GETT

RIDE CANCELLATIONS ANALYSIS

POOJITHA VIJJAPU

INTRODUCTION

Gett is a ride booking company which connects drivers to riders.

We are analyzing ride cancellation data to understand reasons behind cancellations and provide recommendations to reduce overall cancellation rate.

PROBLEM STATEMENT

Why are ride cancellations happening and how can we lower the cancellations?

DATASET DESCRIPTION

- order_datetime - time of the order
- origin_longitude - longitude of the order
- origin_latitude - latitude of the order
- m_order_eta - time before order arrival
- order_gk - order number
- order_status_key - status, an enumeration consisting of the following mapping:
  4 - cancelled by client,
  9 - cancelled by system, i.e., a reject
- is_driver_assigned_key - whether a driver has been assigned
- cancellation_time_in_seconds - how many seconds passed before cancellation

DATA CLEANING

- The date time field has the same date for all the observations. So only extracted only hour of day.
- The data belongs to a small geographical location (a state or city) - there is equal distribution of data throughout this geographical location - hence deleted this field for further analysis.

Other Observations :

- The field cancellation_time_in_seconds has no values for orders which were cancelled by the system.
- Order ETA appears only after driver is assigned.

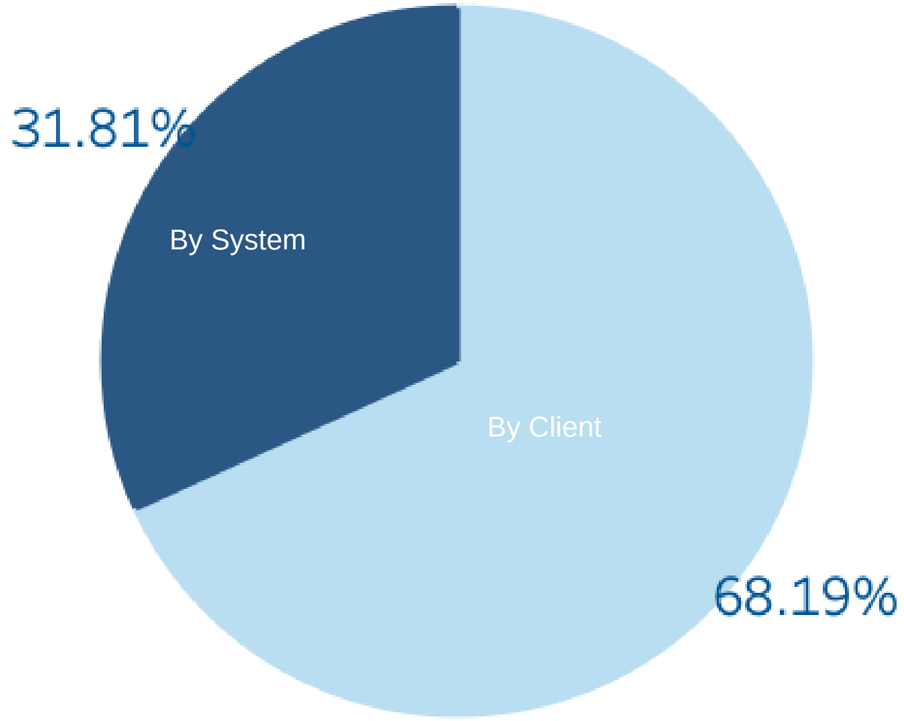
CLEANED AND EXTRACTED DATA :

After cleaning and extracting the important features we are using in this analysis are :

