float64



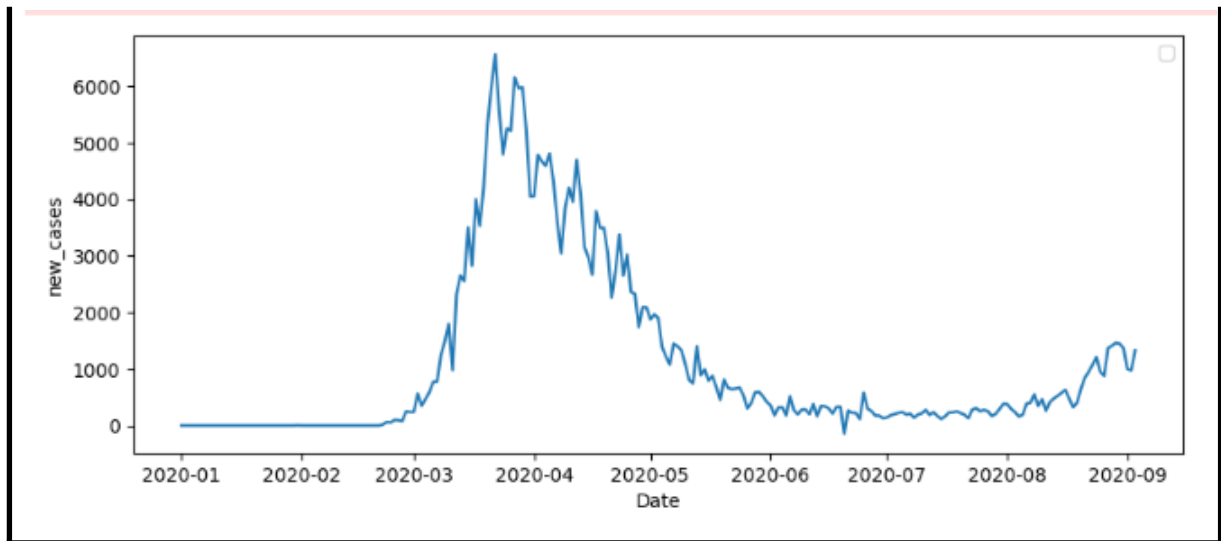
**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**



**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007



**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**





**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**

**EXPERIMENT :- 10**

**Aim :** Loading tips dataset and performing different operations using seaborn library .

**Source Code :**

```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
import seaborn as sns
plt.style.use("fivethirtyeight")
data=sns.load_dataset('tips')
data.head()

#bar plot
sns.barplot(x='sex',y='total_bill',data=data)
plt.show()

#displot
sns.displot(data['total_bill'],bins=[30,40,50,60],kde=True,rug=True)
plt.show()

#scatter plot
sns.relplot(x='total_bill',y='tip',kind='scatter',data=data)
plt.show()
sns.relplot(x='total_bill',y='tip',hue='smoker',data=data)
plt.show()
sns.relplot(x='total_bill',y='tip',hue='smoker',style='sex',data=data)
plt.show()
sns.relplot(x='total_bill',y='tip',hue='smoker',style='sex',size='size',data=data)
plt.show()

#dots
dots=sns.load_dataset("dots")
dots.tail(50)
g=sns.relplot(data=dots,kind='line',x='time',y='firing_rate',col='align',hue='choice',size='coherence',style='choice')
```

