NYC Taxi And Limousine Commission Case Study

Presented by:
Chaithanya Ittamadu Shivakumar

Agenda

- INTRODUCTION
- OPERATIONAL DATABASE
- OBJECTIVE OF NYC EXPLORATION
- DESIGN DECISIONS
- DATA VISUALIZATION
- CONCLUSION

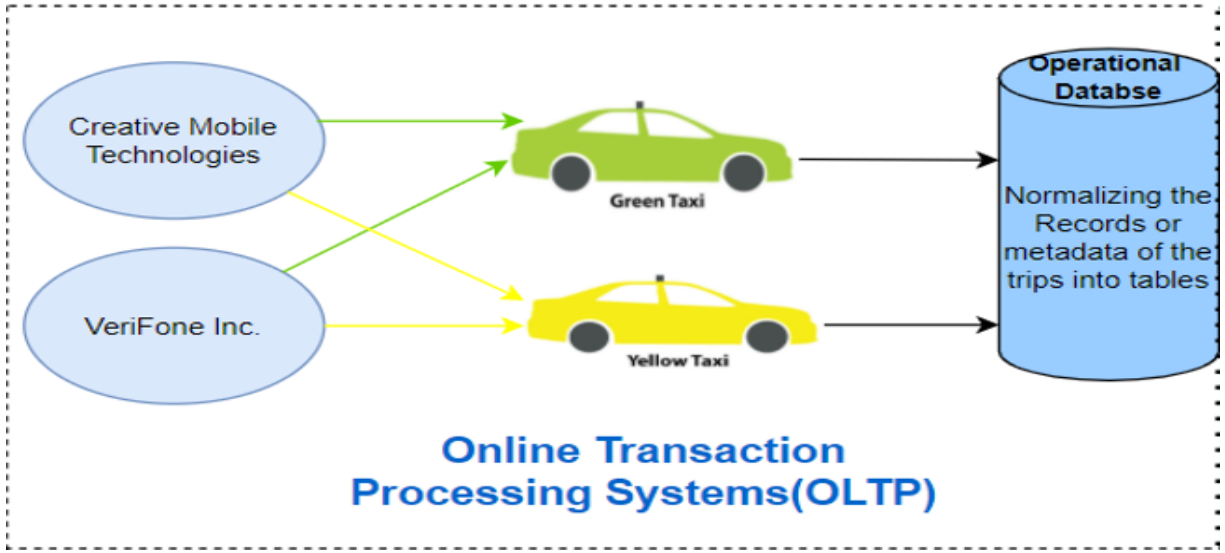
Introduction

- The New York City Taxi and Limousine Commission (NYC TLC) is an agency of the New York City government that licenses and regulates medallion(yellow) and boro(green) taxis.
- The NYC TLC model is of 2020 yellow and green taxi trip records which contain the metadata of the trips across different cities of NYC.

OPERATIONAL DATABASE

Steps in creating operational Database

1. Each month consists of Green and Yellow trips datasets in CSV format. It is then converted into Excel format using internal Excel functions.
2. Import wizard feature from Microsoft SQL Server (2019) is used in uploading the data to the Relational Database for further cleaning of the data.
3. Cleaning of the data is performed on a single table where 12 months of data is imported using the Import wizard.
4. CRUD operations are performed to create databases, tables, establishing relationships between the tables.



OBJECTIVE OF NYC DATA EXPLORATION

The main objective of any cab vendor is to run the business successfully.

How?

1. Analyse business.
2. Increase the profit.

Analysis On NYC Taxi

Dynamic strategies towards trips.

