

# TLC Trip NYC

Prepared by :

Mashael Asiri  
Ajwad Almajnuni







# Agenda

**01**

Introduction

**03**

EDA

**02**

Tools

**04**

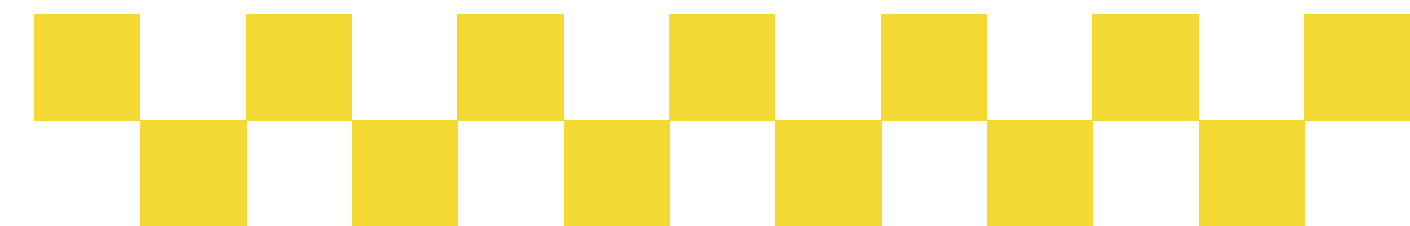
Modeling





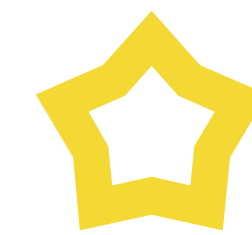
# TLC Trip NYC

The New York City Taxi and Limousine Commission (TLC) is an agency that is responsible for licensing and regulating New York City's taxi cabs. Over 200,000 TLC licensees complete approximately 1,000,000 trips each day.



# Goal

Predict of the fare amount for each trips in the NYC taxi.



Care



Quality



Profit

# Tools





# Dataset

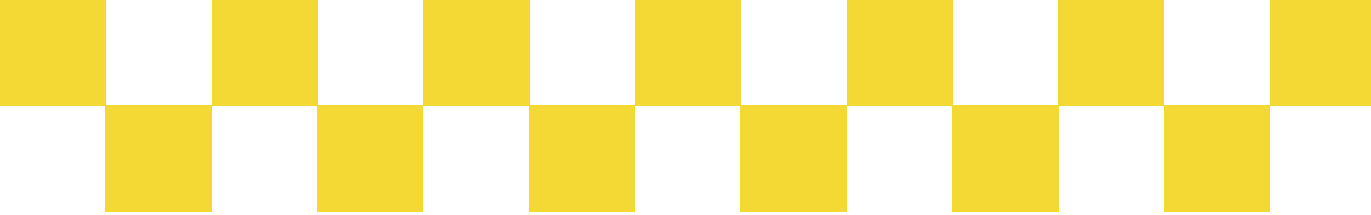
The data used in the attached datasets were collected and provided to the NYC Taxi and Limousine Commission (TLC).



