

Project Plan for "Projet Police"

[Nessim LIAMANI]

June 30, 2024

Contents

1	Understanding the Problem	2
1.1	Theme	2
1.2	Context	2
1.3	Objective	2
2	Data Sources	2
2.1	Primary Data	2
2.2	Additional Data	2
3	Administrative Objectives	2
3.1	Business and Functional Analysis	2
3.2	Project Management	3
4	Technical Objectives	3
4.1	Database Construction	3
4.2	ETL Process	3
4.3	Visualization and Reporting	3
5	Key Questions to Address	3
5.1	Global Situation: Analyze overall trends in police shootings . . .	3
5.2	Ethnic Discrimination: Determine if certain ethnic groups are disproportionately affected	4
5.3	Regional Differences: Compare shootings across different U.S. regions	4
5.4	Temporal Evolution: Identify trends over time	4
5.5	Police Misconduct: Assess the extent of police misconduct	4
6	Detailed Plan and Steps	5
6.1	Analysis and Design	5
6.2	Data Preparation	5
6.3	Data Warehouse Setup	5
6.4	ETL Development	5
6.5	Visualization with Power BI	5
6.6	Documentation and Reporting	6

1 Understanding the Problem

1.1 Theme

Analyzing police shootings in the U.S. to identify potential discrimination and patterns.

1.2 Context

There is a significant concern about police shootings, especially involving ethnic minorities.

1.3 Objective

Create a comprehensive data solution to analyze and report on police shootings, addressing key questions on ethnic discrimination, regional differences, temporal trends, and police misconduct.

2 Data Sources

2.1 Primary Data

- `fatal-police-shootings.csv`: Details on police shootings.
- `Ethnicity Data USA.xlsx`: Ethnic distribution in the U.S.
- `uscities.csv`: Information about U.S. cities.
- `fatal-police-shootings-agencies.xlsx`: Agencies involved in shootings.

2.2 Additional Data

Other relevant datasets as needed for further analysis.

3 Administrative Objectives

3.1 Business and Functional Analysis

- Model the use case diagram for understanding interactions.
- Design the database schema for OLTP and OLAP systems.

3.2 Project Management

- **Deliverables and Tasks:** Identify tasks, assign responsibilities, and estimate time required.
- **Critical Path:** Determine the sequence of essential tasks.
- **Resource Allocation:** Assign team members to tasks.
- **Risk Management:** Identify and plan for potential risks.

4 Technical Objectives

4.1 Database Construction

- **OLTP:** Build `Police_0.Main` for initial data storage.
- **Data Cleaning:** Clean and preprocess data using Python and T/SQL.
- **OLAP:** Build `Police_1.DWH` using a star schema for analytical queries.

4.2 ETL Process

- Develop ETL scripts using Python, T/SQL, and Powershell.
- Schedule regular ETL tasks using Windows Task Scheduler.

4.3 Visualization and Reporting

- Utilize Power BI to create interactive reports and dashboards.

5 Key Questions to Address

5.1 Global Situation: Analyze overall trends in police shootings

1. What is the annual number of fatal police shootings in the United States from 2015 to 2022?
2. What are the most common circumstances (e.g., manner of death, armed status) under which police shootings occur?
3. What is the demographic profile (age, gender, race) of individuals involved in police shootings?

5.2 Ethnic Discrimination: Determine if certain ethnic groups are disproportionately affected

1. What is the rate of fatal police shootings per 100,000 people for each ethnic group?
2. How do the characteristics of police shootings (e.g., armed status, threat level) differ between ethnic groups?
3. What proportion of the total police shootings does each ethnic group represent, and how does this compare to their proportion of the total population?

5.3 Regional Differences: Compare shootings across different U.S. regions

1. Which states or regions have the highest and lowest rates of fatal police shootings?
2. Comparison of police shootings by state, focusing on the number of shootings involving unarmed individuals
3. What are the characteristics of police shootings (e.g., armed status, threat level) in different regions of the United States?

5.4 Temporal Evolution: Identify trends over time

1. How has the number of fatal police shootings changed annually from 2015 to 2022?
2. What are the trends in police shootings for different ethnic groups over time?
3. Are there seasonal or monthly patterns in the occurrence of fatal police shootings?

5.5 Police Misconduct: Assess the extent of police misconduct

1. What percentage of fatal police shootings involve victims who were unarmed at the time of the incident?
2. How many incidents involve officers who had their body cameras turned off or malfunctioning during the shooting?
3. What is the frequency of fatal police shootings involving signs of mental illness, and how are these incidents handled compared to others?

6 Detailed Plan and Steps

6.1 Analysis and Design

- **Use Case Diagram:** Define interactions between the system and users (DOJ analysts, data engineers).
- **Database Schema:** Design schemas for `Police_0_Main` and `Police_1_DWH`.

6.2 Data Preparation

- **Data Import and Cleaning:**
 - Write Python scripts to load and clean data from CSV/Excel files.
 - Use T/SQL scripts for further cleaning and integrity checks.

6.3 Data Warehouse Setup

- **Schema Design:** Define fact and dimension tables for star schema.
- **DDL Scripts:** References to DDL scripts will be included in the appendices.

6.4 ETL Development

- **ETL Process:** Develop ETL scripts using Python, T/SQL, and PowerShell.

6.5 Visualization with Power BI

- **Connect Power BI to `Police_1_DWH`.**
- **Create Reports and Dashboards:**
 - **Global Trends:** Line chart showing trends over years.
 - **Ethnic Discrimination:** Bar chart comparing shooting rates by ethnicity.
 - **Regional Analysis:** Map visualization showing shootings by state.
 - **Temporal Trends:** Line chart showing shootings over time.
 - **Police Misconduct:** KPI indicators for signs of police misconduct.

6.6 Documentation and Reporting

- **Project Report:**
 - Methodology
 - Data sources and preparation
 - Analysis and findings
 - Visualizations and insights
- **Backup of Data Warehouse:** Provide a full backup of the `Police_1_DWH` database.
- **Source Code:** All scripts used for ETL and database management.

By following this structured plan, the project will systematically address the key objectives and deliver comprehensive insights into the police shootings across the United States.