Chicago Taxi Ridership Case Study: Hook
DS 4002: Data Science Project
April 28, 2025

We've all seen movies or clips where someone steps outside and is faced with a bustling street of cars and pedestrians. The person then lifts their hand as they approach the street and a taxi promptly pulls up to guide the actor into the next scene. The movement of raising one's hand while on a curb or street is immediately recognizable as someone trying to hail a taxi; however, as we shift into a technology based society, this hand motion has changed into a click on an app. Across the US, taxi ridership has decreased as more and more people look into using Uber or Lyft to quickly get a ride. Specifically, Chicago has seen a marked decrease in daily taxi pickups from 125,000 at the end of 2014 to approximately 12,500 by 2022 [1]. This decline can be attributed to Chicago formally allowing rideshare apps within the city [2].

Now imagine you are a fellow resident of Chicago, maybe even referring to yourself as a Chicagoan. With your newly acquired Data Science degree, you are working at the Chicago Department of Business Affairs and Consumer Protection specifically focused on public vehicles, like taxicabs. Due to the documented decline of taxi ridership count, you have been tasked with building and employing two time series predictive models to identify the best one that can provide insight into how to best support taxi companies that have been negatively impacted. These models will be built using data from October 2023 to evaluate which model best predicts hourly ridership counts for October 30th, 2023. This prediction will be used to forecast potential trends that suggest lower counts at specific times during the day, allowing taxi companies to shift their operations to minimize any costs and become more competitive against other rideshare applications.

Your deliverables should include a GitHub repository that contains exploratory plots of the Taxi ridership data from the City of Chicago for 2023, a SARIMA predictive model with evaluation metrics, a BSTS predictive model with evaluation metrics, and a short written reflection. Additionally, you should include any graphs/plots/tables of results from the models that help convey the conclusions reached through your analysis of the City of Chicago Taxi Trips data. A link to a GitHub repository containing additional information can be found here: https://github.com/borayd3/DS4002---CS3

References

[1] Badgujar, S. (2022, July 28). *Chicago public taxi data analysis (over 200 million rows!)*. *Medium.*

[2] City of Chicago. (2014, November 18). *Chicago issues first two transportation network provider (rideshare) licenses to Sidecar and Lyft*. Chicago.gov.