**Final Report**

**Safe LA: Enhancing Traffic Safety Through Data-Driven Insights**

**Team 8**

Alexandra Gladkova

Christina Saju

Hao Lun Rong

Kushwanth Sai Kolli

Ladan Asempour

**Capstone Project**

**BIA-5450-0LA**

**Samer Al-Obaidi**

**Executive Summary**

This project aimed to utilize traffic collision data in Los Angeles to identify areas of high risk and factors that contribute to traffic accidents. The end goal was to create a data-driven solution to enhance road safety, reduce accidents, and improve traffic management. The data analysis focused on the years 2023–2024, utilizing publicly available data on traffic collisions to identify trends by location, time, collision types, and demographics.

The Los Angeles City was faced with severe road safety issues, with frequent crashes resulting in injuries, fatalities, and economic inefficiencies. This project addressed these issues by examining key analytical questions that focused on high-density crash locations, most accident-prone hours, and the most affected engaged demographics. The solutions presented were designed to offer actionable information to city planners, law enforcement, and policymakers to allow them to adopt more effective and targeted interventions.

The project used Power BI visualizations and dashboards to display key trends, including time-based trends, geographic clusters, and demographic data. Data processing operations verified that the dataset was clean and well-structured for analysis. Furthermore, the project considered the potential to utilize a robust IT infrastructure to integrate the traffic collision data into an integrated analytics framework.

Outcome testing validated that the dashboards effectively presented the data and responded to the critical analytical questions. The findings of the project will be used to enable informed decision-making to improve traffic safety and minimize congestion in Los Angeles. The project suggested additional optimizations in data collection and integration to enhance real-time analysis and extend the scope of future insights.

**Introduction**

This project will provide a comprehensive analysis of traffic collision data in Los Angeles from 2023–2024. Its primary objective will be to leverage data-driven insights to identify high-risk areas and factors contributing to collisions, ultimately aiming to reduce traffic incidents and fatalities, and create safer roads for all. The project will be structured into several key sections.

“Business Problem Overview” will clearly outline the challenges faced by the City of Los Angeles in ensuring road safety. It will examine the impact of frequent traffic collisions, highlighting associated economic costs and mobility inefficiencies. To address the business problem, the “Analytical Questions” will define the key questions guiding the analysis, such as identifying high-density collision locations, peak collision times, common types of collisions, and the demographics most frequently involved. Then, the project's boundaries will be defined on the “Scope Statement” section, focusing on data from within Los Angeles city and using publicly available traffic collision datasets from 2023 onwards. Through “Data Sources and Flow”  
the primary data sources will be described, including the Traffic Collision Data from Data.gov and additional documents from the City of Los Angeles' open data portal. It will also explain key data entities and how data flows through the system. Next, the data preparation process will be summarized in the “Data Manipulation Process and Data Output” section, including steps such as data cleaning, MO code processing, refining location data, and handling demographic fields. Also, the new solution design and how it fits into the existing IT architecture will be presented on “Solution Design” section. It will focus on developing visualizations and dashboards that address the core analytics questions, and their integration within the current data analysis framework. After that, on the “Solution Implementation and Testing” section, the implementation of dashboards analyzing chronological, geographic, report-type, and demographic factors will be detailed, along with the testing methods that were used to ensure accuracy and effectiveness. At the end the “Solution Optimization” section will reflect on challenges faced during the project and propose potential enhancements to improve the analysis and future data collection processes.

**Business Problem Overview**

The City of Los Angeles faces a significant challenge in ensuring road safety due to frequent traffic collisions. These incidents pose serious risks to public safety, often resulting in injuries and fatalities. Additionally, traffic congestion caused by these collisions leads to delays, economic losses, and inefficiencies in urban mobility. Law enforcement and city planners struggle with resource allocation, making it difficult to implement proactive safety measures. Without a data-driven approach, identifying high-risk areas and predicting future collision trends remains a challenge. A systematic solution is necessary to analyze collision data, detect patterns that will aid in recommending targeted interventions to reduce accident rates and improve overall traffic safety.

The solution must ensure data accuracy by cleaning collision records to improve the reliability of insights. It should incorporate predictive capabilities using statistical and machine learning models to forecast high-risk collision areas and times. Additionally, the solution must provide actionable insights by analyzing trends that can aid in enhancing traffic safety and resource management.

To measure success, the solution should contribute to improved traffic flow by helping authorities implement measures that decrease congestion caused by accidents. Furthermore, the solution should facilitate stakeholder engagement by equipping law enforcement, policymakers, and urban planners with data-driven insights to enhance decision-making.

Resource allocation must be optimized by prioritizing high-risk locations, ensuring interventions focus on areas with the highest density of collisions. Additionally, the solution should analyze time-based risk factors, identifying peak accident periods to inform preventive actions. By meeting these criteria, the solution will contribute to a safer, more efficient, and well-managed urban traffic system in Los Angeles.

**Analytical Questions**

*Which locations have the highest density of collisions?*

* Types of Analytics: Descriptive
* Decision: By identifying the areas where most collisions happen, we can take action to make those locations safer.