1. Analyse frequent pick-up points.

2. Analysing payment modes.

3. Seasonal hail.

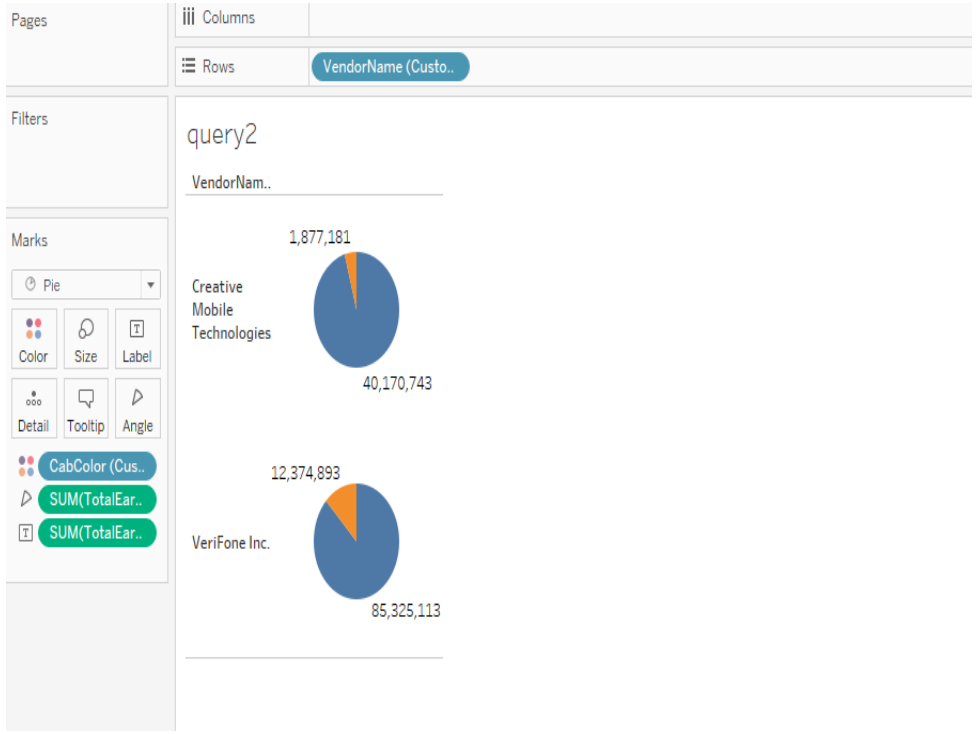
Approach

Dimensional model supports the reporting and analytical needs of a data warehouse system.

Data Visualization

- Tableau is a visual analytics platform used to represent the data in graphical manner which helps to analyze the results and solves the business problems.

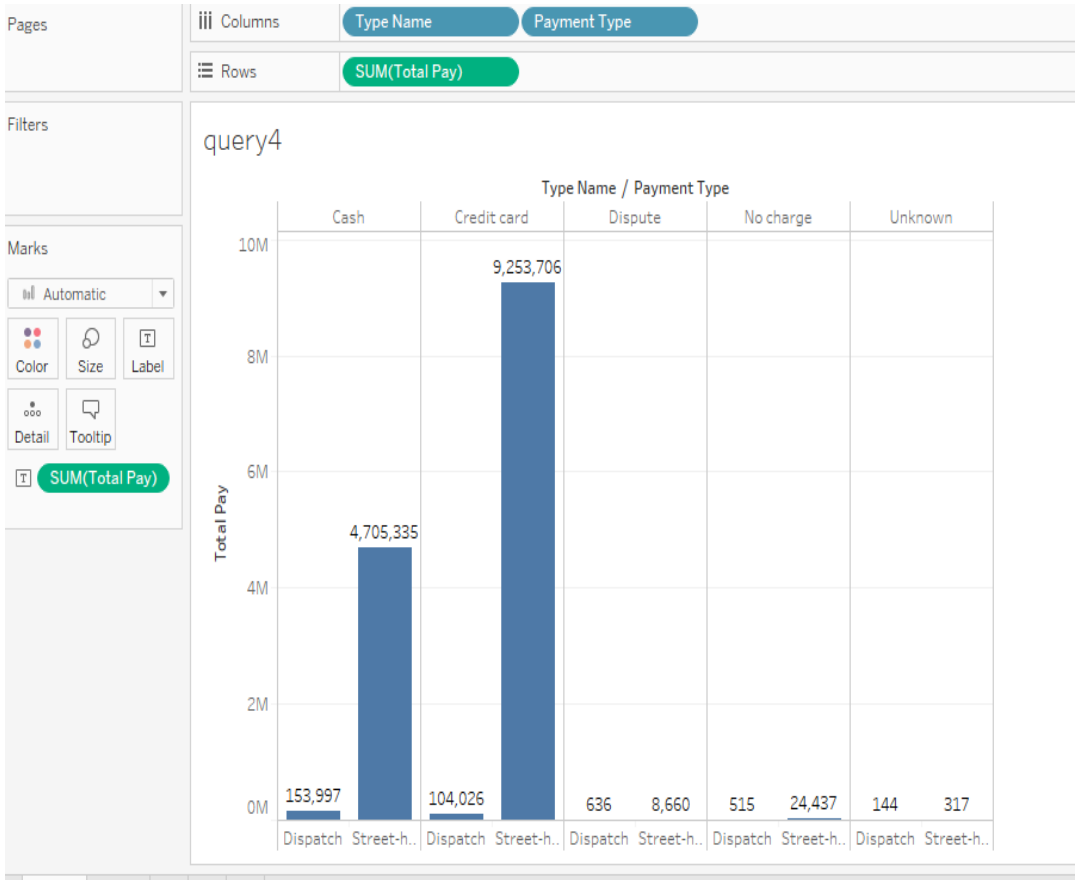
1. <u>Earnings by vendor and their cab color</u>
   - x-axis represents vendor
   - y-axis represents cab color
   - values are represented using pie-chart



