Phase 3: Development Part 1

Data Loading and Data Preprocessing Phase

Step 1: Loading Libraries and Dataset

```
import pandas as pd
import geopandas as gd
import sklearn as sk
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import folium as fl

Dataset = pd.read_csv(r"C:\Users\ns117\Downloads\IBM
NM\DDW_B06ST_3300_State_TAMIL_NADU-2011.csv")
```

Step 2: Exploring the Dataset

```
Dataset.head()
Dataset.describe()
Dataset.columns
```

Explanation: The dataset was loaded, and the first few rows, descriptive statistics, and column names were displayed for initial exploration.

Step 3: Handling Missing Values

```
Dataset.isnull().sum()
```

Output:

```
Table Code                                              0
State Code                                              0
District Code                                           0
Area Name                                               0
Total/ Rural/ Urban                                     0
                                                       ..
Industrial Category - R to U - HHI - Males              0
Industrial Category - R to U - HHI - Females            0
Industrial Category - R to U - Non HHI - Persons        0
Industrial Category - R to U - Non HHI - Males          0
Industrial Category - R to U - Non HHI - Females        0
Length: 69, dtype: int64
```

Explanation: As we can see, there are no null values present in our dataset, so we can directly start implementing our code with the given dataset.Since there are no null values present in our dataset, we don't need to create a new dataset.

Conclusion:

In this phase, we successfully loaded the dataset, explored its initial structure, and scanned for missing values. The dataset is now ready for in-depth analysis, modeling, and prediction in the subsequent phases of the project.