Unit 4: Scooters and E-Bikes in Kansas City: Exposing Usage Patterns and Biases

Student's Name: Avinash Bunga

Course: CIS621 Data Analysis for Business Analytics

Professor: Robert Kao

Date: 09.13.2023

Institution: Park University

Scooters and E-Bikes in Kansas City: Exposing Usage Patterns and Biases

Exploring the Data:

Understanding the Data: Before we dive into the nitty-gritty, let's get to know our dataset a bit better. We've got information about scooter and e-bike trips in Kansas City. To make informed decisions about our city-operated scooter rental system, we need to analyze this data comprehensively.

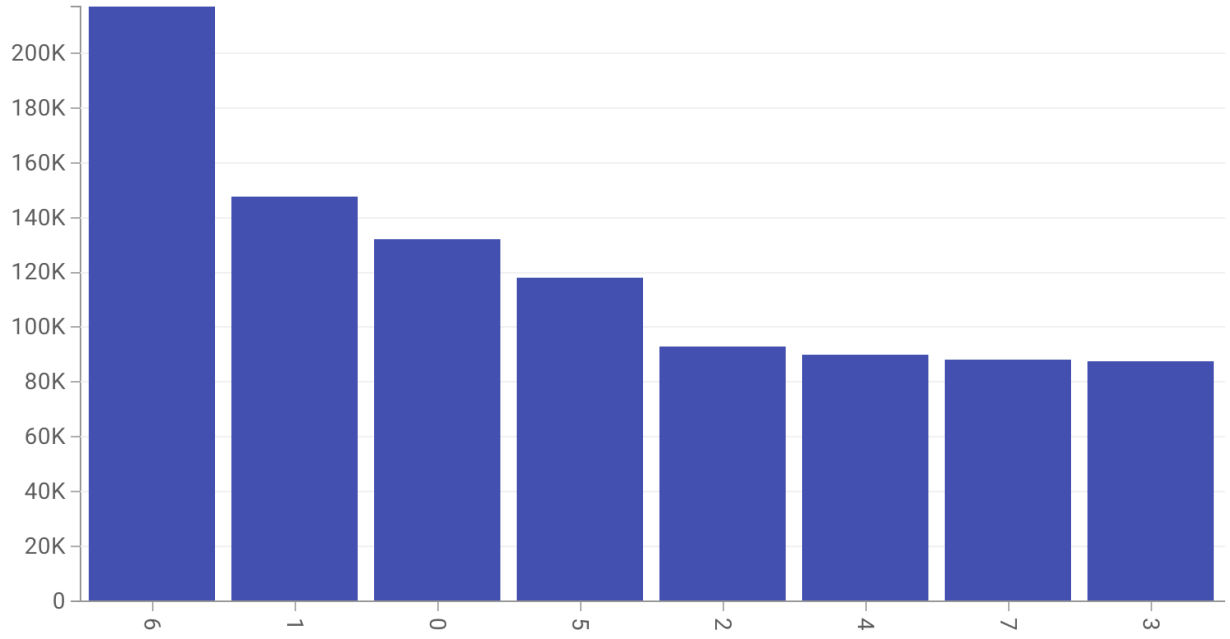
Descriptive Insights:

Total Trips: To kick things off let's start by calculating the total number of trips recorded, We have trip data of 973168 rides.

We can get a sense of how frequently these scooters and e-bikes are used throughout the city from this simple data.

Frequencies of Rides / Days of the Week: Let us comprehend weekly usage patterns in the "Day of Week" column. We'll determine the busiest days so we can manage our service availability accordingly.

[Frequencies of Rides / Days of the Week - Link]



The above Column Chart indicates that day 6 of the week is the busiest, Also considering it to be the weekend.

Day of Week	Count of Rows (Trip/Trips)
6	216,995
1	147,506
0	132,141
5	118,124
2	92,912
4	89,936
7	88,056
3	87,497

Let us calculate the Standard Deviation for the above table

Standard Deviation, σ: 41845.444893792

Count, N: 8

Sum, Σx: 973167

Mean, μ: 121645.875

Variance, σ2: 1751041258.3594

Steps:

