Name: Varshitha Yanamala

Campus ID: AZ70378

Project Proposal: Street Safety Analysis Using 311 Requests, Crime Data, and Traffic Crashes

1. **Title:**
   Deep Dive: Comprehensive Street Safety Assessment and Prediction Using Public Data

2. **Business Problem/Question:**

   How can we assess and predict the safety levels of different neighbourhoods in the City of Chicago by integrating multiple data sources (such as crime reports, 311 service requests, and traffic incidents) to help residents, city planners, and law enforcement make informed decisions about urban safety?

3. **Data Sources:**

   - **311 Service Requests**: Records of non-emergency issues reported by residents, such as streetlight outages and potholes.
   - **Crime Reports**: Data on reported crimes in the city, categorized by type, location, and time.
   - **Traffic Crashes**: Information about traffic incidents, including severity, location, and time.

4. **Modeling Approach:**

   - **Feature Engineering**: After data integration, features such as the number of crime incidents, traffic crashes, and service requests are aggregated to create predictors.
   - **Street Safety Score Calculation**: A scoring system is developed by assigning weights to different incident types, reflecting their severity. This score serves as a foundational metric for further classification.
   - **Classification Model**: A Random Forest Classifier is used to predict safety levels across neighbourhoods

5. **Project Approach:**
   Our approach aligns with the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology. This cyclical method initiates with:

   - Business Understanding
   - Data Exploration
   - Data Preparation
   - Modelling
   - Evaluation
   - Deployment

6. **Significant Assumptions/Constraints:**

- **Data Integrity**: It is assumed that the datasets used are current, complete, and reflect actual conditions in the city.
- **Data Availability**: The analysis depends on the availability of data from the City of Chicago's open data portal. Any delays or gaps in data updates may affect the results.
- **Scoring System Validity**: The Street Safety Score is based on predefined weights for different incidents. It is assumed that these weights accurately reflect the severity and impact of each type of incident.
- **Model Generalization**: The model is trained on historical data, and its accuracy depends on the stability of incident patterns. Sudden changes in crime or traffic patterns may require recalibration.

7. **Adapting Existing Work**

Several preliminary studies have explored safety analysis in urban areas using crime or traffic data. By referencing these foundational works, our project aims to extend the analysis by integrating multiple data sources (crime reports, 311 service requests, and traffic incidents) to develop a comprehensive safety score. This will enable a more detailed and neighbourhood-specific understanding of urban safety challenges.

8. **Deviations from Project Requirements:**
   At this point, I foresee no deviations from the defined project guidelines. Our intent is a comprehensive analysis that remains in strict alignment with all provided directives.