Project Information
Team Members: Arushi Gupta
Purdue Usernames: gupt1077
Path Taken: 2
GitHub Repository Link: https://github.com/ECEDataScience/miniproject-s24-i4rushi.git

Dataset Description
The dataset used in this study is sourced from NYC Open Data and includes daily counts of bicycle traffic on four major bridges in New York City—the Brooklyn Bridge, Manhattan Bridge, Williamsburg Bridge, and Queensboro Bridge—during the year 2016. Additionally, the dataset features weather-related data such as high temperature, low temperature, and precipitation. This dataset serves as a crucial tool for analyzing traffic patterns and supports decision-making related to urban planning and traffic management in the context of cycling.

Analysis Methods
My analysis consists of three primary components:

Firstly, I conducted a correlation analysis to determine which bridges should have sensors installed to best predict overall traffic. This analysis was based on the premise that bridges with traffic patterns most correlated to the total would offer the most accurate predictions of city-wide cycling activity. I expected to find clear indicators showing that certain bridges correlate more strongly with total traffic patterns.

Secondly, I performed a regression analysis using weather data to examine if weather conditions could predict bicycle traffic volumes. This analysis was crucial for urban planning, especially for allocating resources on days predicted to have high bicycle traffic. I hypothesized that warmer temperatures would encourage more bicycle use, while precipitation would deter it.

Lastly, I used a classification approach to predict the day of the week based on the number of bicyclists counted on the bridges. This analysis aimed to uncover any potential patterns in daily bicycle traffic that could assist in resource allocation and enforcement of cycling-related regulations. I anticipated challenges in accurately predicting the day of the week due to potential similarities in daily traffic volumes.

Results and Discussion
The correlation analysis indicated that the Manhattan, Williamsburg, and Queensboro Bridges are the most suited for sensor installation, given their high correlation with the total bicycle traffic. This finding supports my sensor placement strategy by highlighting these bridges as key data points for city-wide traffic estimation.

In the regression analysis, my results confirmed the significant impact of weather on cycling activity. Specifically, higher temperatures were associated with increased bicycle traffic, while precipitation had a strong negative effect, reducing the number of cyclists significantly.

These findings validate my expectation of weather as a major factor influencing cycling behavior.

My attempt to predict the day of the week based on bridge traffic counts resulted in a classification model with only about 23% accuracy. This low accuracy suggests that daily traffic alone may not have distinct enough patterns to determine the day of the week effectively. This outcome highlights the complexity of traffic dynamics and suggests additional variables may be needed to improve prediction accuracy.

Conclusion

My analyses have provided valuable insights into the factors affecting bicycle traffic across NYC's major bridges. By identifying optimal bridges for sensor installation and quantifying the impact of weather, I have laid a foundation for more informed urban planning decisions. However, the challenge in predicting the day of the week from traffic counts alone underscores the need for further research. Future studies might explore additional variables that impact cycling traffic or extend the analysis to other types of traffic and broader geographic areas. This comprehensive approach not only enhances my understanding of urban cycling patterns but also contributes to the broader field of urban traffic management and planning.