USING DATA TO CONTRIBUTE TO CLEANER TRANSPORTATION METHODS

How has bike transportation changed during the investigated period of time and how can it be improved?

Dinara Akmurzina
Marina Zviagina
Magdalena Krupa
Sharyn McPherson
Vanessa Chan
Lamiaa Bahgat

TABLE OF CONTENTS

1.1 INTRODUCTION:

The aim of this group project is to study and understand how bike transportation has changed during an investigated period and explore ways to improve it. Our main goal is to use data provided by MetroBike to promote cleaner transportation methods. This project explores the trends and patterns in MetroBike's bike stations, bike fleet size, changes in bike usage, factors that influence the number of bike rides and subscriptions. By examining these changes and suggesting improvements, we hope to encourage more people to use bicycles as a sustainable mode of transportation.

1.2 AIMS AND OBJECTIVES:

- Study bike station trends
- Analyse changes in bike usage
- Explore factors that contribute to an increase in bike rides and subscriptions
- Propose improvements for bike transportation

1.3 ROADMAP:

- Data Collection: Gathering data on bike stations, bike rides and weather conditions.
- Data Cleaning: Fully clean the collected data to enable accuracy. Adapting missing values, deleting duplicates, correcting errors and formats.
- Data Exploration: Familiarise ourselves with the cleaned data to gain insight
- Data Analysis: Identify trends and patterns in the changes in number of bike stations, bike usage and subscription types.
- Data Visualisation: Employ visuals, for example maps, graphs and charts to help present findings in a clear and concise manner.
- Compare investigated years: Look and compare data from 2014-2022 to identify any significant changes in bike transportation over those years.
- Identify factors contributing to bike rides and subscriptions: Explore the impact of weather conditions, subscriptions fees, student programs, seasonal holidays and COVID-19.
- Suggestions for improvements: Taken from our analysis, suggest strategies to enhance bike transportation.
- Project Report: Consolidate all analysis, findings, suggestions and conclusions into a thorough report that presents the outcomes of the project.
- Project presentation: Create an engaging presentation that summarises our research findings, recommendations and conclusions. Highlight the key takeaways and actional steps aimed to enhance bike transportation.

2.1 BACKGROUND

With the increasing emphasis and concerns for environmental sustainability and the need to mitigate carbon emissions, cities worldwide have been prompted to explore and seek out alternative modes of transportation. One solution that has gained popularity is bike-sharing programs. The city of Austin has adopted its own bike-sharing infrastructure to provide an eco-friendlier way to commute around their city which helps with the contribution to cleaner air and reduced traffic congestion. The bike-sharing program consists of bike stations placed around the city where individuals can rent bicycles for short-term use. The findings from our analysis could aid in the decision-making process for a wide range of individuals and organisations, including:

- Environmental and sustainability organisations: Learn how bike transportation systems can save energy, reduce emissions, and boost public health. Get practical tips to make them more sustainable and popular.

- Bike advocacy groups: Discover the benefits of bike transportation systems for urban mobility and liveability. Find out how to improve their infrastructure, location, and design for cyclists.
- Transportation authorities and officials: Access valuable data and evidence on the performance and impact of bike transportation systems. Learn how to plan, manage, and integrate them with other transport modes.
- Residents, community members, and the general public: Understand how bike transportation systems can improve one's quality of life, health, and well-being. Learn how to access and use them easily and contribute to their success.
- Travel agencies and tourists: Explore how bike transportation systems enriches the travel experience. Learn how to find and use them in different destinations.
- Students, researchers, and academics: Read a comprehensive and rigorous assessment of bike transportation systems from various perspectives. Learn a novel methodology and framework to evaluate and compare them across cities and regions.

2.2 NATURE OF ANALYSIS

The nature of our analysis entails an examination of data related to bike stations and rides. The report will be written at a moderate technical level, ensuring that the robust data analysis is paired with clear and accessible explanations and visuals, making it suitable for a diverse audience.