**19**  
Features



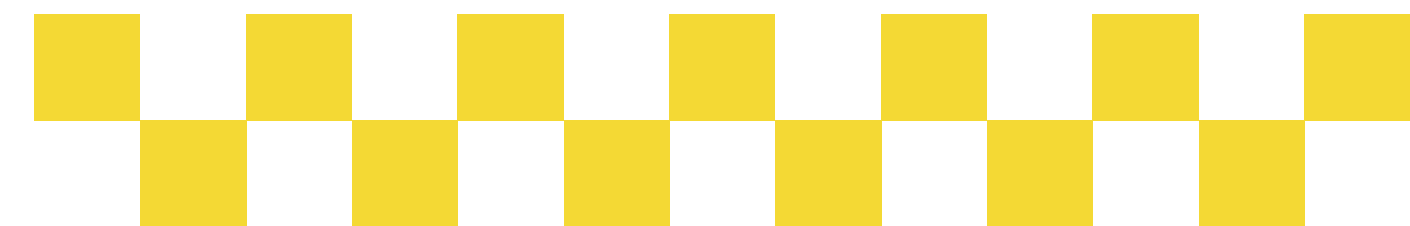
Rows  
**19 M**



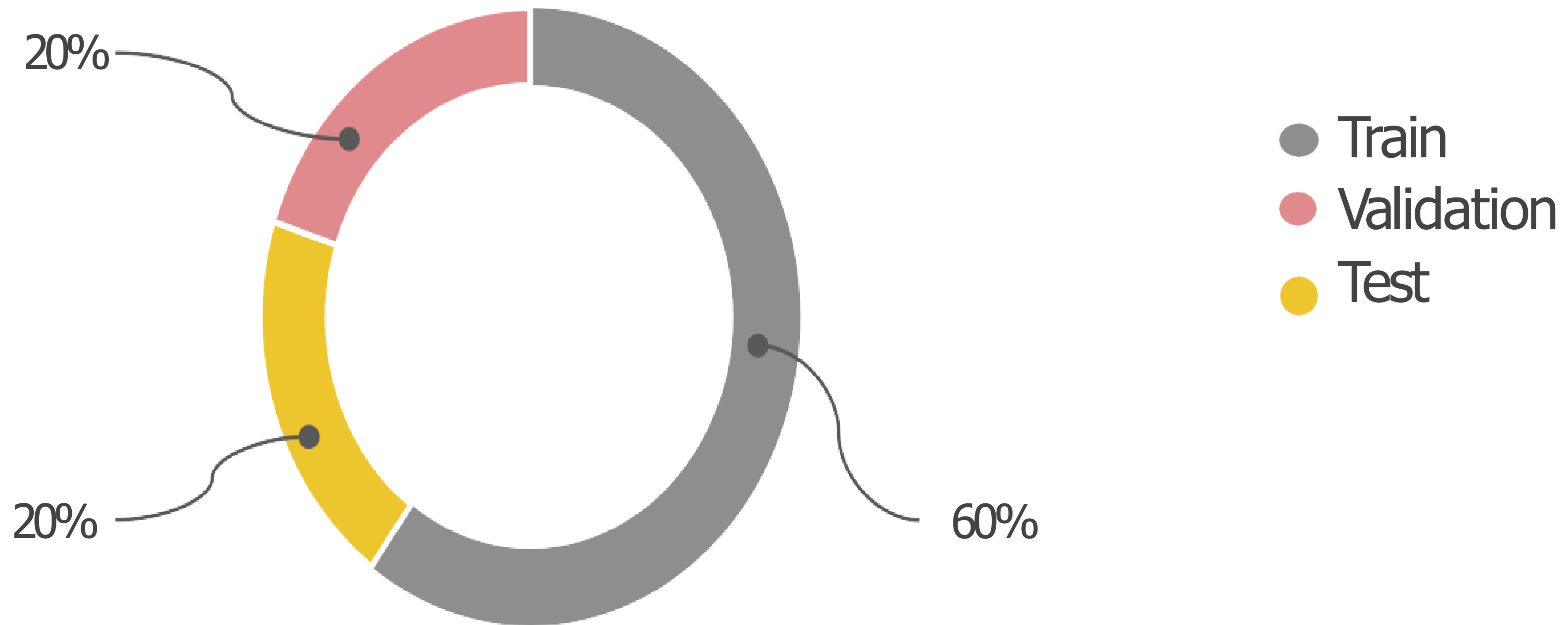
first taxi

# Cleaning Data

- **1** Missing Values
- **2** Drop duplicated data
- **3** Extract Numerical Values.
- **4** Convert categorical values into dummy variables