σ2 = Σ(xi - μ)2 / N

= (87497 - 121645.875)2 + ... + (216995 - 121645.875)*2 / 8

= 14008330066.875 / 8

= 1751041258.3594

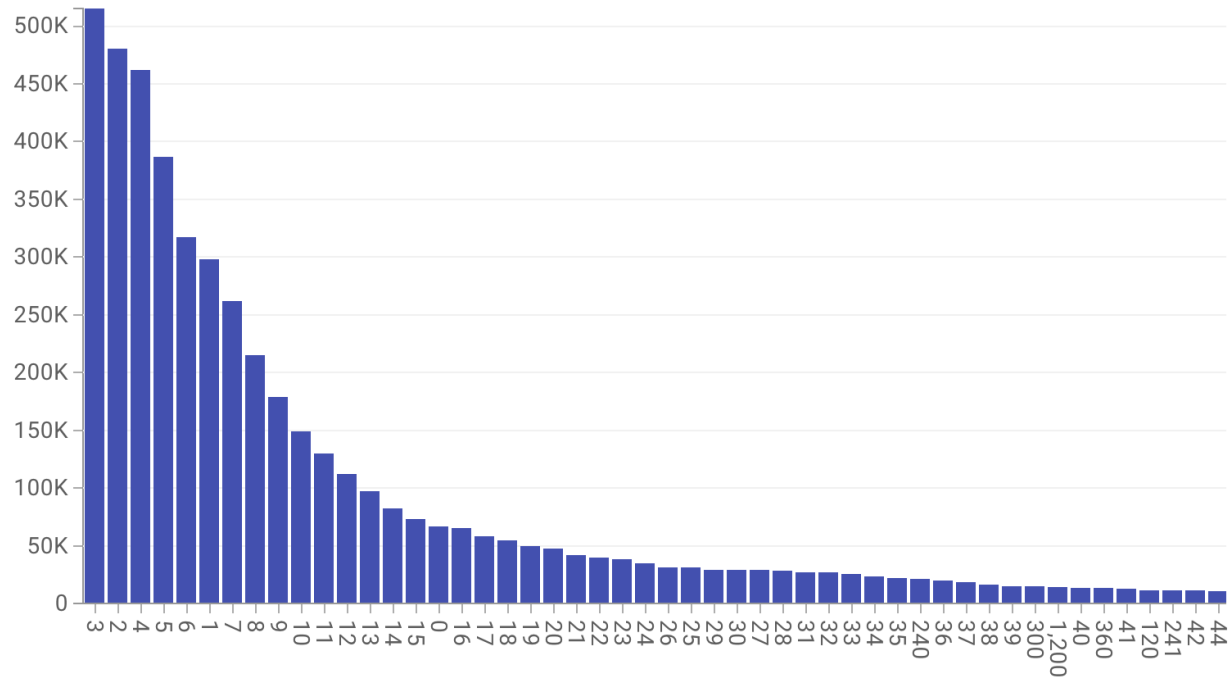
σ = √1751041258.3594

= 41845.444893792

Below are a few more sample charts to understand the data.

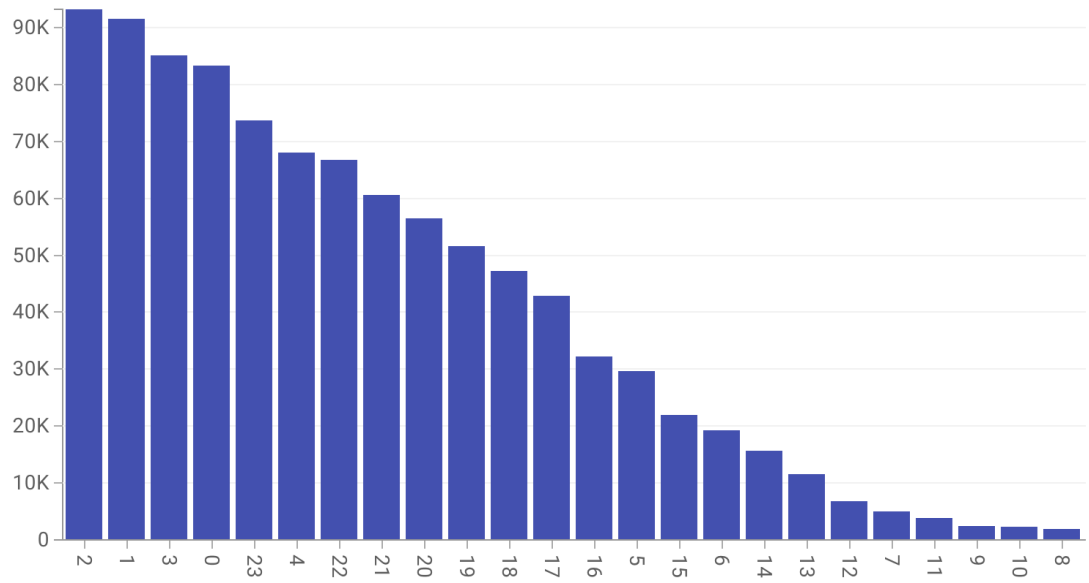
Trip Duration: https://data.kcmo.org/d/hsw6-iq3d

In this case, if we can see the graph is skewed to the right, It clearly shows short trips are preferred more than long trips.



