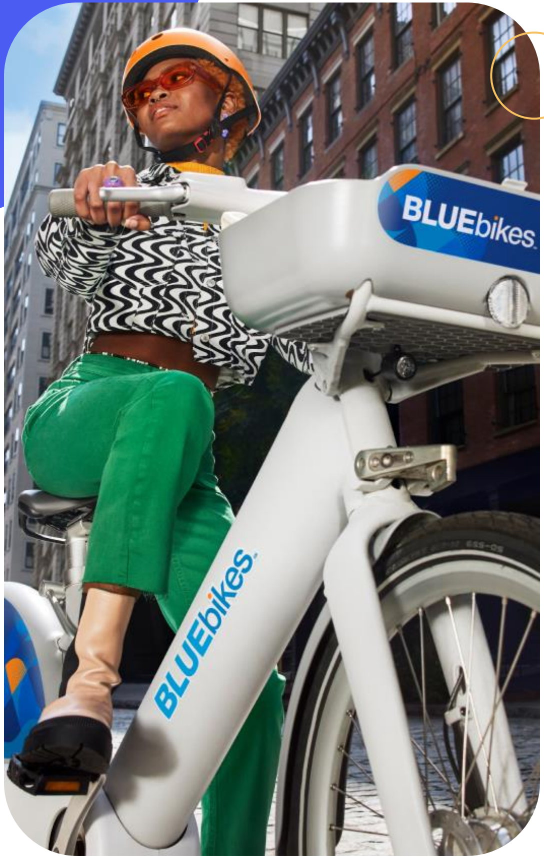


Assessing the Influence of Electric Bicycle Integration on BlueBikes System Usage

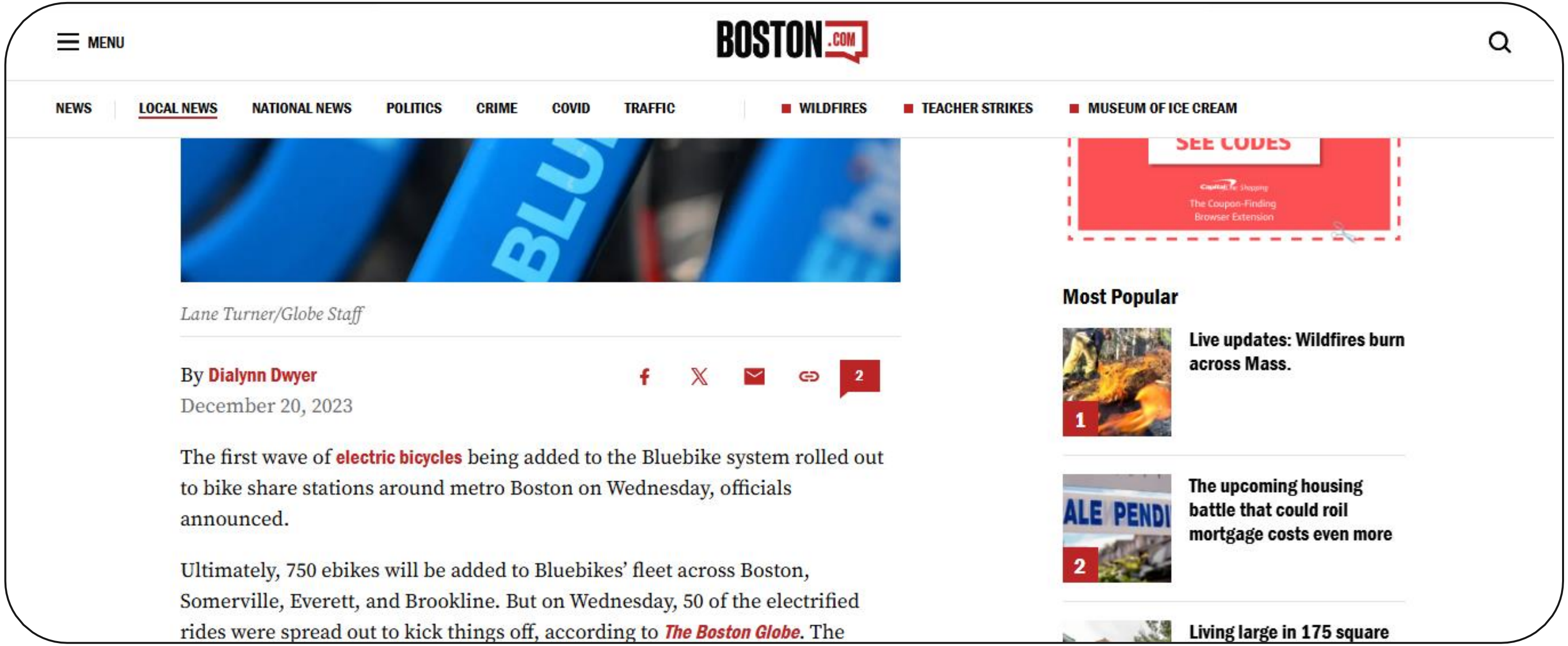
A One-Year Analysis of Ridership Trends and Behavioral Patterns in Greater Boston

