MAXIMISING REVENUE FOR DRIVERS

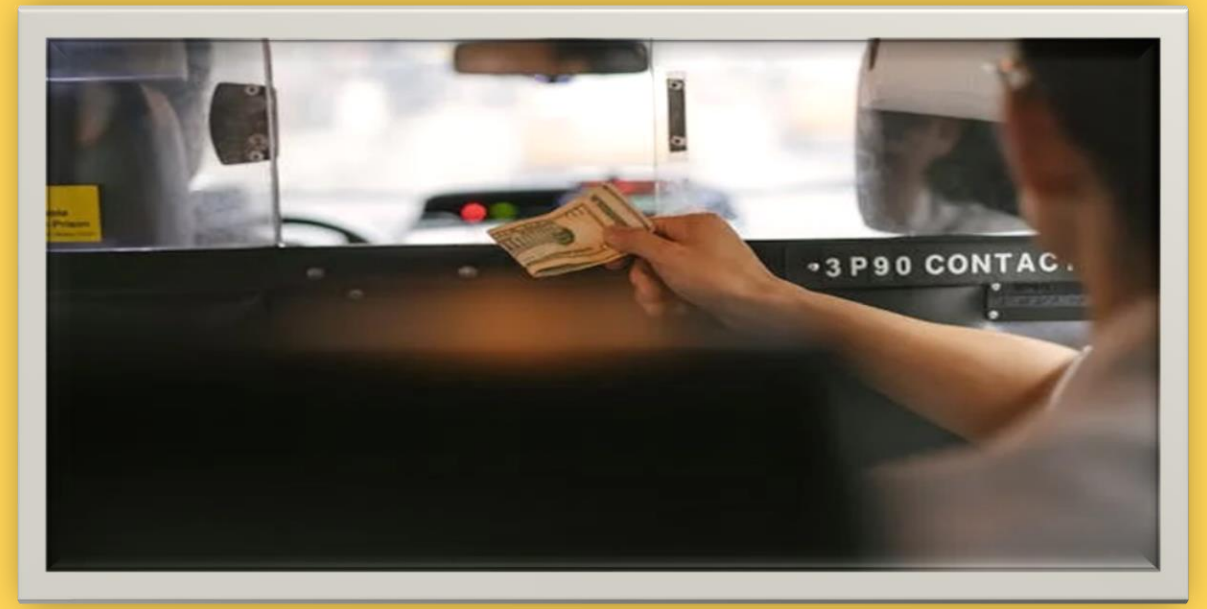
Welcome to the project on NYC Yellow Taxi Trips!

Created by: Shweta Shetty

Agenda

- Problem Statement
- Data Overview
- Methodology
- Analysis and Findings
- Hypothesis Testing
- Regression Analysis
- Recommendations

Project Overview



- The objective of this project is to analyse the impact of payment methods (e.g., card or cash) on taxi fares to provide actionable insights for optimising driver earnings.

- By understanding these dynamics, the goal is to enhance driver satisfaction and overall service success.

Data Overview

For this analysis, I have utilized the extensive NYC Taxi Trips (2017) dataset sourced from Kaggle. Through meticulous data cleaning and strategic feature engineering, I focused specifically on key columns essential for the investigation.

	passenger_count	payment_type	trip_distance	trip_duration_minutes	fare_amount
0	6	1	3.34	14.066667	13.0
1	1	1	1.80	26.000000	16.0
2	1	1	1.00	7.200000	6.5
3	1	1	3.70	31.000000	20.5
4	1	2	4.37	16.716667	16.5
...	...	...	...	...	...
22692	1	1	5.70	18.016667	19.0
22693	1	1	0.89	9.000000	7.5
22694	3	2	0.61	3.266667	4.0
22695	1	1	16.71	41.000000	52.0
22697	1	1	2.36	11.933333	10.5

