

Executive Summary 5: Predicting Taxi Fares for Better Service

How Data Helps Us Estimate Ride Costs

ISSUE / PROBLEM

The goal was to create a reliable way to predict taxi fares in New York City. This helps both the company and riders. Right now, it's hard to know exactly how much a ride will cost before it starts, which can cause problems for everyone.

RESPONSE

We built a model using past taxi ride data. This model looks at things like how far and how long the ride is, the taxi company (vendor), and if it's rush hour. We used a type of math called linear regression to make our predictions. We also made sure to clean up the data, like fixing mistakes and dealing with very high fares, to make our model work better.

IMPACT

Give Riders Better Estimates: People will have a clearer idea of how much their ride will cost.

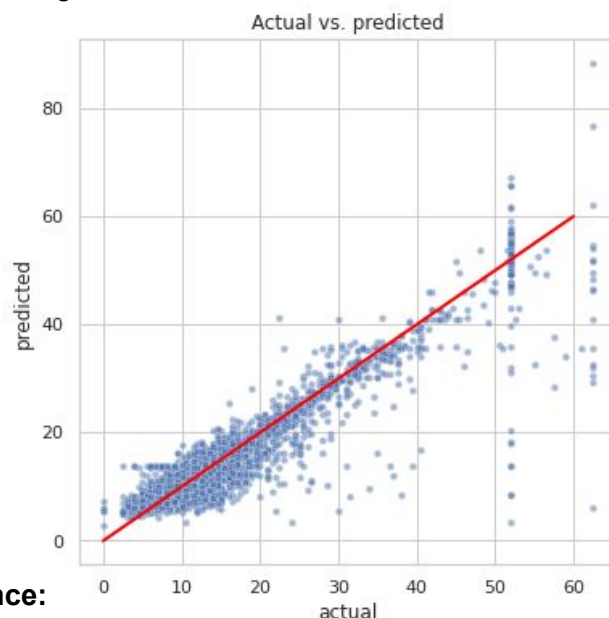
Run Things More Smoothly: We can plan better by knowing how long rides will likely take.

Make Smarter Pricing Choices: We can use the data to set fares that make sense.

Keep Riders Happy: By being more accurate and fair, we can improve customer trust.

Scatter Plot: Predicted vs. Actual Fares

This graph shows how close our predictions are to the actual fares. If the dots form a straight line, it means our model is doing a good job. This is easy to understand and shows the model's accuracy at a glance.



Model Performance:

- R^2 : 0.87 (87% of the fare's changes can be explained by our model)
- MAE: 2.1 (On average, our predictions are off by \$2.10)

Next Steps:

- We should keep collecting data, especially for less common routes.
- We can use this model to create an app that shows riders their fare before they get in the taxi.
- We can look at things like weather and events to see if we can make the model even better.

KEY INSIGHTS

- The distance and time of a ride are the biggest factors in how much it costs.
- The taxi company and rush hour also have some effect on the price.
- Our model is pretty good at predicting fares, with a score of 87% accuracy ($R^2 = 0.87$).
- It is very important to understand how each factor affects the price, and to remember that the numbers have been scaled.
- Understanding the business needs of the company is very important to creating a useful model.
- Looking at the graph of predicted vs actual fares shows how well the model works.