**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**

```
#catplot
sns.catplot(x='day',y='tip',kind='strip',hue='day',data=data)

#violinplot
sns.violinplot(x='time',y='tip',data=data,hue='sex',palette='dark',split=True)
plt.legend()
plt.show()

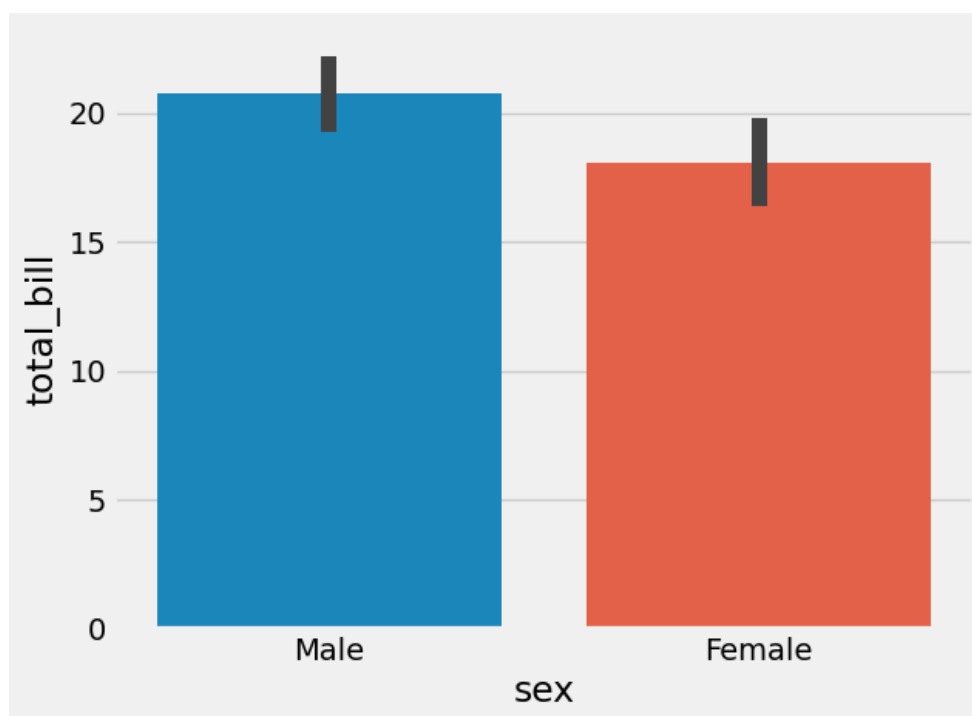
#facegrid
fg=sns.FacetGrid(data,col='time',row='sex')
fg=fg.map(plt.hist,"tip",color="red")

#facetGrid
fg=sns.FacetGrid(data,col='time',row='sex')
fg=fg.map(plt.scatter,'total_bill','tip',color='white', edgecolor='blue')
```

