

|  |  |
| --- | --- |
| **DOMAIN :** | **Data Analytics With Cognos** |
| **PROJECT TITLE :** | **Public Transportation Analysis** |
| **TEAM MEMBERS AND REGISTER NUMBER** | **1.Bathri Narayanan S - 420421104011**  **2.Dinesh G - 420421104016**  **3.Dhilip Kumar M - 420421104015**  **4.VinothKumar S - 420421104303** |

**Table of Contents:**

1. **Introduction**
   * **Briefly introduce the project, its objectives, and the dataset.**
2. **Data Acquisition**
   * **Provide information on how to download the dataset from Kaggle.**
   * **Mention the required libraries and tools for this project (e.g., Python, Jupyter Notebook, IBM Cognos).**
3. **Data Preprocessing**
   * **Load the dataset using Python and the Pandas library.**
   * **Check for missing values and handle them if necessary.**
   * **Perform data cleaning and transformation as needed.**
   * **Save the cleaned dataset for further analysis.**
4. **Exploratory Data Analysis (EDA)**
   * **Use Python to explore the dataset and gain insights.**
   * **Generate basic statistics, such as mean, median, and standard deviation.**
   * **Create visualizations (histograms, scatter plots, etc.) to understand the data distribution and relationships.**
   * **Identify key trends and patterns in the data.**
5. **Integration with IBM Cognos**
   * **Explain how to connect IBM Cognos to your dataset.**
   * **Import the preprocessed data into IBM Cognos for more advanced visualizations.**
6. **Advanced Data Visualizations with IBM Cognos**
   * **Create various visualizations within IBM Cognos, such as dashboards, charts, and reports.**
   * **Utilize Cognos features to perform complex analytics, such as time series analysis and predictive modeling.**
   * **Describe the insights and conclusions drawn from the visualizations.**
7. **Conclusion**
   * **Summarize the key findings from your data analysis.**
   * **Discuss any business insights or recommendations based on the analysis.**
   * **Reflect on the project's success and any challenges faced.**

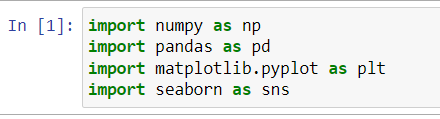
**DataSet Link:**

[**https://www.kaggle.com/datasets/rednivrug/unisys?select=20140711.CSV**](https://www.kaggle.com/datasets/rednivrug/unisys?select=20140711.CSV)