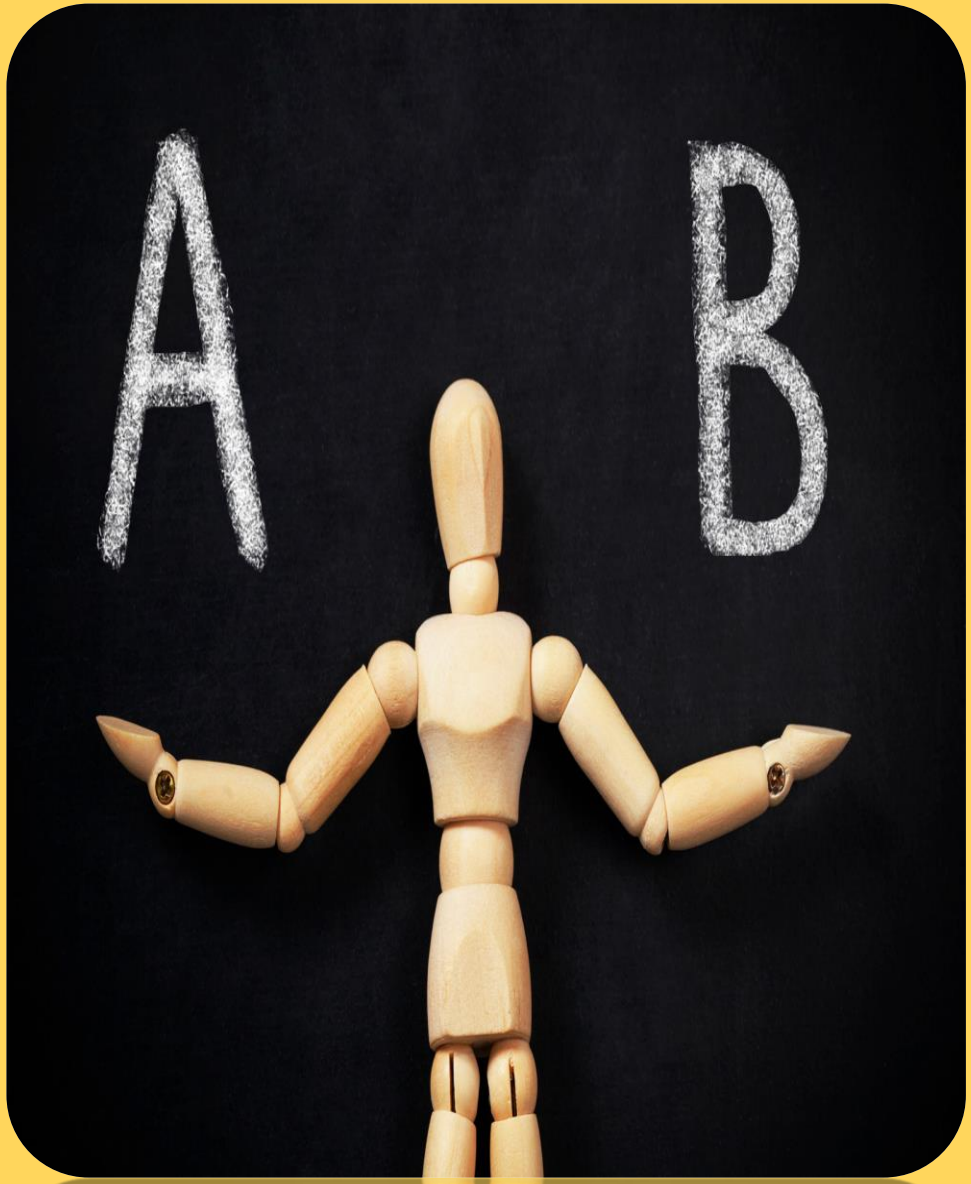
Relevant columns used for this research:

- passenger_count(1-6)
- payment_type(card or cash)
- fare_amount
- trip_distance(miles)
- trip_duration_minutes

Methodology

Step	Description
Descriptive Analysis	Performed statistical analysis to summarise key aspects of the data, focussing on fare amounts and payment types
Hypothesis Testing	Conducted Mann Whitney U test to evaluate the relationship between payment type and the fare amount, testing the hypothesis that different payment methods influence fare amounts
Regression Analysis	Conducted regression analysis to evaluate the relationship between trip distance, trip duration , passenger count , payment type and fare amount.

Hypothesis Testing(A/B Testing)



The project starts with a strong belief that the mode of payment (cash or card) affects the fare amount. To examine this, hypothesis testing (a statistical method) has been conducted.

Null Hypothesis: The **average fare** is the same for both payment methods (i.e., there is no difference in fares between cash and card payments).

Alternate Hypothesis: There is a difference in **average** fares between the two payment methods (i.e., the fares for cash and card payments are different).