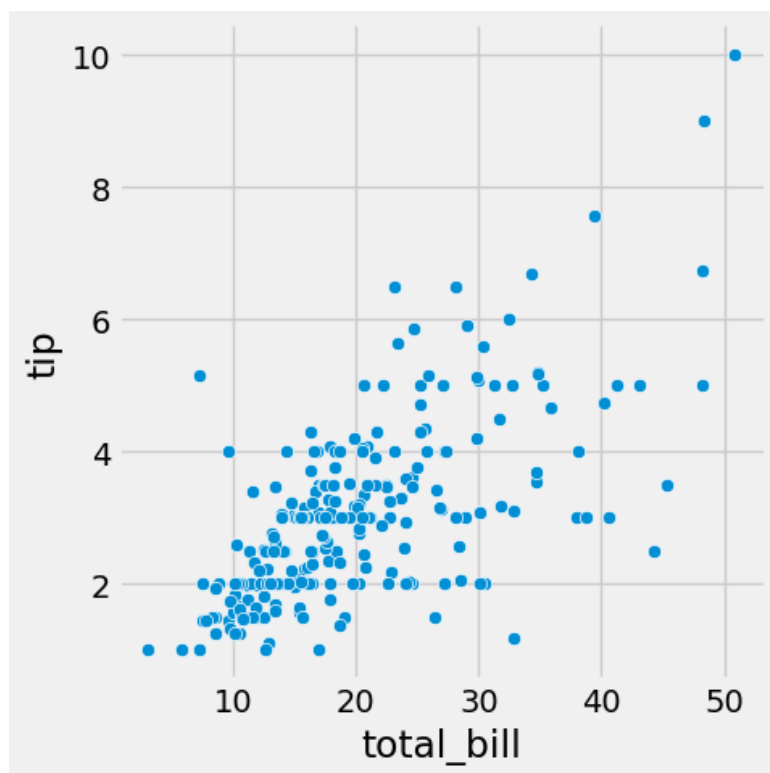
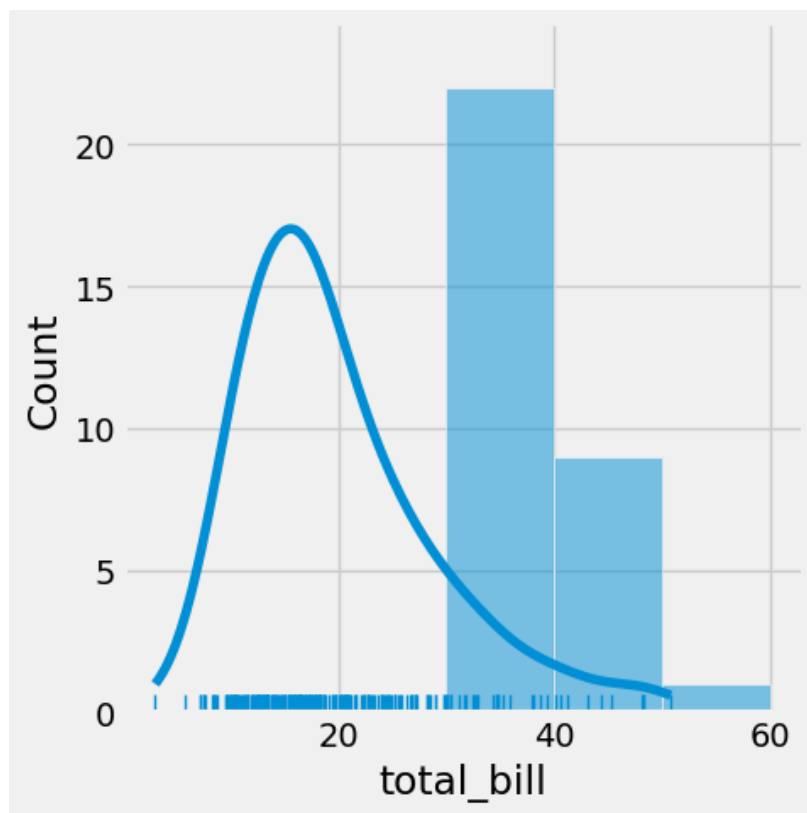
**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007**