3.1 STEPS SPECIFICATIONS

Our team approached each key step of the data analysis process with drive, determination, and a positive attitude. We understood that these qualities were crucial in ensuring a successful and impactful analysis. Here is how we approached each step:

- Framing Question: Through collaborative discussions and brainstorming sessions, our team formulated a concise and measurable question that centred around investigating bike transportation trends and the factors influencing bike stations and rides in Austin. We aimed to explore the dynamics of bike transportation, identify any significant changes over time, and understand the key factors that contribute to variations in bike usage. The formulated question served as a guiding framework for our analysis, providing a clear direction to delve into the intricacies of bike transportation in Austin.

- Data Gathering: To obtain a robust and comprehensive dataset for our analysis, our team engaged in an extensive search process. We explored various online platforms and databases such as Kaggle, government sources and Google Cloud BigQuery etc. Eventually, we identified suitable datasets on Google Cloud BigQuery public dataset that provided the required information for both bike stations and bike rides. Additionally, data was collected from Visual Crossing on weather conditions during the investigated period (2014-2022) to analyse their impact on bike transportation.

- Data processing: To ensure the consistency of all the data that was collected, we implemented preprocessing procedures. These procedures involved performing various tasks which included, removing duplicates, standardising values, sorting data formats, adjusting the columns names, processing datetime columns, checking validity of station_id, excluding incomplete boundary years, addressing outliers and handling missing values by removing missing data as well as imputing data for missing months and incomplete years. We applied a variety of statistical techniques, carefully selecting the most suitable ones based on the nature of the data and research objectives. We immersed ourselves in uncovering patterns, relationships, and correlations within the data. Even in the face of complex analysis tasks, our positive attitude remained steadfast, driving us to maintain our motivation.

- Insights and recommendations: As valuable insights emerged from our analysis, we critically examined the results, taking into account both the quantitative findings and the contextual aspects of the problem at hand. Engaging in constructive discussions and welcoming diverse perspectives, we fostered an environment that facilitated a comprehensive understanding of the derived insights. Our collective effort, driven by a positive attitude, allowed us to extract meaningful interpretations and ensure the relevance and applicability of the obtained insights.

3.2 DATA SOURCES

In this project, we gathered data from different sources to conduct our analysis. Here is an overview of the data sources and how we acquired them:

Bike Data:

To obtain information on bike rides and bike stations in Austin, we accessed the Google Cloud BigQuery public dataset. Two specific datasets were retrieved:

- Bike Rides Data: This dataset contains detailed records of bike rides taken between December 2013 and March 2023, with each row representing an individual journey. We obtained this dataset directly from the Google Cloud BigQuery platform.
- Bike Station Data: This dataset provides comprehensive information on both active and inactive bike stations in Austin. We accessed this dataset from the Google Cloud BigQuery public dataset repository.

We downloaded the bike data on 27th April 2023. It served as a valuable resource for analysis on bike usage patterns and understanding the characteristics of bike stations in Austin.

Weather Data:

Weather data was sourced from Visual Crossing. This dataset encompasses detailed information about weather conditions in Austin, Texas, covering the years 2014 to 2022. We downloaded the weather data on 18th May 2023.

4.1 IMPLEMENTATION AND EXECUTION

Development Approach and Team Member Roles: In our project, we followed a collaborative approach, where each team member selected tasks and took responsibility for the completion of these. We used various platforms such as Trello and Slack to pursue efficiency in communication and coordination. Trello helped organise our task management and gave an easy view of the overall progress. Slack was our central communication channel, using this for real-time discussions and updates. Weekly meetings were held to discuss any issues, ideas or decisions that needed to be made collectively. Having these meetings gave us an added opportunity to bond as a group and work through challenges together.

