Practical No. 14

by Kartik Deshpande

Use the following dataset "covid_vaccine_statewise.csv" and Perform the following data analytics.

1. Describe the dataset.
2. No. of persons by state vaccinated for first dose.
3. No. of persons by state vaccinated for second dose.
4. No. of males vaccinated.
5. No. of females vaccinated.
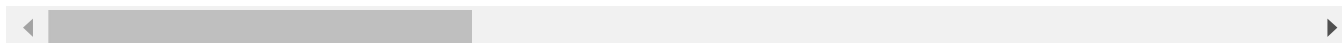
```
In [1]:    import pandas as pd
```

```
In [4]:    df = pd.read_csv("covid_vaccine_statewise.csv")
```

```
In [6]:    df.head()
```

Out[6]:

	Updated On	State	Total Doses Administered	Sessions	Sites	First Dose Administered	Second Dose Administered	Male (Doses Administered)	Ac
0	16/01/2021	India	48276.0	3455.0	2957.0	48276.0	0.0	NaN	
1	17/01/2021	India	58604.0	8532.0	4954.0	58604.0	0.0	NaN	
2	18/01/2021	India	99449.0	13611.0	6583.0	99449.0	0.0	NaN	
3	19/01/2021	India	195525.0	17855.0	7951.0	195525.0	0.0	NaN	
4	20/01/2021	India	251280.0	25472.0	10504.0	251280.0	0.0	NaN	

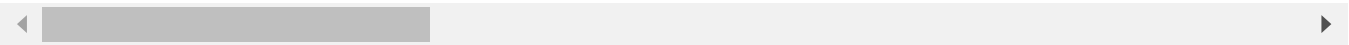5 rows × 24 columns

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

Out[7]:   (7845, 24)

```
In [8]:    df.describe()
```

	Total Doses Administered	Sessions	Sites	First Dose Administered	Second Dose Administered	Male (Doses Administered)	Admini
count	7.621000e+03	7.621000e+03	7621.000000	7.621000e+03	7.621000e+03	7.461000e+03	7.4610
mean	9.188171e+06	4.792358e+05	2282.872064	7.414415e+06	1.773755e+06	3.620156e+06	3.1684
std	3.746180e+07	1.911511e+06	7275.973730	2.995209e+07	7.570382e+06	1.737938e+07	1.5153
min	7.000000e+00	0.000000e+00	0.000000	7.000000e+00	0.000000e+00	0.000000e+00	2.0000
25%	1.356570e+05	6.004000e+03	69.000000	1.166320e+05	1.283100e+04	5.655500e+04	5.2107
50%	8.182020e+05	4.547000e+04	597.000000	6.614590e+05	1.388180e+05	3.897850e+05	3.3423
75%	6.625243e+06	3.428690e+05	1708.000000	5.387805e+06	1.166434e+06	2.735777e+06	2.5615
max	5.132284e+08	3.501031e+07	73933.000000	4.001504e+08	1.130780e+08	2.701636e+08	2.3951

8 rows × 22 columns

In [9]:
```python
df.describe(include='object')
```

Out[9]:

	Updated On	State
count	7845	7845
unique	213	37
top	16/01/2021	Delhi
freq	37	213

In [10]:
```python
avg_firstdose = df["First Dose Administered"].astype("float").mean(axis = 0)
print("Average of First Dose:", avg_firstdose)
```
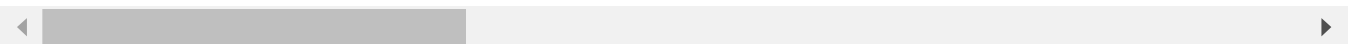
Average of First Dose: 7414415.300354284

In [11]:
```python
df["First Dose Administered"].fillna(value = avg_firstdose, inplace=True)
df.head()
```

Out[11]:

	Updated On	State	Total Doses Administered	Sessions	Sites	First Dose Administered	Second Dose Administered	Male (Doses Administered)	Ac
0	16/01/2021	India	48276.0	3455.0	2957.0	48276.0	0.0	NaN	
1	17/01/2021	India	58604.0	8532.0	4954.0	58604.0	0.0	NaN	
2	18/01/2021	India	99449.0	13611.0	6583.0	99449.0	0.0	NaN	
3	19/01/2021	India	195525.0	17855.0	7951.0	195525.0	0.0	NaN	
4	20/01/2021	India	251280.0	25472.0	10504.0	251280.0	0.0	NaN	

5 rows × 24 columns

In [12]:
```python
avg_seconddose = df["Second Dose Administered"].astype("float").mean(axis = 0)
print("Average of Second Dose:", avg_seconddose)
```

Average of Second Dose: 1773755.2436688098

```
In [13]: df["Second Dose Administered"].fillna(value = avg_seconddose, inplace = True)
         df.head()
```

Out[13]:

	Updated On	State	Total Doses Administered	Sessions	Sites	First Dose Administered	Second Dose Administered	Male (Doses Administered)	Ac
0	16/01/2021	India	48276.0	3455.0	2957.0	48276.0	0.0	NaN	
1	17/01/2021	India	58604.0	8532.0	4954.0	58604.0	0.0	NaN	
2	18/01/2021	India	99449.0	13611.0	6583.0	99449.0	0.0	NaN	
3	19/01/2021	India	195525.0	17855.0	7951.0	195525.0	0.0	NaN	
4	20/01/2021	India	251280.0	25472.0	10504.0	251280.0	0.0	NaN	

5 rows × 24 columns

```
In [15]: first_dose = df.groupby('State')[['First Dose Administered']].sum()
         first_dose.head()
```

Out[15]:

	First Dose Administered
State	
Andaman and Nicobar Islands	6.091235e+07
Andhra Pradesh	1.277347e+09
Arunachal Pradesh	9.349147e+07
Assam	6.300867e+08
Bihar	1.514989e+09

```
In [16]: first_dose = df.groupby('State')[['Second Dose Administered']].sum()
         first_dose.head()
```

Out[16]:

	Second Dose Administered
State	
Andaman and Nicobar Islands	1.476109e+07
Andhra Pradesh	3.694601e+08
Arunachal Pradesh	2.257485e+07
Assam	1.414313e+08
Bihar	2.814331e+08

```
In [17]: male = df["Male(Individuals Vaccinated)"].sum()
         print("The Total Number of Male Individuals Vaccinated are :", int(male))
```

The Total Number of Male Individuals Vaccinated are : 7138698858

```
In [18]: female = df["Female(Individuals Vaccinated)"].sum()
         print("The Total Number of Female Individuals Vaccinated are :", int(female))
```

The Total Number of Female Individuals Vaccinated are : 6321628736