*What time of day and days of the week have the most collisions?*

* Types of Analytics: Descriptive
* Decision: If we know when accidents are more likely to happen, we can be more prepared by placing more officers on the road at those times.

*What types of traffic collisions (e.g., vehicle vs. vehicle, vehicle vs. pedestrian, vehicle vs. property) occur most frequently?*

* Types of Analytics: Descriptive
* Decision: If certain types of collisions happen more often, we can take steps to address them. This can include improving road design, adding better signage, or adjusting traffic laws to make those areas safer.

*Are certain demographics more frequently involved in traffic collisions?*

* Types of Analytics: Descriptive
* Decision: If certain groups (by age, sex, and descent) are involved in more collisions, we could take specific actions:

Age: If younger drivers are more involved, improve driver education.

Sex: If males are more involved, focus on risky behavior awareness.

Descent: If certain ethnic groups are more affected, provide cultural-specific outreach programs or language-supportive safety materials.

**Scope Statement**

The project aims to identify collision-prone areas and times/seasons that can enable enhanced targeted interventions in those regions. It requires traffic collision data from 2023 onwards which we have downloaded from a publicly available website.

This project report will include analysis of collision data to obtain demographics of people having a higher chance of getting into a traffic collision incident, identification of high-risk areas as well as times-of-day and seasons when probability of a collision is very high.

The limitations are a strict focus on traffic incidents within Los Angeles city limits. It excludes the development of new traffic management software and will not involve user research or the collection of data beyond publicly available information. This project will analyze publicly available data on traffic collisions and their patterns to produce reliable and up-to-date results.

This project will give the primary stakeholders the necessary tools to make data-driven recommendations to improve traffic safety and reduce traffic incidents. This project has the potential to greatly increase road safety for all city residents by helping employ demonstrably effective methods, along with substantially improved traffic signal timing in high-risk areas.

**Data Sources, Key Data Entities and Flows**

The primary data source for this project is the “*Traffic Collision Data from 2010 to Present”*, available on Data.gov *(*[*https://catalog.data.gov/dataset/traffic-collision-data-from-2010-to-present*](https://catalog.data.gov/dataset/traffic-collision-data-from-2010-to-present) *).* This dataset contains detailed records of traffic collisions in Los Angeles, including location, time, involved parties, and contributing factors.

To understand the structure and meaning of the dataset’s fields, the official data dictionary and column descriptions provided by the City of Los Angeles were referred to. *(*[*https://data.lacity.org/Public-Safety/Traffic-Collision-Data-from-2010-to-Present/d5tf-ez2w/about\_data*](https://data.lacity.org/Public-Safety/Traffic-Collision-Data-from-2010-to-Present/d5tf-ez2w/about_data)*)*.

Additionally, for incident classification and interpretation of MO (Modus Operandi) codes, the official MO Codes document from the city’s open data portal, accessible in MO\_CODES\_Numerical\_20180627.pdf on the <https://data.lacity.org/> website has been used.

These sources together provide a comprehensive foundation for analyzing traffic collisions, identifying trends, and deriving insights to enhance road safety in Los Angeles.

Based on our business problem, we identified the key data entities:

Collision – represents individual collisions.

Collision MO – represents all the crime factors involved in each of the collisions

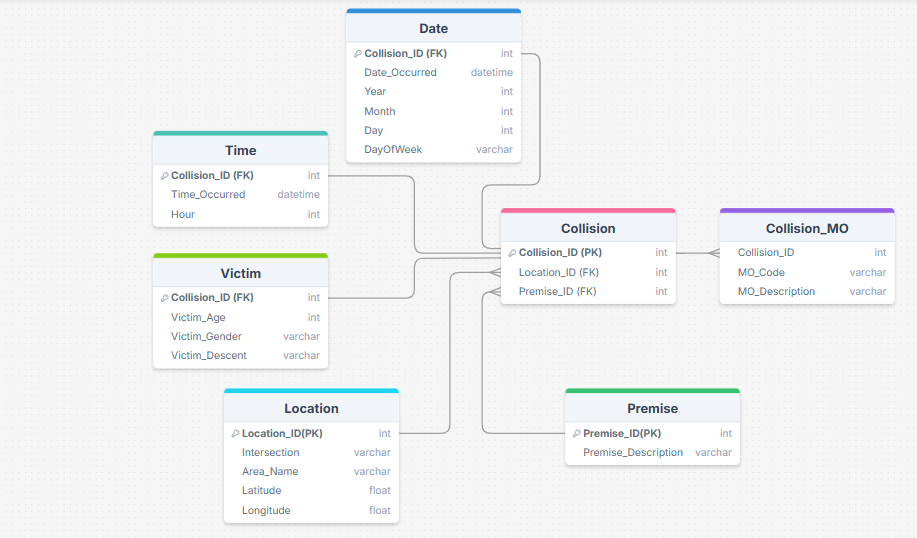
Premise – represents the area description of where the collisions occurred.

Location – represents the location details of the collisions.

Victim – represents the demographics of the primary victims involved in the collision.

Time – represents the time details of the collisions.