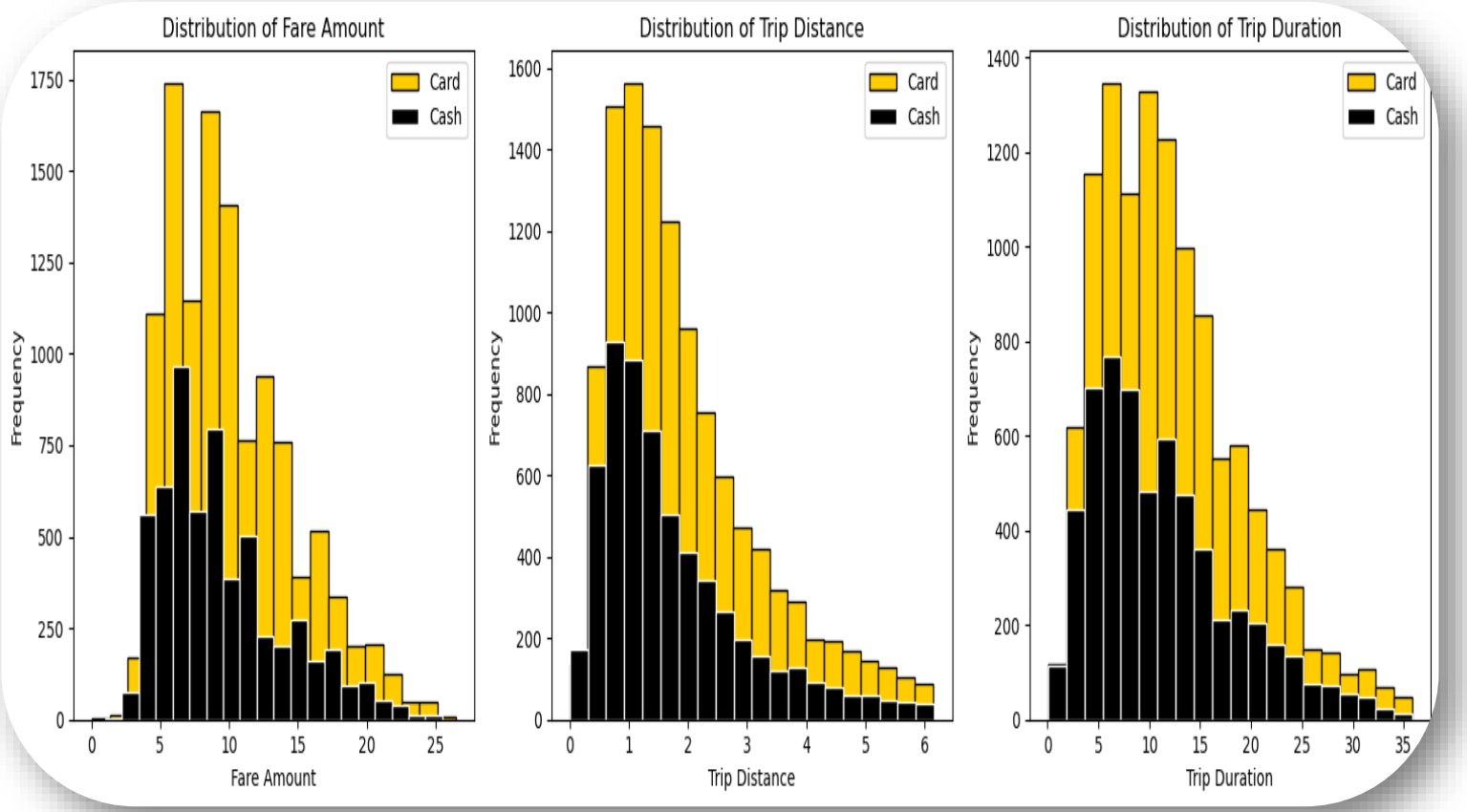
Descriptive Statistics

- To gain a basic understanding of the dataset, descriptive statistics such as mean, median, and standard deviation were calculated.

- Additionally, charts were created to visualize the distribution of fare amounts and trip distances based on the mode of payment.

