US State Police

MAJOR RESEARCH PROJECT PROPOSAL

Crime Analysis

PREPARED BY:

HARBIR SINGH||ABHISHEK SHARMA|| DIVYABEN PATEL || BRIJ KUMAR PATEL

ABOUT US

Intellectual Insights is a service based customer centric data driven organization, formed in 2018 that caters to International clients all across the world. Providing end to end data related solutions from architecture to collection as well as sharing meaningful insights for real world implementation.

The core team of the company is an Georgian college Alumni.
We use sophisticated analytical tools and processes to uncover operational weaknesses, spot bottlenecks, and optimize performance in our client's business operations.

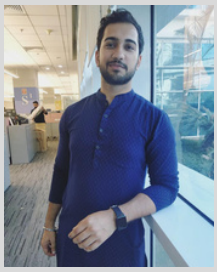
We combine business analytics, data mining, data visualization, data tools and infrastructure, and best practices to help organizations make more data-driven decisions.

THE DEVELOPMENT TEAM

Say hello to our team of experts!

Team Lead

Harbir Singh



Business Analyst

Abhishek Sharma

Developer

Divyaben Patel

Developer

Brij Kumar Patel

INTRODUCTION

In ordinary language, a **crime** is an unlawful act punishable by a state or other authority. The term crime does not, in modern criminal law, have any simple and universally accepted definition, but in the modern day crimes are something that can be observed commonly around.

Even though Authorities employ various mechanisms to regulate certain behaviors in general, no one is able to mitigate this evil completely.

The FBI estimated crime statistics for the nation are based on data received from 11,794 of 18,806 law enforcement agencies in the country and the prediction on the crime rate is an ongoing rise.

The United States Capitol Police (USCP) dates back to 1800 when the Congress moved from Philadelphia to Washington, D.C.
Today, the USCP embodies the best in American policing and serves as a model in security, urban crime prevention, dignitary protection, specialty response capabilities, and homeland security.

PROJECT DESCRIPTION

With the goal of minimizing and analyzing the trends of crime over the United States, USCP requires in depth data analysis to predict and deploy resources accordingly.

The aim of the project is to look for any seasonality or trend from the historical data to understand if certain crimes are happening at some specific region.

And, to study the efficiency of the department to review and improve the handling time of various crimes.

OUR RESEARCH QUESTIONS AND QUERIES

- To check the average time taken to dispatch for different police stations
- To analyze the trend of highest crime types over geographical locations
- Victim analysis over the type of crimes
- To track the spread of Theft and Drug related crimes over the years
- Pareto analysis of the types of crimes

SCOPE OF PROJECT

This project is focused on: -

- Defining requirements for the project including key questions
- Data cleaning to battle the data inconsistency.
- From the prospective of the region, the geographical analytics.
- Using predictive models to determine better technique for the given data.
- From the prospective of the people, the analysis of the average time taken for dispatch
- From the prospective of ML, Time series prediction technique, classification, clustering and MBA.

OUT OF SCOPE

- The dashboard/analysis of the project would not be published on real time.
- The analysis would not be interactive in nature.
- The results would not be published on a web site or app for public review.
- No data pipelines would be created for the process.

OUTCOMES AND BENEFITS

- The client would be able to pin-point the certain crimes to create hot-spots
- Better visualization/dashboard once published on the would help the people to make calculated decisions
- The analysis would help in smart allocation of resources

DELIVERABLES

The deliverables include:

