Optimizing PFW campus shuttle ridership service using statistical method and data analysis

Presenter: Ayomide Ajiboye

Department of Mathematical Science, Purdue University Fort Wayne

Introduction

The Purdue University Fort Wayne campus shuttle program started in October 2022. The program officially took off in January 2023. The program was started through the EMSE (Enrollment Management and the Student Experience) Division, in partnership with the SGA (Student Government Association) based on the request to provide free transportation to and from campus for all PFW/IUFW students. Most of these students live at the following locations: Student Housing, Canterbury Green, Saint Joe, Holiday Inn, The Arch. The campus destinations are Walb Student Union, Kettler Hall, and Doermer School of Business. Figure 1 shows the coordinator of the campus shuttle program, Abbey Wang riding one of the shuttle buses.

Figure 1. Abbey Wang in the Campus Shuttle

Figure 2. Campus Shuttle at PFW

The shuttle runs on days that the University is on session. The shuttle does not run on recognized holidays or campus breaks, including summer breaks. Three university fleet buses driven by PFW retirees are scheduled to shuttle between the housing locations about every 15 minutes, And on Fridays, students sign up for a ride to Kroger on North Clinton Street. PFW, like other educational institutions, recognizes the importance of reliable and efficient campus shuttle service. The shuttle system plays a vital role in addressing the transportation needs of a community, facilitating an easy and comfortable movement of individuals across different locations on and off-campus. The campus shuttle service contributes to the overall accessibility and connectivity of the university.

Data Sources and Software

The data for the analysis was compiled by the coordinator of campus shuttle program, Abbey Wang. The data was ridership data which was collected and converted into a xlsx file which was later converted into a csv file. The data reflects 15 minutes time interval and the number of ridership offered within each of those time slot from 8:00am - 7:30pm from Mondays to Thursdays and Fridays from 8:00am - 1:45pm. The data set also indicated the different destination of the ridership.

The PowerBI and R software package were used for visualization and analysis of the data. Figma was used to design the User Interface for the prototype for the campus shuttle ridership app.

Method

The goal of this study is to provide actionable insights and recommendations for the enhancement of the PFW campus shuttle service. This will help to design effective ways to enhance the shuttle efficiency, reduce wait times, and enhance user satisfaction.

EDA techniques and Statistical models are used to uncover patterns, trends, and correlations within the collected data. Here we see visualizations to gain insights into the shuttle ridership. The goal of the statistical model is to use a multinomial logistic regression model to analyze campus shuttle ridership by estimating the likelihood of students using the shuttle during different times of the day: Morning, Afternoon, and Evening.

Exploratory Data Analysis (EDA)

