Rappi Experimentation & Analitics Senior Case

Agenda

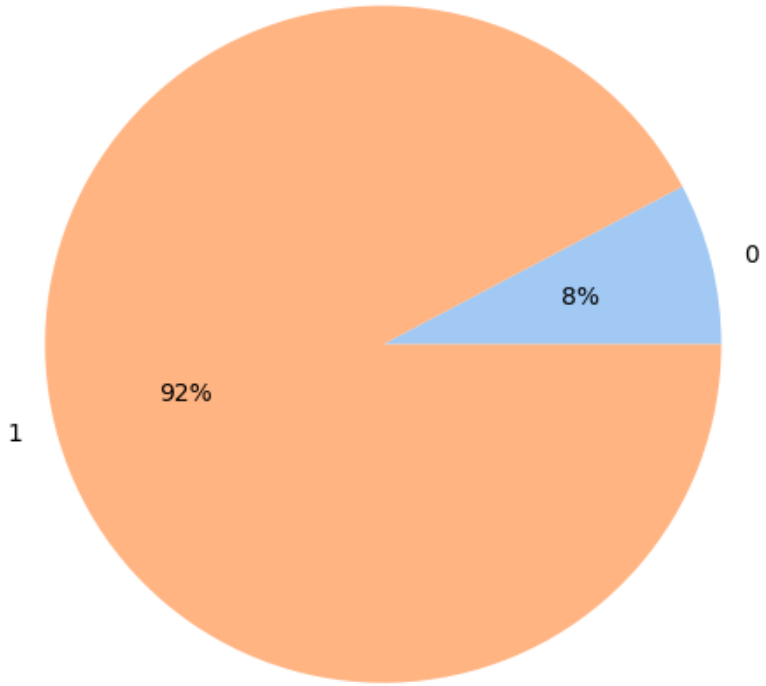
- 1. Exploratory preliminar Analysis
  - 1.1. Bussiness case
  - 1.2. What we learned from the data
  - 1.4. Full statistical prior analysis
- 2. Analytics Models
  - 2.1. Process and experiment setting
  - 2.2. Extra: Analysis per variable
- 3. Insights
- 4. Reccomendations
- 5. Extra: Results on the data mining exercise

1.1. Bussiness Case

- Rappi's Operations team is interested in decreasing the number of orders that are not taken by any courier, due to the fact that they are not attractive enough for couriers.
- Scope: Data taken from the sample of 1 month, September 2017.
  - Included variables: total earning, distance from user to store (km), difference in meters between the store and user altitude, date-time of creation, taken as a binary variable: 1 if taken, 0 otherwise.

The objective is to identify some key drivers that might predict if a given order is taken or not.

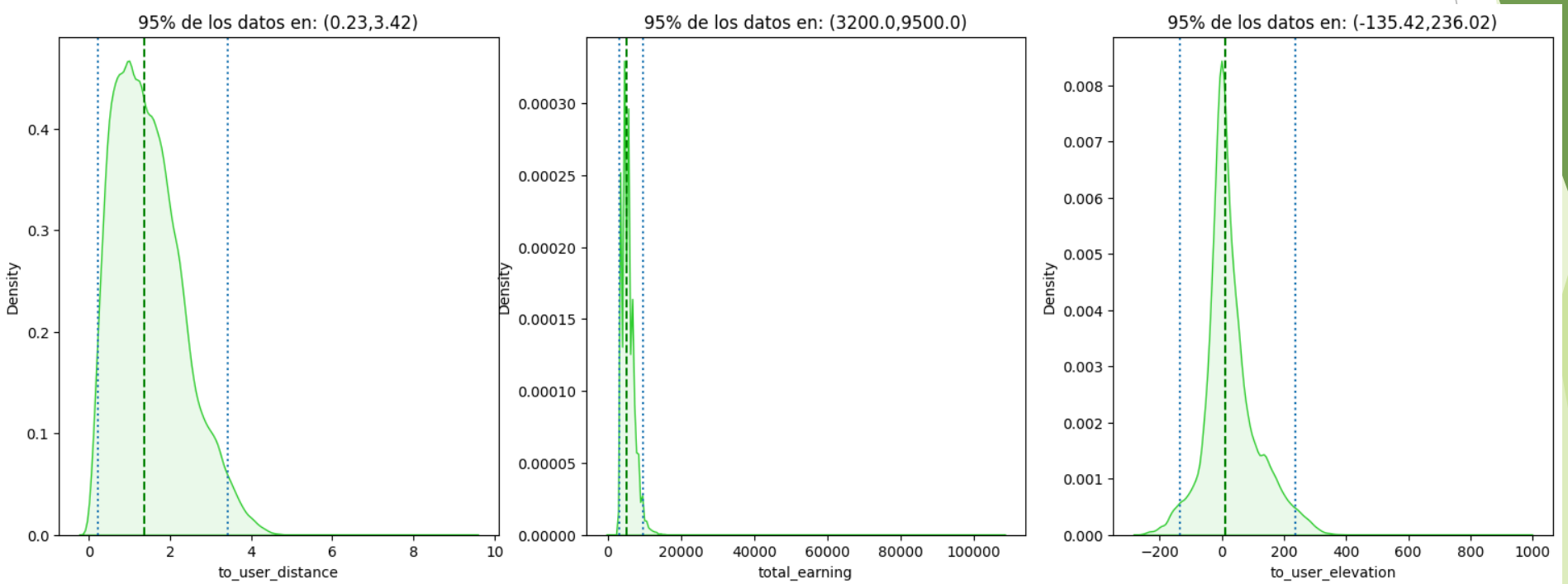
1.2. What we learned from the data



92% of the orders were taken by any Courier, 8% were not taken by any Courier.

