Executive Summary Taxi Fare Prediction Analysis

# Project Overview

**Objective**: Develop a predictive model for taxi fare estimation

**Dataset**: 2017 Yellow Taxi Trip Data

**Business Goal**: Optimize revenue for taxi drivers through better fare prediction

**Methodology**: Linear regression modeling with feature engineering

# Key Findings

* Strong correlation between fare amounts and trip distance (0.9 correlation coefficient)
* Credit card payments associated with higher average fares.
* Fixed-rate trips identified (particularly for airport routes at $52)
* Model explains 82.9% of fare variability (R² score)

# Statistical Analysis

## Payment Method Impact

Credit card payments show significantly higher average fares than cash.

Hypothesis test results :

* P-value < 0.05
* Statistically significant difference between payment methods
* Credit card payments consistently yield higher tips.

## 

## Model Performance

### Metrics

* R² Score: 0.85 (test data)
* Mean Absolute Error: ~$2.40
* Root Mean Squared Error: ~$4.00
* Cross-validation Score: 0.829

### Key Predictors

* Mean distance
* Rush hour status
* VendorID

# Business Recommendations

## Payment Strategy

* Actively promote credit card payments
* Consider implementing card payment incentives
* Develop digital payment options

## Route Optimization

* Focus on high-value fixed-rate routes
* Optimize airport route coverage
* Consider standardized rates for common routes

## Operational Improvements

* Utilize fare prediction model for trip screening
* Focus on rush hour opportunities
* Implement route planning based on historical data

# Conclusion

This presentation structure provides a clear, business-focused summary of the technical analysis while keeping the technical details in the appendix. The visualizations referenced would need to be generated from the actual notebook data for the final presentation.