# SPLITTING THE DATA

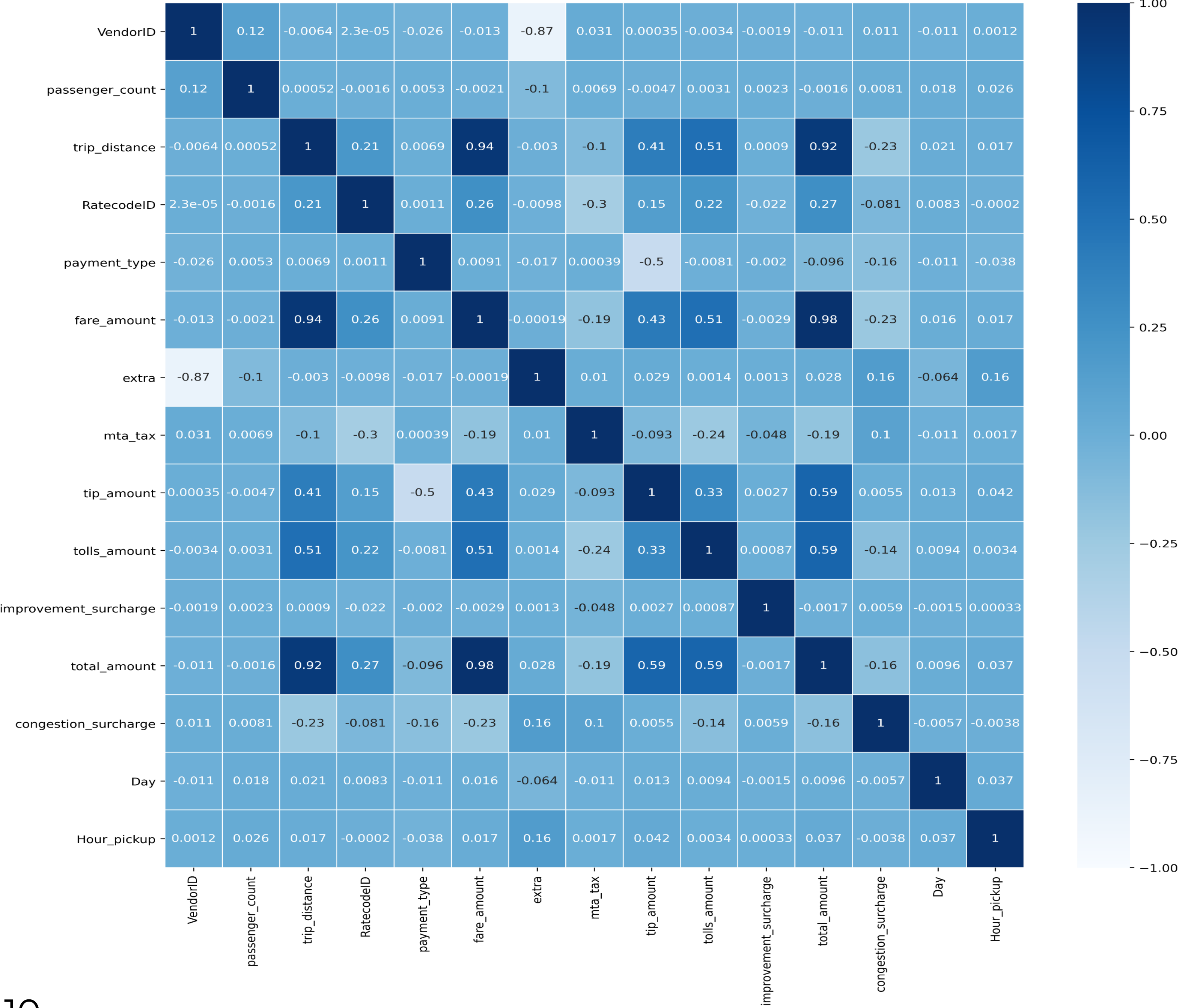




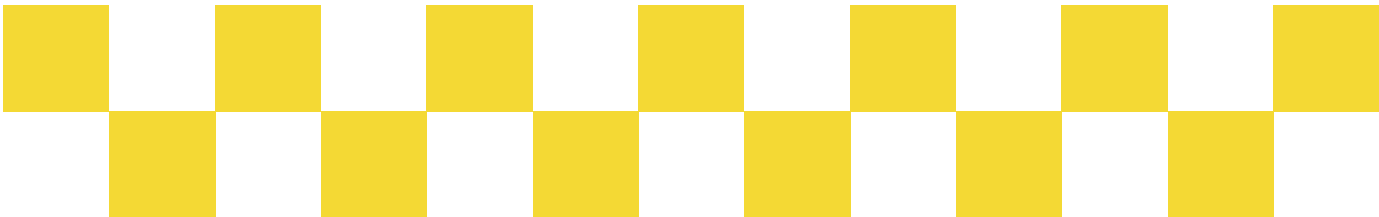


