

## **Air Quality**

### **1.Data cleaning**

```
import pandas as pd
```

```
import numpy as np
```

```
import warnings
```

```
warnings.filterwarnings('ignore')
```

```
df=pd.read_csv('data.csv',encoding ='cp1252')
```

```
df.head()
```

```
df.shape()
```

```
df.info()
```

```
df.isnull().sum()
```

```
df.count
```

```
df.describe()
```

```
df=df.drop(['stn_code','agency','sampling_date','location_monitoring_station'])
```

```
df.column
```

```
df["type"].unique()
```

## Facebook

1.import pandas as pd

import numpy as np

df = pd.read\_csv('dataset\_Facebook.csv', sep = " , ")

import io

2. df.head(10)

3. #Creating Subset

df1 =df [['Page total likes' , 'Category' , 'Post Month' , 'Post Weekday']].loc[0:15]

print (df1)

df2 =df [['Page total likes' , 'Category' , 'Post Month' , 'Post Weekday']].loc[16:30]

print (df2)

df3 =df [['Page total likes' , 'Category' , 'Post Month' , 'Post Weekday']].loc[31:50]

print (df3)

4. #Merge data

merging =pd.concat([df1,df2,df3])

print(merging)

5.#Sort data

sort\_values = df.sort\_values ('Page total likes' , ascending = False)

print(sort\_values)

6. #Transposing data

transposed\_df = df.transpose()

print (transposed\_df)

7. #shape

df.shape

8. #Reshape

```

pivot_table = pd.pivot_table (df,index=['Type' , 'Category'],values='like')
print(pivot_table)

9.#Describe
Df.describe()

10.#Select column
df[['Category', 'Paid']]

11.#Delete row/column
df = df.drop('Paid' , axis=1)

12.#Restore row/column
df = pd.read_csv('your_dataset.csv')

```

```

import numpy as np
from sklearn.impute import
simple Imputer
imputer=simpleImputer
(misssin_values=np.nan,strategy='mean')
Df[cols]=imputer.fir_transform(df[cols])
df.head(5)
df.info()
df['type'].value_counts()
df['type'].replace(("PRO":)),
{"l":z,"Ro":3,"s":4,"RIRUO":5,"R":6}inplace=true)
Df['type']
From sklearn.preprocessing
Import Label Encoder
Labe; encoder = laber Encoder()
Df("state")=label encoder.fit_transform(df("state"))

```

```
Df.head(5)
Df['state'].value_counts()
dfAndra = df((df['state']==0))
df Andhra['location'].value_counts()
from sklearn.preprocessing import One Hot Encoder
onehotencoder = One Hot Encoder
(sparse=False,handle_unknown='error',drop='first')
pd.DataFrame(onehotencoder.fit_transform(df Andhra[['location']]))
```