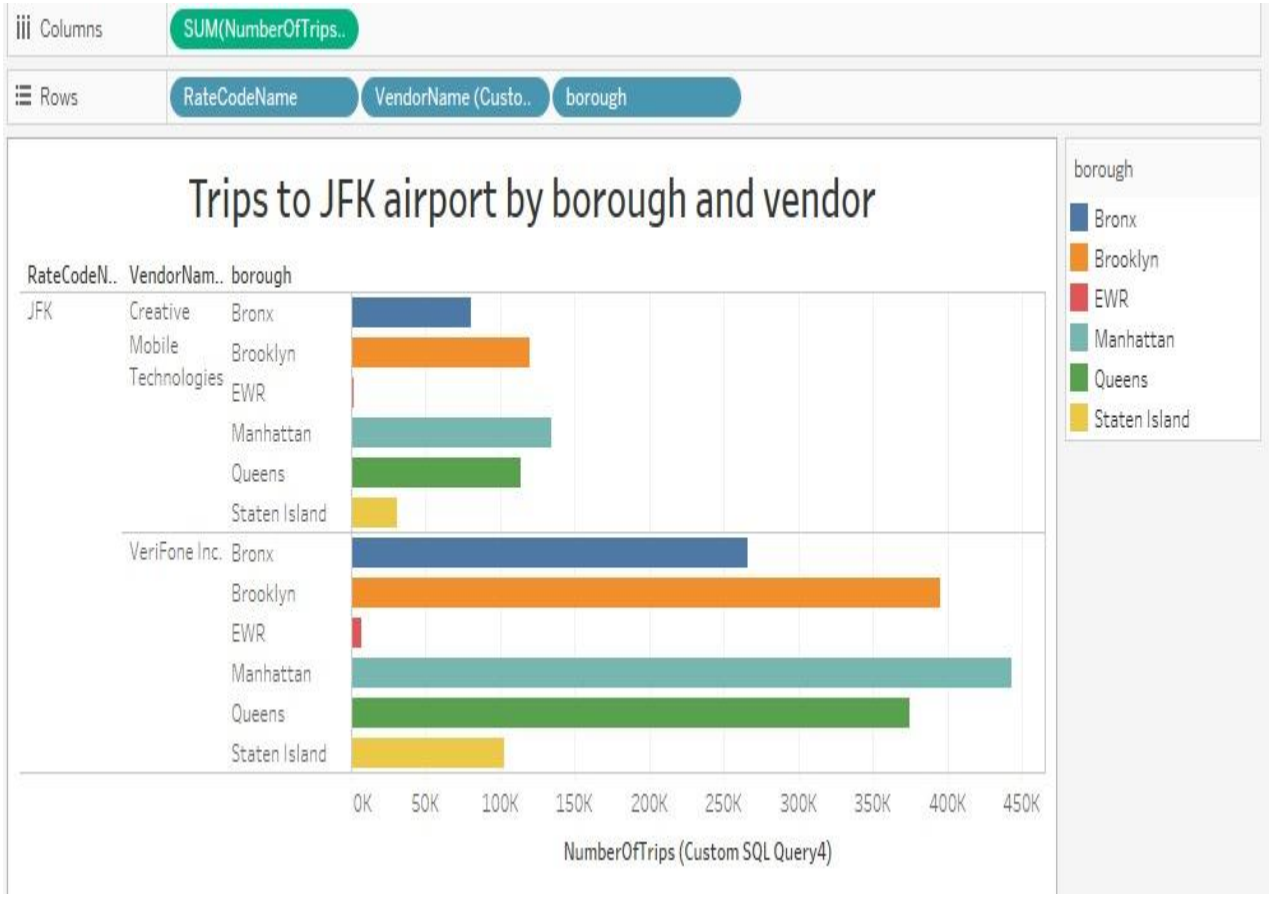
2. <u>Top 10 Pick up zones</u>
   - x-axis represents Zones
   - y-axis represents Number of trips
   - values are represented using bar-chart

Data Visualization contd.

