

Service-Requests-Analysis

Problem Statement:

You've been asked to analyze data on service request (311) calls from New York City. You've also been asked to utilize data wrangling techniques to understand the patterns in the data and visualize the major types of complaints.

Problem Objective :

Perform a service request data analysis of New York City 311 calls. You will focus on the data wrangling techniques to understand the pattern in the data and also visualize the major complaint types. Domain: Customer Service

Analysis Tasks to be performed:

(Perform a service request data analysis of New York City 311 calls)

1. Understand the dataset:

- 1.1 Import the dataset
- 1.2 Visualize the dataset
- 1.3 Print the columns of the DataFrame
- 1.4 Identify the shape of the dataset
- 1.5 Identify the variables with null values

2. Perform basic data exploratory analysis:

- 2.1 Draw a frequency plot to show the number of null values in each column of the DataFrame
- 2.2 Missing value treatment
 - 2.2.1 Remove the records whose Closed Date values are null
- 2.3 Analyze the date column, and remove entries that have an incorrect timeline
 - 2.3.1 Calculate the time elapsed in closed and creation date
 - 2.3.2 Convert the calculated date to seconds to get a better representation
 - 2.3.3 View the descriptive statistics for the newly created column
 - 2.3.4 Check the number of null values in the Complaint_Type and City columns
 - 2.3.5 Impute the NA value with Unknown City
 - 2.3.6 Draw a frequency plot for the complaints in each city
 - 2.3.7 Create a scatter and hexbin plot of the concentration of complaints across Brooklyn.

3. Find major types of complaints:

- 3.1 Plot a bar graph to show the types of complaints
- 3.2 Check the frequency of various types of complaints for New York City
- 3.3 Find the top 10 complaint types
- 3.4 Display the various types of complaints in each city
- 3.5 Create a DataFrame, df_new, which contains cities as columns and complaint types in rows

4. Visualize the major types of complaints in each city

4.1 Draw another chart that shows the types of complaints in each city in a single chart, where different colors show the different types of complaints.

4.2 Sort the complaint types based on the average Request_Closing_Time grouping them for different locations

5. See whether the average response time across different complaint types is similar (overall)

5.1 Visualize the average of Request_Closing_Time

6. Identify the significant variables by performing statistical analysis using p-values and chi-square values.

7. Perform a Kruskal-Wallis H test

7.1 Fail to reject H_0 : All sample distributions are equal

7.2 Reject H_0 : One or more sample distributions are not equal