# **Interim Report on Transit Windsor Capstone Project**

**Name:- Abhishek Rana**

**Student ID:- 790459**

**Group:- 3**

## **Executive Summary:-**

The Transit Windsor Capstone project is a comprehensive and collaborative effort aimed at analyzing the transit system's ridership data, rider behavior and transit timeline and routes to identify opportunities for optimization and improvement and make data-driven recommendations for the transit authorities to excel their services. The project followed a methodology that involved brainstorming about the problems Windsor transit is facing, data collection, cleaning and preparation of data, exploratory data analysis, predictive modeling, and visualization and reporting. The project has been successful in achieving its objectives so far. The data includes ridership and operational data which have been consolidated into a single database. This has allowed for a more comprehensive analysis of the transit system's performance and trends.

The key findings of the project included potential opportunities for route optimization and predictive modeling. Analysis of ridership data identified the most popular routes and times of day when they are heavily used, which could help identify opportunities to optimize schedules and allocate resources more efficiently. The team has developed interactive and intuitive data visualizations to present the findings. Also, a user-friendly and interactive website/platform has also been created for data visualization. The website allows users to interact with the data and gain valuable insights into the transit system's performance and schedule.

The insights gained from the project could have significant implications for the transit system, its riders, and the community as a whole. By optimizing routes, schedules, and resources, the transit system could improve its efficiency, reduce costs, and provide better service to its riders. The insights gained from the project could also inform future planning and decision-making in the transit field, potentially leading to more effective and sustainable transit systems.

Overall, the project has made significant progress towards achieving its goals. The insights and recommendations gained from the analysis of the transit system's performance can be invaluable to the transit authority in their efforts to optimize their services. The project's visualizations and website provide an intuitive and interactive platform for stakeholders and other interested parties to engage with the data, understand the trends and performance of the transit system, and gain insights to inform decision-making. This project serves as an excellent example of the value of data analytics in the transportation industry and provides a roadmap for future projects in this area.

## **Progress:-**

The Transit Windsor Capstone project has made significant progress since its inception. The team has successfully completed the stages of data collection, consolidation, and cleaning, allowing for more meaningful and comprehensive analysis of the transit system's performance. In the first month from the project initiation, the team worked on brainstorming the ideas and problems, SWOT Analysis for the project which helped us get a better understanding of what exactly we are trying to achieve. Based on the brainstorming and SWOT, we came up with problem statement and started gathering the data for our project. In the second month of the project, we worked on data cleaning and data exploration and along with that we have also done descriptive analysis of the problem. The third month into the project, team started working on different Machine Learning models, improving accuracy of these models and the final dashboard with which the users will be able to interact. Lastly a website was created to incorporate the project details, and integrate the dashboard and repository.

In terms of analysis, the team has been focused on exploring the ridership and rider’s behaviors, identifying trends, and understanding the factors driving the transit system's performance. Several ML models were tried for predictive analysis. For coding ML models we have used Python language. After working on Linear Regression, Tree Regression, Random Forest and Gradient Boosting models, we found that the Random Forest model is most accurate for prediction purpose for this project. Data visualizations have been created, allowing for easier interpretation and communication of the findings. For creating visualizations, we have used Tableau software which helped us create interactive and dynamic graphs and dashboard. These visualizations include bar charts, line graphs, and maps, all of which provide different perspectives on the data and insights.

Another critical aspect of the project has been the development of the website, which allows users to get information about the project and they can also interact with the data and explore the visualizations in a user-friendly and intuitive way.

Talking about the work done by me, I actively contributed in the initial phase of the project which consisted brainstorming, problem statement formulation and SWOT Analysis. I was the SPOC for data collection from Prof. John and Windsor Transit. Once we as a team finalized the data for analysis, I have done in-depth statistical analysis to identify relevant features, identifying the features importance for our analysis. Apart from the all the visualization aspect has been handled by me. I have worked on finding initial trends from the data using the visualizations in Tableau for descriptive analysis. Also, I have designed and developed a user-friendly Tableau dashboard, visualized and explored Transit Windsor data.

Moving forward, the team plans to get hands on the new and updated data which can be incorporated in the predictive Machine Learning models for better prediction accuracies, work on new and enhanced visualizations and enhance website’s functionality. The team will also explore additional data sources and incorporate more advanced data analysis techniques to gain further insights into the transit system's performance.

