Bike Rentals, WELCOMEBIKE

Bike Rental analysis

## presented by

Tonatiuh De Leon Carvajal

Ezekial Curran

Isaac Bojorquez

1. Background

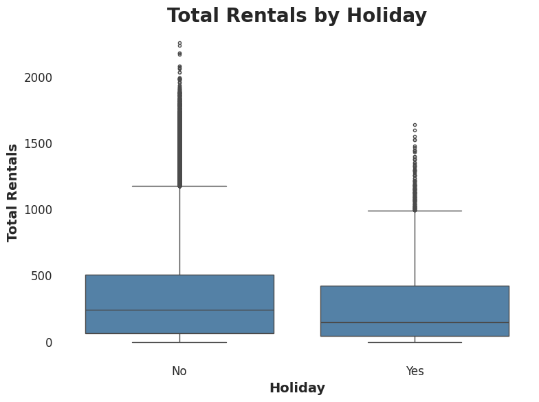
WELCOMEBIKE is new the to the U.S. and has begun operations in the D.C. area. We wanted to see how many bikes we can realistically expect to rent out in any given day and hour. Most foreign companies seem to fail in this market. Using the data provided we used a neural network to predict our expected rentals. We took into account factors like the weather, holidays, and seasons.

1. Methodology

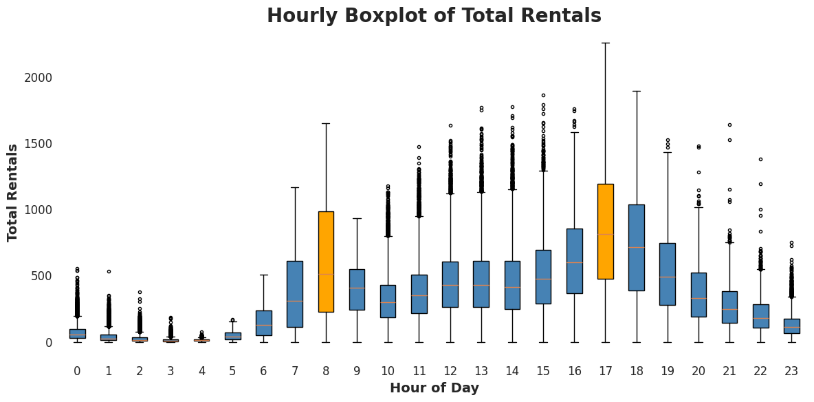
We used the historical bike rental data and enriched it with weather and holiday information. We then trained a small neural network to learn patterns and predict the daily rental counts. The model is small because the relationships found in the dataset are not necessarily overly complex. We used mathematical transformations such as taking the logarithm version of the target variable to stabilize variance. Our model learned better when we transformed all of our data to be in a range from 1 to 0 (mix-man scaling) or in categories.

A graph of a graph showing the temperature of a bike rental

AI-generated content may be incorrect.We saw that the weather is correlated to the amount of rentals per day. Temperature and precipitation were critical drivers in daily bike rentals. Warmer days showed higher rental numbers. While rental went down with chiller days.



Another factor that we saw was that major holidays brought overal rentals lower. This is because our clientele is mostly people that use our bikes to go to work rather than tourists. This is why people are likely to stay at home during holidays and they tend to not use our bikes that much in those days.



This is further being shown in this boxplots. They represent the distribution of rentals per hour of the day for all training data. We can see that the busiest hours for our bikes are 8 a.m. and 5 p.m. which are related to office entry and exit times. Most of our rentals have a seasonality in that sense.

The model shows that we should align bike availability and staffing to fit high demand periods like warm weekends and lower them during poor weather days or off-peak periods. We should also use the forcast data to run promotions to boost rentals on days that are predicted to be slower.

For the test results on the mini holdout the company had for testing purposes, the model performed with a median absolute error for the model was 39 which means that on average our predicted rental count is off by 39 bikes per day. The mean absolute error was 72, this means that if we predict 400 bikes will be rented on a given day we will actually see anywhere between 328 and 472 bikes. Our R^2 metric value was .92 which tells us that we have a decently accurate model and that it accounts for 92% of the day-today fluctuations in bike rentals are taken into account by our predictors.

A graph of a graph

AI-generated content may be incorrect.

Working with neural networks is prone to be subject to variability. When you initialize one model they can differ from others even if they are built the same depending on the computer or software you are working with. This is why most of our models ranged with an R^2 “accuracy” between 87% and 92% which are decently good for our purposes. From all the models that were “trained”, the best one was saved with an R^2 of around .92.

The model may be sensitive to abrupt changes in demand due to unexpected events like special events or extreme weather. Another limitation could be that it does not take into account real-time traffic or special events like marathons, elections, etc. We should constantly be refining and re-training this model given that it is time based and new information affects future performance in that sense.

1. Python Notebooks

Below are Github Gist links to the notebooks we used during this case study:

* [Module 4 Team 7 - Predictions CSV](https://docs.google.com/spreadsheets/d/1yvEZ1SQj3idVLsZJnlvX09SSBP_lbCWAqJJ8i3dnpOg/edit?gid=0#gid=0)
* [Module 4 Team 7 - Coding Notebook](https://colab.research.google.com/drive/14XAT6XtCwETQg0qNgUkkldX3S2bqCKOj#scrollTo=nmzb8jqa2M3k)