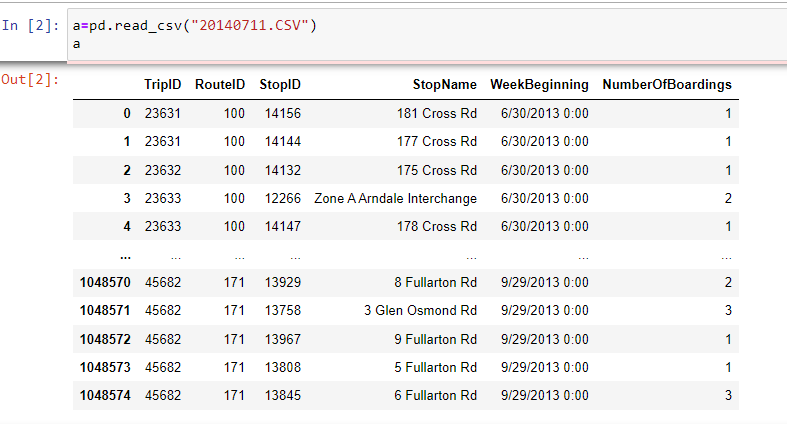
**Program:**

**Data Preprocessing:**

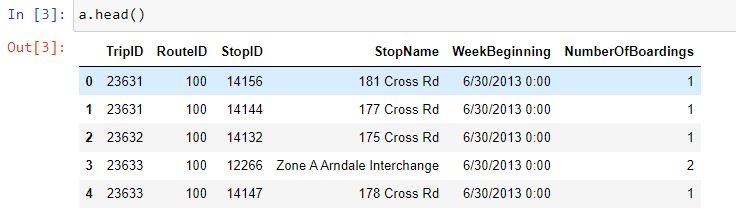
**1.Import the Library Function:**

****

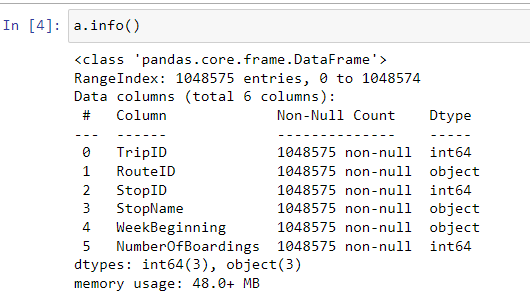
**2. Load the Dataset:**

****

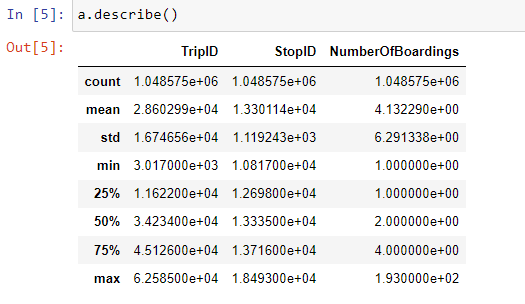
**3. Head of the Dataset:**

****

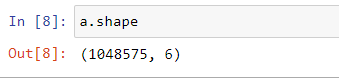
**4. Info of the Dataset:**

****

**5.Describe the Dataset:**

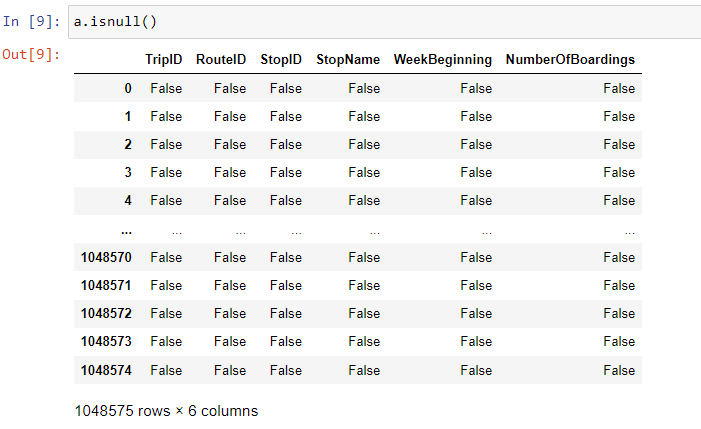