NatDave (2024)

Presentation Outline

1. Project Overview & Goals
2. Exploratory Data Analysis
3. Statistical Tests
4. Results and Conclusions
5. Strengths and Weaknesses

Project Overview



MENU

BOSTON.COM

NEWS | LOCAL NEWS | NATIONAL NEWS | POLITICS | CRIME | COVID | TRAFFIC | WILDFIRES | TEACHER STRIKES | MUSEUM OF ICE CREAM

Lane Turner/Globe Staff

By Dialynn Dwyer

December 20, 2023

The first wave of electric bicycles being added to the Bluebike system rolled out to bike share stations around metro Boston on Wednesday, officials announced.

Ultimately, 750 ebikes will be added to Bluebikes' fleet across Boston, Somerville, Everett, and Brookline. But on Wednesday, 50 of the electrified rides were spread out to kick things off, according to *The Boston Globe*. The

Most Popular

1 Live updates: Wildfires burn across Mass.

2 The upcoming housing battle that could roil mortgage costs even more

Living large in 175 square

What has changed since the introduction of e-bikes into the Bluebikes fleet?

Project Goals

1 **Trip Duration Analysis:**
how much further or shorter riders are willing to ride based on bike type

2 **Usage Patterns:**
based on time of day, day of week, user type (casual vs. member), etc.

3 **Station Popularity:**
which stations are most/least frequently used based on bike type

