URBAN HARMONY:
VISUALIZING ILLEGAL PARKING COMPLAINTS IN NYC OVER THE YEARS

CA682 Data Visualization Report

Submitted By:

23264020 - Ayesha Imran (Ayesha.imran2@mail.dcu.ie)
22262455 - Abdullah Safder (abdullah.safder2@mail.dcu.ie)

Submitted To:

Dr Suzanne Little
Associate Professor

December 1, 2023

DECLARATION

1. We understand what palagiarism is and we are aware of the University's policy in this regard.
2. We declare that the work hereby submitted is our own original work.Where other people's work has been used (either from a printed source, Internet or any other source),this has been properly acknowledged and referenced in accordance with the SHSPH's requirements.
3. We have not used work previously produced by another student or any other person to hand in as our own.
4. We have not allowed, and will not allow,anyone to copy our work with the intension of passing it off as his or her own work.

Signature – Ayesha Imran
Abdullah Safder

Date- December'1 2023

Visualisation Video File Link

ABSTRACT

"Urban Harmony: Visualizing Illegal Parking Complaints in NYC 2010 till 2023" tackles the persistent challenge of illegal parking in New York City through advanced data analysis. The project employs innovative preprocessing techniques and exploratory analyses to uncover patterns and insights within illegal parking complaints. Utilizing Bar Charts, Line Graphs, and Heat Maps, the project not only discerns temporal and categorical dynamics but also delves into the contributing factors.The specific focus on months with elevated complaint volumes, notably May to October, aims to dissect seasonal variations influenced by factors such as increased tourism. This targeted approach seeks to address the core of the problem and identify potential solutions. The project's insights serve as a practical resource for urban planners, policymakers, and residents, offering a detailed understanding of the multifaceted factors influencing illegal parking. By fostering informed interventions, "Urban Harmony" strives to contribute to a more harmonious and well-managed urban landscape in New York City, addressing the intricate challenges associated with illegal parking.

DATA COLLECTION

My data is provided in tabular format by NYC OpenData Social-Services NYC Data .The following are the ways that I believe my data reflects the three components of Big Data. Firstly, It updates regularly: The website offers downloads for their monthly updated backup. Moreover, the backup is stored over an online serviced Database and contains various columns, presenting various information including Created Date, Agency Name, Complaint Types, Descriptor, Location Type and many more. With a monthly update frequency and automated processes in place, the dataset remains a dynamic and reliable source for monitoring and addressing citizen concerns, contributing to the ongoing improvement of city services. Also, the New York Service Record dataset takes over 21GB of storage of data spanning from 2010 till Present. From the raw data we have extracted, the Illegal Parking Complaints Record for these years which involved various columns but our main focus was over the Complaint Type their Descriptors and the Complaint Dates. After filtering the data it included a total of 348555 rows.

DATA EXPLORATION,PROCESSING,CLEANING AND/OR INTEGRATION

Data Cleaning and Processing

What did you need to do to prepare the dataset(s) to create your graph/chart?

Acquired the NYC Service Requests dataset which includes information on various types of service requests in the city.

Complaint Types

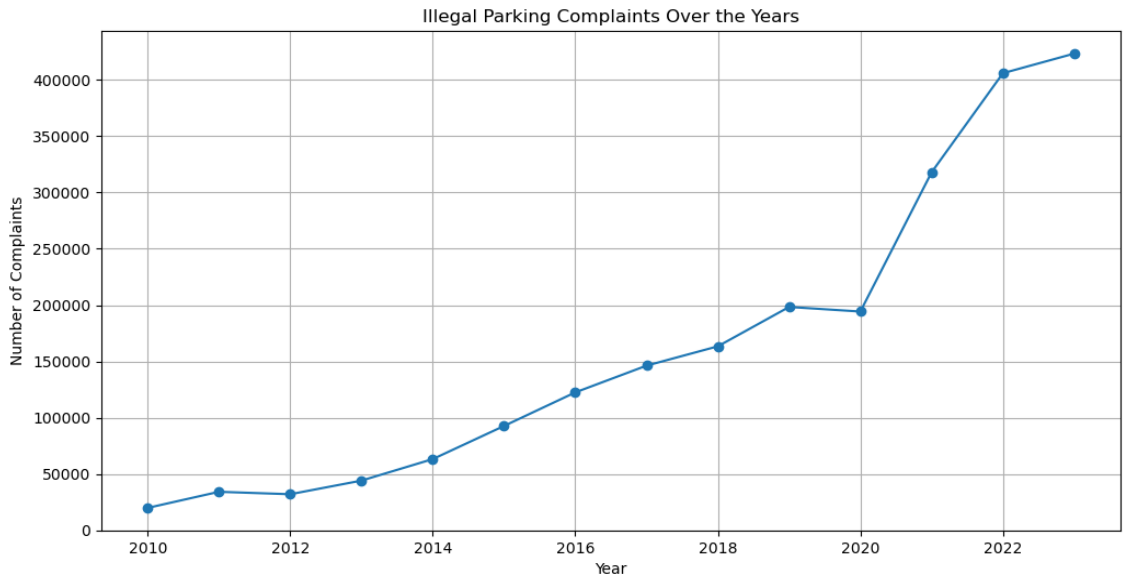
I used the MySQL database loaded my data there and applied different queries over the Complaint Type column to filter data to get me just one particular type of complaints data that was Illegal Parking.