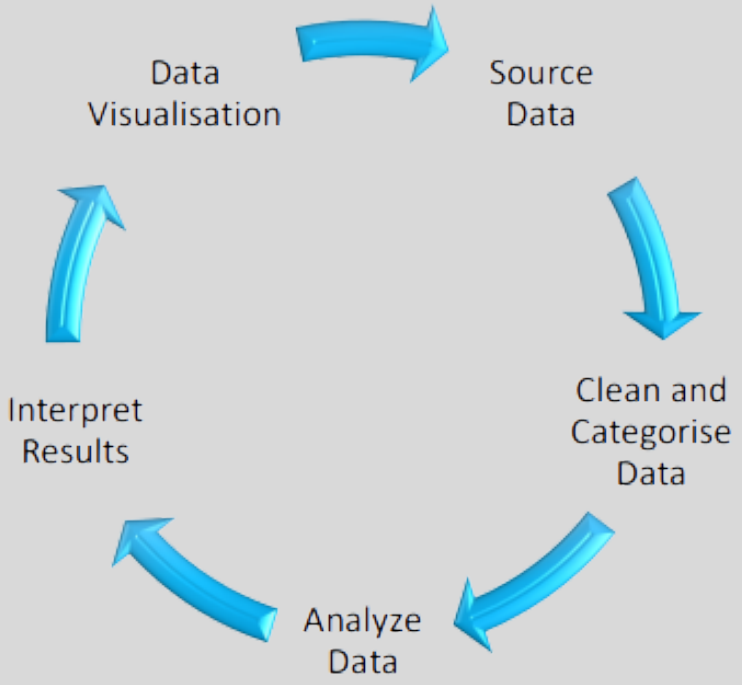
- A comprehensive report and an in-person slideshow presentation of the approach taken for the project, the observations, Analysis and conclusions.

- Knowledge transfer of all research, analysis, visualizations, analysis models, testing, codes, formulas, algorithms, technology, documentation and work product

EXECUTION PLAN

- **Sources** :
  The US state police open-source data
- **Types** :
  Transactional data, manually input
- **Structure** :
  Comma Separated Values (CSV)
- **Format** :
  Traditional, Microsoft Excel file columns and rows
- **Volume** :
  1. About 300,000 rows of data across 30 different variables

DATA ANALYSIS PLAN



STEP1: SOURCE DATA

The data was obtained from United States Capitol Police (USCP) department's portal.
It contains information on the type of crime, incident date, location, street address, latitude and longitude, dispatch date time.
The dataset contains over 300,000 rows of incidents occurred over a span of 5 years ranging from 2017 to 2022.
And, has a mix of various data types.

STEP2: CLEAN AND CATEGORIZE

DATA STEP3: ANALYZE DATA

STEP4: INTERPRET RESULTS

STEP5: DATA VISUALIZATION

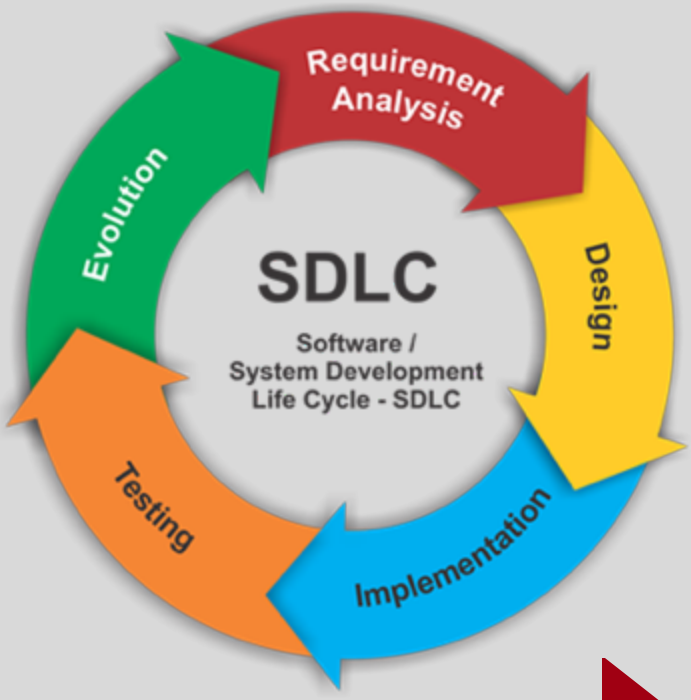
TOOLS AND LIBRARIES

- Jupyter Notebook
- Spyder
- Microsoft Excel
- Power BI
- Matplotlib
- Pandas
- Microsoft SQL Server Management Studio

EXECUTION PROCESS

Software Development Life Cycle (SDLC)

- Requirements
- Design
- Development
- Testing
- Release



POTENTIAL CHALLENGES

Even though some challenges may develop, we plan to provide the Client with completed product. Still, these potential challenges would impact our Process and schedule:

1.Timeframe
The Team will exercise its best efforts to respect deadlines and deliver the research in the right time frame.