# Visualization

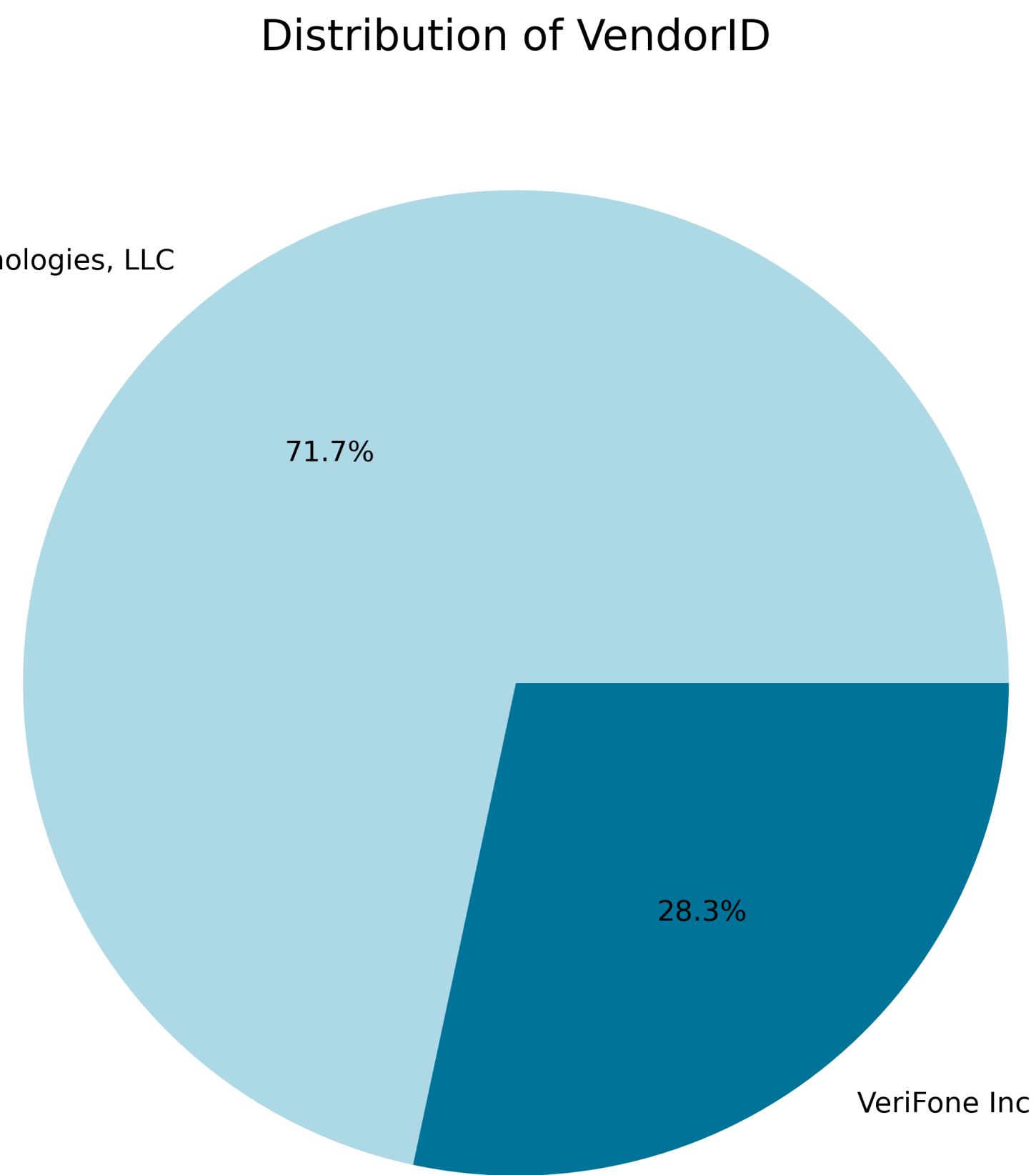