All in all, the Windsor Transit Capstone project made good progress towards achieving its objectives, providing valuable insights into the transit system's performance, and identifying areas for improvement. The team is committed to continuous improvement for the project while working and delivering meaningful and actionable recommendations to the transit authority.

## **Data and Metrics:-**

Data and metrics are essential components of the Windsor Transit Data Analytics project. The project relies on various data sources, including transit schedules, ridership data, and demographic data, to analyze transit service performance and identify areas for improvement. The source of our data is Windsor Transit and City of Windsor Open data. The team has analyzed different transit routes and time periods to identify patterns and trends. The team has also used ridership data to evaluate transit service demand, analyzing data on passenger volumes by route and time of day to identify opportunities for service optimization. Demographic data has been used to understand transit service demand in different areas of the city and to identify potential opportunities for service expansion or route changes. We have taken into account the metrics such as accuracy, precision, recall and F1 score for our models.

In addition to these metrics, the team has used various data visualization techniques to make data more accessible and understandable for stakeholders. The team has used interactive maps, charts, and graphs to display data and analysis, enabling stakeholders to explore the data and insights in a more intuitive and user-friendly way.

In conclusion, data and metrics have played a critical role in the Windsor Transit Data Analytics project, providing a basis for analysis, evaluation, and decision-making. The team has used a range of metrics and visualization techniques to make sense of complex data and communicate insights to stakeholders effectively. The use of data and metrics has been key to the project's success, providing a foundation for evidence-based decision-making and enabling stakeholders to make more informed decisions about transit service delivery and optimization.

## **Changes:-**

Since the initial proposal for the Windsor Transit Data Analytics project, the team has made a few moderate level changes to the project plan and goals. These changes were necessary to ensure the project's success and to address unforeseen challenges that arose during the project's implementation.

One significant change was the inclusion of additional data sources in the analysis. Initially, the project team planned to use only data provided by transit to evaluate transit service performance. However, as the project progressed, the team identified the need to incorporate demographic data to gain a better understanding of transit service demand in different areas of the city and for that we have incorporated the city of Windsor open data. This change was necessary to ensure that the analysis was comprehensive and provided insights into all relevant factors affecting transit service performance.

The team also made some changes to the project timeline to accommodate unforeseen challenges, such as data availability issues and technical difficulties with data processing and analysis tools. These changes were necessary to ensure that the project was completed within the agreed timeframe and that the analysis was robust and reliable.

In summary, the changes made to the project plan and goals were necessary to ensure that the analysis was comprehensive, robust, and reliable, and that it provided stakeholders with the insights they needed to make informed decisions about transit service delivery and optimization. These changes were made in response to unforeseen challenges and to ensure that the project met the needs of stakeholders and was completed within the agreed timeframe.

## **Challenges and Solutions:-**

The Windsor Transit Data Analytics project has encountered several challenges throughout its progress, and the team has taken steps to overcome these challenges, including:

* Technical skills: The team members come from diverse backgrounds, and not all members possess the technical skills required for the project. The team has addressed this challenge by offering group discussions and peer learning to improve technical skills and knowledge.
* Communication: Communication has been a challenge in the project. With team members working from their homes, it has been challenging to ensure effective communication and collaboration. The team has addressed this challenge by setting clear communication protocols, scheduling regular team meetings, and using collaborative tools such as Microsoft Teams.
* Data availability: From the start of the project Data availability was a challenge for the team. Some data sources have been difficult to access, and data collection has been time-consuming. The team has addressed this challenge by reaching out to professors and data providers and developing strategies for obtaining missing data.
* Resource constraints: The project has faced resource constraints, including time, and personnel. The team has addressed this challenge by prioritizing tasks, optimizing workflows, and collaborating effectively to maximize project output.

Overall, the team has taken a proactive and strategic approach to addressing the challenges faced in the project. The efforts of the team have been critical to the project's success, ensuring data accuracy, effective collaboration, and timely project delivery.

## **Results:-**

In evaluating our results for the Windsor Transit Data Analytics project, we have made significant progress towards achieving our objectives. Our primary objective was to evaluate transit service performance and identify opportunities for improvement. To achieve this objective, we collected and analyzed data on transit schedules, ridership, and demography. The analysis revealed several key findings, including areas where transit service demand was high but service levels were low, indicating the need for service improvements. We also identified areas where transit service was underutilized, indicating opportunities to optimize service delivery and reduce costs. These findings are significant and will help inform future transit service planning and optimization efforts in Windsor.