**Output :**

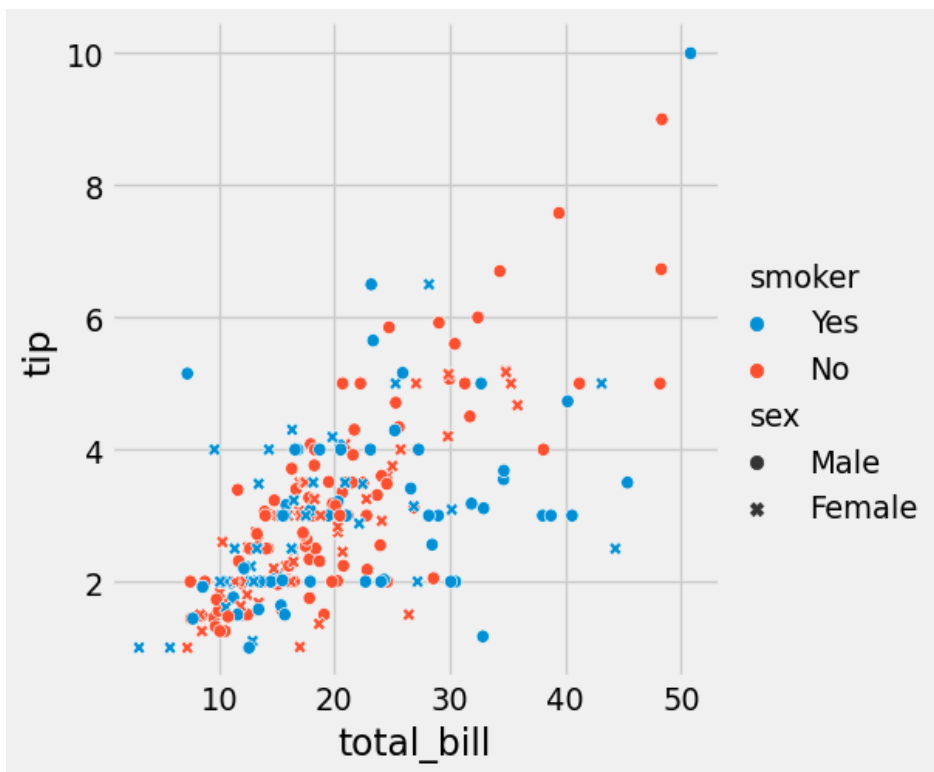
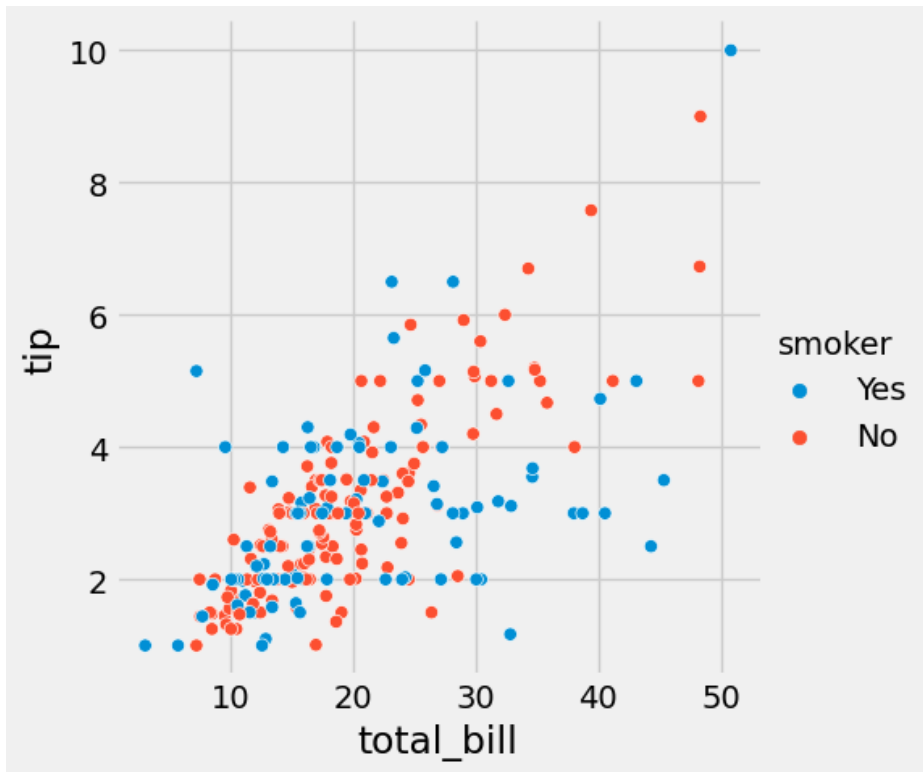
	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4