The attributes of each of the entities and the flow of data between them is given in the diagram below.



**Brief Overview of Data Manipulation Process and Data Output**

Data cleaning was performed on the dataset to ensure accuracy and relevance in analysis. First, the dataset was filtered to include only accidents from the years 2023 and 2024.

Several columns were dropped, including DR Number, Date Reported, Area ID, Reporting District, Crime Code, and Premise Code, as they did not provide direct value for the analysis and hence were deemed unnecessary or redundant.

Before:

A screen shot of a computer

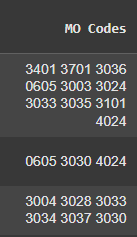
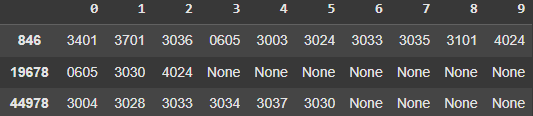
AI-generated content may be incorrect.

After:

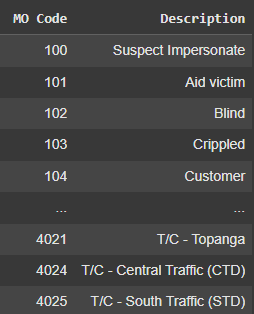
A screen shot of a computer program

AI-generated content may be incorrect.

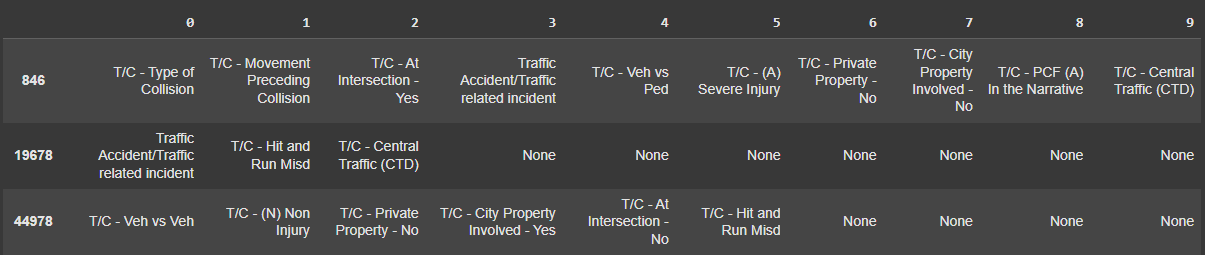
The MO\_Codes column, originally stored as a single string with multiple values seperated by space, was split into separate columns (MO\_Code01 to MO\_Code10) to allow for better individual analysis of contributing factors.

Before:  After: 

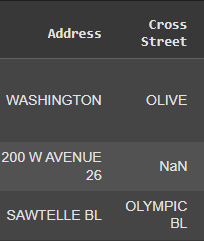
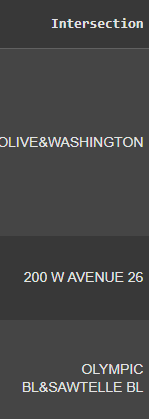
The MO Codes were then cross referenced with another CSV file (MO\_Code\_Ref), containing the description of all the MO Codes used by LAPD. Using SQL (using a specialized package to run SQL queries in Python), the MO Codes from the collision dataset were matched to their corresponding definitions in the MO\_Code\_Ref CSV.



Once the definitions of the MO Codes were cross-referenced, all the codes were remapped into their respective descriptions. This provided more meaning to the common MO Codes present in traffic collision reports. The descriptions were also cleaned to ensure no commas present as well as to shorten some descriptions that were unnecessarily long.



For the Location Data, an intersection column was created by combining the main address of the collision with the nearest intersection. Both columns were joined, separated by a comma to create the intersection column with entries containing the first street and second street names.

Before:  After: 

To ensure consistency of data and not confuse entries due to the order of the streets (example: 1st and 3rd street is the same as 3rd and 1st street), the street combinations were sorted in alphanumeric order before joining for each entry in the dataset. Any extra spaces were also deleted.

The Date\_Occurred column was converted into a standard DateTime format (%m/%d/%Y) to maintain consistency and facilitate easy time-based queries.

Similarly, the Time\_Occurred column was reformatted into the HH: MM 24-hour format to ensure consistency in time-related analysis.

The demographic fields required careful handling of missing values. The Victim\_Sex and Victim\_Descent columns were cleaned by filling missing values with X (unknown) to ensure uniformity.

The Victim\_Age column was binned into predefined age categories to standardize age-related analysis and improve interpretability.



The final dataset is now structured and cleaned, making it ready for in-depth analysis. The MO\_Codes have been mapped to their descriptions, allowing for better understanding of contributing factors in each accident. The data is formatted correctly, ensuring smooth integration with visualization tools for reporting and analysis.

Key insights can be drawn from this structured data, including trends related to victim demographics, high-risk geographical zones, and prevalent causes of collisions. The dataset is optimized for exploratory data analysis (EDA), predictive modeling, and decision-making processes that aim to enhance traffic safety.