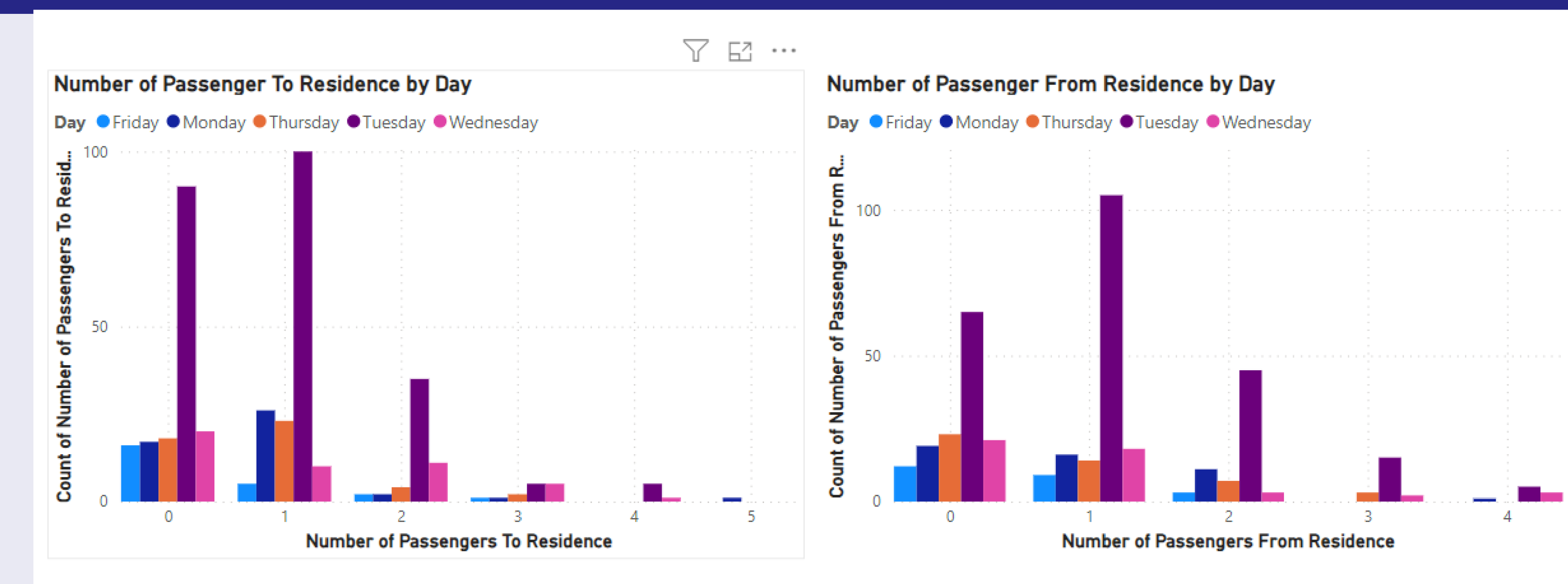

Figure 3. Average Ridership by Day

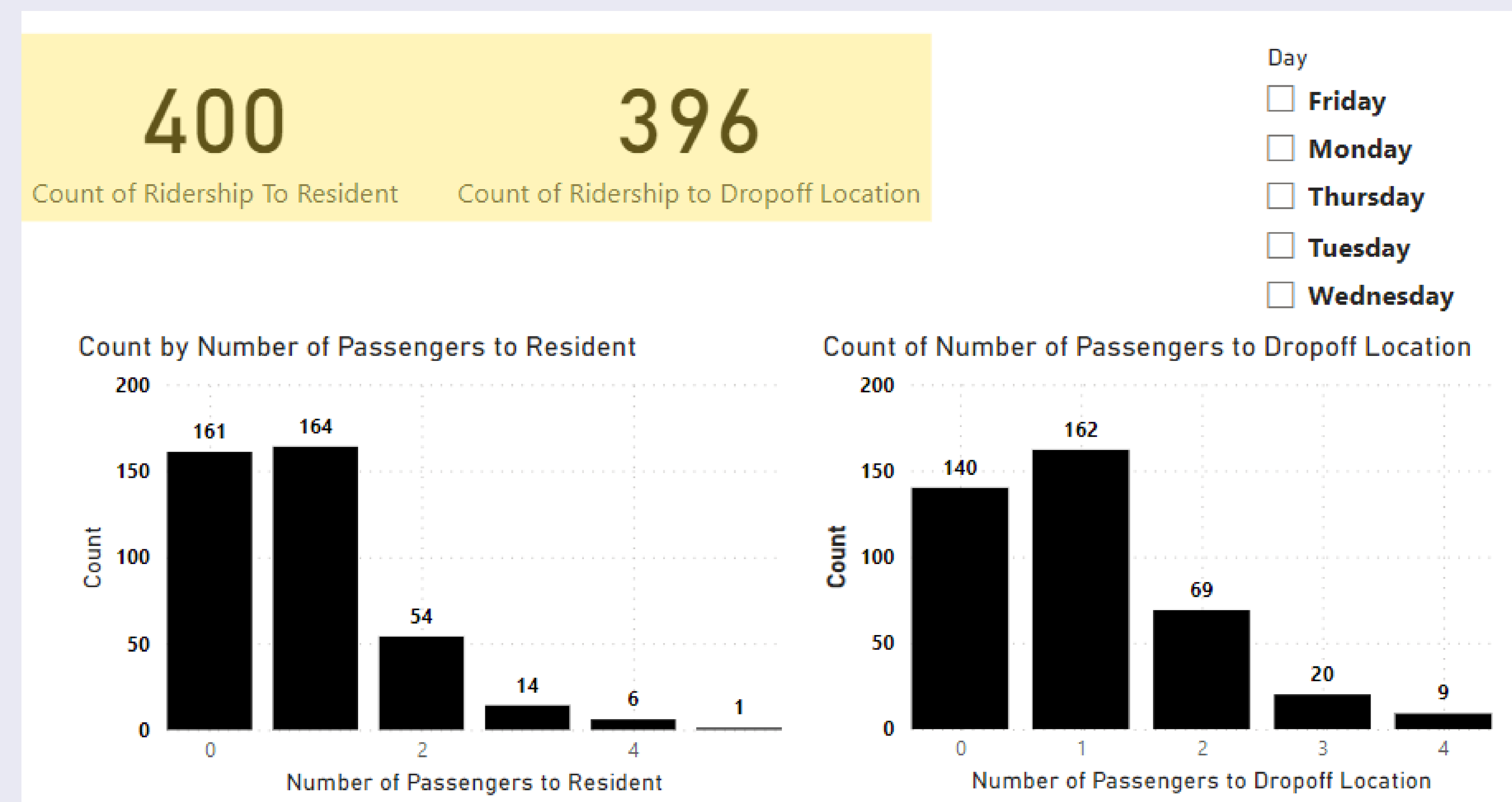

Figure 4. Average Weekly Ridership

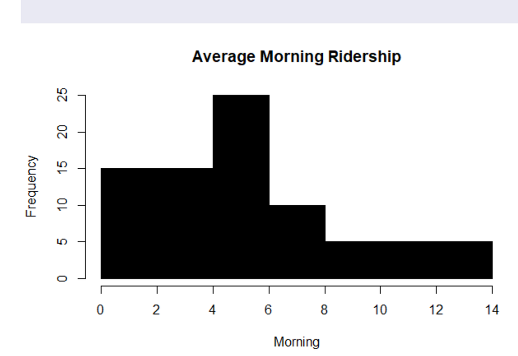

Figure 5. Average Morning Ridership

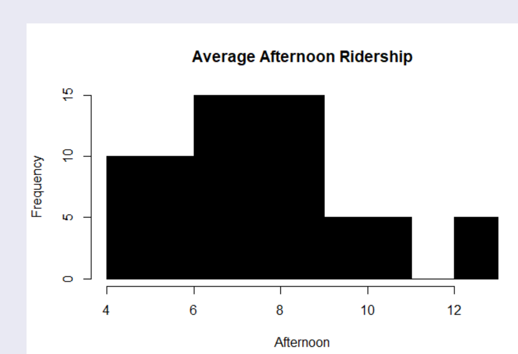

Figure 6. Average Afternoon Ridership

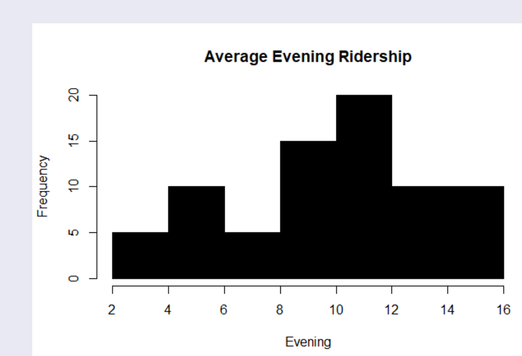

Figure 7. Average Evening Ridership

Figure 8. Ridership Average For Times of the Day

Times of the day	Morning	Afternoon	Evening
Count of Ridership	88	126	153
Percentage	23.98	34.33	41.69

Table 1. Times of the Day Ridership

Statistical Modeling

$$\ln\left(\frac{P(Y_i = k_1)}{P(Y_i = k_2)}\right) = \beta_{10} + \beta_{11}x_{11} + \ldots + \beta_{1n}x_{1n}, i = 1, 2, \ldots, n \quad (1)$$

$$\ln\left(\frac{P(Y_i = k_3)}{P(Y_i = k_2)}\right) = \beta_{20} + \beta_{21}x_{21} + \ldots + \beta_{2n}x_{2n}, i = 1, 2, \ldots, n \quad (2)$$

This is a multinomial logistic regression where Y_i is the response variable, x_{1i} and x_{2i} represent the predictor variables, β_{1i} and β_{2i} represent the parameters (or coefficients). k_1, k_2, and k_3 represent Morning, Afternoon and Evening respectively.

Optimization Strategies

Figure 9. First prototype for the Campus Shuttle App

Figure 10. Second prototype for the Campus Shuttle App

Analysis and Results

Zero ridership means that the rides given by drivers from picked up (residence) location to drop off location and/or vice versa is to no (zero) passengers. Figure 4 tells that there are many more times drivers get zero riderships when they go to residential locations