****

**6.Shape of the Dataset:**

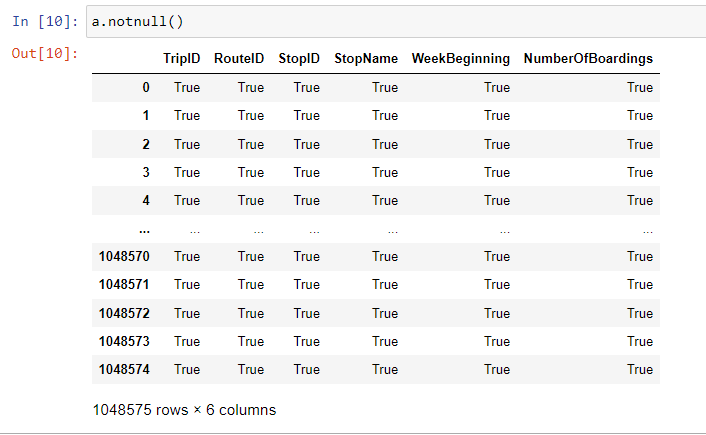
****

**7. Handling Missing Data:**

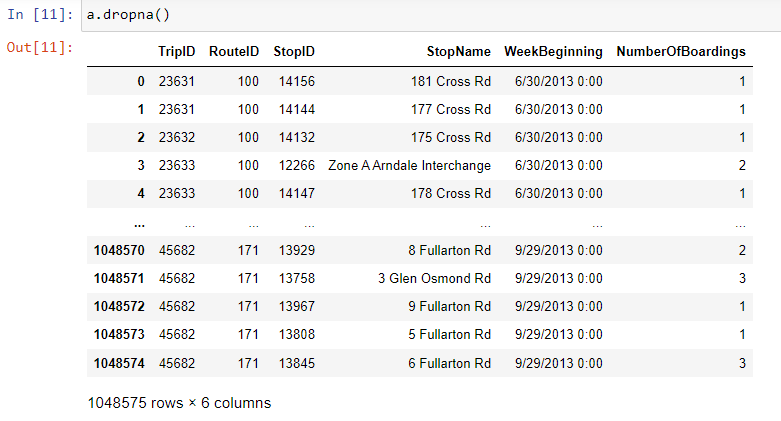
**i) isnull()**

****

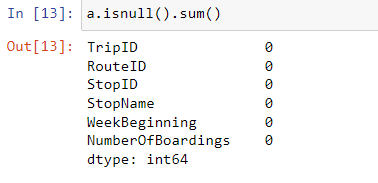
**ii) notnull()**



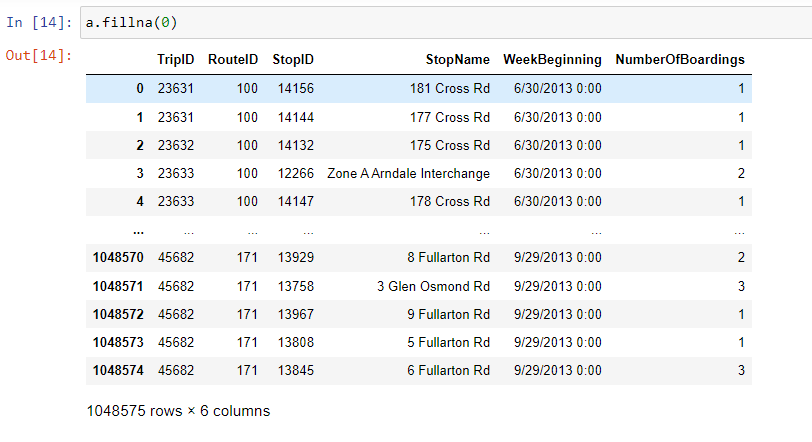
**iii) Dropna()**



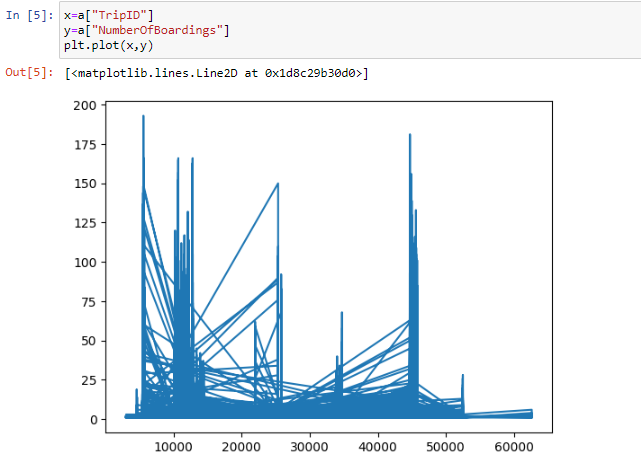
**iv) isnull().sum()**

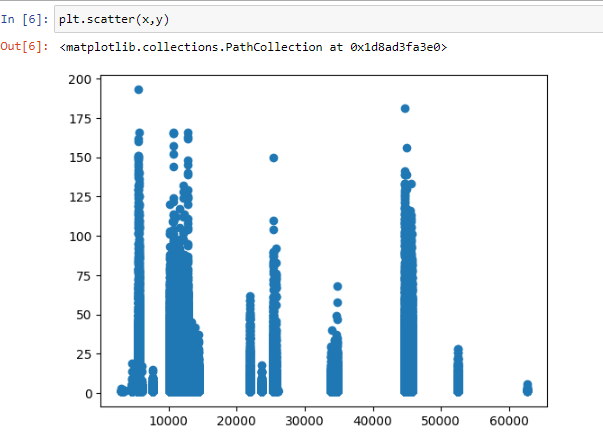
****

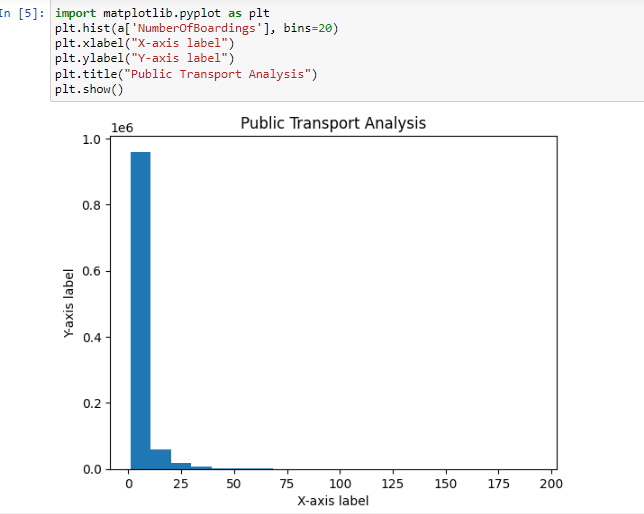
1. **Fillna()**



1. **Data Visualization:**

****

****

****