**BDAT 1011 – Major Research Project (MRP)**

**Final Report – Project Closeout**

**Group # 5**

***Instructor: Roshan Sahu***

**Name** **& Student ID:**

1. **Namrakumar Patel (200535046)**
2. **Gaurav Gaurav (200578193)**
3. **Samarjeet Singh (200558169)**

**Group #: 5**

**Submission Date**: 11/12/2023

Final Report – Project Close Out

# Project Summary

* Toronto Police Data Portal is an invaluable resource that provides a comprehensive and enriched perspective on data from different time periods.
* Data offers detailed insights into a wide range of criminal activities, presenting a vast array of statistical information.
* Main objective is to provide the Toronto Police Department with a comprehensive report on key crime indicators.
* Our team is embarking on a highly intriguing and challenging analysis project that centres on the Toronto Police Department. With the advent of the digital age, copious amounts of data have been made publicly accessible on their website, notably in the form of the "major crime indicators" dataset. Our primary objective is to utilize this resourceful and valuable information to conduct a comprehensive and in-depth analysis of the Toronto Police Department's operations, uncovering meaningful insights, and deriving valuable conclusions.
* Analysis will provide essential insights into the complex dynamics of criminal activities over time.
* To be sure, this project will be no mean feat, given the complexity and sheer volume of the data to be analyzed. Nevertheless, we are keen and determined to dig deep into the numbers, applying various analytical techniques, and using state-of-the-art data mining tools and technologies to unearth actionable and meaningful results.
* Main goal is to provide more security for such occasions when criminal record is high.
* We are confident that our endeavour will shed light on various aspects of the Toronto Police Department's modus operandi, ranging from crime trends, hotspots, patterns, and response times to name but a few. Our aim is to offer valuable insights to stakeholders and interested parties, including law enforcement officials, policymakers, and researchers, among others, with the hope of improving public safety and security in the City of Toronto.
* A comprehensive report and an in-person slideshow presentation of the approach taken for the project, the observations, and conclusions.
* The analysis will involve an in-depth examination of various crime indicators, including but not limited to, homicide, assault, sexual assault, robbery, break and enter, theft, and motor vehicle theft. The results of the analysis will be presented in a clear, concise, and comprehensive manner, highlighting the key findings and insights obtained.
* It is an independent study, will be done by us.

# Communications Plans

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **Role** | **Send/Receive Email** | **Weekly Meetings** | **Documentation** |
| Roshan Sahu | Instructor | Yes | Optional Attendee | Grading |
| Namrakumar Patel | Project Manager | Yes | Attend, Book, Minutes, Agenda | Project Management Presentation components |
| Samarjeet Singh | Data Analyst | Yes | Attend | Data gathering, Data Documentation |
| Gaurav Gaurav | Developer | Yes | Attend | Coding, Developing solution |

# Application

**Resources:**

The data utilized in this project has been carefully collected from the reputable Toronto Police Department's extensive data library. This library serves as a valuable resource containing a wide range of datasets covering various topics and subjects. It is a rich source of information that can be utilized to meet the specific needs and preferences of the professor. The datasets are organized in a structured format, presented in rows and columns, making them easily understandable and applicable. Access to this valuable data repository can be obtained through the Toronto Police Department's website, conveniently located at Data.torontopolice.on.ca.

**Information Sources**

<https://en.wikipedia.org/wiki/Crime_in_Toronto>

[Toronto Police Service Public Safety Data Portal](https://data.torontopolice.on.ca/)

<https://www.kaggle.com/datasets/kapastor/toronto-police-data-crime-rates-by-neighbourhood>

**Techniques**

* **Descriptive Analysis:** Descriptive Analysis summarizes data describing what it is or what it shows. It is essential to identify the data characteristics before in-depth Analysis.
* **Discovery Analysis:** Discovery Analytics consolidates data for evaluation. Once raw data is converted, we drill down to identify factors for further Analysis.
* **Ranking Analysis:** This technique involves ranking items based on specific criteria. Ranking analysis can be useful for making decisions based on the relative importance of different factors.
* **Trend analysis**: This technique involves analyzing data over time to identify patterns or trends. Trend analysis can be useful for identifying changes in the data, predicting future behavior, and understanding how different factors affect the data.
* **Comparative study:** This technique involves comparing two or more sets of data to identify similarities and differences. A comparative study can help to understand how different variables affect the data and can be useful for making decisions based on the results.
* **SWOT analysis:** SWOT stands for strengths, weaknesses, opportunities, and threats. SWOT analysis is a strategic planning technique that involves evaluating the internal and external factors that affect an organization or project. It can be useful for identifying areas for improvement, understanding the competitive landscape, and making strategic decisions.