**New Solution Design and Fit into the Existing IT Architecture**

**Existing IT Architecture:**

The IT architecture for analyzing traffic collisions in Los Angeles includes key components such as data sources, software tools, and integration mechanisms to compile the components into a structured dataset.

**Major Components:**

* **Data Sources:** The Traffic Collision Data from 2010 to Present from Data.gov (<https://catalog.data.gov/dataset/traffic-collision-data-from-2010-to-present>) , contains details regarding location, time, involved parties, and contributing factors. The official data dictionary (<https://data.lacity.org/Public-Safety/Traffic-Collision-Data-from-2010-to-Present/d5tf-ez2w/about_data>) and MO Codes document (<https://data.lacity.org/>) are available to interpret and classify data.
* **Software Tools:** Python has been utilized for data exploration, processing, cleaning, and Power BI was used for statistical and visual analysis to identify trends in collision data.

**Integration and Ecosystem:**

The components all work in together to analyze collision data. The data is downloaded manually from Data.gov, cleaned, and processed using Python. Insights are derived from the data to inform traffic safety decisions and visualized using Power BI. APIs are not utilized in the data ingestion and integration process but integrates data through manual downloads and processing.

**Connections:**

* **Data Flow:** Data is downloaded from Data.gov and processed using Python.
* **Security and Privacy:** Data cleaning techniques and privacy guidelines ensure the protection of sensitive information.

This IT setup enables data-driven decision-making for City of Los Angeles Department of Transportation (LADOT) and Law Enforcement (LAPD) and Traffic Monitoring Agencies, enhancing road safety and urban mobility in Los Angeles.

**Solution’s design**

The solution design for this project consists of developing interactive dashboards to answer key components of our analysis. This includes the key analytical questions that were proposed earlier which include:

* *Which locations have the highest density of collisions?*
* *What periods of the day and days of the week have the most collisions?*
* *Are certain demographics more frequently involved in traffic collisions?*
* *What types of traffic collisions (e.g., vehicle vs. vehicle, vehicle vs. pedestrian, vehicle vs. property) occur most frequently?*

By answering these questions, patterns, trends, and other variables that contribute to the reported accidents are hoped to be uncovered, which will allow for stakeholders to have better access and understanding to make more informed decisions.

***Solution 1. Identifying areas with high collision density***

To identify locations in Los Angeles with the greatest density of traffic accidents, we plan to create a geography-driven dashboard based on visualization instruments such as ArcGIS maps and bar charts. The dashboard will graphically represent clusters of accidents in different neighborhoods and districts and assist in determining locations where specific interventions or infrastructural development must be given top priority. Cards will highlight peak spots with the most incidents, and supporting statistics will distinguish between street-level and non-street locations, i.e., parking lots or private property. Filters to dynamically switch the view by area or year will also be included in the solution. The analysis will give a geographic foundation for the detection of traffic hotspots in the city.

***Solution 2. What time of day and days of the week have the most collisions?***

This dashboard will explore temporal collision patterns to uncover the time intervals with the highest collision frequencies. Using bar charts, line graphs, and donut charts, we aim to explore trends in hours of day, days of week, and months of year. Visualizations to be done are average daily collisions, peak collision hour, weekday vs. weekend analysis, and general breakdown by time-of-day. This solution will help to determine if collisions peak at certain rush hours, weekends, or months. From this analysis, we aim to determine temporal risk patterns and provide a framework for possible scheduling realignments for enforcement or awareness campaigns.

***Solution 3. Are certain demographics more frequently involved in traffic collisions?***

This solution is aimed at the demographic structure of victims involved in the collisions. The dashboard will include visualizations dividing information by age, gender, and ethnicity. Tree maps, matrix tables, and bar charts are the intended visual pieces to find the most affected groups and to compare by category. For example, we will look at which demographic age category or ethnic category has the most collisions, and the gender distribution by demographics. Slicers will be used to filter out individual groups interactively. From this information, we would like to develop a better understanding of vulnerable populations and facilitate demographic-specific interventions or educational campaigns.

***Solution 4 - What types of traffic collisions (e.g., vehicle vs. vehicle, vehicle vs. pedestrian, vehicle vs. property) occur most frequently?***

To understand the traffic collision composition in LA, this dashboard will sort out the types of collisions reported, such as vehicle vs. vehicle, hit and run, and vehicle vs. property. We plan to have card visuals to display top three collision types, along with bar charts and donut charts for distributional data. In addition, we will graphically display the most frequent MO codes police officers use when reporting the circumstances of crashes. This will provide insight into how crashes really occur and what types of crashes are most likely to be crash-for-injury-prone. This proposed analysis will complement enhanced risk factor and crash scene understanding that needs to be investigated.

