DA Problem Statement

Mentorness Internship Program

Problem Statement:

This internship project aims to conduct a comprehensive analysis of Chicago traffic incidents using Power BI, leveraging a dataset containing detailed information about crashes. The dataset includes various attributes such as crash date, speed limits, weather conditions, and injuries sustained. The goal is to visualize and analyse traffic patterns, identify contributing factors to accidents, and provide actionable insights for enhancing road safety and traffic management in Chicago.

Dataset Description:

1. CRASH_RECORD_ID: Unique identifier for each crash record.
2. CRASH_DATE_EST_I: Indication if the crash date is estimated.
3. CRASH_DATE: Actual date and time of the crash.
4. POSTED_SPEED_LIMIT: Posted speed limit at the crash location.
5. TRAFFIC_CONTROL_DEVICE: Type of traffic control device at the crash location.
6. DEVICE_CONDITION: Condition of the traffic control device.
7. WEATHER_CONDITION: Weather conditions at the time of the crash.
8. LIGHTING_CONDITION: Lighting conditions at the crash location.
9. FIRST_CRASH_TYPE: Primary crash type.
10. TRAFFICWAY_TYPE: Type of traffic way where the crash occurred.
11. LANE_CNT: Number of lanes at the crash location.
12. ALIGNMENT: Street alignment at the crash location.
13. ROADWAY_SURFACE_COND: Condition of the roadway surface.
14. ROAD_DEFECT: Defects in the road at the crash location.
15. REPORT_TYPE: Type of crash report.
16. CRASH_TYPE: Type of crash (e.g., hit and run).
17. INTERSECTION_RELATED_I: Indicates if the crash is intersection-related.
18. NOT_RIGHT_OF_WAY_I: Indicates if the crash occurred outside the right of way.
19. HIT_AND_RUN_I: Indicates if the crash was a hit and run.
20. DAMAGE: Extent of damage in the crash.
21. DATE_POLICE_NOTIFIED: Date and time when the police were notified.
22. PRIM_CONTRIBUTORY_CAUSE: Primary contributing cause of the crash.
23. SEC_CONTRIBUTORY_CAUSE: Secondary contributing cause of the crash.
24. STREET_NO: Street number of the crash location.
25. STREET_DIRECTION: Street direction at the crash location.
26. STREET_NAME: Street name at the crash location.
27. BEAT_OF_OCCURRENCE: Police beat where the crash occurred.
28. PHOTOS_TAKEN_I: Indicates if photos were taken at the crash scene.
29. STATEMENTS_TAKEN_I: Indicates if statements were taken at the crash scene.
30. DOORING_I: Indicates if dooring was involved in the crash.
31. WORK_ZONE_I: Indicates if the crash occurred in a work zone.
32. WORK_ZONE_TYPE: Type of work zone, if applicable.
33. WORKERS_PRESENT_I: Indicates if workers were present in a work zone.
34. NUM_UNITS: Number of units involved in the crash.
35. MOST_SEVERE_INJURY: Most severe injury reported.

36. INJURIES_TOTAL: Total number of injuries reported.
37. INJURIES_FATAL: Number of fatal injuries reported.
38. INJURIES_INCAPACITATING: Number of incapacitating injuries reported.
39. INJURIES_NON_INCAPACITATING: Number of non-incapacitating injuries reported.
40. INJURIES_REPORTED_NOT_EVIDENT: Number of injuries reported but not evident.
41. INJURIES_NO_INDICATION: Number of cases with no indication of injuries.
42. INJURIES_UNKNOWN: Number of cases with unknown injuries.
43. CRASH_HOUR: Hour of the day when the crash occurred.
44. CRASH_DAY_OF_WEEK: Day of the week when the crash occurred.
45. CRASH_MONTH: Month when the crash occurred.
46. LATITUDE: Latitude of the crash location.
47. LONGITUDE: Longitude of the crash location.
48. LOCATION: Geographical coordinates of the crash location.

Project Objectives:

1. Data Exploration:

- Explore the dataset to understand the distribution of variables and identify potential trends.

2. Visualization and Analysis:

- Create informative Power BI visualizations to analyse crash patterns and trends.
- Explore relationships between various factors and the occurrence of accidents.

3. Geospatial Analysis:

- Utilize geographical information to map and analyse crash locations in Chicago.

4. Contributory Factor Analysis:

- Identify primary and secondary factors contributing to crashes.
- Analyse the most common causes of accidents.

5. Injury Analysis:

- Explore the severity and distribution of injuries in reported accidents.

Deliverables:

- Interactive Power BI dashboards with visualizations and insights.
- Summary reports on crash patterns, contributing factors, and injury analysis.
- Recommendations for improving road safety based on the analysis.

This project provides interns with hands-on experience in using Power BI for real-world traffic data analysis, offering insights that can contribute to Chicago's road safety initiatives.