4.2 TOOLS AND LIBRARIES

The following tools and libraries were used in our project to help facilitate different aspect of data analysis and visualisation:

1. CSV: To read and write our datasets in the CSV format.
2. Pandas: To manipulate and analyse our data efficiently.
3. Numpy: For various numerical operations and computations.
4. Datetime: To perform time-series analysis and extract valuable insights
5. Scipy.stats: To perform statistical analysis
6. ArcGIS: To create and manage geographic data.
7. Seaborn: To create informative and visually appealing charts and graphs.
8. Matplotlib: To generate visualisations for our data and communicate our findings.

9. Folium: To enhance visual representations of geographic information into interactive leaflet maps.
10. Branca.element: To create elements that can be embedded in figures
11. Plotly: To generate interactive visuals
12. Logging: To set severity of logging messages to ignore warnings during the analysis process
13. Requests: To make API requests in fetching data
14. Warnings: To disable the display of warning messages
15. Os: To interact with operating system while saving files and figures

4.3 ACHIEVEMENTS, CHALLENGES AND DECISION TO CHANGE

Throughout the implementation process of our analysis, our team experienced notable accomplishments, encountered challenges, and made crucial decisions to adapt and enhance our approach. Here is an overview of our implementation journey:

- **Achievements:**

Within the data gathering stage our team successfully gathered comprehensive datasets that were relevant to our project question, securing strong groundwork for our analysis. During our data cleaning stage, our team was meticulous and rigorous throughout this procedure to ensure enhanced data quality for our analysis. Analysis techniques such as correlation analysis and time-series analysis were implemented. Despite our team's limited experience with data analysis, we approached the project with enthusiasm and determination. We successfully identified trends in bike usage and compiled actionable recommendations to help drive a positive change within the bike transportation system. These accomplishments showcase our team's dedication to deliver high-quality analysis and project findings

- **Challenges:**

During our implementation, we faced challenges that impacted our data analysis. Ensuring data of high quality was crucial. It was necessary that incomplete, inaccurate, or inconsistent data were dealt with. As one of the datasets consisted of more than 1.8 million rows of data, we found that Data Cleaning was an iterative process. We would move on to the Data Exploration stage and discover problems that required further cleaning.

Missing data necessitated appropriate imputation methods to minimise bias. Data consistency posed challenges due to variations in formats, naming conventions, and units, requiring standardisation efforts. Data transformation, such as converting addresses to geographical coordinates, required additional steps for interpretation. Insufficient data, like the absence of bike type information, limited the analysis's depth and specific insights. Complexity of analysis demanded specialised skills and expertise in statistical models and algorithms.

Avoiding excessive analysis was essential to maintain efficiency and focus on key research questions. Limited resources, including computing power and skilled personnel, posed constraints on the depth and scope of analysis. Time restraints called for efficient time management and prioritisation. Overcoming these challenges involved careful considerations to ensure reliable and meaningful results within the given constraints.

Decision to change:

Our main research question was adapted to better align with our dataset and accomplished analysis. This modification allowed us to narrow the scope of our analysis and have a more comprehensive understanding of the data gathered. We carefully evaluated our data and aligned our analysis techniques accordingly to ensure meaningful results.

5.1 RESULT REPORTING

The analysis of bike usage data revealed several key findings that provide valuable insights into the transportation patterns and factors influencing bike ridership. These findings can inform strategies for improving bike transportation and enhancing the overall biking experience.

- **Bike usage trends:**

As seen in Fig 1, there was an overall increase in bike usage from 2014 to 2022. A significant spike in bike rides was observed in 2018. This can be attributed to the University of Texas' partnership with MetroBike, offering free membership to its students (The University of Texas at Austin, 2018). As a result, the annual rides in 2018 almost doubled that of 2017. The introduction of free memberships for students contributed to a substantial increase in bike rides, demonstrating the success of the bike-share program in attracting riders.

However, in 2019 we see a sharp decrease in the annual rides. This can be due to the pilot study in 2019 allowing electric scooters and electric bikes on certain parkland trails. The influx of electric scooters from companies such as Lime, Scoot, and Bird, resulted in a shift in transportation method, causing the bike ridership to drop. The pilot started in January 2019 and ran through September 2019 (The City of Austin, 2020). Moreover, MetroBike ended its free student membership in 2019, resulting in a significant drop in rides from students (Barton, 2019).