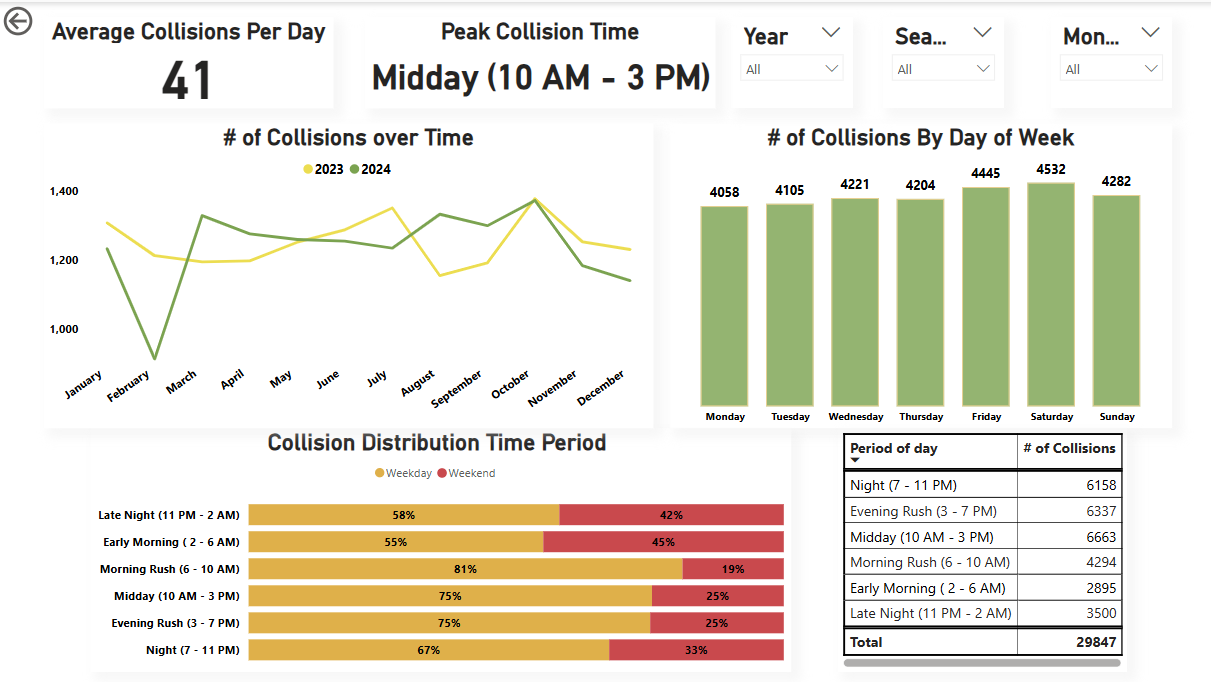
The proposed solutions seamlessly compliment the IT architecture which is currently being used. Taking advantage of the fully explored and engineered dataset and giving insights and meaning through not only being able to see the numbers through tables, but various types of visualizations to understand not only, location, time, identification, as well as risk factors police have documented associated with these collisions.

By transforming the processed data which has been a part of our framework into meaningful insights, the solution provides a comprehensive way for stakeholders to assess road safety concerns in LA.

**Solution Implementation and Outcome Testing**

The Solution Design for this project consists of multiple interactive dashboards composed of various visualizations including Fact Cards, Bar Charts, Stacked Column, Line Graphs, Tree Maps among others. In this section, each dashboard, which answers one of our analytical questions, along with the components within each, will be reviewed.

*Analysis of Chronologic Factors*



The solution design of this dashboard consists of multiple cards and visualizations which assess various temporal factors. These include time of day, days of the week, and month, as well as the distribution of collisions comparison per years.

The dashboard highlights key insights such as the average collisions per day as well as the distribution of collisions by time period. Several trends such as monthly distribution of collisions, comparisons between the weekdays and weekends were also visualized. The dashboard is complimented with various slicers allowing for selection of different months, seasons, and years. The visualizations created are listed as follows:

Average Collisions per Day Card: This visualization emphasizes the amount of car collisions recorded by the LAPD daily throughout the past 2 years. It was created by utilizing the card visualization and inputting a DAX measure in the field, which takes the count of collisions divided by the number of days listed over the 2-year span.

Peak Collision Time Card: Insight into the time period which records the most collisions is given by this card. This was created by using the card visual and inputting a DAX measure that returns the time period with the highest record of collisions.

Collisions Over Time Chart: The line chart shows the count of collisions recorded each month, with separate lines representing each year, respectively, to highlight the contrast in recorded collisions. This was created by utilizing the line chart, inputting the Month and Collisions into the X and Y axes, respectively, and inputting the Year into the legend.