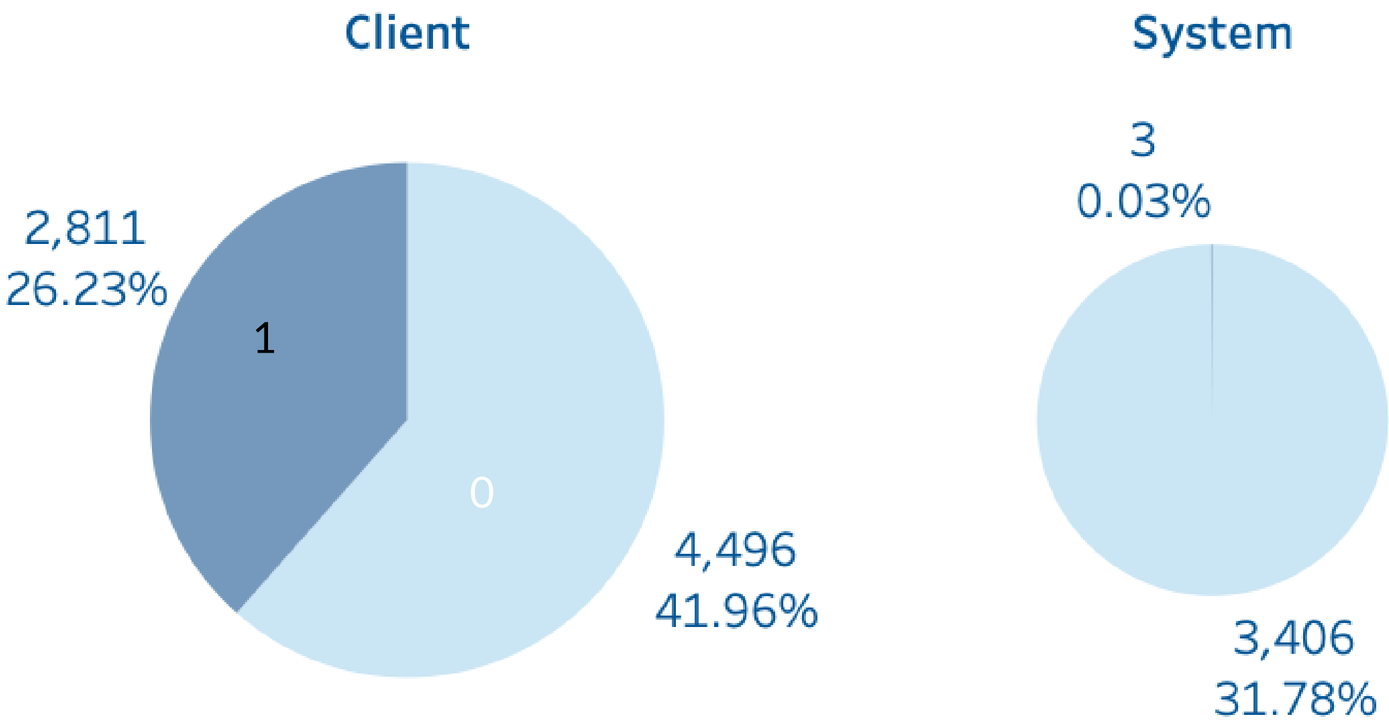
- Order Status Key - *cancellation status of an order whether it is cancelled by the system or the client*
- Is Driver Assigned - *tells whether a driver is assigned or not*
- Cancellation time in seconds - *seconds passed before cancellation*
- Order ETA - *Estimated time of arrival for a driver (only driver assigned rides)*
- Hour of the Day

Orders by Cancellation Type

~68% orders are cancelled by client and ~31% are cancelled by the system.