than when they take passengers to University drop off locations. Generally, we see that the (non-zero) riderships on Tuesdays is very much different from that of other days of the week. Based on Figure 3, we see that the highest number of zero riderships occur mostly on Tuesdays. While the number of zero riderships on other week days are about the same. On Tuesdays, everything seems to be higher. Figure 8 shows that the times of the day that has highest zero ridership is Morning. We also see that the times of the day with the highest riderships is the evening time. These are confirmed by our model below.

$$\ln\left(\frac{P(Evening)}{P(Afternoon)}\right) = -1.48 - 14.67x_{11} + 14.90x_{12} + 14.98x_{13} + 14.36x_{14} + 15.66x_{15} \quad (3)$$

$$\ln\left(\frac{P(Morning)}{P(Afternoon)}\right) = 2.08 - 0.14x_{21} - 0.28x_{22} - 0.42x_{23} - 0.40x_{24} + 0.40x_{25} \quad (4)$$

where x_{11} and x_{21} represents the Total ridership variables for Evening and Morning respectively, x_{12} and x_{22} represent Number of Passengers on Monday variables for Evening and Morning respectively, x_{13} and x_{23} represent Number of Passengers on Tuesday variables for Evening and Morning respectively, and so on. The coefficients represent the change in the log-odds of being in a particular category (Morning or Evening) compared to the reference category (Afternoon) for a one-unit change in the corresponding predictor variable, given that all variables remain constant. This model has an accuracy level of 74.5% (with Residual Deviance: 235.7398 and AIC: 259.7398)

ajibaa02@pfw.edu