Bluebikes System Data

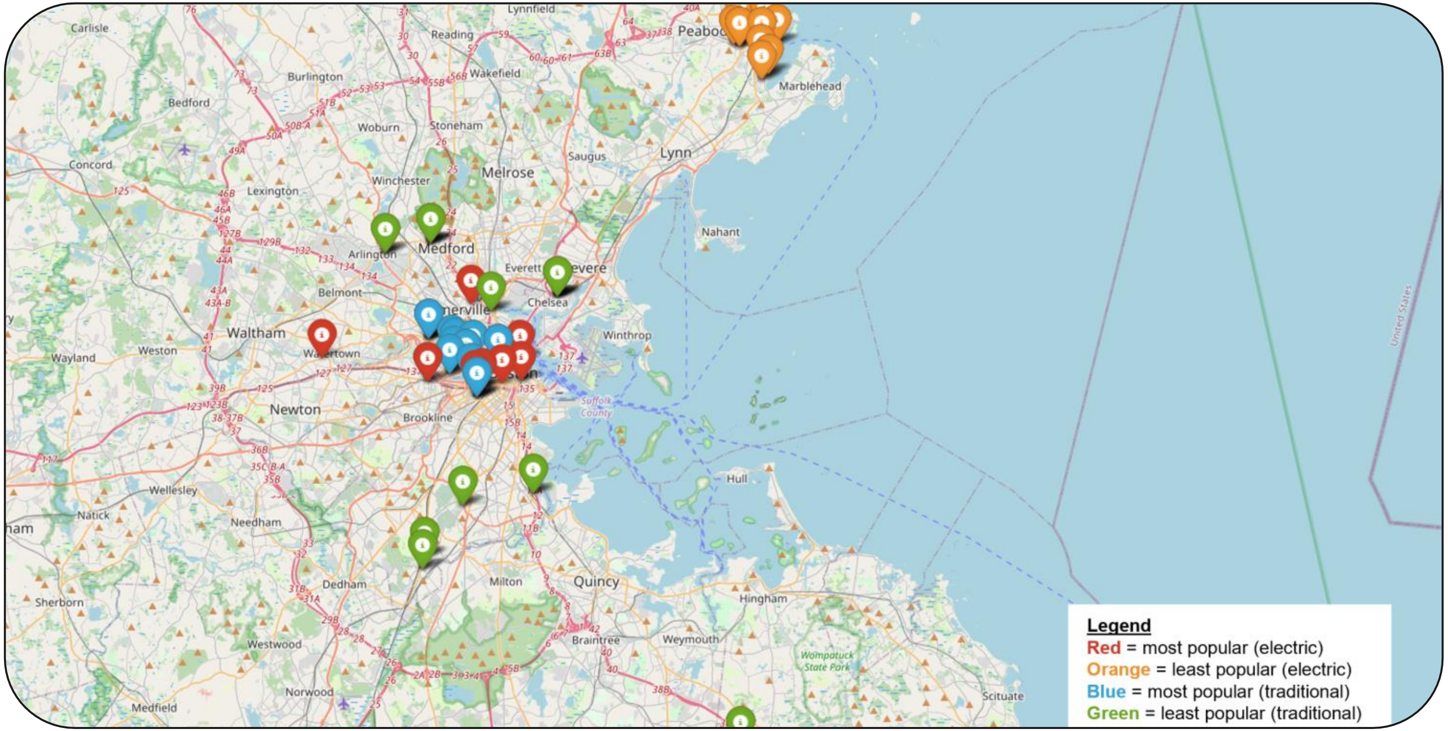
Feature	Description
ride_id	IDs assigned to each ride
rideable_type	whether bike is electric or non-electric
started_at	when the ride begins (time and date)
ended_at	when the ride ends (time and date)
start_station_name	the name of the station where the ride starts
start_station_id	ID assigned to the station where the ride starts
end_station_name	the name of the station where the ride starts
end_station_id	ID assigned to the station where the ride ends
start_lat	latitude coordinate of the station where the ride starts
start_lng	longitude coordinate of the station where the ride starts
end_lat	latitude coordinate of the station where the ride ends