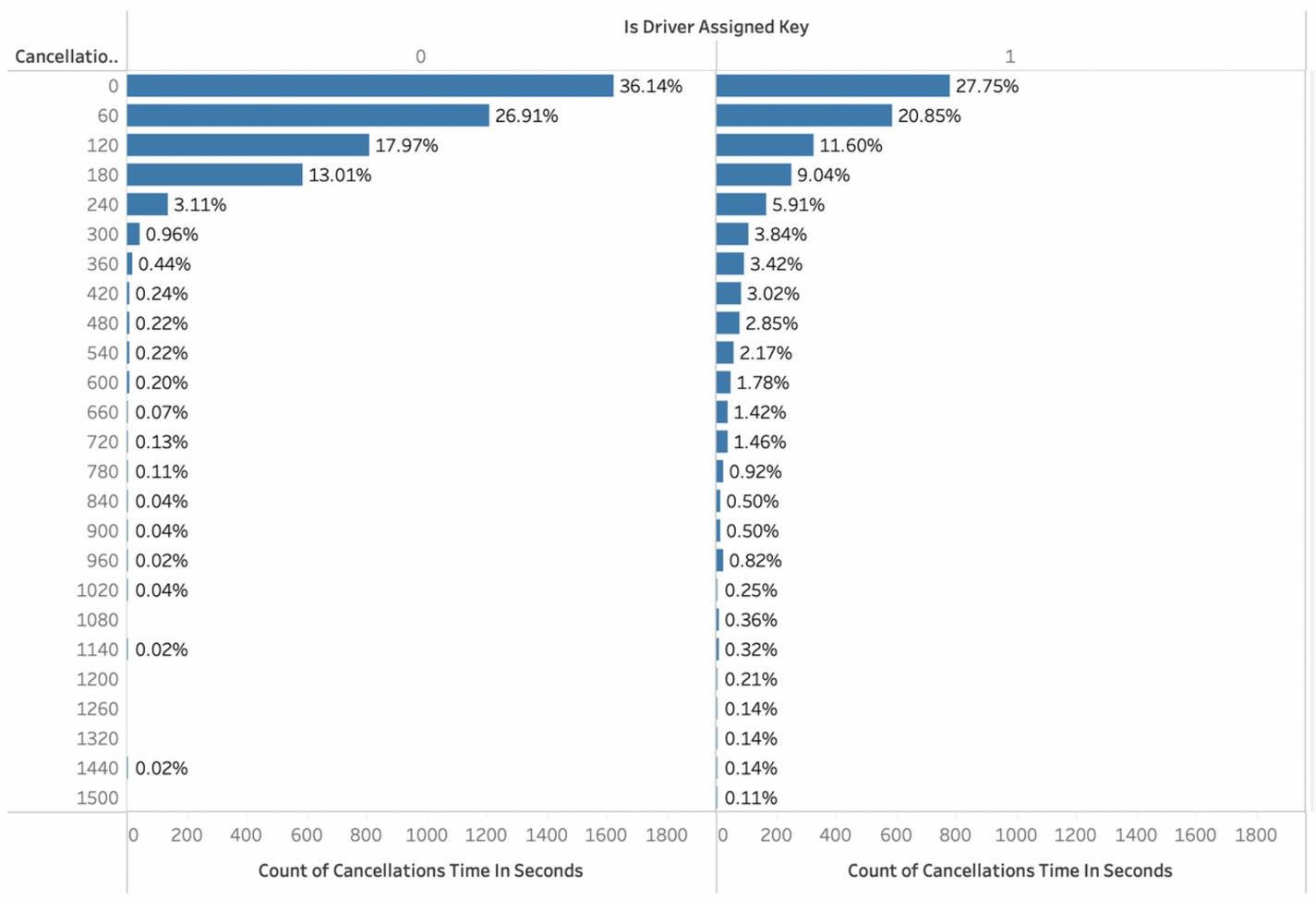


Driver Assignment



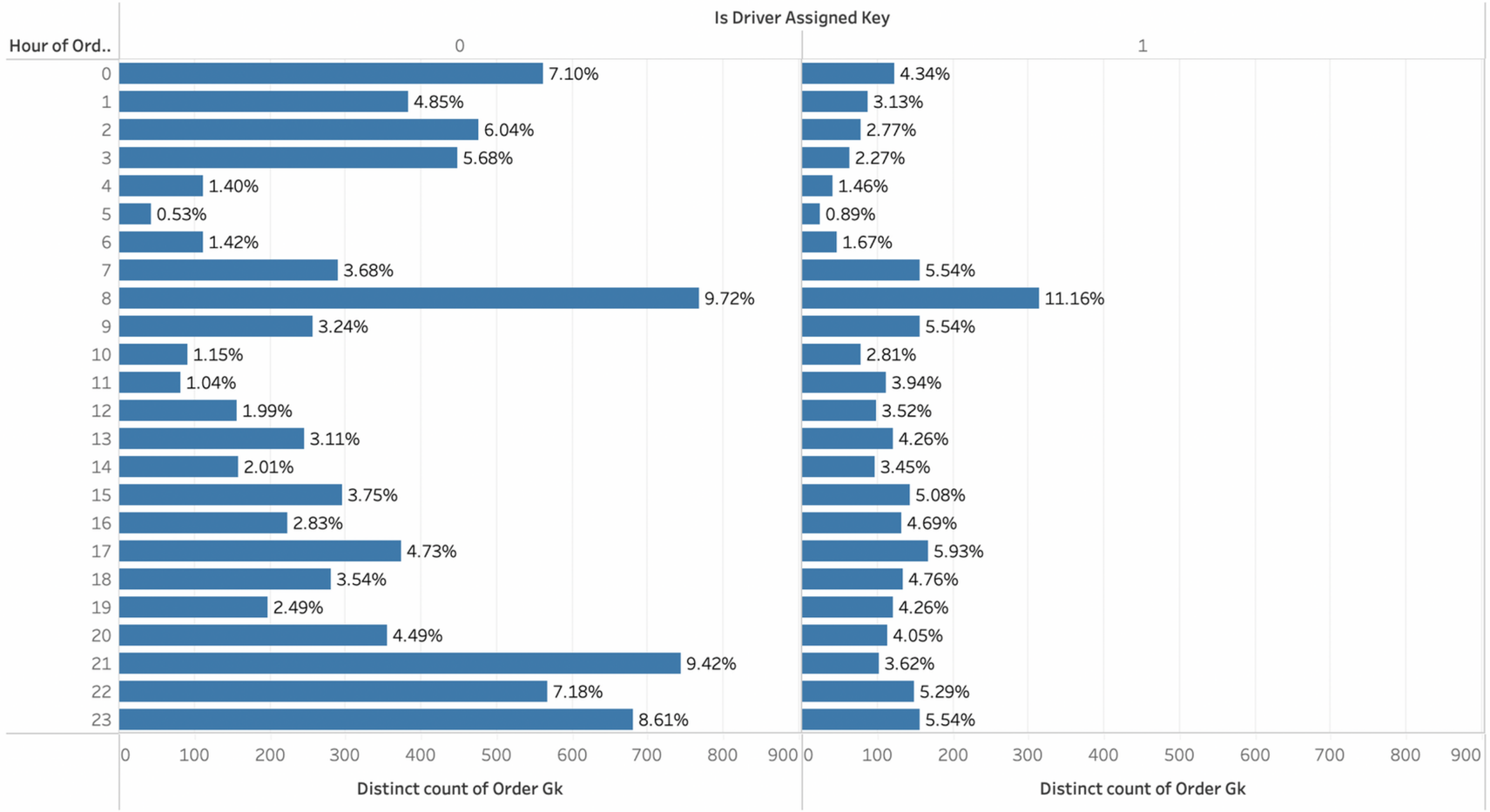
- For cancellations by client : Cancellations after the driver is assigned is less by at least 40%.
- Most of the cancellations by the system happen before driver assignment.