How did you choose the attributes and data subset to visualise?

Upon loading the filtered data into Python, a meticulous data preprocessing phase was executed to handle missing values and rectify inconsistencies, particularly focusing on null values and duplicate entries. The "Date Created" column was then filtered to extract the year of each complaint, facilitating a temporal analysis. To ensure uniformity in data types, the complaint type data underwent comprehensive conversion. In further refining the dataset, a close examination of the "Descriptor" column was conducted. This investigation aimed to streamline and categorize descriptors, enhancing manageability and creating a more meaningful set of categories for similar descriptors.

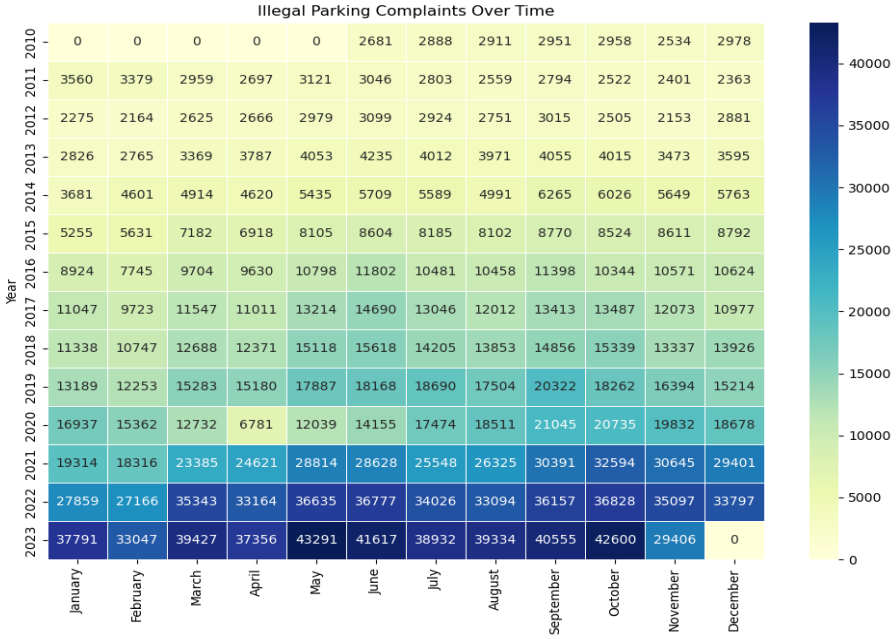
Data Exploration

Subsequently, an Exploratory Data Analysis (EDA) was initiated to comprehend the distribution of data. This analysis provided valuable insights into the patterns and characteristics of illegal parking complaints, offering a nuanced understanding of their fluctuations over the years. By identifying key patterns, the EDA phase laid the groundwork for more in depth Analysis.



In our initial data exploration, we employed a line graph to analyze the trend of illegal parking complaints in New York City from 2010 to the present. The overarching pattern revealed a consistent increase in the number of complaints over the years, signifying a growing trend in reported parking issues. Notably, there was a deviation in this trend from 2019 to 2020, where the number of parking complaints remained relatively stable.

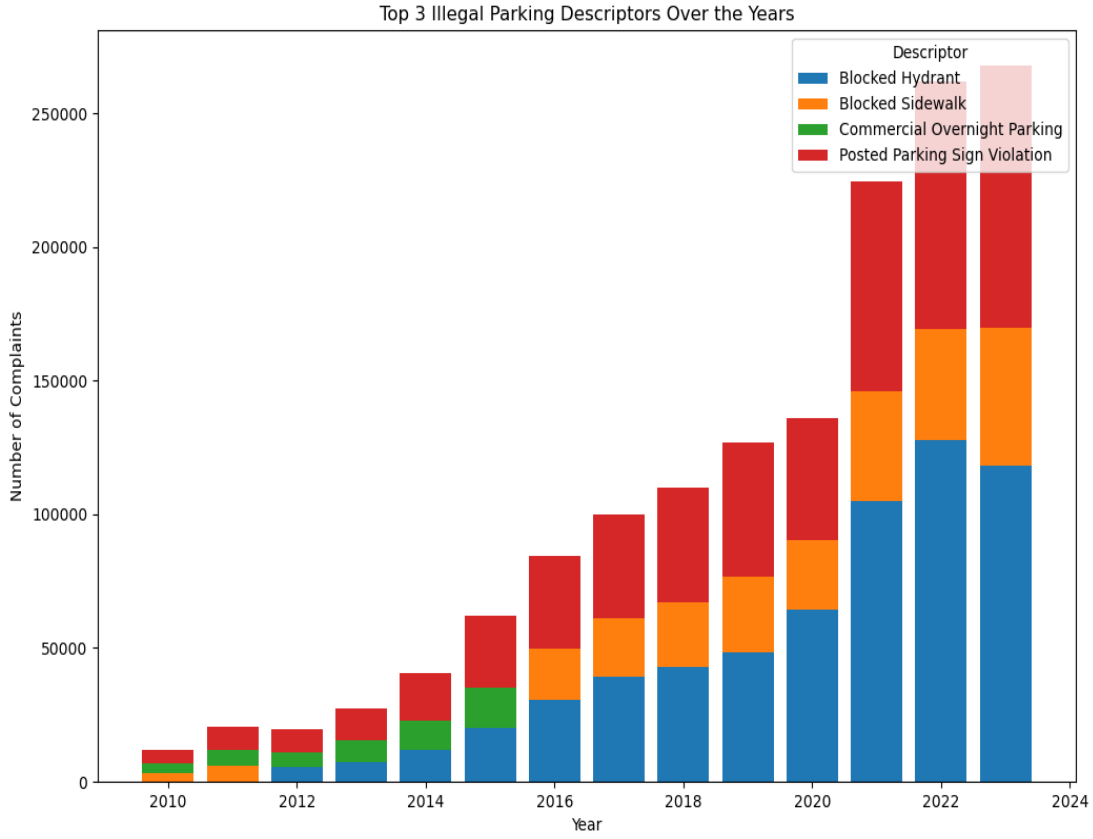
Interpreting these findings, the rising trend observed from 2010 suggests a correlation with the expanding number of vehicles in the city. This surge in the volume of cars likely contributed to the notable increase in parking complaints throughout the specified period. However, the stabilization observed in 2019-2020 prompts further investigation into potential factors influencing this temporary plateau in reported illegal parking incidents.