# Coloration Data

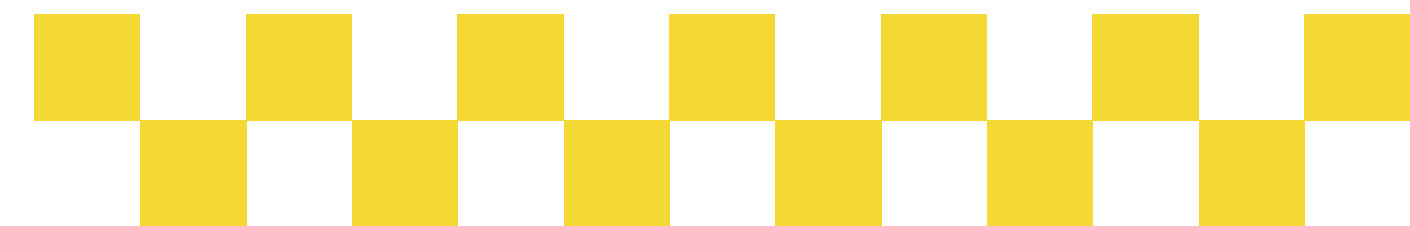
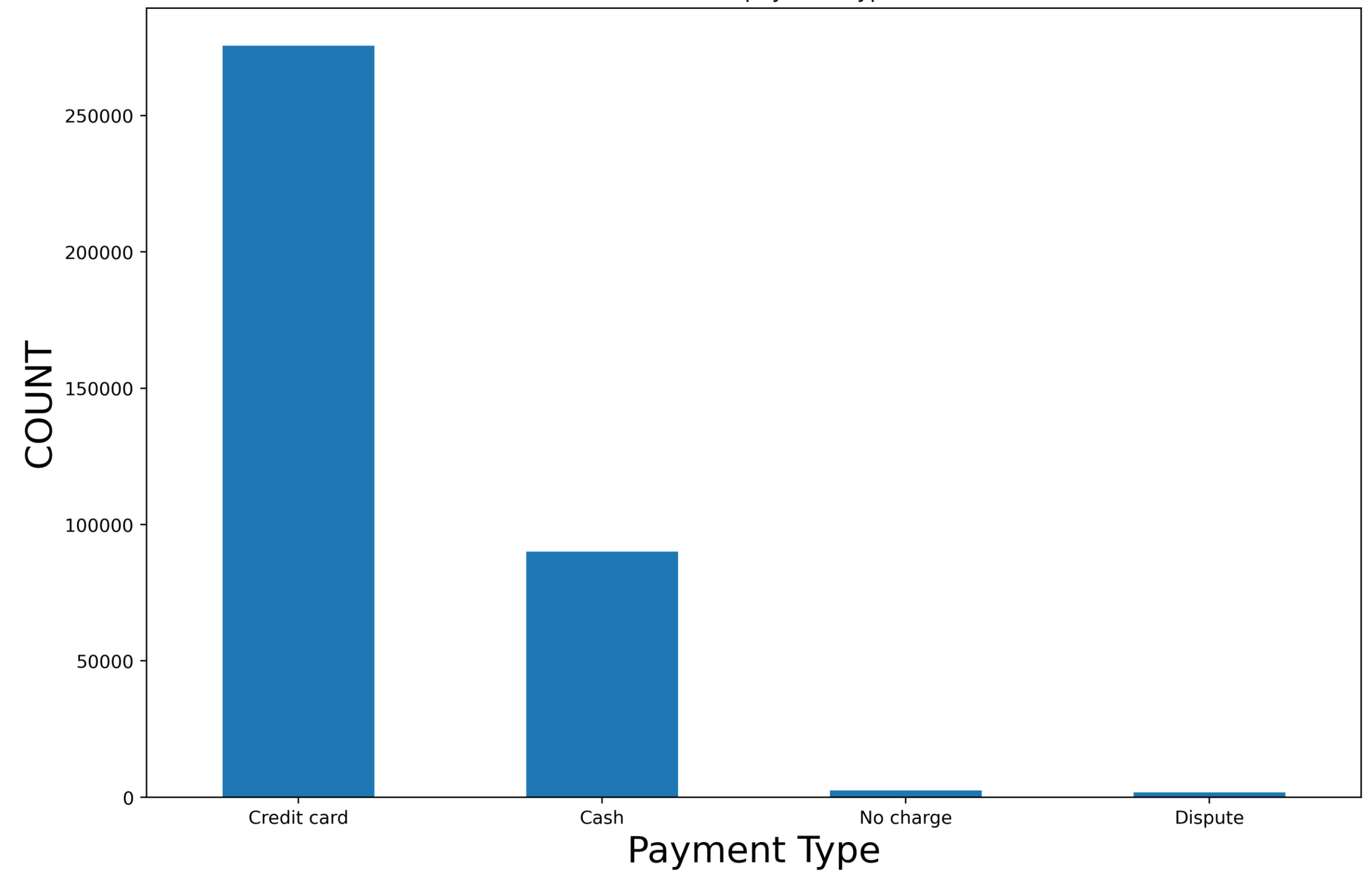