bluebikes.com/system-data

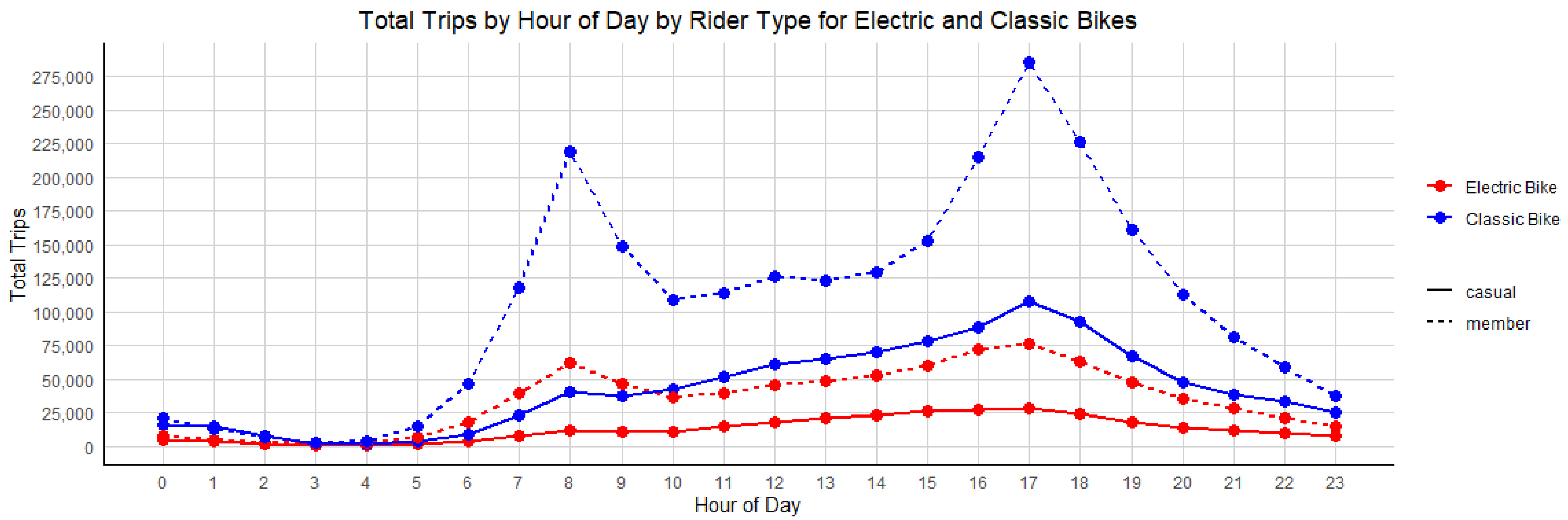
Data Cleaning and Preprocessing

1. Missing values **(< 0.01%)**
2. Temporal feature engineering
3. Station mapping validation
4. Same station trips **(< 5 mins)**
5. Date-time columns