To better understand the trends in parking complaints over the years, a heat map was created to highlight the months with the highest complaint volumes annually. Analysis revealed that, on average, the months from May to October consistently exhibited the highest number of complaints during these years. This pattern aligns with the summer months, characterized by increased tourist activity, contributing to the heightened busyness of New York City. The influx of tourists during this period is identified as a significant factor influencing the surge in parking complaints. The comprehensive heat map provides valuable insights into the seasonality of parking issues, shedding light on the impact of tourism-related factors on the dynamics of complaints throughout the year

VISUALISATION

The chosen chart type, a descriptive stacked bar chart, provides a comprehensive visual representation of the top three illegal parking descriptors from 2010 to the present. This visualization effectively compares the frequency of complaints for each year, offering valuable insights into the dynamic nature of illegal parking issues over time.



Choice of chart types

We chose a stacked bar chart to visualize parking complaints and their top three descriptors each year. This simple yet impactful choice accommodates both numeric (complaint counts) and categorical (top descriptors) variables. It offers a clear view of trends, allowing easy distinctions between different descriptors and their respective counts, providing a comprehensive overview of illegal parking issues annually.

Design choices

I started from this tutorial: [Stacked Bar Charts: pandas] and made the following changes:

- **Color Palette:** We have used a set of colors to represent the top three descriptors for parking complaints from 2010 to now. This choice is intentional because the specific descriptors change each year. .Assigning unique colors to each descriptor allow viewers to quickly identify shifts in the most prevalent issues over the years.
- **Background:** To ensure that the descriptive colors pop out for viewers, we chose a white background. This design decision maximizes contrast and readability, emphasizing the key information presented in the stacked bar chart.
- **Bar Orientation:** We selected a vertical orientation for the bars to allow a comparison between different descriptors. This layout accommodates the numeric variable (complaint counts) on the X - axis, providing a clear view of the distribution of counts for each descriptor.

- **Simplicity:** The design prioritizes simplicity to avoid visual clutter. Each bar communicates a wealth of information, capturing both the count of complaints and the top descriptors, without overwhelming the viewer.

Tools and libraries used

MySQL was used to make queries to extract relevant data on the data.Python was used to do all the cleaning and processing of the data.

Python,*matplotlib,pandas,numpy was used to create the visualization and create unique effects.*

CONCLUSION

In grappling with the dataset for "Urban Harmony: Visualizing Illegal Parking Complaints in NYC 2010 till 2023," the primary challenge laid in cleaning and processing the data due to numerous unverified and missing entries. While the resulting charts were really enagaging, the majority of the project time was devoted to animating the visuals for the visualisation,leaving other aspects rushed.For the main bar chart, the aim was to enhance clarity by using a stacked bar chart to visualize and our intention was to introduce additional colors to distinguish the top three descriptors for illegal parking for each year. This approach provided a straightforward yet insightful representation of the data, highlighting key trends and patterns over the specified period.One aspect that could be improved is the addition of an interactive filter for users to explore the types of illegal parking complaints, such as those related to short-term parking, blocked driveways, or other specific categories. This would offer a more tailored exploration, providing a deeper understanding of the nuanced dynamics of illegal parking incidents.Overall, this project can serve as a valuable tool for both government agencies and the public to collaboratively address illegal parking challenges in New York City.

REFERENCES

DataSet:311 Service Requests from 2010 to Present

Stacked Bar Charts with Python’s Matplotlib

Easy way to create a Bar Chart