**Tools:**

The Toronto Police Department relies on a range of powerful tools to effectively analyze and visualize crime data, allowing law enforcement officials to better understand crime patterns and trends in their community. These tools include:

**A green square with a white x

Description automatically generatedMicrosoft** **Excel**

Microsoft Excel is a spreadsheet that features calculation and graphing tools, pivot tables, and programming capabilities for functions, queries and reports. Excel is used for data storage, data preparation, analysis and visualizations of charts, graphs and pivot tables. The selection of Excel as a tool is based on its reliability and universal acceptance as a fundamental analytics application.

**A yellow rectangular objects with a white background

Description automatically generatedPower BI**

Power BI is a cloud-based business analytics service by Microsoft that provides data warehousing, data preparation, data discovery, interactive visualizations, and business intelligence capabilities. Power BI will be used for data storage, data preparation, analysis, and visualizations in interactive visualizations and dashboards. Again, the selection of Poser BI is based on its reliability and its acceptance as a reliable analytics, visualization, and business intelligence application.

**Google** **Data Studio**

**A close up of a logo

Description automatically generated** A powerful web-based reporting and visualization tool that enables the analysis team to create interactive dashboards and reports from a variety of data sources, including Major Crime Indicator (MCI) data. Google Data Studio provides a simple, yet sophisticated way to analyze complex data sets, with visually engaging reports that can be shared with a variety of stakeholders.

A blue hexagon with a magnifying glass and a white circle with a graph in the center

Description automatically generated**Big Query**

A cloud-based data warehouse that provides real-time analysis of large datasets. Our team can use Big Query to analyze MCI data quickly and efficiently, enabling them to identify patterns and trends in crime that might be difficult to detect otherwise. By providing real-time access to the data, Big Query ensures that law enforcement officials can respond to emerging threats in a timely manner.

A logo of a cloud platform

Description automatically generated**Google Cloud Platform (GCP)**

A comprehensive suite of cloud-based services and tools designed for data analysis, storage, and management. This includes Big Query and Google Data Studio, as well as a host of other services that can be used to analyze MCI data. By leveraging the power of GCP, the Toronto Police Department can access cutting-edge technologies and tools that help them stay ahead of emerging trends in criminal activity.

**Execution Process**

Requirements Phase

* Define and document the requirements and expected outcomes and benefits of the project.
* Identify sources of data and relative information
* Collect and prepare the data for Analysis.
* Descriptive Analysis
* Discovery Analysis
* Select and document Techniques, Tools, and visualization methods.

# Thinking

**Potential Challenges**

Data quality issues: The accuracy and completeness of the data can be a challenge, as data may be incomplete, inaccurate, or inconsistent. This can lead to unreliable or biased results.

Data complexity: Major crime indicator data can be complex, with a large number of variables and multiple data sources. Analyzing this data requires advanced statistical techniques and data processing tools, which may be difficult to use for those without specialized training.

Privacy concerns: The use of sensitive personal information in crime data analysis can raise privacy concerns, and appropriate measures need to be taken to ensure that personal data is protected.

Data integration: Major crime indicator data can come from multiple sources, such as law enforcement agencies, courts, and correctional facilities, making it difficult to integrate and analyze the data effectively.

Bias and discrimination: Analysis of major crime indicator data can be subject to bias and discrimination, particularly if there are underlying biases in the data collection or analysis process. This can lead to incorrect conclusions or unfair treatment of certain groups.

## **Changing crime patterns:** Crime patterns can change over time, and major crime indicator data may not always reflect current crime trends. The analysis must be kept up-to-date to ensure its relevance.

## **Risk and Issues Management**

* Risk and Issues Management is an integral component of our Material Requirements Planning (MRP) strategy. Our focus is on identifying, assessing, and mitigating potential risks and issues that may arise during the Project process. By promptly addressing these challenges and implementing proactive measures, we aim to minimize disruptions, optimize inventory levels, and ensure the seamless execution of Project activities. Our commitment to efficient and reliable Project outcomes remains unwavering.

**Timeframe**

* Our dedicated team is fully committed to ensuring the efficient completion of all Problem Statement activities and deliverables within the designated time frame, ensuring timely and satisfactory outcomes.