Another shift in ridership happened due to the COVID-19 and ‘stay at home policy’ in 2020. The lifting of the pandemic related measurements in August-October 2020 slowly led to the increase in rides in 2021 (Fig 1) (Villalpando, 2021).

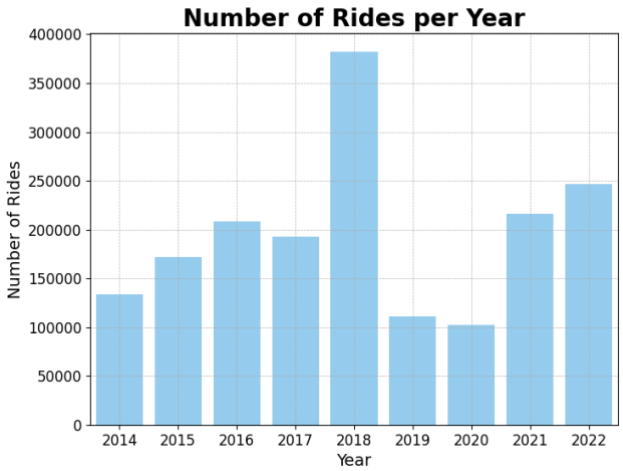

Fig 1. Number of Rides per Year

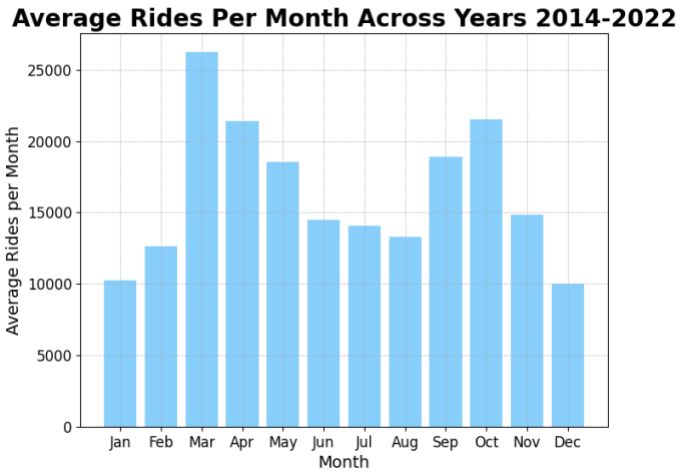

Fig 2. Average Rides per Month across Years 2014-2022

- **Seasonal and Time Patterns:**

The analysis of bike usage revealed significant spikes in rides during specific months, with March being the busiest (Fig 2). There was a correlation between temperature and bike usage, with the mean temperature in March, 19℃, identified as the optimum for rides in Austin. This emphasises the importance of considering weather conditions when promoting ridership and planning infrastructure. Saturdays were found to be the busiest day, followed by Fridays and Sundays. The majority of rides occurred in the afternoon, between 12 pm and 5 pm.

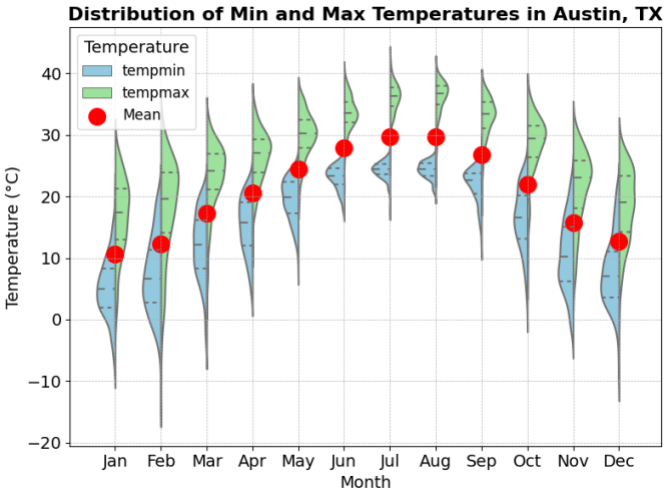

Fig 3. Distribution of Min and Max Temperatures in Austin, TX

- **Bike Fleet and Station Trends:**

The study of bike fleet trends indicated a doubling in size from 2014 to 2022, reflecting an expansion in bike-sharing services. Additionally, there were notable bike station closures observed in the Uptown/North District compared to other districts, suggesting a need for re-evaluation and reallocation of resources (Fig 4).