3. Total payment by trip and payment type
   - x-axis represents payment mode
   - y-axis represents cab total pay
   - values are represented using bar-chart

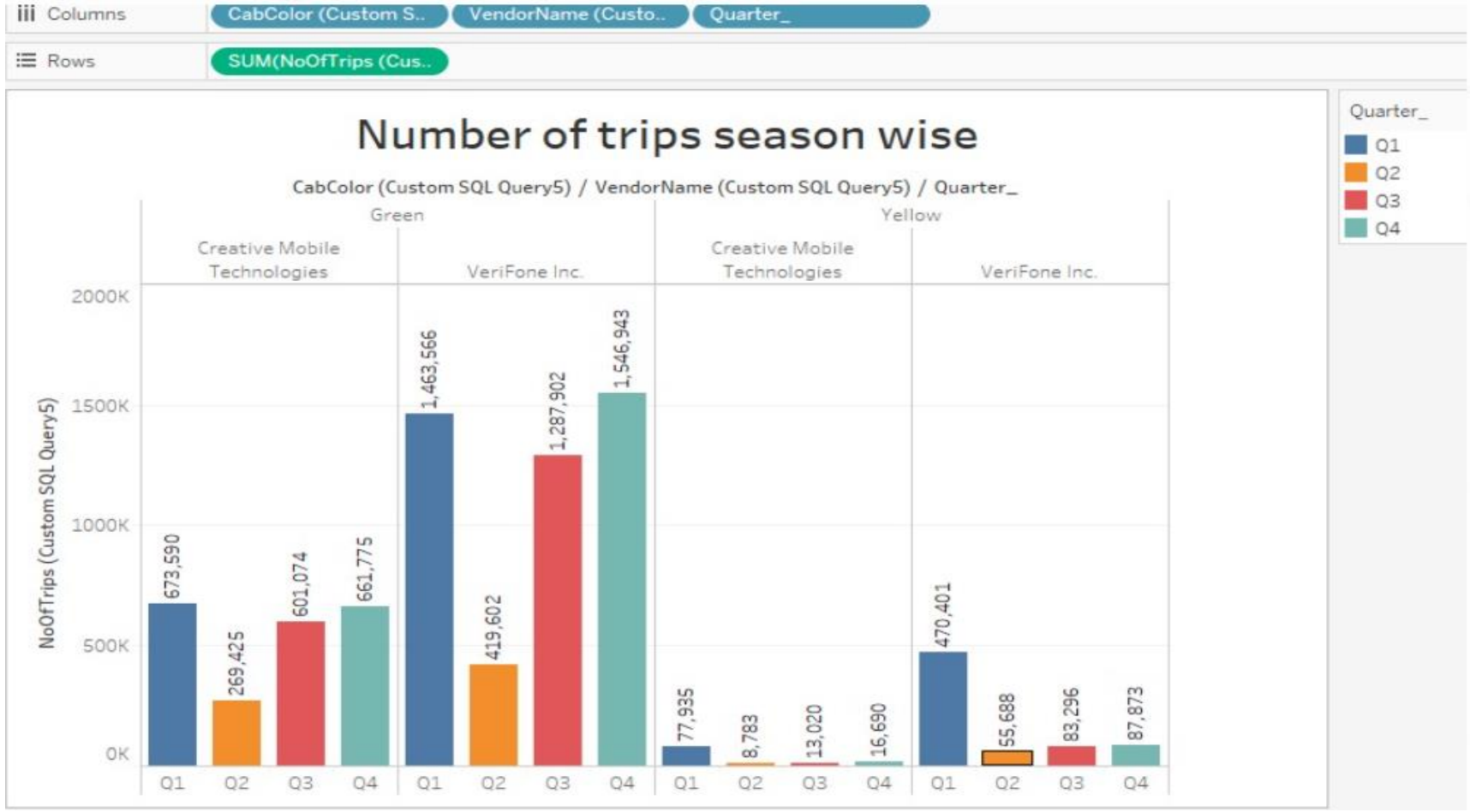


4. Trips to JFK airport by borough and vendor
   - x-axis represents number of trips
   - y-axis represents rate code name, vendor name and borough
   - values are represented using horizontal bar

