Delay to Data: Revolutionizing TTC Timeliness Through Analytics

1. **Introduction and Background**

The Toronto Transit Commission (TTC) bus network serves thousands of commuters daily, making it a critical component of public transportation in Toronto. However, bus delays significantly impact the quality of service and commuter satisfaction. The goal of this project is to analyze bus delay data from 2023 to current and identify areas for improvement, aiming to reduce delays by 70% in the upcoming years.

The motivation behind this project arises from the need to optimize public transit efficiency and reliability. By examining historical data on bus delays, we aim to uncover patterns and factors contributing to frequent service interruptions. The insights derived from this analysis will help transit authorities implement data-driven strategies to minimize delays and enhance commuter experiences.

1. **Objectives and Goals**

The primary objectives are:

1. **Identify Key Factors:** Determine the primary causes of bus delays and their frequency.
2. **Geographic Analysis:** Map the hotspots of bus delays across Toronto to identify problematic areas.
3. **Temporal Analysis:** Examine how delays vary across weekdays, weekends, seasons, and years.
4. **Route Efficiency:** Identify which bus routes experience the most delays.
5. **Impact of Incidents:** Understand the types of incidents that lead to significant delays.

By answering these questions, the project aims to support transit authorities and policymakers in making data-driven decisions to improve bus services and enhance commuter satisfaction.

1. **Datasets**

* **Source:** Toronto Open Data Portal ([Link to Dataset](https://open.toronto.ca/dataset/ttc-bus-delay-data))
* **Collector:** City of Toronto
* **Funding:** Municipal government
* **Dataset Description:**
  + The dataset contains records from **2023 to current**.
  + It includes **bus delay incidents**, **delay duration**, **route number**, **location**, **type of incident**, and **direction**.
  + The combined dataset across the years has **less than one million records**.
  + The data spans the entire **Toronto area**, allowing for city-level analysis.
* **Limitations:**
  + The data does not include **subway or streetcar delays**.
  + Some locations may lack precise geographic coordinates.
  + Data accuracy depends on TTC's internal reporting.

1. **Visualization Plan**

The project will utilize the following visualizations:

1. **Map:**
   * Display geographic hotspots of delays.
   * Visualize routes with the highest frequency of delays.
2. **Bar Chart:**
   * Show top 10 routes with the most delays.
   * Breakdown of delay incidents by day of the week.
3. **Table:**
   * Summarize delay durations by incident type.
4. **Line Chart:**
   * Trend of delays over the years (2014 to 2024).
5. **KPI:**
   * Display the total number of delay incidents and the average delay duration.
6. **Lollipop Chart:**
   * Compare the average delay per route to visualize route performance.
7. **Heatmap:**
   * Analyze delays by hour of the day and day of the week.
8. **Doughnut Chart:**
   * Proportion of incidents by delay type.
9. **Waterfall Chart:**
   * Display the difference in delays between weekdays and weekends.
10. **Scatter Plot:**

* Correlation between delay duration and gap duration.

1. **Contributions**

Since this is a group project, all tasks including data extraction, cleaning, visualization, analysis, and reporting will be carried out by Giovanny, Tanay and Gurmehak.

**Skill Demonstration:**

The final product will demonstrate proficiency in:

* Connecting to the dataset using with the help of an external database like PostGres and then connect it to Tableau.
* Applying various filters (date, top N, data source).
* Creating calculated fields, table calculations, and using LOD calculations.
* Using dynamic parameters for interactive visuals.
* Combining multiple data sources when necessary.
* Publishing interactive dashboards and creating a storytelling experience.

1. **Conclusion**

The project aims to offer data-driven insights to reduce TTC bus delays by 70%. By analyzing historical data, identifying problematic routes, and understanding delay patterns, we will provide actionable recommendations for transit authorities. This comprehensive analysis will aid the TTC in enhancing service reliability and improving commuter satisfaction.

References

City of Toronto. (n.d.). TTC Bus Delay Data. Retrieved from Toronto Open Data Portal (https://open.toronto.ca/dataset/ttc-bus-delay-data/)