To showcase our final results, we have created a website which has a Tableau Dashboard consisting of various visualizations integrated in it. The dashboard provides users the options to look for different routes in Windsor on which Transit is operating. The user can select the stop using the Stop Name filter and based on the filter, all the buses for that stop are shown. Additionally, users can select the bus and the complete route for the selected bus will be shown and all the expected arrival times which are predicted based on ridership analysis by time will also be shown.

**Snapshot of Dashboard**

Graphical user interface, application

Description automatically generated

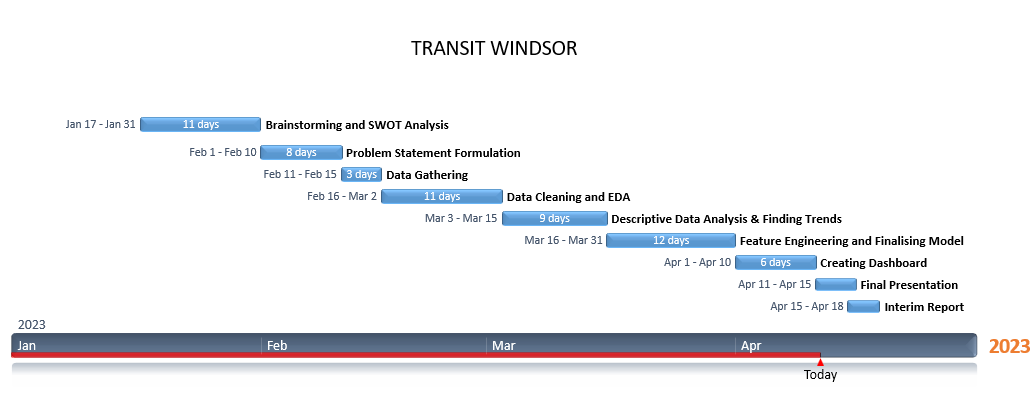
**Tableau Dashboard URL** <https://public.tableau.com/app/profile/abhishek.rana6106/viz/TransitWindsorCapstoneProject/Dashboard1?publish=yes>

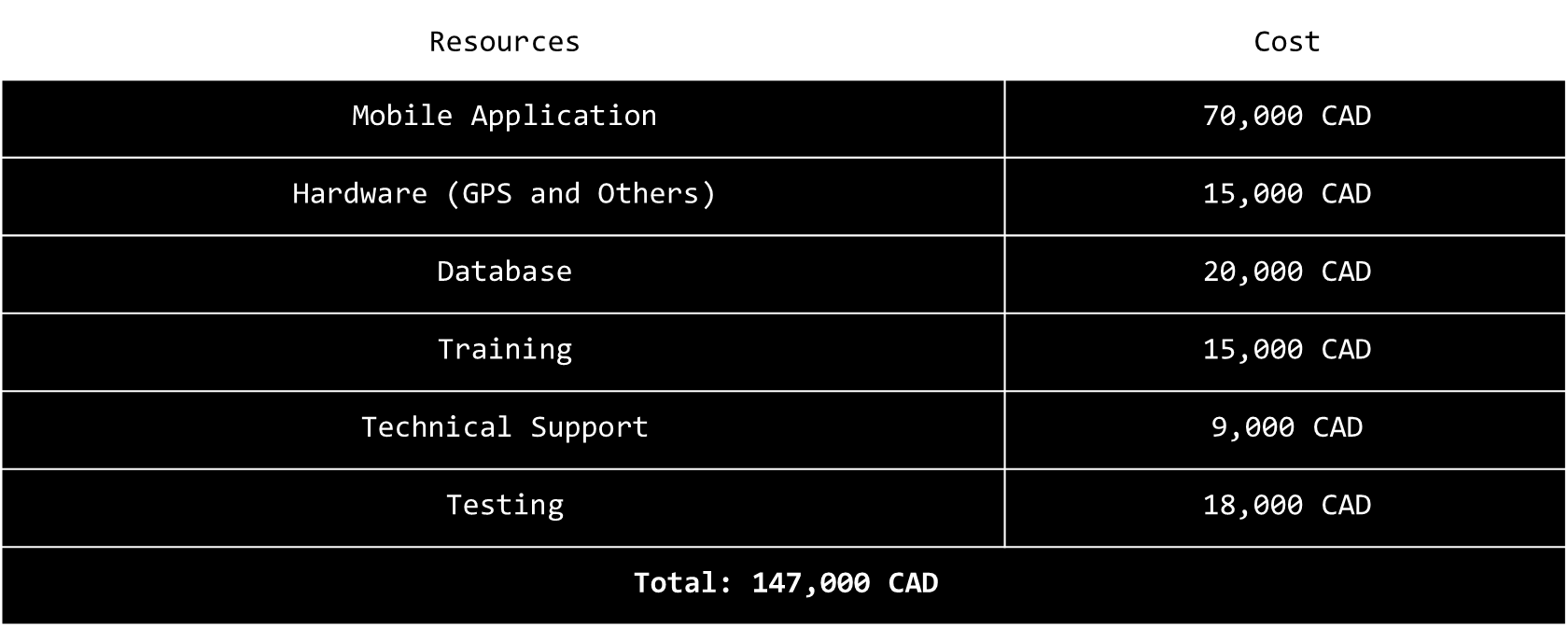
**Github Repository URL** <https://shubham-sihag.github.io/Transit-Windsor/>