Creative Mobile Technologies, LLC



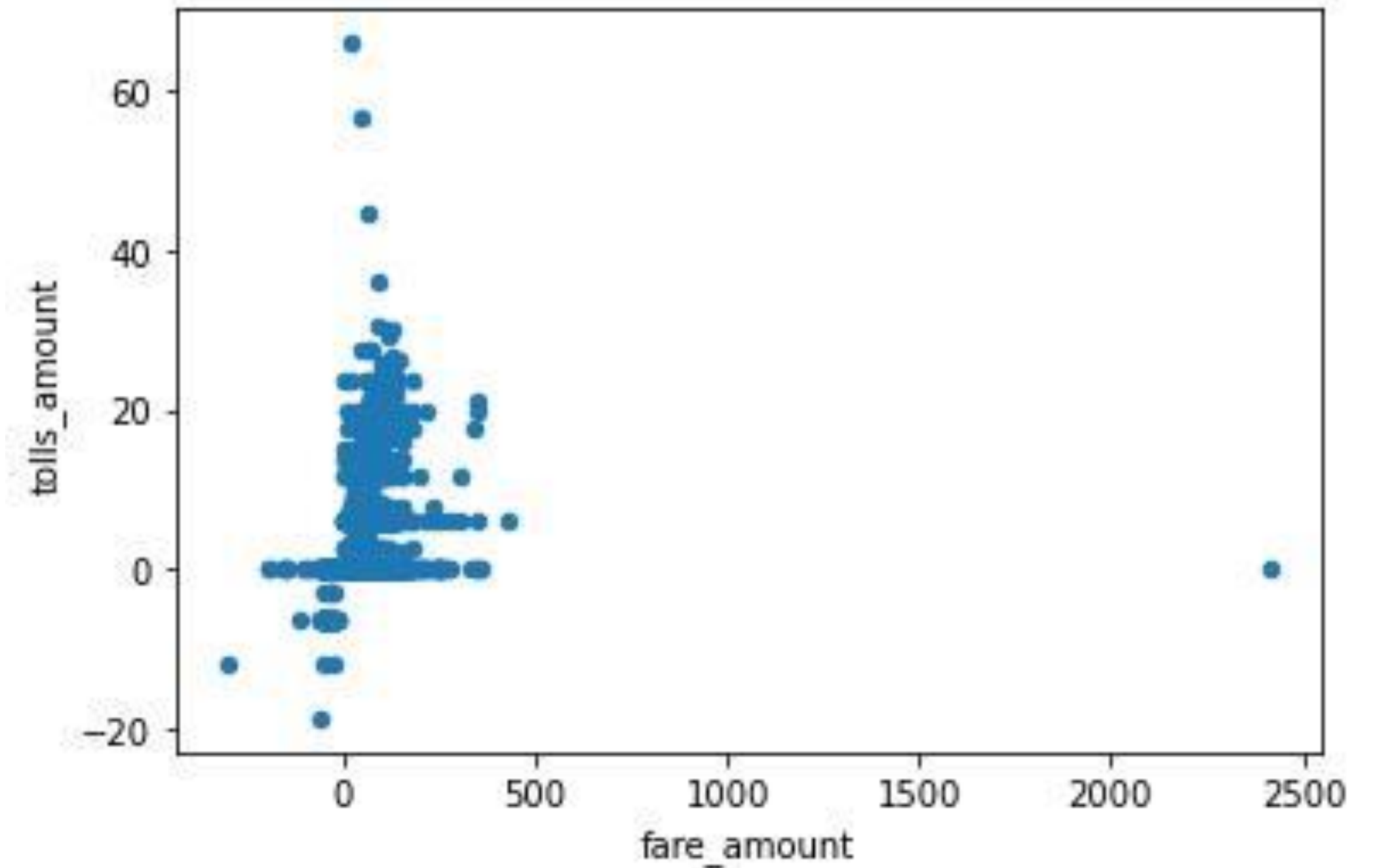
