**SER 517 – Team 29**

**Sprint 2 Report (Group) 16th February 2024**

**Members:**

1. Bindiya Vundavalli
2. Vedasree Bodavula
3. Jayachandra Poluri
4. Sai Venkat Naresh
5. Mohana Narasimha Reddy Attunuru
6. Vamsi Krishna Yadav

**Sprint 2 Planning**

* Sprint 2 planning report – [Team 29 Spring 2 Planning report](https://docs.google.com/document/d/1KOmfltiXXTgDDqH7tB-t9lrIfuOvB3NxJudqgZBXH_Q/edit)
* Sprint 2 Taiga – [Sprint 2 Taiga](https://tree.taiga.io/project/msbodavula-ser-517-capstone/taskboard/sprint-2-11544)
* Sprint 2 meeting minutes – [Sprint 2 meeting minutes](https://docs.google.com/document/d/1TYOaIx9jHUdiUZV0BKqvv98fXiQElEJ007mMZK2kf6M/edit)

**Sprint 2 deliverable**

* Youtube link - [Video link](https://drive.google.com/file/d/1NyXMf3gk5zwEE5RCAmx2e7b5z_g_KU_I/view?usp=drive_link)
* Google drive link - [Team 29 google drive](https://drive.google.com/drive/folders/1Z-6YQLamB4hHLcqKAZTXPgpN7I4TSMA7?usp=drive_link)
* Taiga link - [Team 29 Taiga](https://tree.taiga.io/project/msbodavula-ser-517-capstone/timeline)
* Github link – [Team 29 github](https://github.com/Bindiyaa5/SER517team29)
* Github Usernames :
* Bindiya Vundavalli (Username - Bindiyaa5)
* Vedasree Bodavula (Username - [MsBodavula](https://github.com/MsBodavula))
* Jayachandra Poluri (Username - [JayachandraPoluri](https://github.com/JayachandraPoluri))
* Sai Venkat Naresh (Username - [SaiVenkatNaresh](https://github.com/SaiVenkatNaresh))
* Narasimha Reddy (Username - [narasimhareddy04](https://github.com/narasimhareddy04))
* Vamsi Krishna Yadav (Username - [VamsiKrishnaYadavLoya](https://github.com/VamsiKrishnaYadavLoya))

**Documentation:**

* **What do you think is the value you created in this deliverable?**
  + - The initial step involved cleaning both the datasets - bicycle sharing system (BSS) and scooter sharing system (SSS) within the specified time frame (June 2022 to January 2024). This ensured a unified dataset that encompasses a comprehensive view of urban mobility through different ride-sharing modes.
    - Data pre-processing and cleaning was done to identify any missing values, outliers, or inconsistencies that could skew the analysis. This step was crucial for maintaining the accuracy of the insights drawn from the data.
    - Missing values were addressed either by imputation, where appropriate, or by removing records that could not be reliably corrected.
    - To make a comparative analysis and ensure consistency, data normalization techniques were applied. This was particularly important for demographic information and timestamps, ensuring that all data points were in a consistent format.
    - Any errors identified during the quality assessment were corrected, and the data was cleaned to remove duplicates and irrelevant information.
    - The cleaned dataset was used to analyze trends over time, including seasonal variations and the impact of the pandemic
* **Do you think you worked enough and that you did what was expected of you?**
* The thorough process of data preprocessing and cleaning ensured that the analysis was based on accurate and reliable data. By addressing missing values, correcting errors, and standardizing data formats, the foundation was laid for a robust analysis.
* We have started to analyze the of usage trends, demographic insights, and the impact of external factors like weather and geography on ride-sharing
* By aligning the analysis with the objectives of understanding urban mobility patterns, optimizing fleet distribution, and enhancing user satisfaction, the work directly supported strategic decision-making for the Divvy system
* By considering feedback on the initial project scope, we have incorporated both bicycle and scooter sharing data, along with an examination of external factors and have performed a detailed analysis
* The analysis generated valuable insights into urban mobility, offering actionable recommendations for service improvements.
* Hence, through diligent data handling, initial analysis, and strategic focus, we think we have worked to deliver what was expected of us
* Would you say you met the customers’ expectations? Why, why not?
  + Based on the comprehensive approach taken in data preprocessing, cleaning, and analysis, as well as the strategic focus on understanding and improving urban mobility through the Divvy system, we think that that we have likely met the customer's expectations. Since the timeframe between the day a dataset was finalized and the end of the sprint was very less, we couldn’t perform in-depth analysis. But the inclusion of a broad timeframe, consideration of pandemic effects, and integration of both bicycle and scooter sharing data further ensures a comprehensive understanding of urban mobility trends, meeting the customer's expectations for a basic analysis of the data.

**Sprint Retrospective:**

* What went right?
* We have ensured high data quality through meticulous cleaning and preprocessing, enabling accurate analysis.
* We have employed started to employ statistical analysis to uncover meaningful insights and patterns in urban mobility.
* What went wrong?
* Initial planning did not fully account for the complexity and volume of the datasets.
* Faced unexpected delay in starting with data cleaning as deciding the final dataset got delayed
* How are we going to do better?
* Adopt advanced data management techniques to efficiently handle complex datasets and streamline the preprocessing phase.
* Expand the analytical scope to include additional transportation modes and external factors for a more holistic understanding of urban mobility.
* Implement an iterative review process to continuously refine methodologies, address challenges promptly, and adapt to new insights

**Sprint 2 in Taiga:**

A screenshot of a computer

Description automatically generated

A screenshot of a graph

Description automatically generated