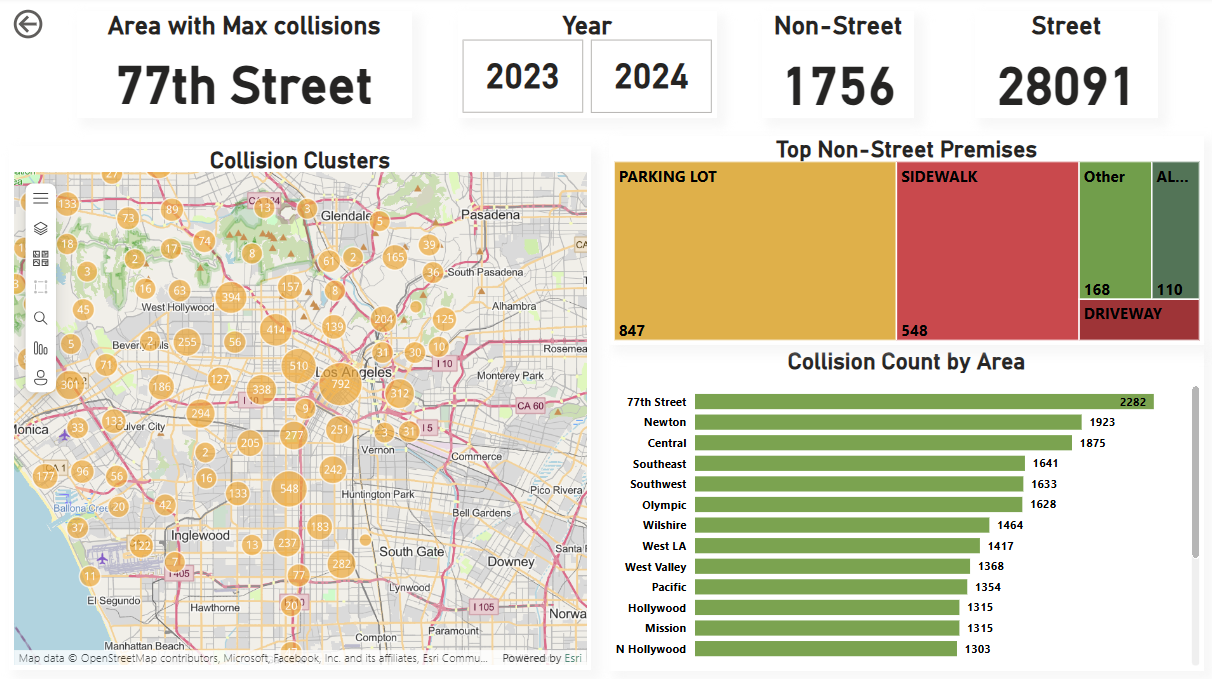
# of Collisions by Day of Week: This visualization shows the total records of collisions grouped by the day of the week they are recorded, providing insight into which days tend to have more collision occurrences throughout the week. This was created by utilizing the column chart, and inputting the Day of the Week and collision counts as the X and Y axes respectively.

Collision Distribution by Time Period: This visualization compares the proportion of collisions during weekends and weekdays, falling within different time periods. It highlights day-specific trends, such as weekend nights being most prone to vehicle collisions, while weekdays show higher collision frequencies during the daytime and evening rush hour periods. The visualization was created by using a stacked bar chart, with the time period and collision count as the X and Y axes, respectively, and the weekend or weekday classification used as the legend.

Period of Day Distribution: This visualization shows the number of collisions categorized by the time period in which the collision was reported. It was created using a matrix chart, with the Period of the Day as the row headers and collision counts as values.

Year, Season and month Slicers: These slicers are meant to create interactive filters by utilizing the slicer visual and inputting the respective category into the Field section.

*Analysis of Geographic Factors*



The solution design of this dashboard consists of multiple cards and visualizations that assess various geographical and location-based factors. These include information on neighborhoods, premises, or surrounding descriptions, as well as maps containing clusters reflecting the frequency of collisions. The dashboard highlights key areas where collisions occur more frequently and provides descriptions of the environment or premises where these collisions take place.

The visualizations created are listed as follows:

Area with most Collisions Card: This visual gives information on the area with the most collisions recorded. It was created by using the card visual and inputting a DAX measure that returns the area name with the highest collision count value.

Non-Street and Street Premise Card: This visual gives information on the count of collisions which happened on the road versus the count of collisions on other premises such as parking lots or sidewalks. It was created by using the card visual and inputting the Collision Count grouped by Location as the Data, and a Dax code which categorizes all the street premise and filters the rest as “other” into the categories input.