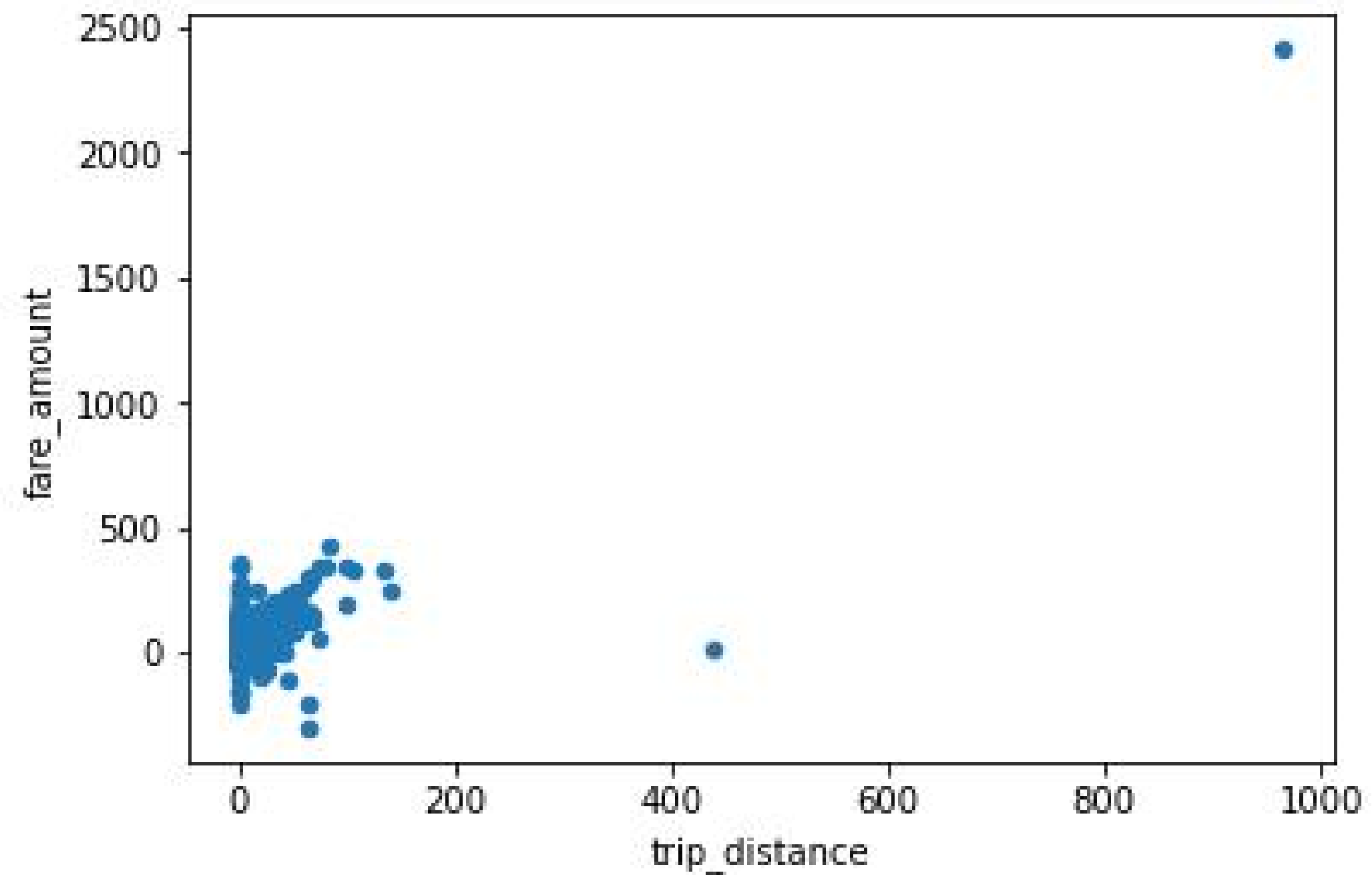
VeriFone Inc

