Lab – 02

Advanced Data Science

Aim- To perform marginal, conditional and joint probability from python.

Requirement – Python,Colab,Statistics

Code –

Q.1

Q . What is the probability of a country experiencing a COVID-19 outbreak, given tha it is located in a AFRO region?

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

```
array(['EMRO', 'EURO', 'AFRO', 'WPRO', 'AMRO', 'SEARO', 'Other'],
      dtype=object)
```

```
afro_region = df[df['WHO_region'] == 'AFRO']
afro_region.head(2)
```

	Date_reported	Country_code	Country	WHO_region	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
2588	2020-01-03	DZ	Algeria	AFRO	0	0	0	0
2589	2020-01-04	DZ	Algeria	AFRO	0	0	0	0

```
total_afro_region = len(afro_region)
total_afro_region
```

```
64700
```

Loading...

```
total_afro_region = len(afro_region)
total_afro_region
```

```
64700
```

```
afro_region_outbreak = afro_region[afro_region['New_cases'] > 0]
afro_region_outbreak = len(afro_region_outbreak)
afro_region_outbreak
```

```
33583
```

```
Q_no_1 =  (afro_region_outbreak / total_afro_region)
Q_no_1
```

```
0.5190571870170015
```

Q.2

Q . What is the probability of a country reporting more than 100 new COVID-19 cases in a day, given that it has a high population density?

```
total_countries = len(df)
total_countries
```

Python

```
306678
```

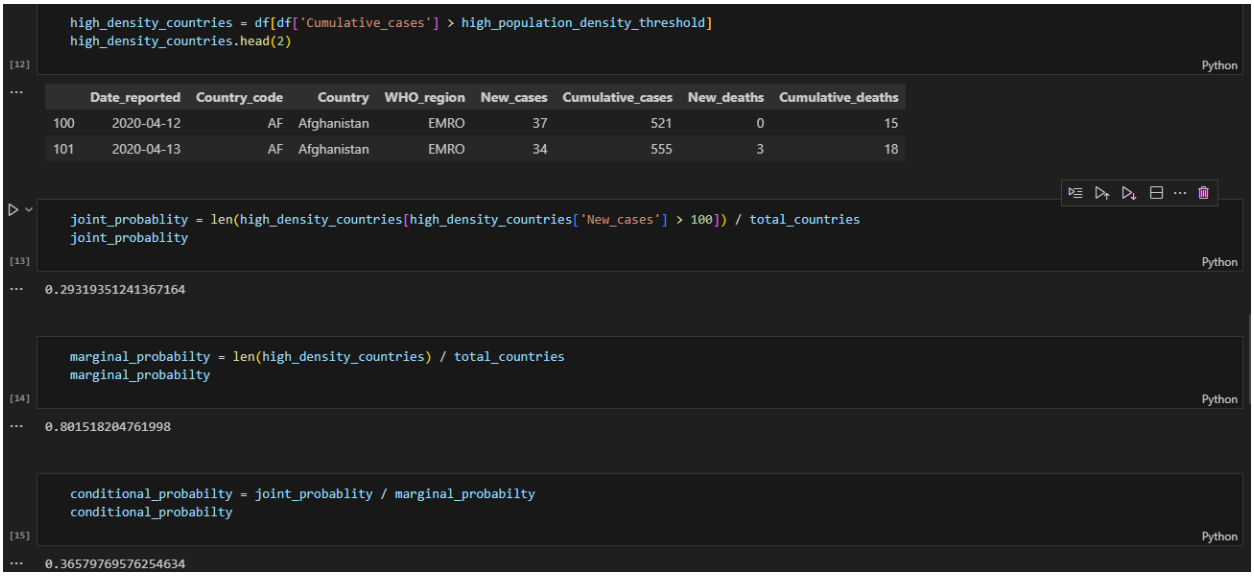
```
high_population_density_threshold = 500
```

Python

```
high_density_countries = df[df['Cumulative_cases'] > high_population_density_threshold]
high_density_countries.head(2)
```

Python

	Date_reported	Country_code	Country	WHO_region	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
100	2020-04-12	AF	Afghanistan	EMRO	37	521	0	15
101	2020-04-13	AF	Afghanistan	EMRO	34	555	3	18



```
high_density_countries = df[df['Cumulative_cases'] > high_population_density_threshold]
high_density_countries.head(2)
```

[12] Python

	Date_reported	Country_code	Country	WHO_region	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
100	2020-04-12	AF	Afghanistan	EMRO	37	521	0	15
101	2020-04-13	AF	Afghanistan	EMRO	34	555	3	18

```
joint_probablity = len(high_density_countries[high_density_countries['New_cases'] > 100]) / total_countries
joint_probablity
```

[13] Python

```
0.29319351241367164
```

```
marginal_probabilty = len(high_density_countries) / total_countries
marginal_probabilty
```

[14] Python

```
0.801518204761998
```

```
conditional_probabilty = joint_probablity / marginal_probabilty
conditional_probabilty
```

[15] Python

```
0.36579769576254634
```

Q.3

Q. What is the joint probability of a country reporting more than 100 new COVID-19 cases in a day and having a high population density?

```
joint_probability = len(afro_region[afro_region['New_cases'] > 100]) / total_countries
joint_probability
```

[6] Python

```
0.03313899268940061
```

Q.4

Q. What is the marginal probability of a country reporting more than 100 new COVID-19 cases in a day?

```
margibnal_probablity = len(afro_region) / total_countries
margibnal_probablity
```

[17] Python

```
0.2109704641350211
```

Q.5


Q. What is the conditional probability of a country having a high population density, given that it has reported more than 100 new COVID-19 cases in a day?
conditional_probabilty = joint_probability / margibnal_probablity
conditional_probabilty
[18]
Python
0.1570788253477589