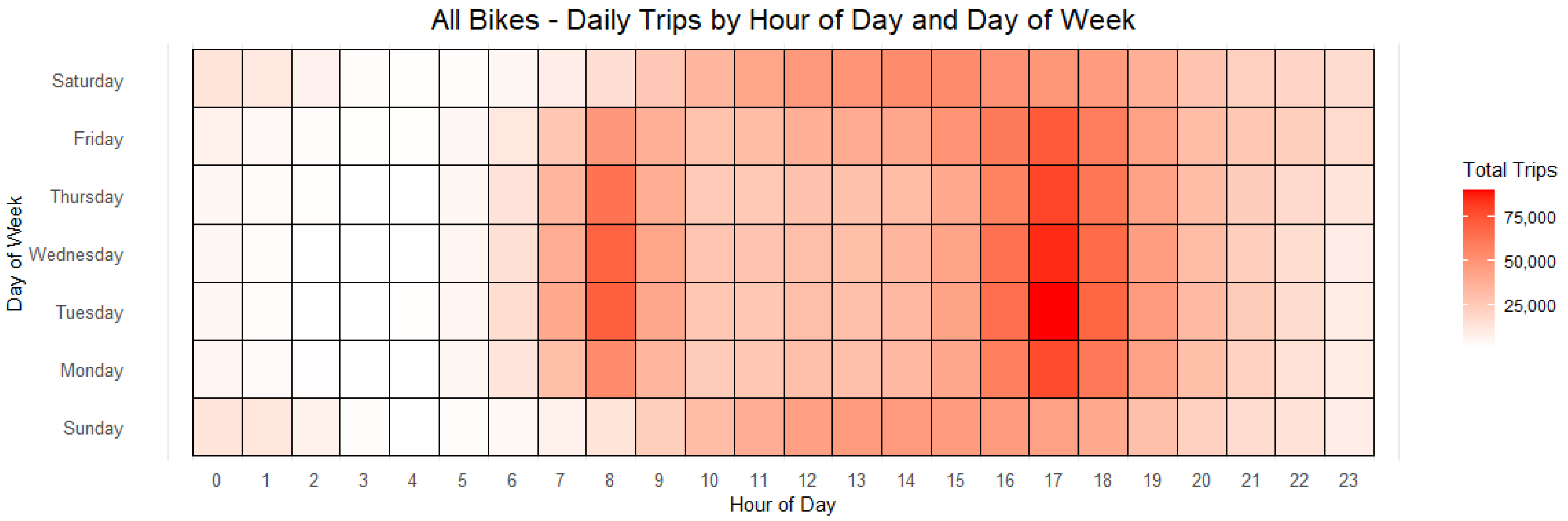
Station Popularity Analysis



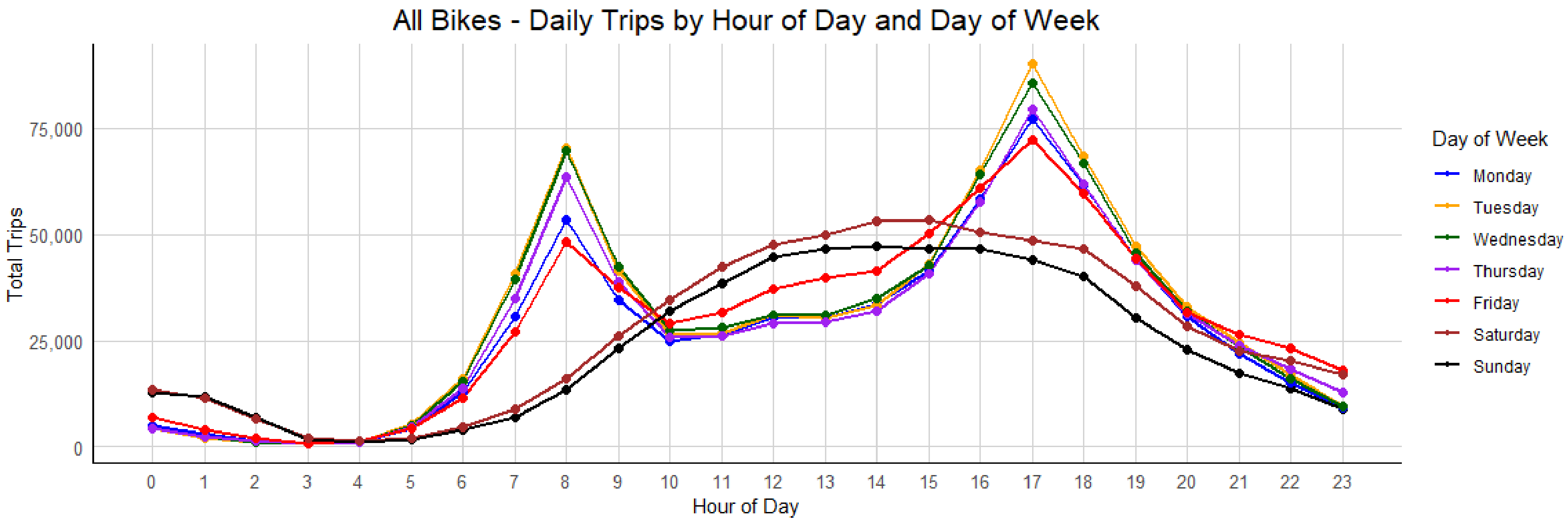
Daily Fluctuations in Bike Trips



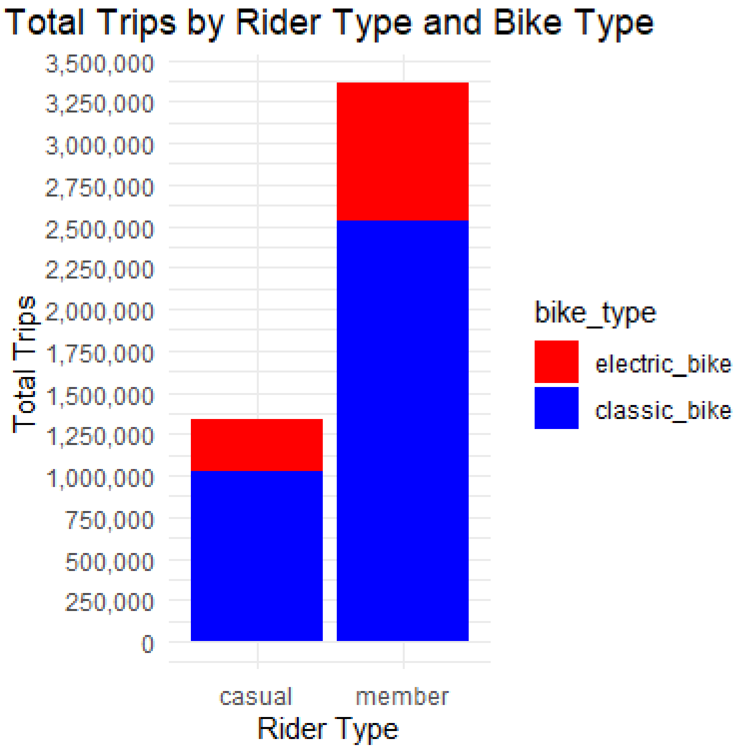
Trips by Hour of Day & Day of Week



Trips by Hour of Day & Day of Week



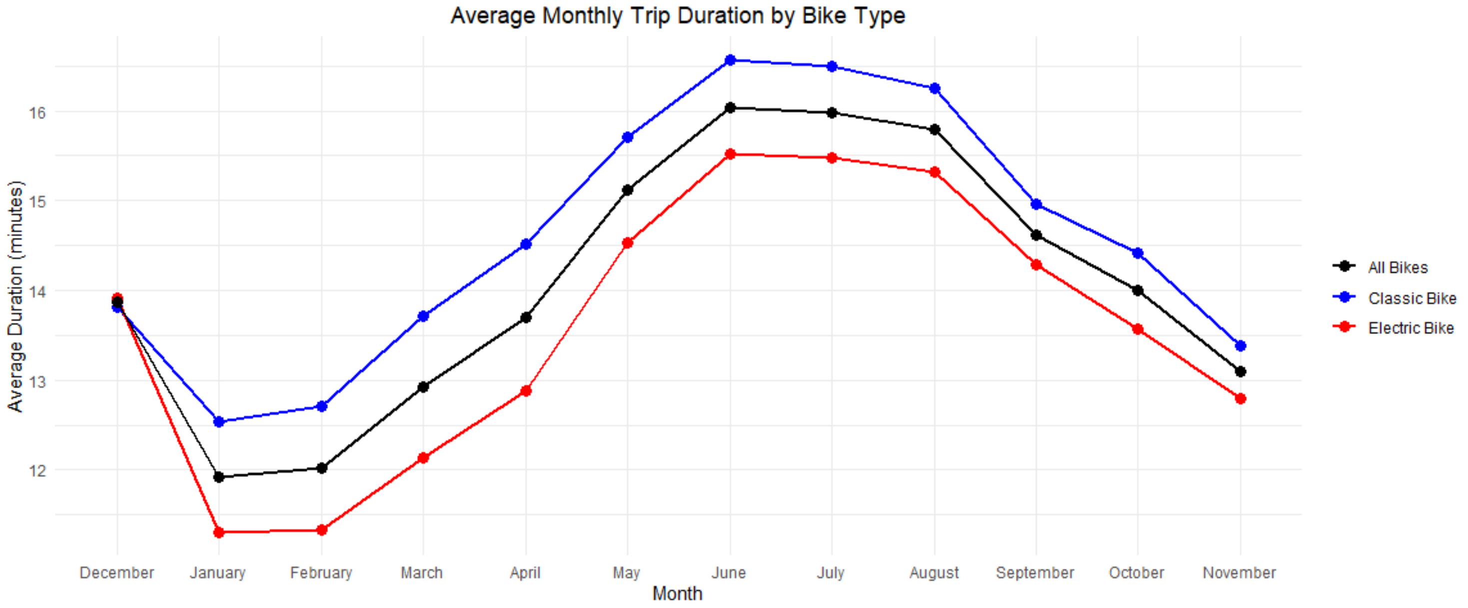
Analysis of Rider Type by Bike Type



Top Station Pairs