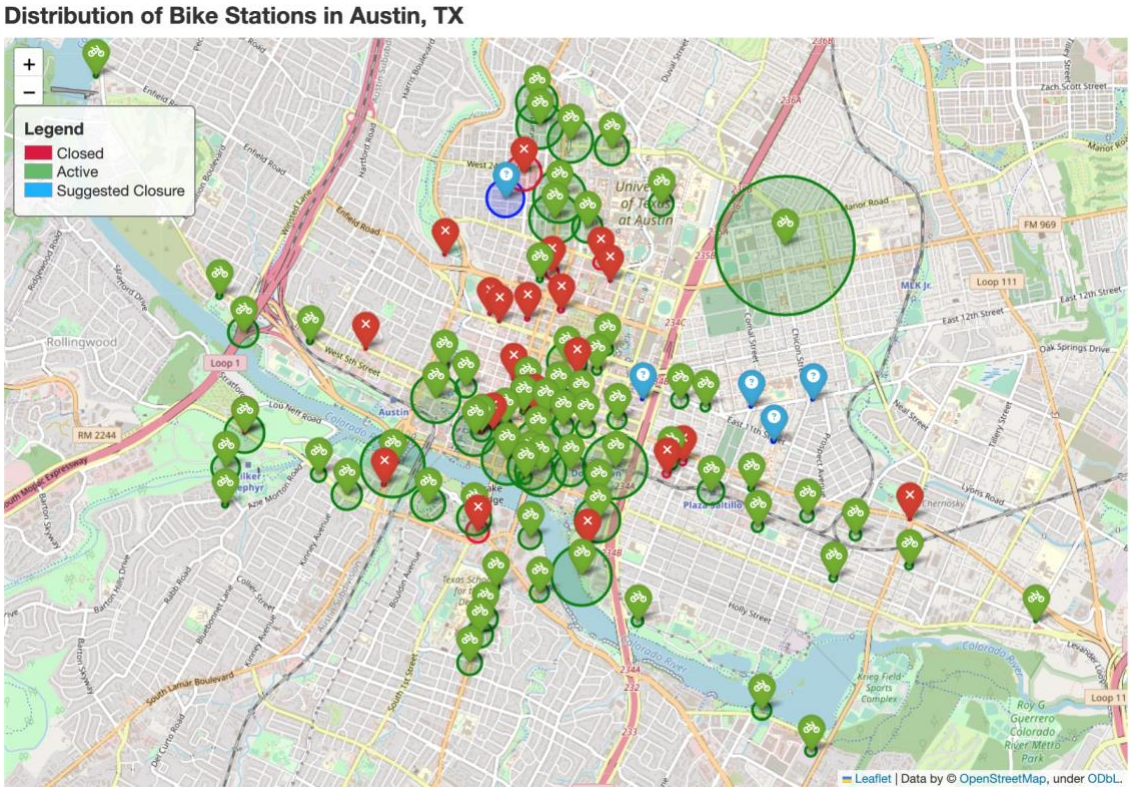

Fig 4. Distribution of Bike Stations in Austin, TX

- **Shift in Ridership due to COVID-19:**

The COVID-19 pandemic had a notable impact on bike ridership patterns. There was a decrease in the number of rides, and the popularity of bike stations shifted from the University to parks and outdoor recreational areas as people sought alternative forms of exercise and leisure activities.