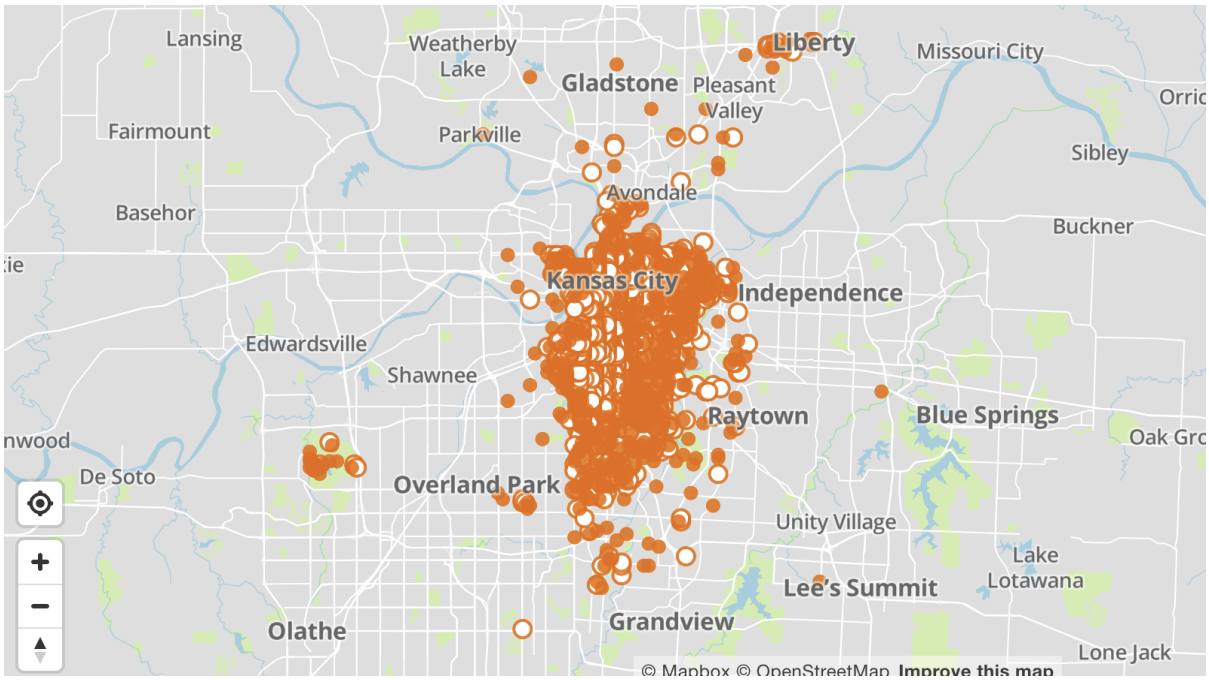
Hourly Trends: https://data.kcmo.org/d/reur-dayf

Moving on, we'll delve into the "hour" column to pinpoint the hours of the day when microtransit services experience the most demand. This knowledge will guide us in ensuring that our services are accessible when people need them the most.



