

NYC Yellow Taxi Tip Analysis — Predictive Factors from 2017 Trip Data

Overview

This project analyzed tipping behavior in NYC Yellow Taxi rides using 22,652 cleaned records from 2017. The goal was to explore which ride variables most impact tip amount, especially payment type and party size.

Objective

Identify key factors that influence tipping and recommend variables for future predictive modeling.

Results

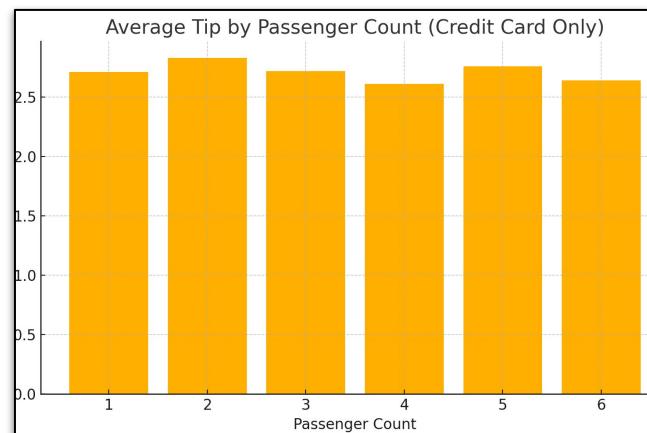
Average tip for **credit card**: \$2.73

Average tip for **cash**: \$0.00

Party size slightly correlates with tip — highest average tip for 2 passengers

Major vendors show consistent pricing behavior

Anomalies exist in **fare_per_mile** and trip duration



Next Steps

- Flag and remove invalid trips (e.g. 0 miles, extreme durations)
- Engineer new features like time of day, zone names, or surge pricing
- Train and evaluate regression or classification models using **payment_type** and **passenger_count** as features