5.2 RECOMMENDATIONS

- Public Initiatives: Increase the frequency of public-level initiatives like Bike to Work Day (Ghisallo Cycling Initiative, 2023) and collaborate with local organisations and government agencies to raise awareness about biking benefits and advocate for infrastructure improvements.
- Business Area Modification: Assess bike usage patterns in different business areas and make necessary modifications to enhance bike accessibility. This may include adding bike lanes, installing bike racks, and providing secure storage facilities to encourage more people to choose biking as a mode of transportation.
- Subscription Programs: Collaborate with businesses to offer subsidised or incentivised bike-sharing subscriptions for employees, making bike transportation more accessible and convenient.
- Station Optimisation: Evaluate bike station popularity and usage and consider closing less popular stations. Allocate resources to support high-demand stations, focusing on bike and dock maintenance, expanding the bike fleet (including electric bikes), and extending services to communities beyond the downtown area.
- Weather Considerations: Account for weather conditions by providing station covers or shelters to protect bikes from rain. Offer free or easily accessible raincoats or protective gear at bike stations to encourage ridership even during inclement weather.
- Seasonal Promotions: Introduce seasonal promotions to boost ridership during off-peak seasons. Offer discounted fares, incentives, or special events to attract riders when demand is lower. Collaborate with local businesses to create joint promotions that highlight the benefits of biking and make it an attractive transportation option.

6. CONCLUSION

In conclusion, our group project on bike transportation analysis has been a collaborative and successful endeavour. Throughout the implementation and execution of the project, we adopted a collaborative approach, assigning specific roles and responsibilities to each team member. This allowed us to leverage our collective skills, knowledge, and expertise, leading to a comprehensive and insightful analysis.

Based on our analysis, we uncovered key findings related to bike transportation, such as the low usage of bikes for work commutes, the overall increase in bike usage over the years, the impact of the MetroBike and University partnership on ridership, and the shift in ridership patterns due to COVID-19. We identified factors contributing to decreases and increases in bike rides and subscriptions, such as weather, COVID-19, location factors, and subscription fees.

Our project provided actionable recommendations to improve bike transportation, including focusing on less busy days and times for promotional activities, offering special discounts during working days, and implementing special offers for companies. These recommendations aim to encourage more bike usage and subscriptions, ultimately promoting cleaner and sustainable transportation options.

As a group, we are proud of our achievements and the valuable insights we have generated through our analysis. Our findings and recommendations can serve as a valuable resource for transportation authorities, policymakers, bike-sharing operators, and other stakeholders. By implementing the suggested improvements, we can contribute to enhancing bike transportation infrastructure and policies.

REFERENCES:

1. The City of Austin (2020) *E-bikes and E-scooters on Austin Trails Pilot Program* [online]. [Accessed 26 May 2023]. Available at: https://www.austintexas.gov/page/e-bikes-and-e-scooters-austin-trails-pilot-program

2. Villalpando, N. (2021) *'I will wake up from this': A year of Austin's heartbreak, resilience with COVID* [online]. [Accessed 26 May 2023]. Available at: https://eu.statesman.com/in-depth/news/healthcare/2021/03/22/austin-covid-one-year-pandemic-texas-timeline/6921773002/

3. Ghisallo Cycling Initiative (2023) *Bike to Work Day ATX 2023* [online]. [Accessed 26 May 2023]. Available at: https://ghisallo.org/blog/bike-to-work-day-atx-2023/#:~:text=National%20Bike%20to%20Work%20Day,40%20fueling%20stations%20around%20Austin.

4. Barton, J. (2019) *Austin B-Cycle ends free student membership after University refuses funding operating costs* [online]. [Accessed 26 May 2023]. Available at: https://thedailytexan.com/2019/01/23/austin-b-cycle-ends-free-student-membership-after-university-refuses-funding-operating/

5. The University of Texas at Austin (2018) *Bike Sharing Launched at UT Austin* [online]. [Accessed 26 May 2023]. Available at: https://sustainability.utexas.edu/news/bike-sharing-launched-ut-austin