AIR QUALITY ANALYSIS AND PREDICTION IN TAMILNADU

DEVELOPMENT PART-1:

CODE:

```
import pandas as pd
import numpy as np
data=pd.read_csv("airquality_analysis.csv")
print(data.head())
```

==========

```
   Stn Code Sampling Date       State  ...   NO2 RSPM/PM10 PM 2.5
0        38      01-02-14  Tamil Nadu  ...  17.0      55.0    NaN
1        38      01-07-14  Tamil Nadu  ...  17.0      45.0    NaN
2        38      21-01-14  Tamil Nadu  ...  18.0      50.0    NaN
3        38      23-01-14  Tamil Nadu  ...  16.0      46.0    NaN
4        38      28-01-14  Tamil Nadu  ...  14.0      42.0    NaN
[5 rows x 11 columns]
```

```
print(data.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2879 entries, 0 to 2878
Data columns (total 11 columns):
 #   Column                      Non-Null Count  Dtype
---  ------                      --------------  -----
 0   Stn Code                    2879 non-null   int64
 1   Sampling Date               2879 non-null   object
 2   State                       2879 non-null   object
 3   City/Town/Village/Area      2879 non-null   object
 4   Location of Monitoring Station  2879 non-null   object
 5   Agency                      2879 non-null   object
 6   Type of Location            2879 non-null   object
 7   SO2                         2868 non-null   float64
 8   NO2                         2866 non-null   float64
 9   RSPM/PM10                   2875 non-null   float64
 10  PM 2.5                      0 non-null      float64
dtypes: float64(4), int64(1), object(6)
memory usage: 247.5+ KB
```

```
print(data.describe())
```

```
          Stn Code          SO2          NO2    RSPM/PM10  PM 2.5
count  2879.000000  2868.000000  2866.000000  2875.000000     0.0
```

mean	475.750261	11.503138	22.136776	62.494261	NaN
std	277.675577	5.051702	7.128694	31.368745	NaN
min	38.000000	2.000000	5.000000	12.000000	NaN
25%	238.000000	8.000000	17.000000	41.000000	NaN
50%	366.000000	12.000000	22.000000	55.000000	NaN
75%	764.000000	15.000000	25.000000	78.000000	NaN
max	773.000000	49.000000	71.000000	269.000000	NaN

print(data.isnull())

	Stn Code	SO2	NO2	RSPM/PM10	PM 2.5
count	2879.000000	2868.000000	2866.000000	2875.000000	0.0
mean	475.750261	11.503138	22.136776	62.494261	NaN
std	277.675577	5.051702	7.128694	31.368745	NaN
min	38.000000	2.000000	5.000000	12.000000	NaN
25%	238.000000	8.000000	17.000000	41.000000	NaN
50%	366.000000	12.000000	22.000000	55.000000	NaN
75%	764.000000	15.000000	25.000000	78.000000	NaN
max	773.000000	49.000000	71.000000	269.000000	NaN

	Stn Code	Sampling Date	State	...	NO2	RSPM/PM10	PM 2.5
0	False	False	False	...	False	False	True
1	False	False	False	...	False	False	True
2	False	False	False	...	False	False	True
3	False	False	False	...	False	False	True
4	False	False	False	...	False	False	True
...	...	...	...	...	...	...	...
2874	False	False	False	...	False	False	True
2875	False	False	False	...	False	False	True
2876	False	False	False	...	False	False	True
2877	False	False	False	...	False	False	True
2878	False	False	False	...	False	False	True

[2879 rows x 11 columns]