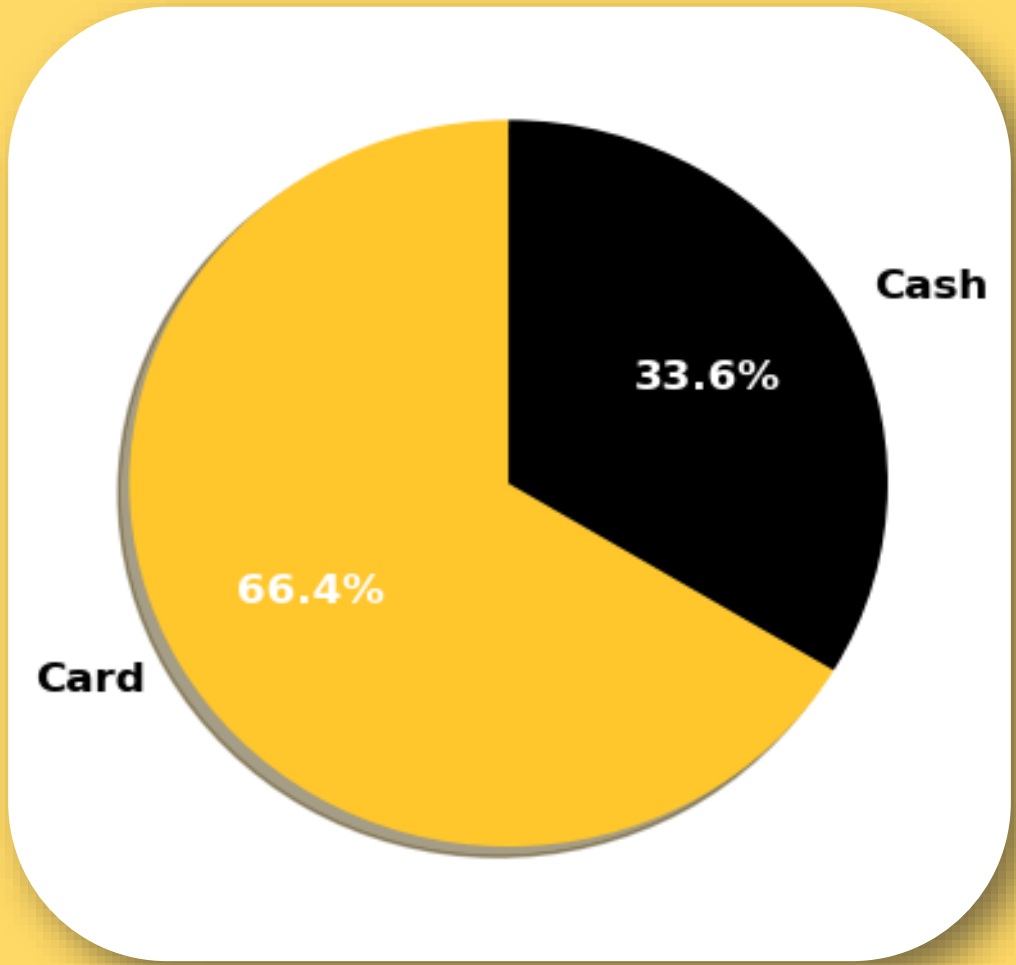
Trip Journey Insights

- Customers who pay with a card generally show slightly higher **average** trip distances, trip duration and fare amounts compared to those who pay with cash.
- This suggests a preference for using cards for transactions involving higher fare amounts and longer trip distances.



	Payment Type	Mean	Standard Deviation
Fare Amount	Card	10.13	4.49
	Cash	9.47	4.47
Trip Distance(Miles)	Card	1.95	1.27
	Cash	1.74	1.24
Trip Duration(Minutes)	Card	12.12	6.92
	Cash	11.27	6.95