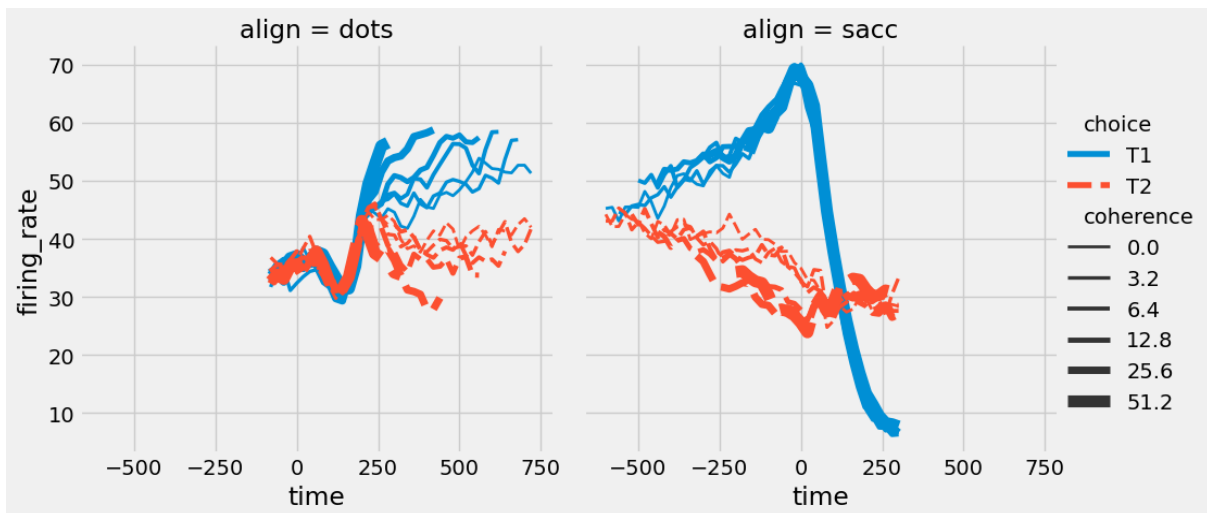
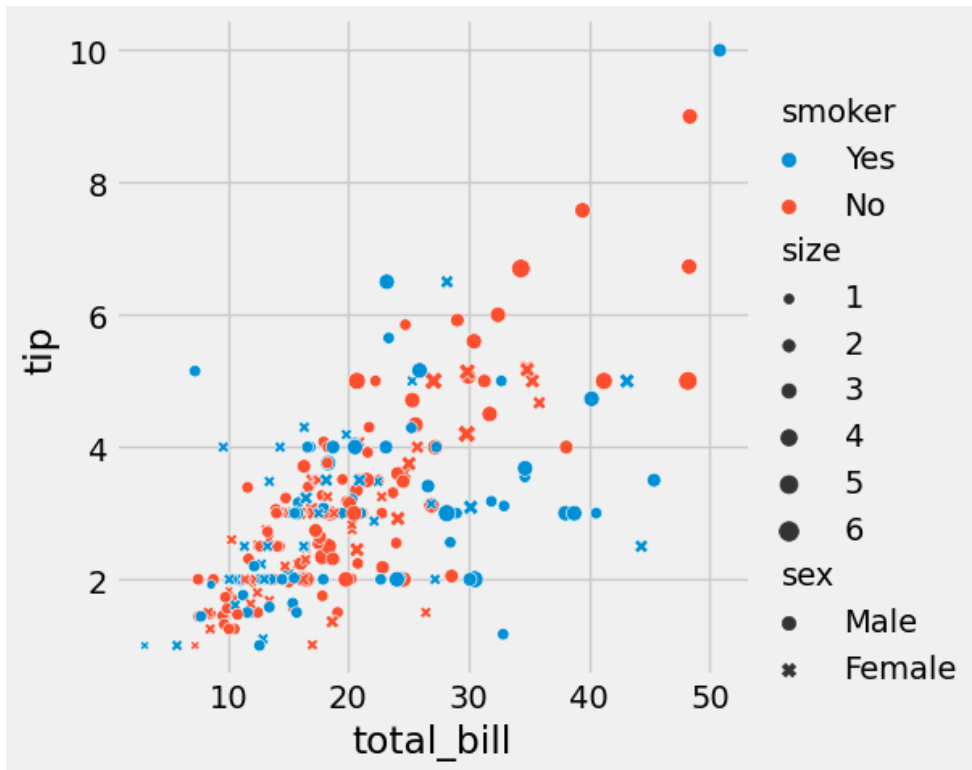
**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007