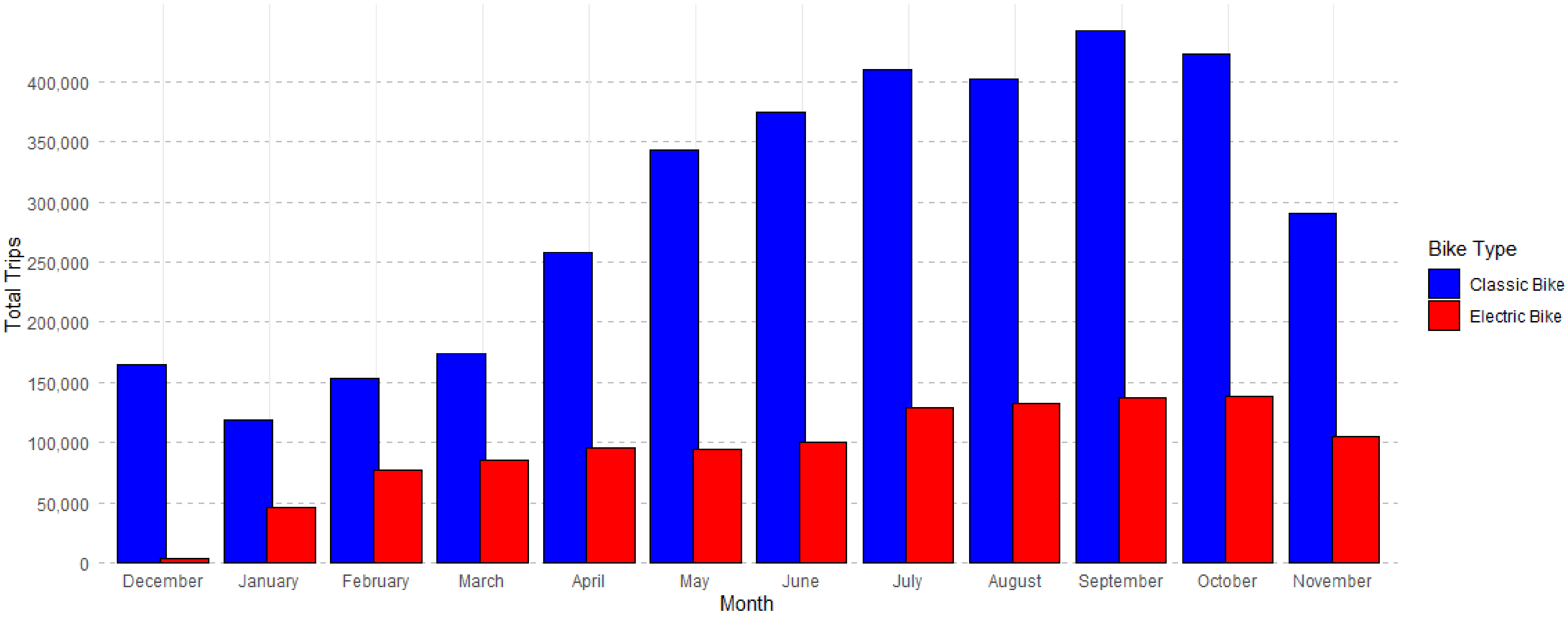
start_station_name <chr>	end_station_name <chr>	Count <int>
MIT at Mass Ave / Amherst St	Beacon St at Massachusetts Ave	5148
MIT at Mass Ave / Amherst St	MIT Vassar St	5132
Beacon St at Massachusetts Ave	MIT at Mass Ave / Amherst St	4585
MIT Vassar St	MIT at Mass Ave / Amherst St	4558
MIT Vassar St	MIT Stata Center at Vassar St / Main St	4542
MIT at Mass Ave / Amherst St	Central Square at Mass Ave / Essex St	4339
MIT Vassar St	Ames St at Main St	4326
Central Square at Mass Ave / Essex St	MIT Pacific St at Purrington St	4279
MIT Stata Center at Vassar St / Main St	MIT Vassar St	3944
MIT Pacific St at Purrington St	MIT Stata Center at Vassar St / Main St	3812