2.Communications
The Client is committed to supporting the Project Team to achieve the objective and provide additional resources if required.

CODE OF CONDUCT

Ethical Considerations
• We acknowledge our strict duty of confidentiality.
• We will ensure the observance of ethical practices demonstrating respect, honesty and dignity.
• We will thoroughly follow the Client's directives and analyze the results as directed by them without deviation from the objectives of the research.
• Value is always present in our thinking.
• Validated data will solely form the basis of our observations, conclusions, recommendations and decision-making.
• We undertake to use holistic technical strategies and repeatable processes.
• We will thoroughly explain and document our development in detail for the Client. We will be transparent in all our activities.

Moreover, we will check and recheck our research for quality, accuracy, completeness, and integrity before presentation to maintain the results' validity and credibility.
• We will proactively circumvent unethical behaviour, such as exaggerating the results of our research.
• We will not falsely interpret, fabricate, embellish or otherwise misrepresent the technology to validate our product or change or omit details.
• We will be forthright and accountable if we make mistakes.
• We will tell the truth, even if it is bad news.
• Failure to observe deadlines is considered a breach of ethics.
• We have no conflict of interest and will not benefit from this project.

INFORMATION MANAGEMENT

• We acknowledge that the details are an asset of any organization. The consequences and repercussions of unethical conduct when dealing with an organization's data can be significant and affect an organization's reputation, relationships and, ultimately, its revenues. Even the perception of unethical data handling has the power to undermine both internal and external trust.
• We give assurance that all the information provided is in our trust. It will remain private and protected from damage or alteration unless authorized.
• We will sign a Non-Disclosure Agreement that is acceptable to the Client and adhere to its conditions.
• We will notify the Client when discovering any sensitive information at any stage, namely personal identifying information or confidential information. We will make recommendations for addressing the issues and act according to their instructions.
• Original data and other information entrusted to us is stored in a secure location, such as SharePoint, and will remain unaltered.
• Copies of original data will be used for cleaning, discovery, manipulation and Analysis.
• Naming conventions and version controls will segregate datasets and align each with the applicable Analysis and documentation.
• All analyses will only be conducted through secure applications.

• All our workstations are protected by a password that is not known to anyone other than me. No one has access to our workstations, including colleagues, family members and friends.
• All data sharing is secured from potential risks using encrypted channels, such as SharePoint.
• Discussions about the data, Analysis and observations will only be amongst client-authorized collaborators, our course instructors and the Client's designated representatives.
• All information entrusted to us will only be used for its intended purposes unless specifically approved by the Client

MILESTONES

Milestone	Complete by
Acceptance of Project Charter	December 15, 2022
Requirement phase	January 26, 2023
Design phase	February 16, 2023
Development phase	March 9, 2023
Testing phase	March 23, 2023
Release phase	April 6, 2023
Handover of the deliverables	April 18, 2023
Project close-out	April 18, 2023

CONCLUSION

Upon completing the project, we will hand over a comprehensive written report of our project and the interface created to the Georgian college. These interfaces can be readily used by the client. Our development and protocols would support daily operations and backend productivity of the company.
We value the opportunity to apply our skills, experience, and learning to the project and our career goals while working with and supporting Georgian College – Design, Visual Arts, and Computer Studies.
Thank you

Georgian College Big Data Analytics Team
HARBIR SINGH || ABHISHEK SHARMA || DIVYABEN PATEL || BRIJ KUMAR PATEL