

MGMT 590: Web Data Analytics Final Group Project

EXECUTIVE SUMMARY CHICAGO DIVVY BIKES



Team 11

tanvi dalvi

vedanti gulalkari

xindi liu

The project idea was influenced from the News Article from May 5, 2022 stating that ‘Divvy became the First U.S. Bikeshare System to Incorporate E-Bike Charging Stations’[1]. Divvy attributes the success of having all-time high subscriber count (~5400 members) to the E-Bike initiative, electrified charging stations being considered a ‘game changer’. Our objective is to analyze the verified factors responsible for increase in membership of Chicago Divvy Bikes. Given memberships lead to increase in revenue, we aim to propose a marketing strategy, based on our factor analysis, to maximize revenue for Divvy Bikes in the 2023.

**Hypothesis:** Did the introduction of new fleet e-bike result in increased rides and memberships for divvy in 2021-22.

[2]90% of Divvy riders are annual subscribers. On rough assumptions, Divvy has negative unit economics - it loses money on each bike due to low ridership and its pricing model. The bulk of its revenue for the city, however, comes from advertising and sponsorships. These factors make our analysis significant.

Primary source of data to analyze the correlation in E-Bike Rides and the membership is the publicly available historical rides data from Divvy Website from years 2020, 2021, and 2022. The reviews data was scraped from the TripAdvisor website, iOS App Store reviews webpage and Google Play Store reviews webpage.

Apart from the missing values and duplicate values processing, the historical rides data required cleaning to rectify the start time and end time of the rides as they were incorrectly captured by the application.The EDA indicated three major insights. Firstly, the apparent rise in usage of electric bike might be misleading. Secondly, classic bike availability at the less crowded stations is very low. The members mostly opt for classic bikes and shorter duration rides, riding mostly on weekdays. The casual users mostly opt for electric bikes and longer durations rides, riding mostly on weekends.

As per the logistic regression model we see that a unit rise in electric bikes will lead to 1.91 times increase in memberships. The random forest model analyzes the effect of ride length, day of week, month of the year, bike type and length of the ride on memberships and the electric bike turns out to be the more important feature as compared to classic bike or docked bikes.

More than 50% reviews on Trip Advisor and iOS have negative sentiment. Besides, major topics of concern as understood from the natural language processing done on the reviews, are ‘application imprecision’ and ‘payment fiascos’.While the models proves that the electric bikes are not responsible for hike in memberships, the reviews analysis hint to the features that can be taken into consideration to maximize Divvy Bikes revenue.

Hence, the reviews imply that the application did not accurately capture the bike start and end, location, and time details. The negative reviews mentioned payment as a major issue faced by majority of the customers. It was also noticed that Divvy did not modernize on classic bikes as much as it did on electric bikes.

Marketing Strategies that Divvy can consider areproviding attractive offers to consumers who are members and ride long distances, integrating with smart wearable applications and collaborating with the fitness centers and other famous Chicago businesses by displaying their ad on Divvy Bikes. Divvy should also offer seasonal memberships as seasonality majorly affects number of rides of the consumers.

**Limitations:** The inference that, investment in application updates will generate more revenue, has not been verified and will require further analysis like A/B Testing. The approach does not state the distribution of funds allocation in Bikes Innovations, Application Improvements and Marketing by analyzing how much each factor will account for revenue generation. Besides, user level data is not made available by Divvy, only ride level data is available. Also, the revenue data needs to be approximated by using the pricing model, not available publicly.