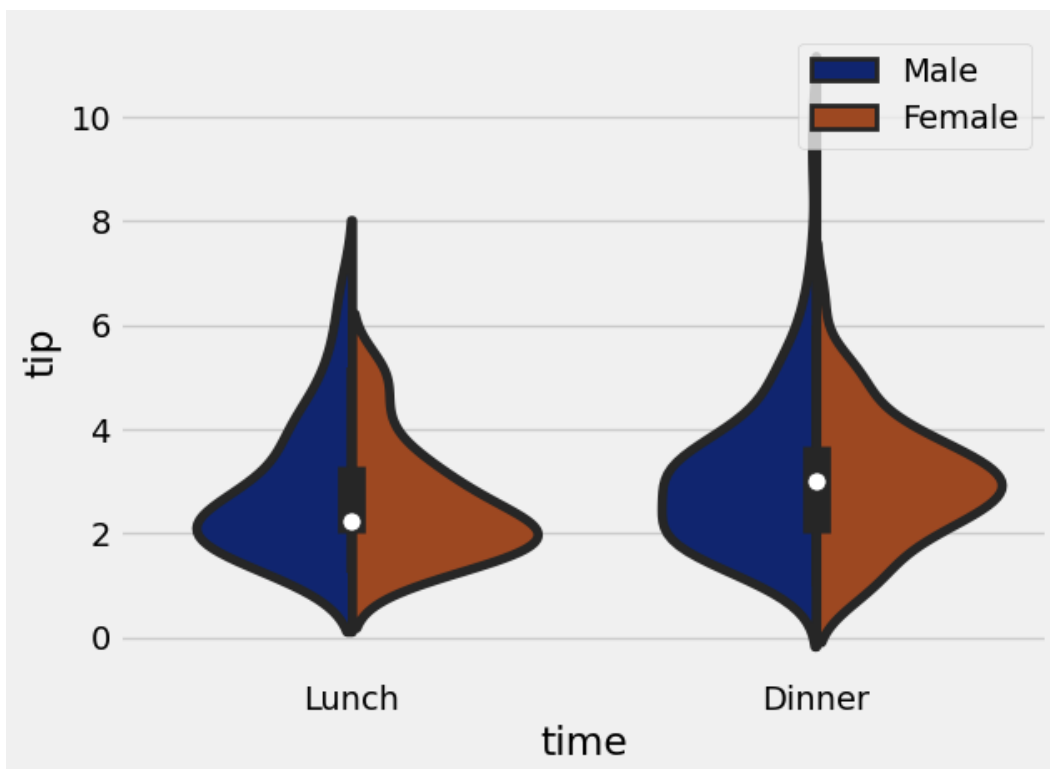
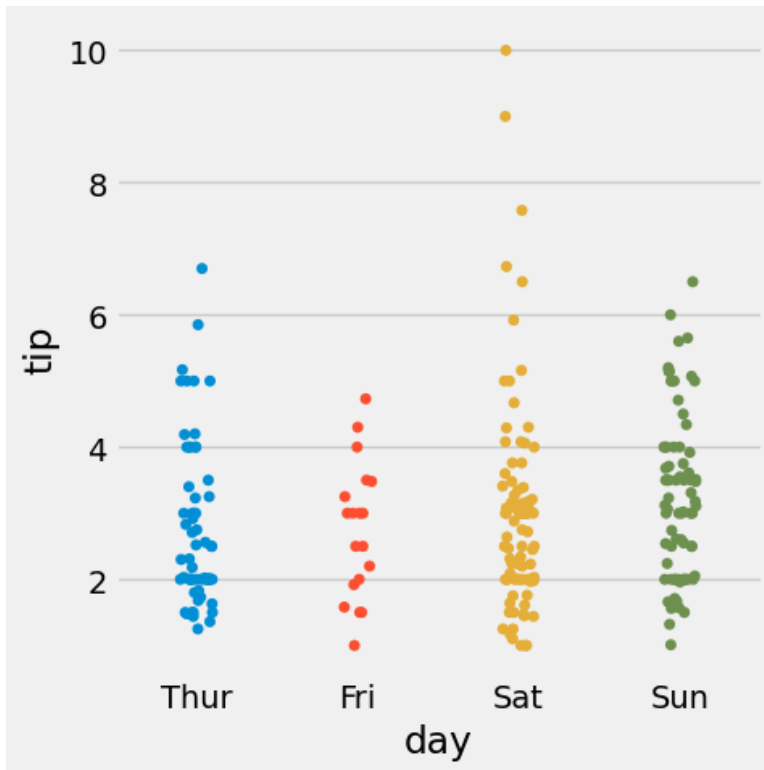
**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
DATA ANALYTICS LAB, 109591CS, ROLL NO.:300111021007



**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.:300111021007**



**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
DATA ANALYTICS LAB, 109591CS, ROLL NO.: 300111021007



**BHILAI INSTITUTE OF TECHNOLOGY, DURG**  
**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING(DS)**  
**DATA ANALYTICS LAB, 109591CS, ROLL NO.:300111021007**