Preference Of Payment Types

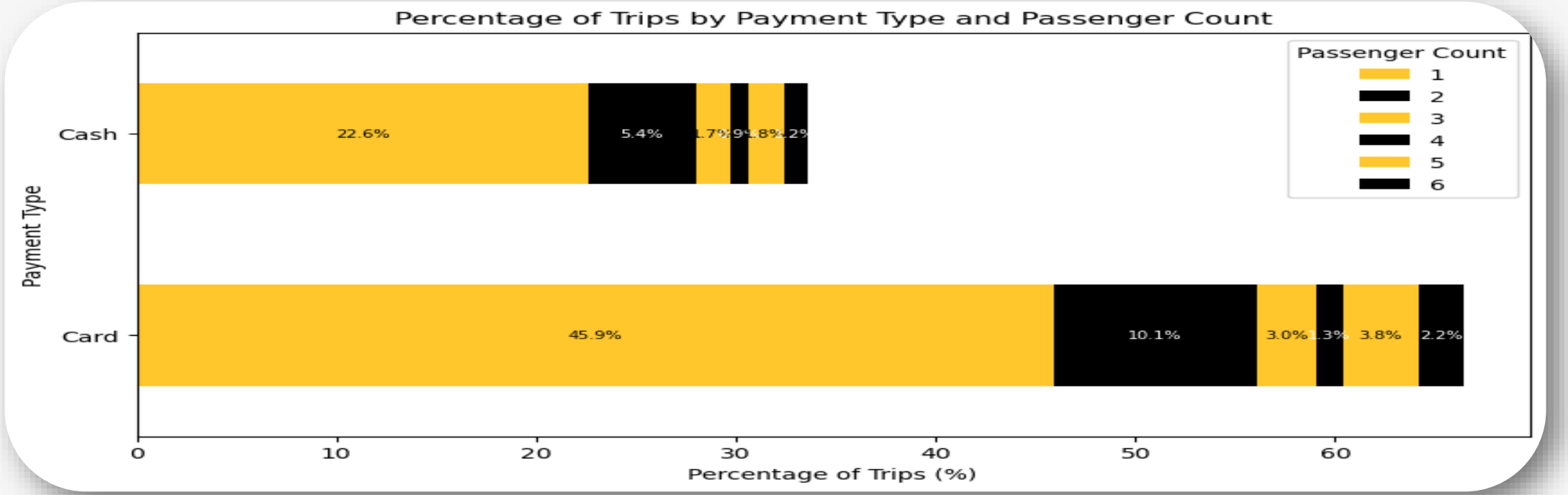


- The proportion of customers paying with cards is significantly higher than those paying with cash, with card payments accounting for 66% of all transactions compared to cash at 34%

- This indicates a strong preference among customers for using card payments over cash, potentially due to convenience, security or incentives offered for card transactions

Passenger Count Analysis

- Among card payments, rides with a single passenger(Passenger Count =1) comprise the largest proportion consisting of 45.9% of all card transactions.
- Similarly, cash payments are predominantly associated with single-passenger rides, making up 22.6% of all cash transactions.
- There is a noticeable decrease in the percentage of transactions as the passenger count increases, suggesting that larger groups are less likely to use taxis, or may opt for payment methods.
- These insights emphasize the importance of considering both payment method and passenger count when analysing transaction data, as they could provide valuable insights into customer behaviour and preferences.



Hypothesis Testing Results

Based on the descriptive statistics, the Mann-Whitney U test was conducted to calculate the p-value (or probability value). This test helped in determining whether to reject or accept the null hypothesis.

Conclusion: The null hypothesis is rejected since the p-value is less than alpha(0.05). This indicates sufficient evidence to conclude that there is a statistically significant difference in fare amounts between 'Card' and 'Cash' payments.

What Is Regression Analysis?

Regression Analysis:

Regression analysis is a statistical method used to understand how different factors influence an outcome. It helps predict outcomes by analyzing the relationships between variables.

Rationale:

- After conducting initial descriptive statistics such as mean calculations and visualisations like histograms and pie charts to analyze preferences between card and cash payments, it became evident that the payment method significantly influences fare amounts.

- However, fare amounts are influenced by multiple factors beyond payment type, such as trip distance, duration, and passenger counts.

- Therefore, conducting regression analysis allows for a comprehensive assessment of how these various factors collectively impact fare amounts. This deeper analysis helps understand each factor's relative importance and guides strategic decisions to enhance driver earnings and service efficiency.

Regression Analysis

Objective:
To understand the importance of each independent variable (trip distance, trip duration, passenger count, and payment type) in predicting the dependent variable (fare amount).

Observations:
Significant Variables:
Trip Distance: Positively impacts fare amount.
Trip Duration: Positively impacts fare amount.
Passenger Count: Slight negative impact on fare amount.

Insignificant Variable:
Payment Type: No significant difference between cash and card payments after controlling for other factors.

Conclusion:
Drivers: Focus on optimising trip distance and duration to maximise fares.
Fare Strategy: Payment method does not need to be a primary concern for fare amounts.

Recommendations

Image by Freepik

Target Longer Trips:
Encourage customers to choose taxis for longer trips, as trip distance has a strong positive impact on fare amounts. For example, consider offering promotions or special discounts specifically for longer trips.

Promote Both Payment Methods Equally:

- Since there is no significant difference in fare amounts between cash and card payments, focus on providing a seamless payment experience for both methods.
- Ensure that customers have the flexibility to choose their preferred payment method.

Enhance Customer Experience:
Improve service quality by focusing on factors such as driver professionalism, vehicle cleanliness, and safety, which can indirectly influence customer satisfaction and encourage repeat business.

THANK YOU