**Project Team**

* While we have full confidence in our team, we recognize that unforeseen circumstances such as exams, illness, or personal events may arise. In the event of any potential delay in the project's progress, we will promptly notify the Professor and provide a clear plan outlining how our team intends to fulfill the project's requirements.

**Model Selection**

* The process of problem statement model selection requires careful consideration and can be time-consuming. If a misfit is identified, there may be limited time available to choose an alternative model and complete the process. However, our dedicated team is committed to exerting their best efforts to ensure the selection of an appropriate model that aligns with the requirements of this project.

**Communications**

* Our project team is dedicated to achieving the objective and stands ready to provide additional resources if needed, ensuring effective communication throughout the project.

**Scope Creep and Changes**

* Scope creep may arise due to inadequate definition, documentation, or control of a project's scope, as well as additional elements introduced by the professor or the team. To mitigate this risk, the team will diligently document each requirement during the planning phase and allocate resources accordingly. In the event of scope changes, the team will document the requirements, assess the impact on the project, inform the Professor, provide recommendations, and act based on the Professor’s instructions.

**Process Of Visualization:**

A diagram of a flowchart

Description automatically generated

# Completion Criteria

Identify the criteria required to complete the project.

|  |  |  |
| --- | --- | --- |
| **Item** | **Completion Criteria** | **Complete** |
| Data Analysis | Client approval of the analysis | Yes |
| Website | Completion of items listed under “Website” in Project Charter, confirmed by client | Yes |
| Dataflow Implementation | Client approval of the Dataflow Implementation | Yes |
| Data Preprocessing and Cleaning | Client approval of the Data Preprocessing and Cleaning | Yes |
| Real-time Result Visualization | Client approval of the Real-time Result Visualization | Yes |
| Reporting and Exporting | Client approval of the Reporting and Exporting | Yes |
| Error Handling and Logging | Client approval of the Error Handling and Logging | Yes |
| Documentation and Training | Client is satisfied with the Documentation and Training | Yes |
| Compliance and Legal Considerations | Client is satisfied with the Compliance and Legal Considerations | Yes |
| Testing | Client is satisfied with the Testing | Yes |

# Project Close-Out Package

This section details all items that are contained in the Project Close-Out Package. Detail all deliverable items referred to in the Project Charter.

|  |  |  |
| --- | --- | --- |
| **Item** | **Method of Delivery** | **Complete** |
| Application Code | Saved to SharePoint | Yes |
| Database | Data has been derived from Toronto Police Department official site. | Yes |
| Dataflow Implementation | Prefect dataflow is sent to the client | Yes |
| Real-Time Result Visualization | Power Bi dashboard is shared via URL to the client through email | Yes |
| Future Trend Analysis | Forecast data for upcoming 2 years. | Yes |
| Documentation and User Guide | All documentations are sent to the client | Yes |
| Data Backup and Security | A link of Data sent is sent to the client all codes for extracting data are shared with the client | Yes |
| Project Deployment Package | Power bi report, Csv file and all codes are sent to the client | Yes |
| Testing and Quality Assurance Reports | All the defect reports, Code Coverage reports, Security Test Reports, Test Execution reports are sent to the client. | Yes |
| SharePoint | Contains every report and our Visualizations file. | Yes |

SharePoint Link: [<https://georgiancollege.sharepoint.com/sites/TorontoPoliceDepartment>](https://georgiancollege.sharepoint.com/sites/TorontoPoliceDepartment)

PowerBi Report: [<https://app.powerbi.com/links/LEgwdDCN6x?ctid=da9a94b6-4681-49bc-bd7c-bab9eac0ad3c&pbi_source=linkShare>](https://app.powerbi.com/links/LEgwdDCN6x?ctid=da9a94b6-4681-49bc-bd7c-bab9eac0ad3c&pbi_source=linkShare)

# Document Approvals

This section lists the significant stakeholders related to the project.

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Role** | **Name** | **Signature** | **Date** |
| Client | Toronto Police Department | Roshan Sahu (Instructor) | 11/12/23 |
| Project Manager | Namrakumar Patel | Namrakumar | 11/12/23 |
| Student 1 Role | Namrakumar Patel  (Team Leader) | Namrakumar | 11/12/23 |
| Student 2 Role | Gaurav Gaurav  (Developer) | Gaurav | 11/12/23 |
| Student 3 Role | Samarjeet Singh  (Data Analyst) | Samarjeet | 11/12/23 |