



NYC Taxi Revenue Optimization

Statistical analysis of payment methods and trip characteristics to maximize driver earnings through data-driven insights

Research Objectives

1 Payment Method Impact

Determine if cash vs. credit card payments significantly affect fare amounts through hypothesis testing

2 Trip Characteristic Analysis

Identify which factors - distance, duration, passenger count - most strongly predict higher fares using regression analysis



01

Format Standardization

Convert date formats for consistent analysis

03

Outlier Detection

Apply IQR method to eliminate statistical anomalies

02

Quality Control

Remove null values, duplicates, and negative entries

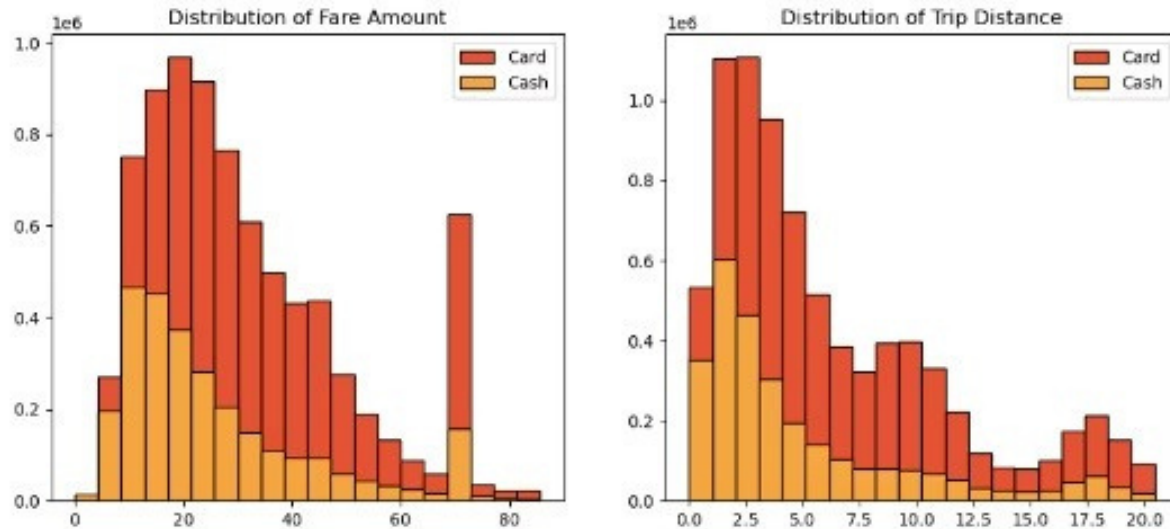
04

Final Dataset

Refined from 38 million to 11 million clean, analyzable records

Payment Method & Trip Characteristics

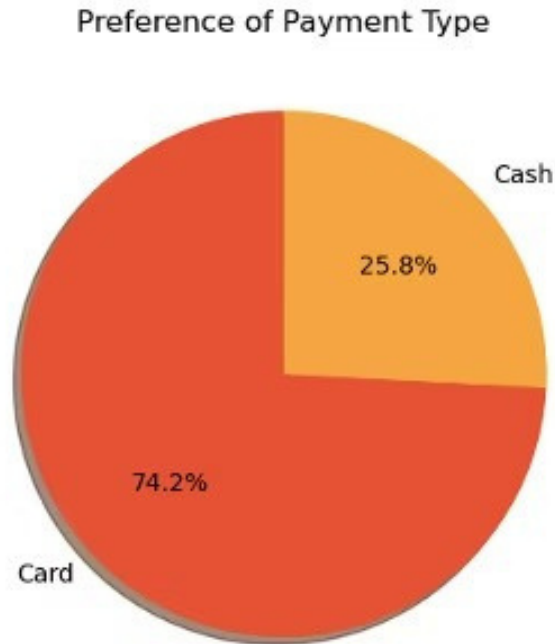
Understanding the distribution of payment types across fare amount and trip distance reveals key patterns in rider behavior and potential revenue opportunities.



