Analyzing SFO Police Department Incident Reports for Social Good

1. Problem Description

We chose the police department incident report dataset of SFO because it offers a significant opportunity to use data science for social good. By analyzing this rich data set, we can gain important insights into crime patterns, help enhance public safety and inform police resource allocation. Using predictive analytics, we can look at areas where crime is most likely to occur, leading to more proactive policing and potentially reducing crime rates.
To analyze crime incident data from San Francisco, focusing on incident categories, locations, times, and trends.
CSV file titled 'processed_Police_Department_Incident_Reports__2018_to_Present.csv' is data source.

2. Significance of the Problem

- Understanding crime patterns aids in enhancing public safety and resource allocation for law enforcement.
- The insights can inform policy makers and community leaders in their decision-making process.

3. Database Normalization:

- **Dataset:** Police Department Incident Reports (2018 to Present).
- **Tables:**
  - Incidents: Stores details about each incident.
  - Location: Contains location-related information for each incident.
  - Report: Holds information about reports filed for incidents.
  - Supplementary_Info: Additional information related to incidents.
  - Time: Time-related data for incidents.
- **Normalization:** Follows Third Normal Form (3NF) for efficient data structure, ensuring atomicity and eliminating redundancy.

4. Methodology

- **Database Creation:** Utilized SQLite to create a database named 'sfo_city_crime.db'.
- **Data Modeling:** Designed multiple tables (Incidents, Location, Report, Supplementary_Info, Time) to organize data efficiently.
- **Data Extraction:** Employed Python and SQL for data manipulation, ensuring relevant information is extracted for analysis.
- **Visualization:** Implemented plots using Plotly for visual representation of trends and patterns.

5. Results and Analysis

- **Trend Analysis of Incident Categories Over Time:** Demonstrated the variation in different types of incidents over the years.
- **Incident Frequency by Day of the Week:** Provided insights into which days have higher incident rates.
- **Incident Frequency by Time of Day:** Identified peak hours for various incident categories.
- **Geographical Analysis:** Mapped incidents based on latitude and longitude to pinpoint hotspots.

- **Yearly Incident Comparison:** Highlighted the fluctuation in incident numbers across different years.
- **Incident Distribution by Police Districts:** Showed which police districts experience higher incident rates.
- **Neighborhood Analysis:** Explored incident categories across different neighborhoods.
- **Hourly Incident Analysis:** Correlated incident types with specific hours of the day.
- **Supervisor Districts Incident Trends:** Assessed how incidents are distributed across different supervisor districts.
- **Online vs. Direct Report Analysis:** Compared the number of incidents reported online versus in-person.
- **Category-Specific Time Analysis:** Focused on a specific category ('Drug Offense') to understand its temporal trends.

6. Conclusions

- The analysis offers a comprehensive view of crime incidents in San Francisco, highlighting key areas and times of high activity.
- Patterns observed can assist in strategic planning for law enforcement and community safety initiatives.
- The dynamic nature of crime incidents underscores the importance of continuous monitoring and analysis.