Cancellation Time in Seconds



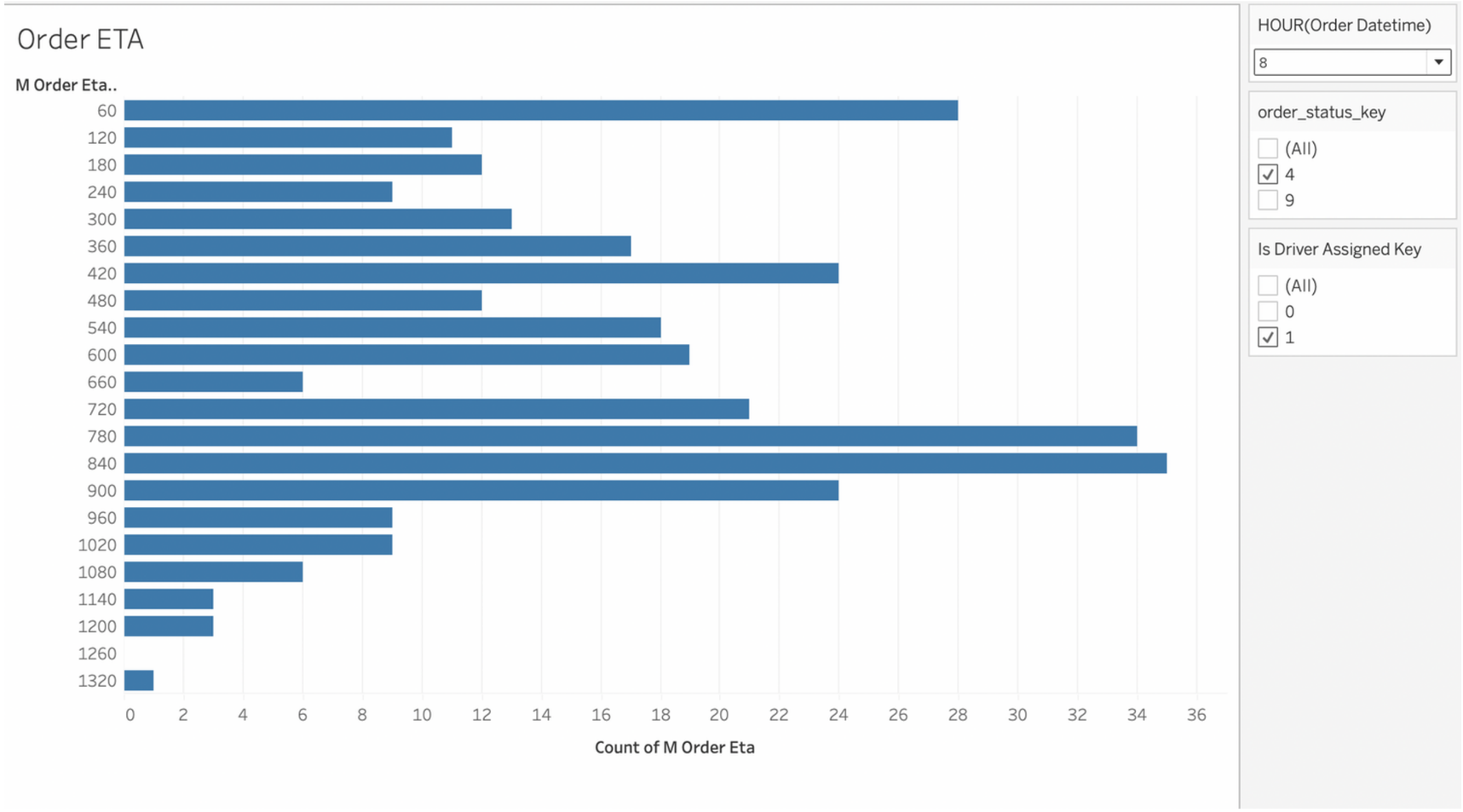
- Before the driver is assigned : ~80% of the cancellations happen in the first 3 min.
- After the driver is assigned : ~60% of the cancellations happen in the first 3 min.
- First 3 min play a key role in determining if the ride will be cancelled or not.

Hour of the Day



- Before the driver is assigned : Cancellations are high in 8th, 21st, 22nd and 23rd hours.
- After the driver is assigned : The 8th hour has the highest cancellations.

Estimated Time of Arrival



The average ETA across all the orders is 441s.

The Order ETA's are very high for the 8th hour.

RECOMMENDATIONS

- Since most of the cancellations are by client - focus should be more on improving client experience.
- Almost 80% of cancellations by client, before driver assignment are happening in the first 3 min - client could be engaged by being informed about the status of their request.
- The cancellations in the 8th hour could imply more demand during the office hour - hence the supply during this time should be increased.
- The cancellations in 8th hour even after the driver is assigned is also high and the order ETA is also high - there could be a better mapping of clients to riders.