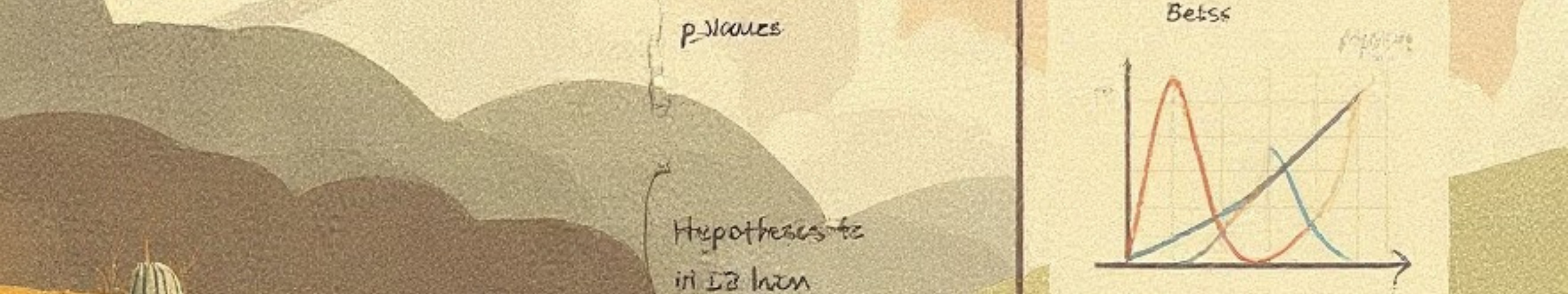
- Credit card transactions generally account for higher average fares and longer trip distances
- Riders opt for cashless payments on more significant journeys.
- Cash payments tend to correspond to shorter, lower-value trips

Payment Method Distribution

Understanding passenger payment preferences is crucial for optimizing driver revenue and improving service efficiency



- **Dominant payment method:** Card payments account for three-quarters of all transactions
- **Key insight:** Passengers overwhelmingly prefer cashless payment options



Hypothesis Testing Framework

Null Hypothesis (H_0)

There is no difference in average fare between customers who pay by card and customers who pay by cash

Alternative Hypothesis (H_1)

There is a difference in average fare between customers who pay by card and customers who pay by cash

Statistical Test Results

450.5

T-Statistic

Extremely high statistical
significance

0.00

P-Value

Strong evidence against
null hypothesis

8M

Card Sample

Robust dataset for analysis

2.8M

Cash Sample

Sufficient data for comparison

\$5.54 More



Card Payments

\$31.34 average fare

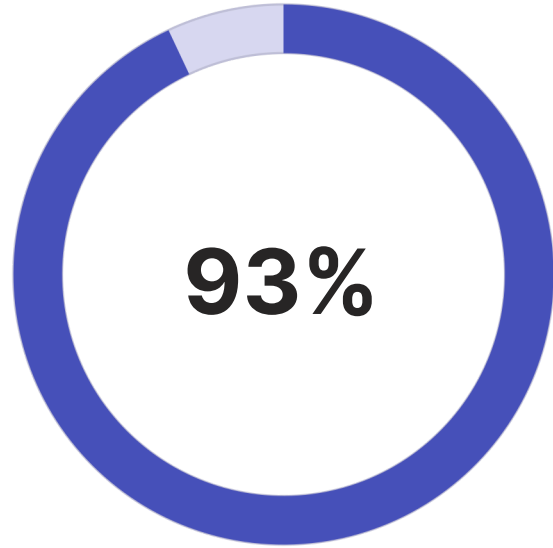


Cash Payments

\$25.80 average fare

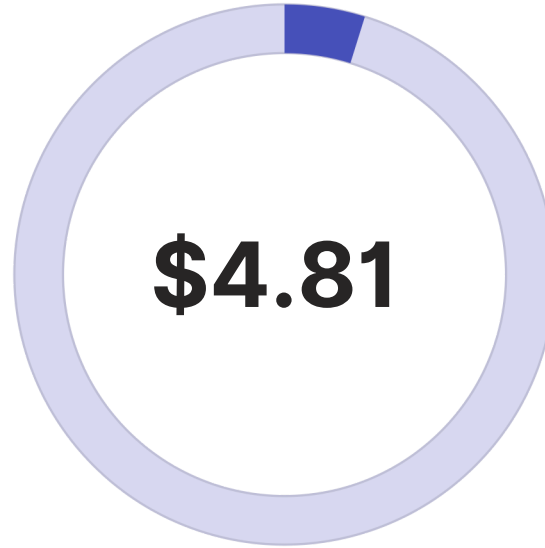
21.5% higher earnings when customers pay by card 4 statistically significant with $p < 0.05$

Regression Model Performance



R-Squared Accuracy

Model explains 93% of fare variance



RMSE

Root MeanSquare Error



MAE

Mean AbsoluteError

Distance, duration, and passenger count capture virtually all pricing factors with exceptional predictive accuracy.

Actionable Insights for Drivers

Payment Strategy

Encourage card payments to increase average fares by 21.5% for a potential \$5.54 per trip revenue boost

Trip Selection

Trips with longer distances and more passengers tend to generate higher fares, as confirmed by 93% prediction accuracy