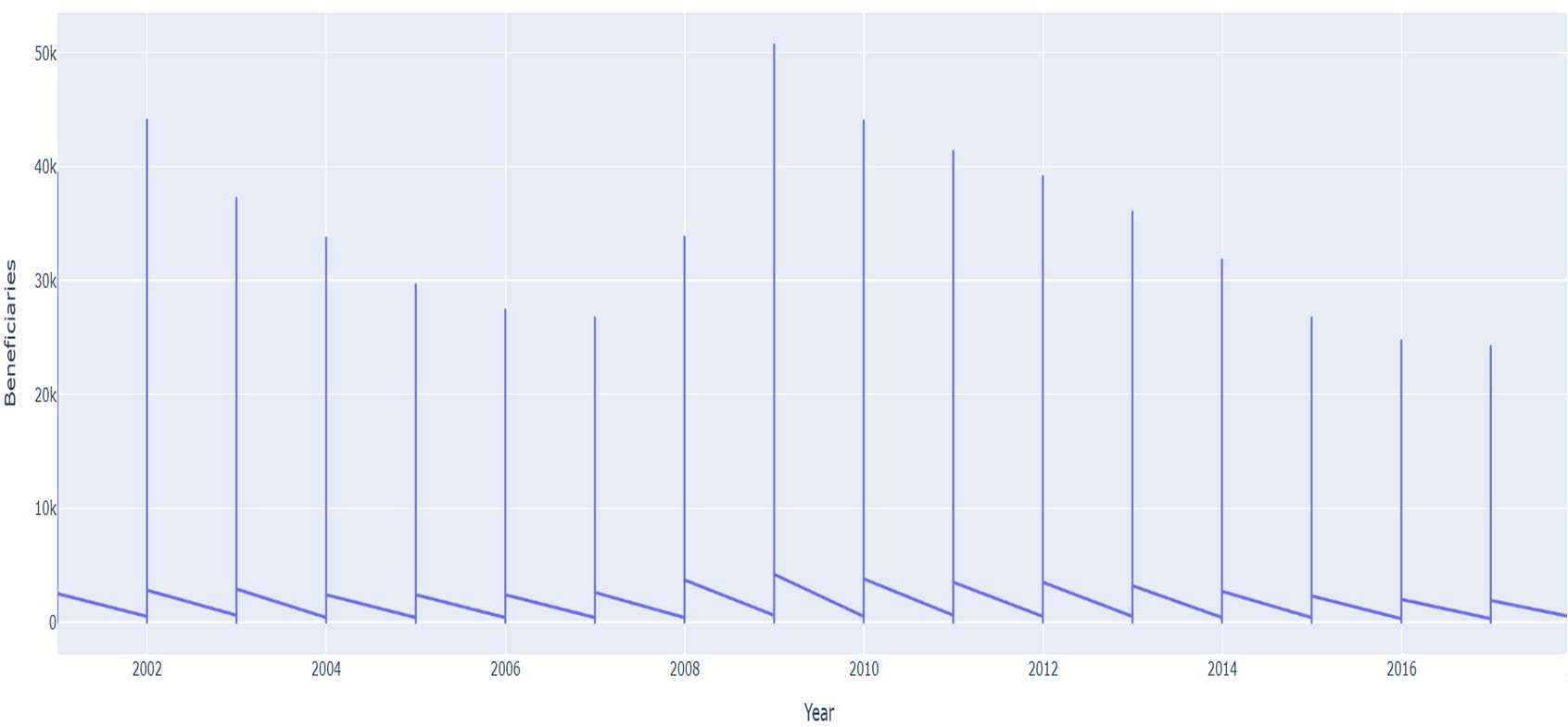
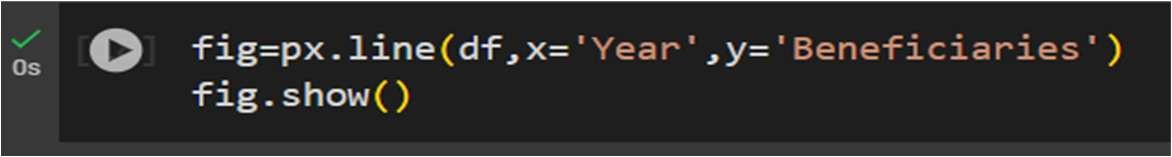
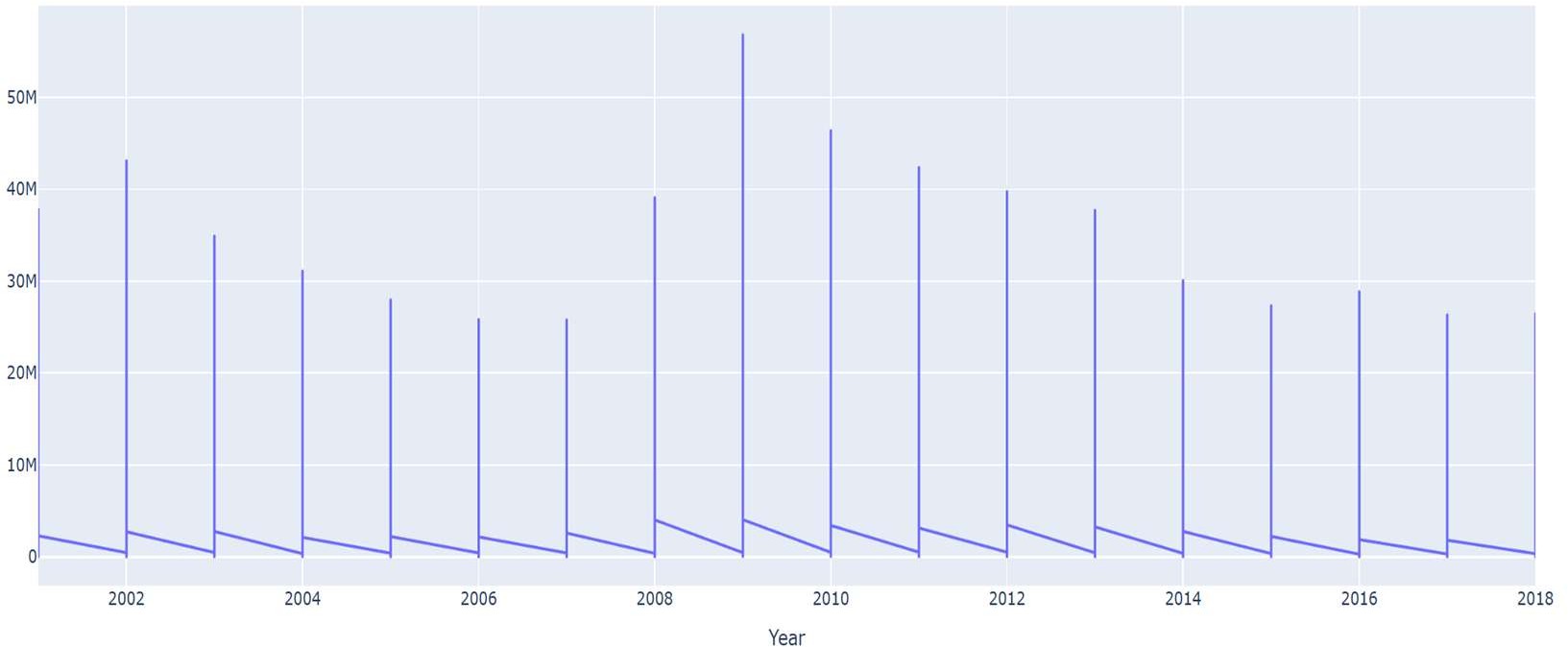
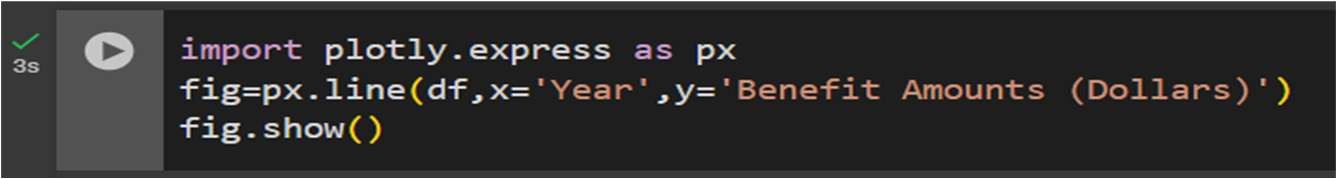
# Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 15 June 2025 |
| Team ID | SWTID1749896042 |