Trip Durations by Bike Type

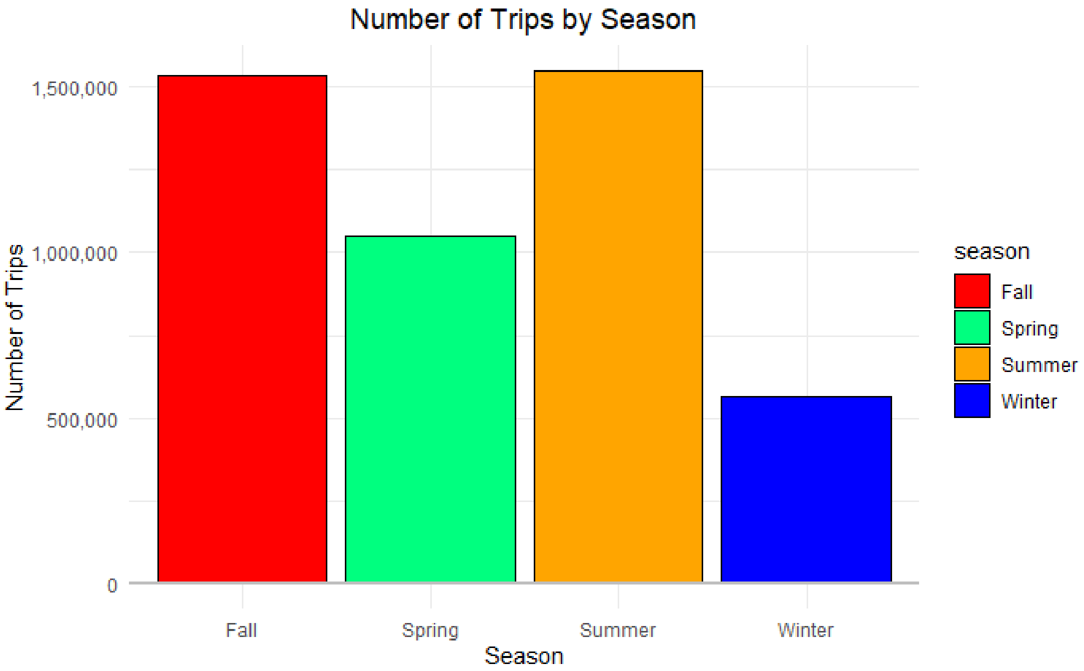


Electric bikes: 13.84 minutes, Classic bikes: 15.00 minutes

Monthly Variations in Ridership



Seasonal Variations in Ridership



Statistical Tests

1 **Hypothesis Test (t-test):**
Compare trip durations between electric and classic bikes

2 **Analysis of Variance:**
Analyze the impact of seasonality on trip duration for e-bikes and classic bikes

3 **Testing of Proportions:**
Do we have more than a quarter of all the trips taking place during the summer?

Hypothesis Testing for Trip Durations (Two-sample t-test)

Null hypothesis: $H_0: \mu_{electric\ bike} = \mu_{classic\ bike}$
Alt. hypothesis: $H_1: \mu_{electric\ bike} \neq \mu_{classic\ bike}$

---

t-statistic: -80.85
p value: approximately zero
alpha: 0.05
95% CI: [-1.23, -1.17] minutes

Conclusion:
Reject the null hypothesis

Seasonality vs. Trip Duration (ANOVA test)

Null hypothesis: $H_0: \mu_{winter} = \mu_{spring} = \mu_{summer} = \mu_{fall}$
Alt. hypothesis: $H_1: At\ least\ the\ mean\ trip\ duration\ for\ one\ season\ is\ different$

F-statistic e-bike:	4,061.92
F-statistic classic-bike:	8,647.42
P value for e-bike:	approximately zero
P value for classic bike:	approximately zero

Conclusion:
Reject the null hypotheses for both bike types

Hypothesis Test for Proportion of Summer Trips

Null hypothesis: proportion of summer trips = ¼ of all trips
Alt. hypothesis: proportion of summer trips > ¼ of all trips

---

t-statistic: 366.1187
p value: approximately zero
alpha: 0.05

Conclusion:
Reject the null hypothesis

Results

1. Avg. trip duration of **e-bikes < traditional bikes**
2. **Seasonality** affects trip durations
3. More than a quarter trips **(25%)** during summer
4. **Peak** usage during **commuting** hours
5. **Members**, more rides; **casual riders**, less

Conclusions

1. Stations near **universities**
2. Investment in **suburban** areas
3. Adapting to **seasonality**
4. **Dynamic** pricing models
5. Expand **e-bike** fleet

Limitations

No demographic and weather data

Low-performance computational resources

Proposed Next Steps & Future Work

1. **External Datasets**
   Weather records, demographic profiles, etc.

2. **Geospatial Analysis**
   Population density, public transit hubs, etc.

3. **Benchmarking**
   Compare results with other bike-sharing systems

Thank you!
Questions?