**Metro Bike Analysis for City of Austin**

1. Data Source
   1. MetroBike Trips
      1. Sourcing
      2. Collection
      3. Contents
      4. Limitations
   2. MetroBike Kiosk Locations
      1. Sourcing
      2. Collection
      3. Contents
      4. Limitations
2. Questions for analysis

**Reason for analysis:**

I chose this dataset because I didn’t know Austin had a bike share program and I was curious to see where in the city people were using this service.

**MetroBike Trips**

**Sourcing:**

Data provided by the City of Austin and Austin MetroBike sharing program. The City of Austin is a government entity and it works with the non-profit Bike Share of Austin. This is a trusted data source.

**Collection:**

This is administrative data and is collected electronically from the bike kiosks as bikes are checked out and returned. Passes can be purchased at kiosks, through the MetroBike website, and through the BCycle App. If you use a single pass there is no time limit for using the bike but they encourage you to check in a bike at most every 60 minutes to avoid fees. This could impact the maximum trip duration.

**Relevance:**

This will tell me cross streets, date, time, duration, and the type of membership people used to access the bikes and go on a trip.

**Contents:**

MetroBike Trips from 2013 to 2023. Each row is a MetroBike trip (note trips less than 2 minutes are not included in dataset) and includes the following:

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Description** | **Data Type** |
| **Trip ID** | A unique identifier for each MetroBike trip | Qualitative – nominal |
| **Membership type** | Membership type of user | Qualitative – ordinal |
| **Bicycle ID** | The MetroBike's ID that was rented during the trip | Qualitative – nominal |
| **Bike Type** | The type of bicycle used for the trip | Qualitative – nominal |
| **Checkout Datetime** | The date and time that the bike was checked out from the MetroBike kiosk. | Numeric – continuous |
| **Checkout Date** | The date the bike was checked out from the kiosk. | Numeric – continuous |
| **Checkout Time** | The time the bike was checked out for the trip. | Numeric – continuous |
| **Checkout Kiosk ID** | The Kiosk ID of the Kiosk that was used to checkout the bicycle. | Qualitative - nominal |
| **Checkout Kiosk** | The location of the checkout Kiosk based on its nearby cross streets. | Qualitative - nominal |
| **Return Kiosk ID** | The ID of the Kiosk that was used to return the bicycle. | Qualitative – nominal |
| **Return Kiosk** | The location of the Kiosk to return the bicycle. Based on the location's the nearby cross-streets. | Qualitative – nominal |
| **Trip Duration**  **Minutes** | The total number of minutes that the bicycle was checked out. | Numeric – Continuous |
| **Month** | The calendar month on which the trip began | Qualitative – ordinal |
| **Year** | The year in which the trip began | Qualitative – ordinal |

**Limitation/Bias/Ethics:**

This dataset has data from the beginning of the program. The program itself has had stakeholders come and go and there are some conventions that seem to be introduced at different times. This means that there is likely more noise in the data that must be handled. For example, there are kiosks that existed in the early years of the program that are no longer included in the location data so they can’t be used. Complete information is also missing for mobile stations set up for special events. None of the information is identifiable to a certain person although some membership types identify whether someone is a student at the local university. Because data is collected electronically there will likely be data coming from glitches or other electronic mishaps.

**MetroBike Checkout/Return Kiosk Locations**

**Sourcing:**

Data provided by the City of Austin and Austin MetroBike sharing program. The City of Austin is a government entity and it works with the non-profit Bike Share of Austin. This is a trusted data source.

**Collection:**

This administrative data is collected from inventory of MetroBike run by the City of Austin.

**Relevance:**

This will tell me kiosk names, IDs, and GPS coordinates to allow me to conduct spatial analysis of the bike trips data.

**Contents:**

MetroBike kiosk locations from. Each row is a MetroBike kiosk and includes the following:

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Description** | **Data Type** |
| **Checkout/Return Kiosk Name** | kiosk name from trip dataset | Qualitative – nominal |
| **Checkout/Return Name Current** | Current name of kiosk | Qualitative – nominal |
| **Checkout/Return Kiosk ID Current** | Current kiosk ID | Qualitative – nominal – binary |
| **Checkout/Return Kiosk Status** | Whether the kiosk is in use currently (active vs inactive) | Qualitative – nominal |
| **Checkout/Return Kiosk Geographic Location** | GPS coordinates for kiosk | Qualitative – nominal |
| **Checkout/Return Kiosk Dock Count** | Number of docks for bikes at Kiosk | Numeric – discrete |

**Limitation/Bias/Ethics:**

The kiosks in this dataset contain all the names of the current kiosks but not historical data from the kiosk locations. This means that all the trips with kiosks that are no longer in operation can’t be used for analysis. Additionally, kiosk IDs have been updated. Will need to map all current kiosk names and IDs to the bike trips dataset.

**Questions for analysis:**

1. When (time, date, month) do the most trips take place?
   1. Most trips take place between 12 and 5pm.
   2. Monday through Thursday trips are consistent. Trip counts pick up on Friday, peak on Saturday, and Sunday has about the same number of trips as Friday. March and October are peak usage months (this is likely due to two major festivals happening at that time SXSW and ACL respectively). January and December are the lowest usage months. 2018 had the highest usage. 2019 and 2020 had the lowest usage and trips counts appear to be the same for these years so COVID doesn’t appear to have impacted this. 2021 had nearly double the usage as 2020 and 2022 is the second highest usage year behind 2018
      1. Does this follow the typical school schedule?
         1. This does not appear to follow the typical school schedule.
      2. What time (time, date, month) do kiosks at different locations get the most use?
      3. Does time of year impact trip duration?
         1. May, June, and July have the longest average trip times compared to other months. They are a full 2 minutes longer than the next closest month which is December.
2. Which membership type is most popular
   * 1. What is the typical trip duration for each membership type?
        1. Most trips between 2 and 5 minutes are taken by student memberships
        2. Most trips more than 5 but 30 minutes or less are taken by walk ups followed closely by students then annual
        3. Most trips longer than 30 minutes are taken by walk ups
     2. Are there kiosk locations that are more popular for different membership types
3. Where are kiosks located around the city?
   1. Most are located in 78701 but the highest traffic kiosks are located near the university campus
4. How long is the typical trip?
   1. The typical trip is between 20 and 28 minutes.
      1. Does this change based on:
         * 1. time of day

longest average trip times 29-30 minutes happen at 3 and 4 in the morning. Trip times are 17 to 20 minutes from 6-8. There is a 4 minute jump in trip time from 8 to 9AM and a 5 minute jump from 9 to 10AM.

* + - * 1. day of week

average trip time Monday through Thursday is 18-20 minutes

average trip time on Friday is 22 minutes, Saturday is 29 minutes and Sunday is 28 minutes

* + - * 1. month of year

May, June, and July have the longest average trip times compared to other months. They are a full 2 minutes longer than the next closest month which is December.

* + - * 1. checkout kiosk location

Hypotheses:

* 1. If someone has an annual membership they take trips Monday through Friday and their trip times are shorter
  2. If someone has a day pass or pay as you go pass trips are longer and more likely to happen on the weekend
  3. If you increase the number of bikes available at highly desirable kiosk locations you will increase the number of trips taken.