| Project Title | Unemployed insurance beneficiary forecasting |
| Maximum Marks | 6 Marks |

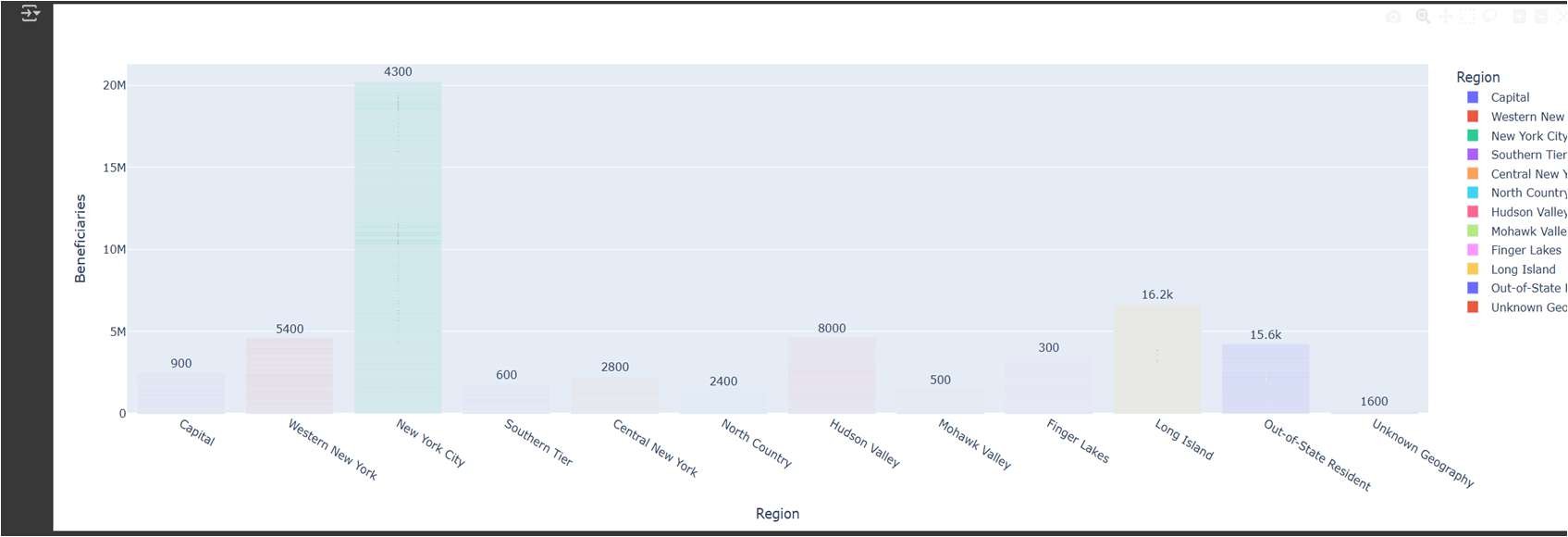
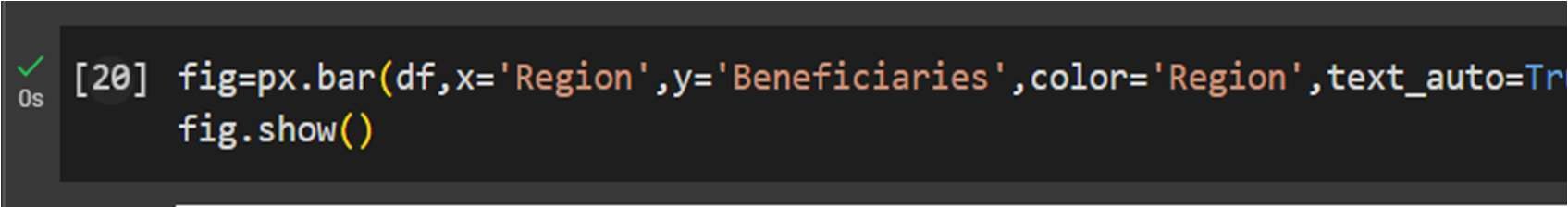
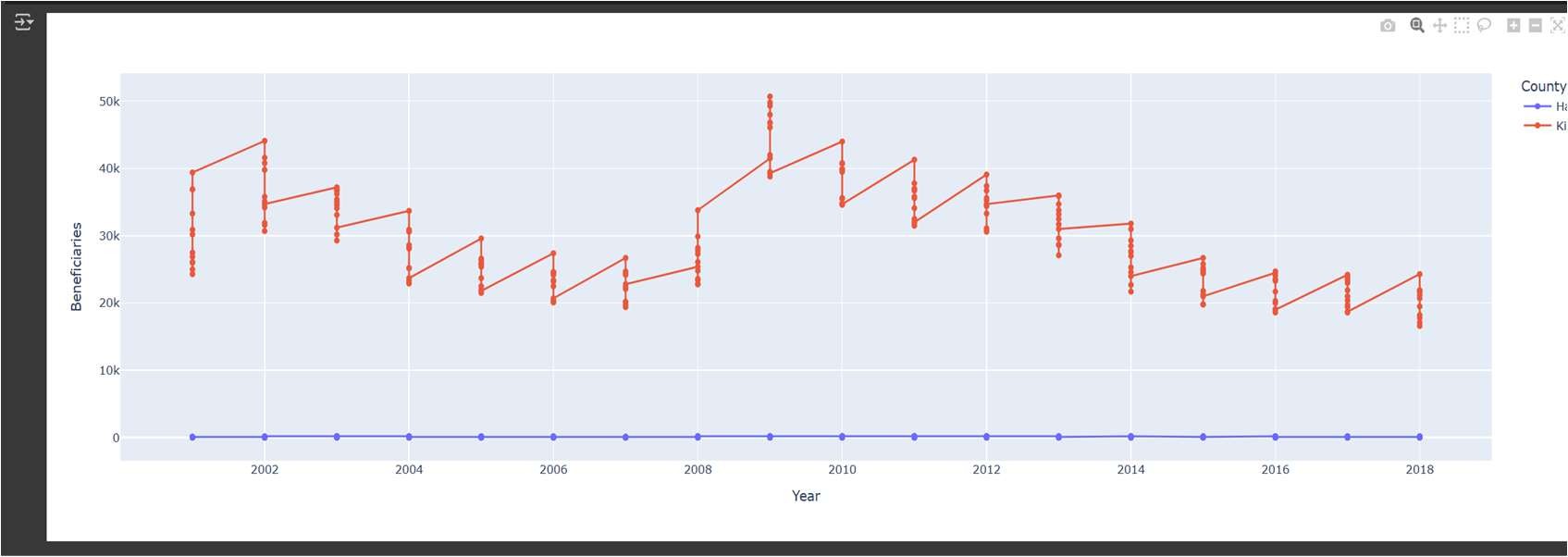
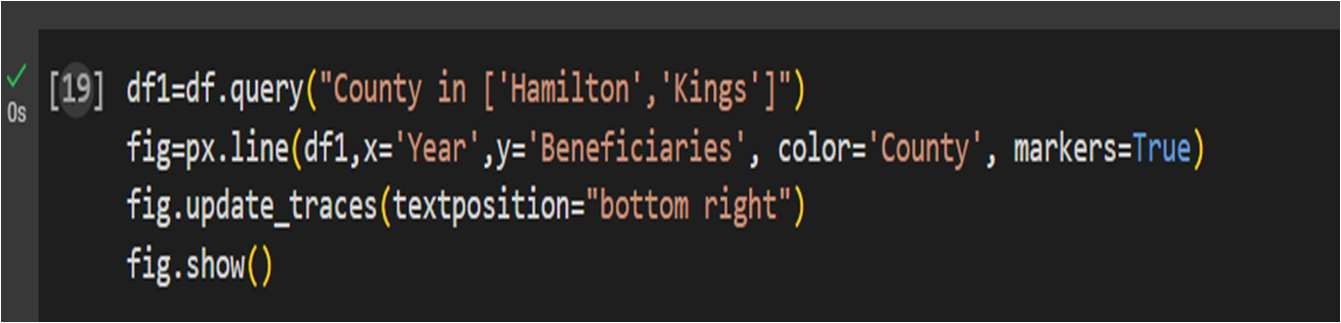
**Data Exploration and Preprocessing Report**

In this project, I explored and prepared data to understand how unemployment beneficiaries and benefit amounts change over time. I used Python to check for missing values and duplicates to keep the data clean. I also combined the Year and Month columns into a single date column and grouped the data by month to make it easier to analyse. These steps helped me create a clear and organised dataset that can be used for accurate time series forecasting and further analysis.