We did encounter some challenges that impacted our ability to fully achieve our objectives. One of the main challenges we faced was data availability issues. We encountered difficulty accessing some of the data sources we needed, which delayed our analysis and limited the scope of our evaluation. To address this challenge, we worked with data providers to identify alternative data sources and improve data sharing practices.

Overall, while we have made significant progress towards achieving our objectives, we recognize that there can still be work done to further improve and enhance the functionalities and the overall product. We will continue to work to address any shortcomings and ensure that our analysis is as comprehensive, robust, and reliable as possible. We plan to expand our analysis to include additional data sources and refine our methodology to improve the accuracy of our results.

## **Timeline and Budget:-**





## **Next Steps:-**

For the next phase of the Windsor Transit Data Analytics project, we have identified several key next steps to build upon our progress so far and achieve our objectives:

* Increasing Model accuracy: We plan to refine our machine learning models and improve their accuracy. This will involve training, testing and fine-tuning our models using a new and updated dataset to ensure they are more robust and accurate.
* LSTM model implementation: We plan to implement an LSTM model to analyze time-series data, such as transit ridership over time. This will allow us to identify trends and patterns in transit demand and make more accurate predictions about future demand.
* Adding new features and functionalities: We plan to add new features and functionalities to our dashboard to make it more user-friendly and comprehensive. This may include features such as real-time transit updates, trip planning tools, and personalized recommendations based on user preferences.
* Conducting benchmarking studies: We could conduct benchmarking studies to compare the performance of Windsor's transit system to similar systems in other cities. This would provide a broader perspective on Windsor's transit system and help identify best practices and areas for improvement.
* Conducting cost-benefit analysis: We could conduct a cost-benefit analysis of different transit service improvements to help prioritize investments and identify the most effective strategies for improving transit service.

## **Conclusion:-**

To conclude our Windsor Transit Data Analytics Capstone project, I can say that the team has made significant progress in analyzing transit ridership data, identifying key metrics, and creating visualizations to inform decision-making. We have encountered challenges along the way, including data issues and difficulties in implementing certain analytical techniques and enhancing model accuracies, but we have taken steps to overcome these challenges and continue to make progress.

Moving forward, we plan to focus on implementing machine learning models, such as LSTM models, to improve our predictive capabilities and enhance the accuracy of our forecasts. We also aim to further develop the visualizations and dashboard to improve usability and add more features and functionalities. In addition, we plan to collaborate with stakeholders to gain additional insights and feedback, conduct benchmarking studies, and conduct a cost-benefit analysis to prioritize investments and identify the most effective strategies for improving transit service.

Based on our progress so far, we anticipate completing the next phase of the project within the original timeline and budget. We recommend continued collaboration and communication with stakeholders to ensure alignment with their needs and priorities, as well as ongoing efforts to overcome any challenges that arise. Overall, we are excited about the potential impact of this project on the Windsor Transit system and the community it serves.

Finally, I would like to extend my sincere gratitude to all the stakeholders involved in the Windsor Transit Data Analytics Capstone project for their support and collaboration throughout the project. I appreciate the valuable insights and feedback provided by professors, team members, and other stakeholders, which have helped to inform our analysis and recommendations.

I would also like to recognize the contributions of the project team, for their hard work and dedication to achieving the project's goals. Their expertise and commitment have been critical to the project's success so far. I and my team look forward to continuing our collaboration with stakeholders as we move into the next phase of the project and strive to make further improvements to the Windsor Transit system. Thank you again to everyone involved for their contributions and support.