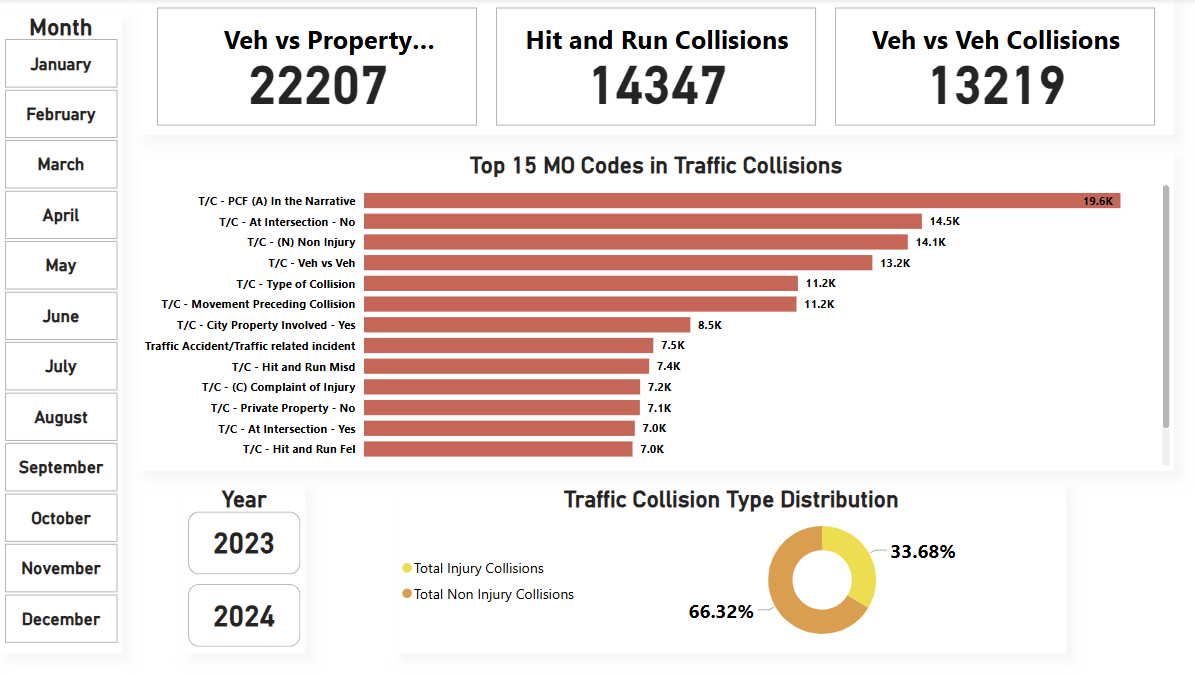
Top Non-Street Collisions Premises: This visual shows the premises that most commonly report collisions, excluding streets, as streets have the highest record of collisions on premises by a large margin. This provides more value into other environments or scenarios where people may be at risk of collisions. It was created by inputting a DAX code that reports the Top 4 non-street premises, with the rest grouped as "Other" into the Category box. The Value box was filled with the Collision Count Per Location.

Collision Clusters Map: This visual shows the collision density for the recorded collisions across LA. The size of the cluster indicates a higher density of collisions within the area. The visual was created by using the ArcGIS visual and inputting the Latitude and Longitude data into it respectively.

Collision Count by Area: This visual shows the number of recorded collisions based on the area or district name. It was created by utilizing a bar chart, with the Y axis representing the Area Name and the X axis representing the Collision Count per Location.

Year Slicers: These slicers are meant to create interactive filters by utilizing the button slicer visual and inputting the year into the Field box.

*Analysis of Crime Report Factors*



The solution design of this dashboard consists of multiple cards and visualizations that assess the factors and descriptions that were entailed in the collision. This gives a closer look at the MO codes that were listed in the collision report. As well as assessing the traffic types during the time of collision. The visualizations created are listed as follows:

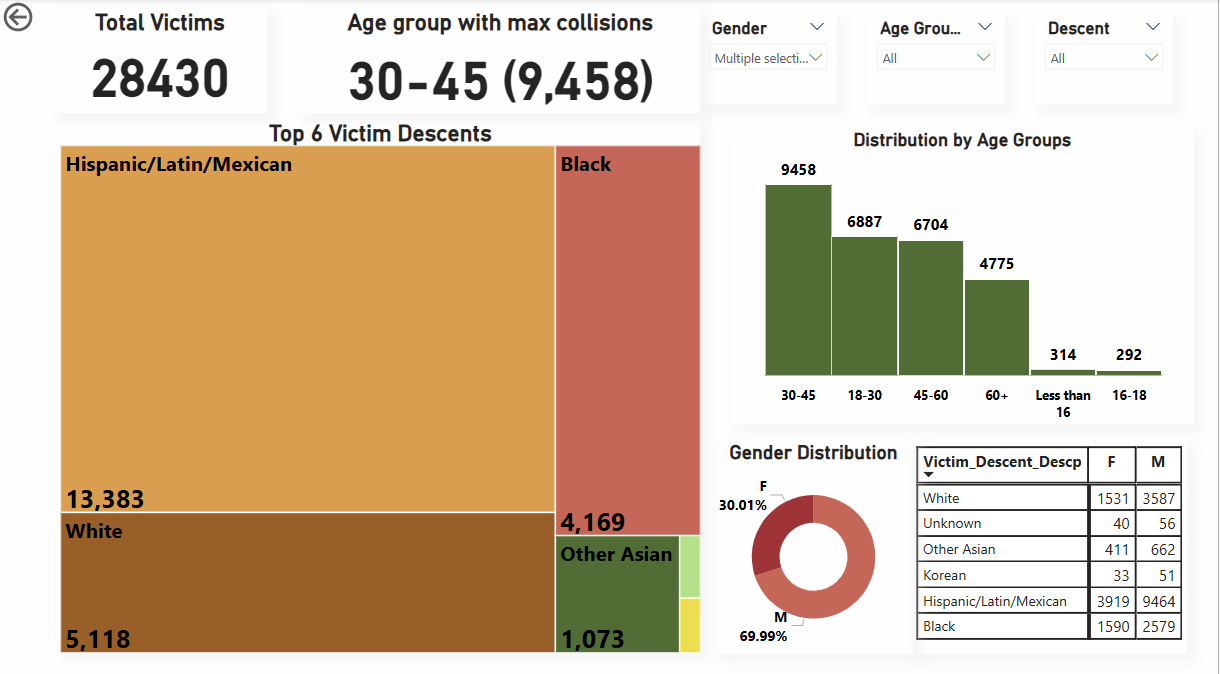
Collision Type Cards: The top 3 visualizations, list the count of each collision type. Which is done by listing the Total Vehicle vs. Property, Total Hit and Run, and Total Vehicle vs Vehicle. This was done by using the card visual and putting the DAX measure to return the filtered data for each.