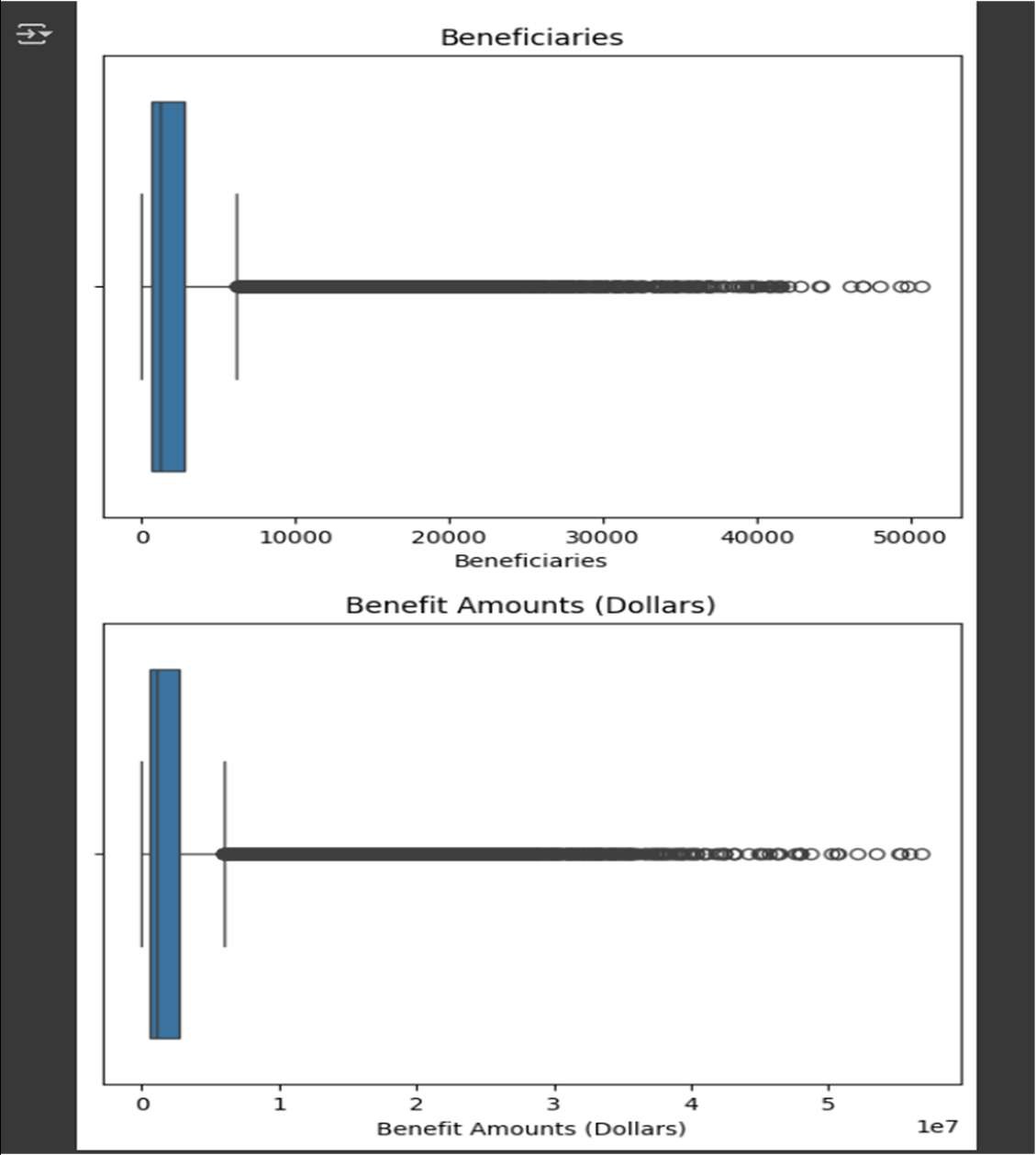
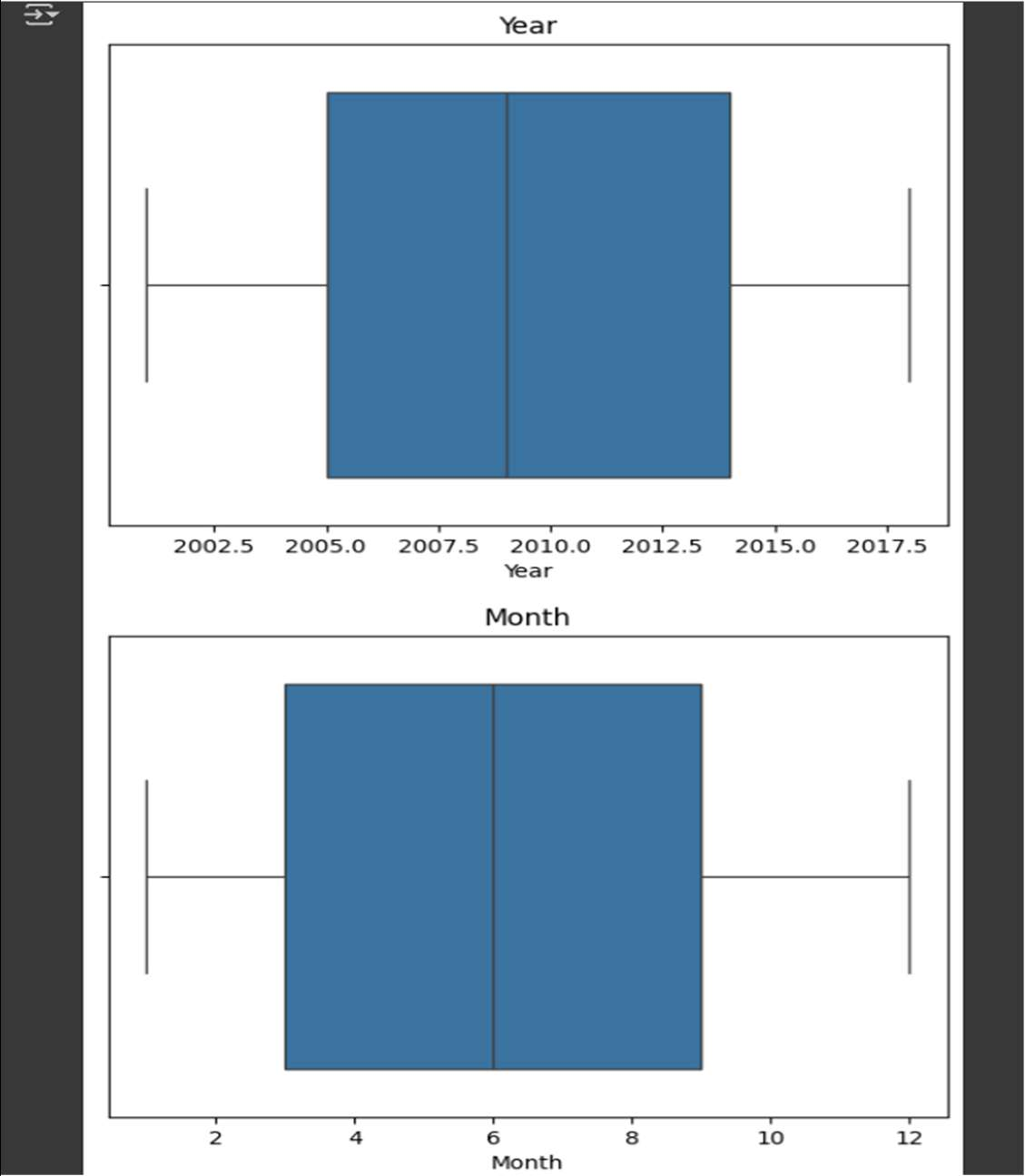
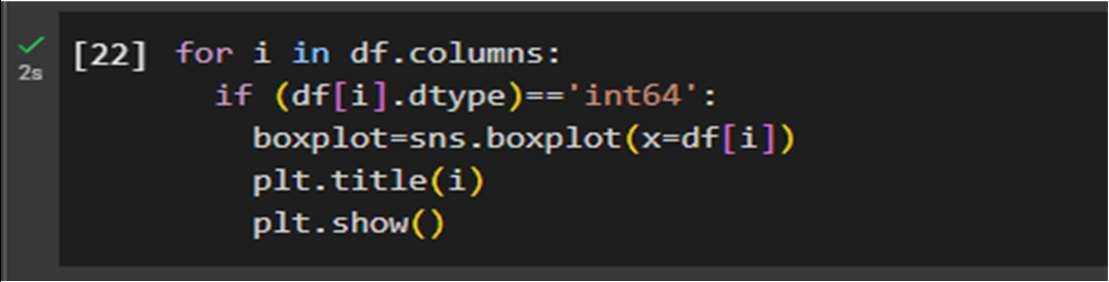
|  |  |
| --- | --- |
| **Section** | **Description** |
| **Data Overview** | Dimension:    Descriptive statistics: |



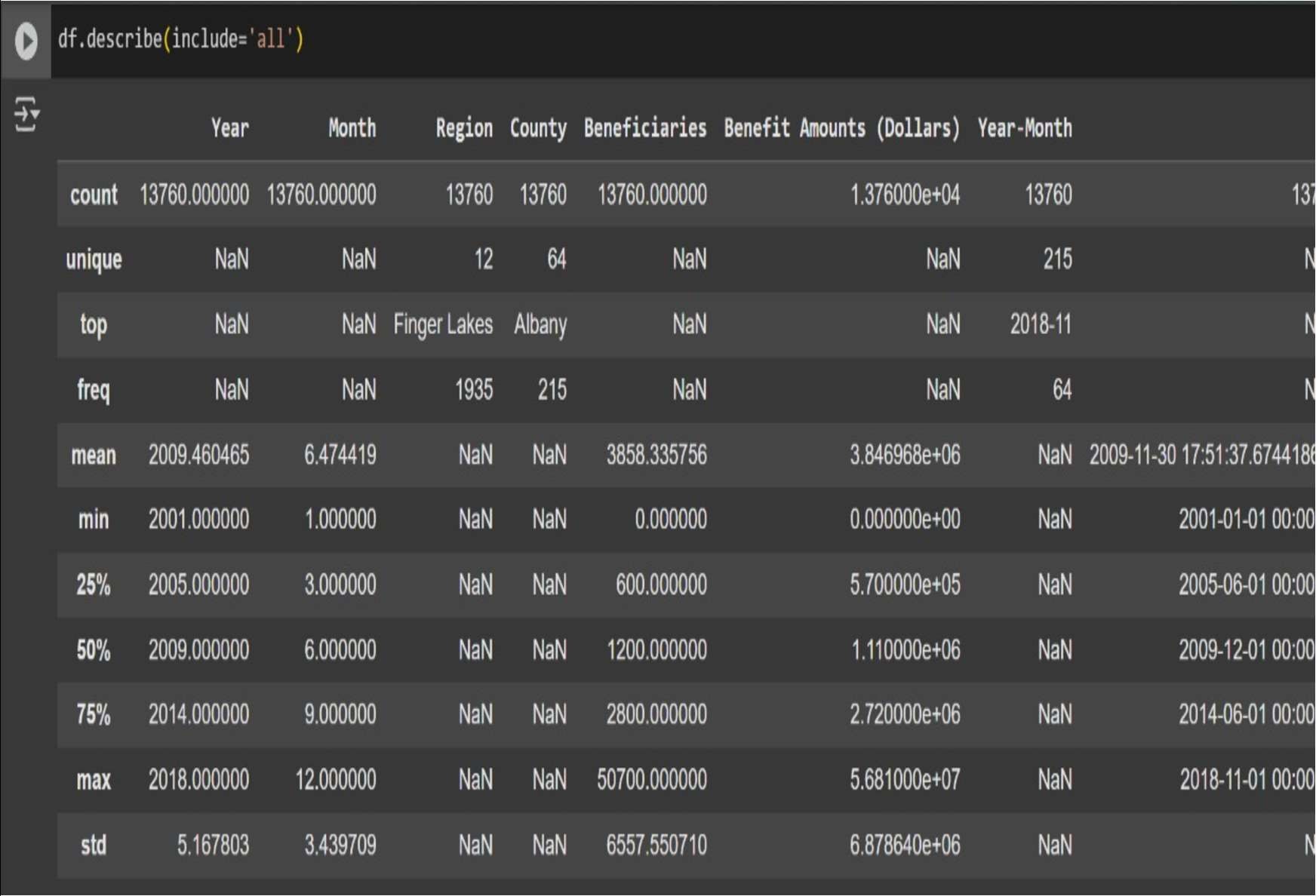
# Univariate Analysis



**Bivariate Analysis**



**Multivariate Analysis**



**Descriptive Analysis**

|  |  |
| --- | --- |
| **Data Preprocessing Code Screenshots** | |
| Loading Data |  |
| Checking for Missing Data |  |
| Data Transformation |  |
| Feature Engineering |  |

|  |  |
| --- | --- |
|  | **Feature Engineering**   * Created a new datetime feature 'ds' by combining Year and Month. * Aggregated Beneficiaries and Benefit Amounts monthly to prepare structured time series features for model training. |