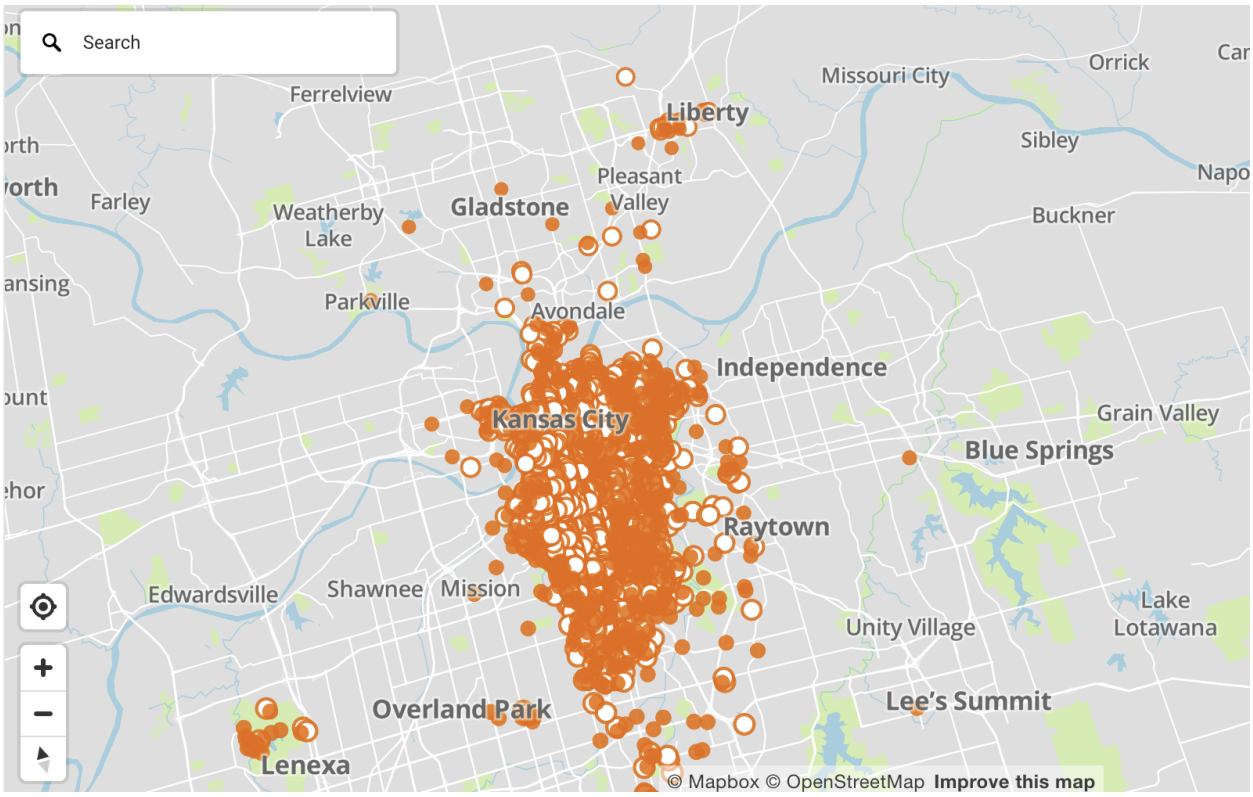
Popular Starting Points: Cities have their favourite spots for scooter adventures. By analyzing the frequencies of trips originating from different locations (Start Location), we'll uncover these hotspots and gain insights into where micro transit journeys typically begin.

https://data.kcmo.org/d/kzaw-9iu4



Top Destinations: Just as there are popular starting points, there are also frequently visited destinations. We'll calculate the frequencies of trips ending at different locations (End Location) to reveal where riders are typically headed.

https://data.kcmo.org/d/7cdq-bgtr



Using the above charts it is clear that the central part of the city has the highest demand and usage.

Addressing Bias in Data Collection

Sampling Bias: Our dataset might lean toward specific neighbourhoods or demographics in terms of Start Location. To ensure a balanced representation, we should aim for diversity in our data collection, covering various locations and user groups.

Time-Period Bias: Seasonal variations can significantly impact ridership. Our data must span all seasons and months to avoid skewing our insights based on a particular time frame.

User Profile Bias: The dataset should ensure it reflects the diverse spectrum of users in our city. If we notice a dominant group, like tourists or a specific age demographic, we'll need to consider the broader population's needs. For instance, having details of the age groups of the users will help the city government understand the actual usage are provide them with better supply according to the demand.

Data Collection Method Diversity: Lastly, we need to broaden our data collection methods beyond app tracking. This will prevent us from overlooking those who don't use apps or have privacy concerns.

conclusion:

By conducting this analysis with meticulous attention to detail and addressing potential biases, we're not just crunching numbers; we're gaining a profound understanding of microtransit usage in Kansas City. This knowledge is our compass, guiding us in making informed decisions about deploying our new city-operated scooter rental system in a way that best serves our community's needs.

Reference

OpenDataKC.(2022).*Microtransit (Scooter and Ebike) Trips* - Link

ennifer Gunner.(2021).*Examples of Standard Deviation and How It's Used*.M.Ed. Education - Link

Pritha Bhandari.(2023).*How to Find Interquartile Range (IQR) | Calculator & Examples* - LINK