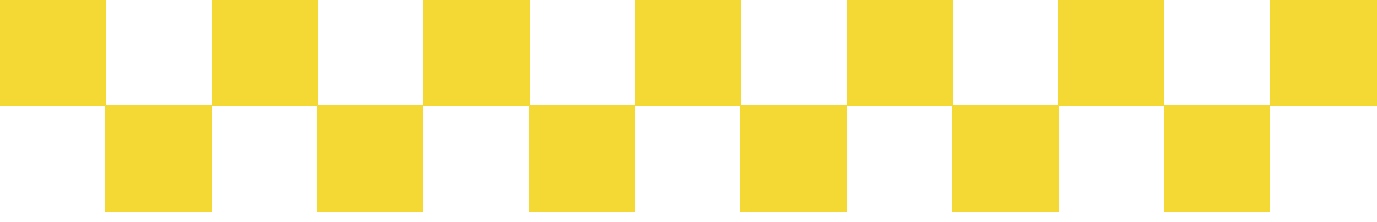
Distribution payment type



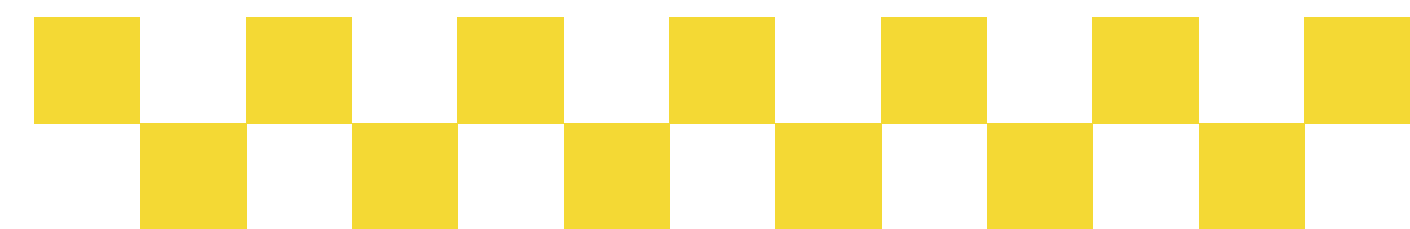


# Relationship between fare amount , Trip distance and tolls amount





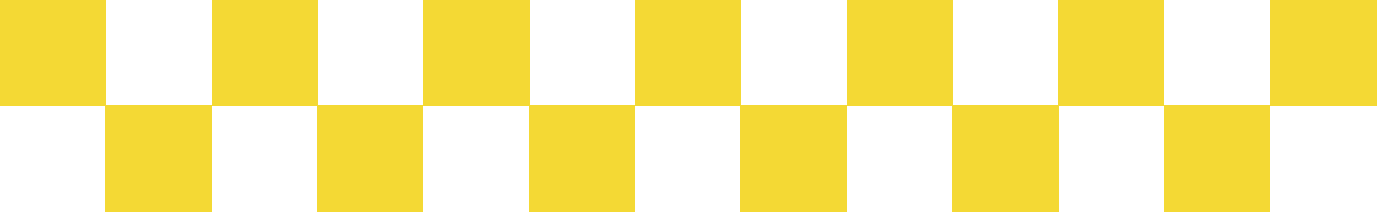
# Modeling



# Modeling Result

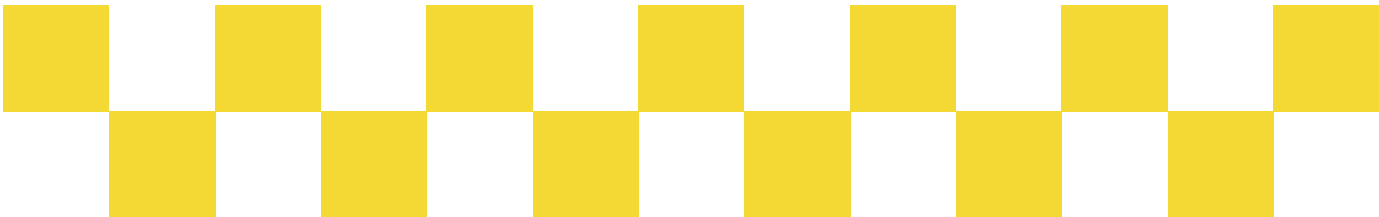
Model \ R <sup>2</sup>	Train	Validation	Test
Linear regression	0.88	0.74	0.16
Polynomial Regression	0.89	0.82	0.84
Ridge regression	0.88	0.74	0.16
Lasso regression	0.88	0.78	0.18





# Modeling Result

	R Squared	MAE
Linear Regression	3.71	1.46
Polynomial Regression	3.60	1.34
Ridge	3.71	1.46
Lasso	4.88	1.70





**Thanks , Any Questions ?**