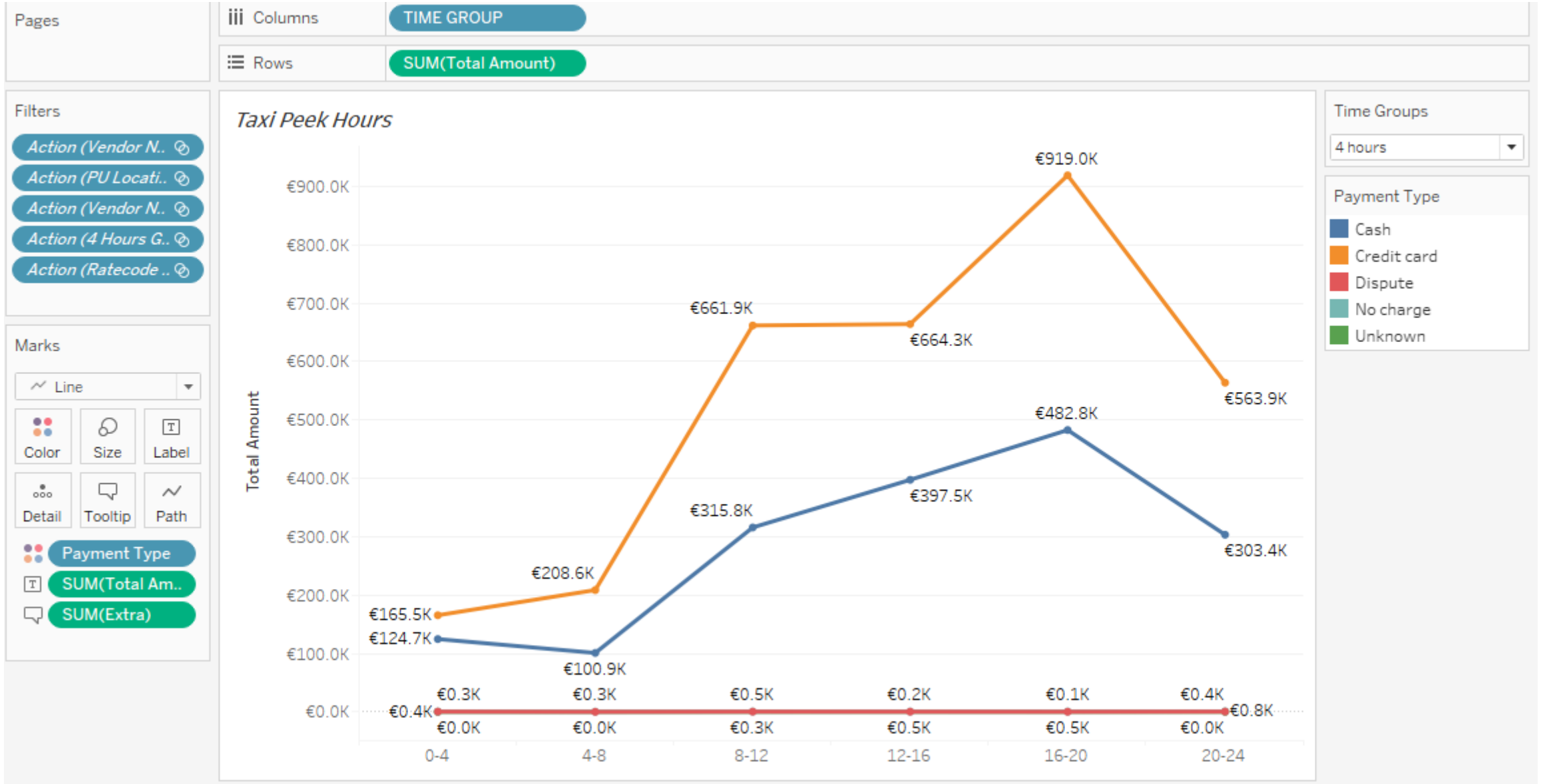


Data Visualization contd.

5.Seasonal pick-ups based on Vendor and Cab colour.



6.Finding Peek time using groups



Conclusion

- Storing, transforming, and managing the data is no problem in today's world. Nonetheless, the significant data which can raise from the information is an understanding each merchant wishes to relate to predicting the business
- The NYC DW, reports, and visualization pictures the analysis of New York Taxi through different pick-up stations, seasonal hailing, number of people traveling from different cities, and vendor earnings which gives better insights for decision making.

THANK YOU