Top 15 MO codes in Traffic Collisions: This shows the 15 most common MO codes used by the police when recording collision incidents. This was done by using a bar chart and inputting the Description of the MO code, and count of the MO codes for the Y and X axes, respectively.

Traffic Collision Type Distribution: This visualization shows the proportion of collisions where people faced injuries versus collisions with no reported injuries. It was created by utilizing the donut chart, and inputting DAX codes into the Values box, counting the collisions where injury was reported as well as another DAX code where injury was not reported.

Button Slicers: Numerous slicers were created for the months and year to enable interactive filters, using the button slicer visual and inputting the respective category into the Field section.

*Analysis of Demographic Factors*



The solution design of this dashboard consists of multiple cards and visualizations which assess various demographic factors consisting of the Age, Gender, and Ethnicity. The visualizations help identify the most affected age groups, the most affected ethnicities, and compares the gender distribution between them. Slicers are implemented to allow users to filter by specific groups for interactive visualization. The visualizations included are listed below:

Total Victims Card: This visualization displays the information about the total amount of collisions victims within our dataset, and it is done by taking the card visual and inputting the count of the collision records.

Age Group with Max Collisions Card: This visualization displays the age brackets into which the victims' ages have been categorized, and the age bracket with the highest number of recorded collisions is selected. This was done by creating a measure which counts the collisions grouped by the age column and selecting the one the highest value of counts.

Top 6 Victim Descents: This tree map shows the 6 highest ethnicities/descents which are involved in collisions, with others being grouped together as "other." This visualization was created by using the tree map and inputting a DAX code to return the 6 highest descents and returning "other" for anything lower than the top 6 values.

Distribution by Age Groups: This visualization shows the number of recorded collisions corresponding to the age bracket the victims are in. The visualization was created by utilizing the column chart and inputting the age group and collision record columns in the X and Y axes respectively.

Victim Profile Matrix: This visualization shows the 5 most common ethnicities that are involved in vehicle collisions and are separated by gender to show the distribution of each ethnicity. This was done by utilizing the matrix table, and inputting descent into the rows, victim gender into the columns, and the collision records as the values. The filter was done manually by using the filters tab and selecting top 5 for the filter type.

Gender Distribution: The ratio of men to women involved in collisions is shown in this visualization using our dataset. It was created by utilizing the donut chart, using the victim gender as the legend and the count of collision records in percentage as the value.

Gender, Age Group, and Descent Slicers: These slicers are meant to create interactive filters by utilizing the slicer visual and inputting the respective category into the Field section.

**Outcome Testing**

During the implementation process, the dashboards were tested to make sure they displayed the data correctly and met the project's goals. The accuracy of each visualization was first checked. For example, the values for the "Average Collisions per Day" and "Peak Collision Time" cards were manually calculated and compared to confirm that the visualizations displayed the correct results. Different methods for building the visualizations were tested, such as using DAX measures for real-time calculations compared to pre-aggregated data. The effectiveness of different chart types was also checked to see which one presented the data most clearly.

The dashboards were also compared to the original project goals. It was confirmed that the visualizations answered the key questions and that the interactive filters made it easy for users to explore the data. In the end, the dashboards successfully communicated the insights.

**Potential Solution Optimization**

Several strategies were employed throughout the process of refining our dashboards from the previous solution design report, along with additional implementations that can be made. The first area was the Data Model, as we had previously not implemented a star schema for the dataset. Multiple tables were separated and then connected with various primary and foreign keys to make the relationships of the tables. This optimized performance when doing our visualizations as we are creating clear and concise dashboards which contain only information from specific tables for the designated dashboard.

DAX measures were also implemented to increase performance time for the visualizations. Multiple DAX formulas were reused across separate charts in a dashboard, rather than creating the same calculations multiple times throughout the table.

Visualizations were also improved by simplifying or narrowing down the options to improve visual clarity. Many graphs were limited to the top “X” amount of a column or constraint, to improve visual clarity from potentially 40 or 50 classes to about the top 6 or 7. The questions we aimed to answer were also iteratively refined as additional visualizations were created throughout the dashboard. This approach allowed each visualization to address a broader scope of questions without becoming repetitive. Furthermore, these visualizations included interactivity features such as the slicers to allow for simple yet effective analysis for potential stakeholders and other viewers.

Future optimization for the project includes technological improvements, particularly in data collection and the data model. The analysis could be further enhanced by updating the database, including real time integration of the information as well as proposing that the LAPD expands their police reports to include several other factors as previously mentioned such as weather or road conditions during the incident.