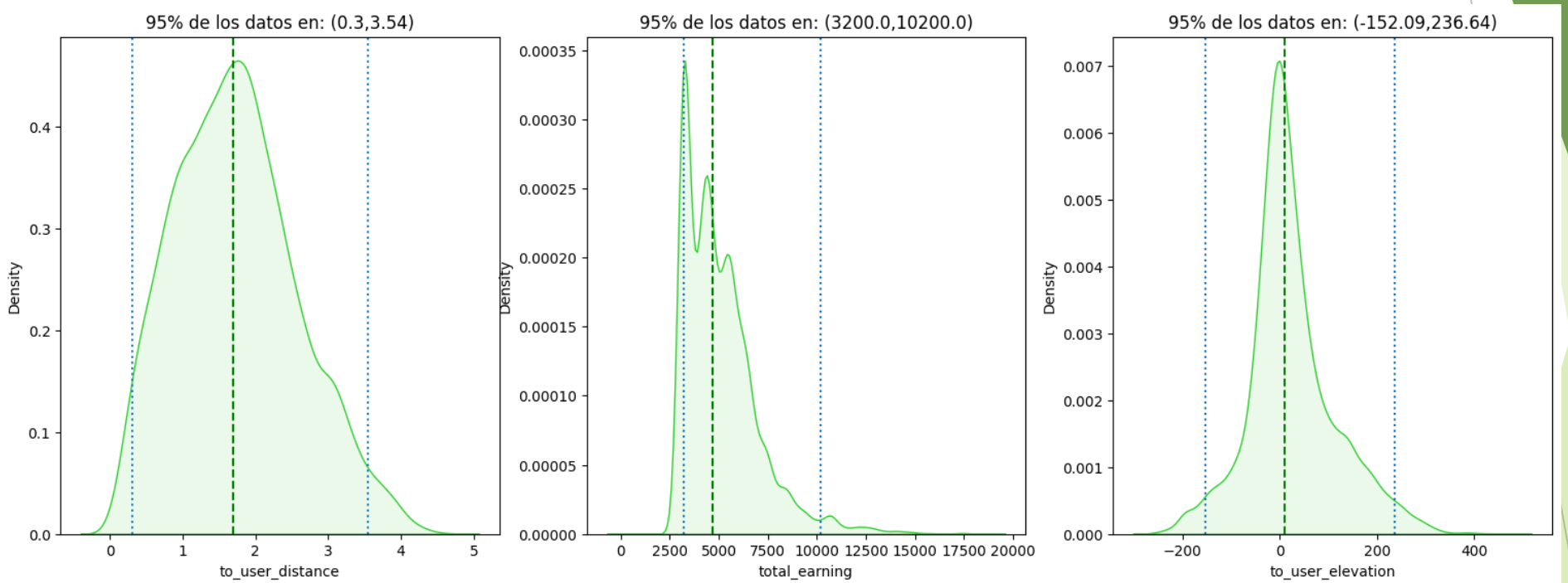
Taken (1)	Not taken (0)
115860	9689

1.2. What we learned from the data



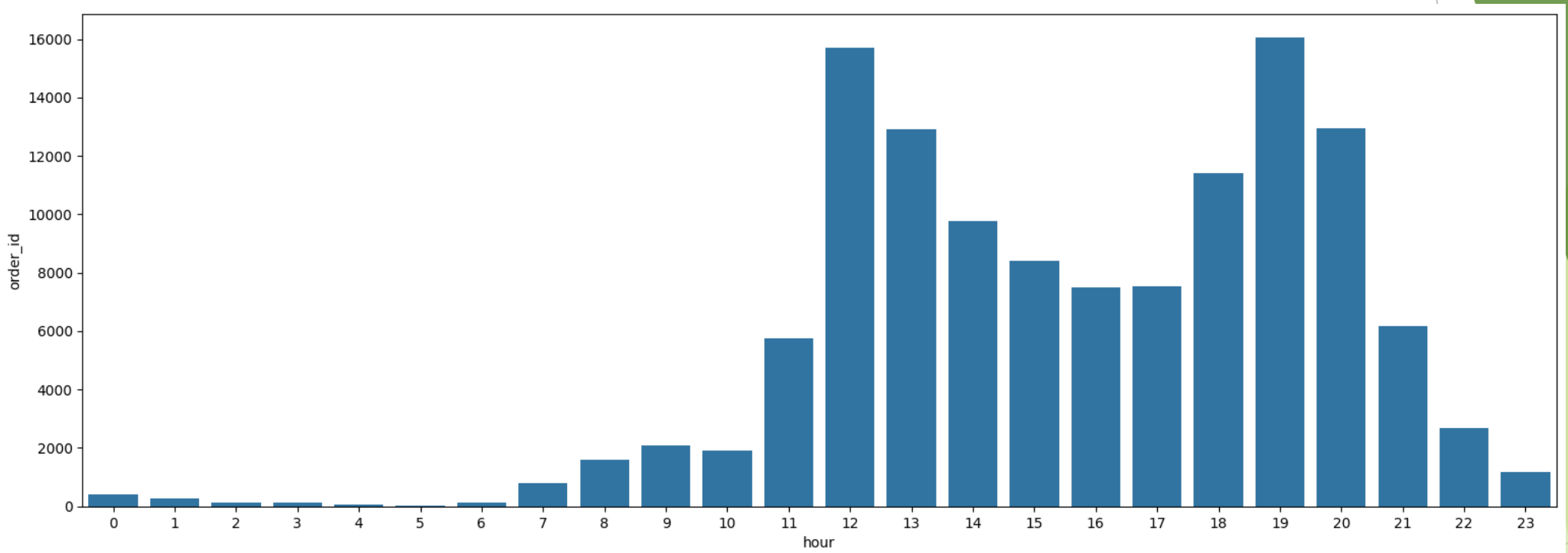
95% of the clients are between 0.23km and 3.42km away from the store.
95% of total earnings of a Courier is between $3200 and $9500
95% of user elevation to the store is between -135.42m and 236.02m

1.2. What we learned from the data



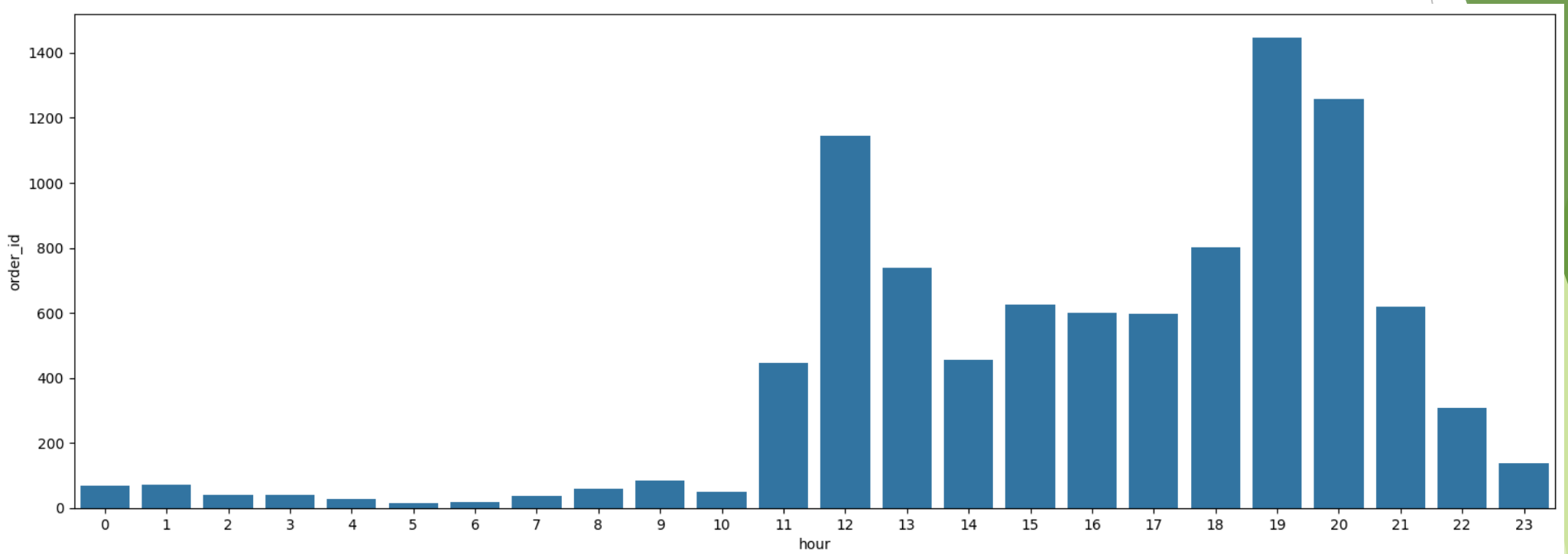
Checking only non taken orders:
95% of the clients are between 0.3km and 3.54km away from the store.
95% of total earnings of a Courier is between $3200 and $10500
95% of user elevation to the store is between -152.09m and 236.64m

1.2. What we learned from the data



The peak hours in terms of orders created usually matches the lunch time and the dinner time, with steady levels in the afternoon and the early night (around 20 to 21 hours).

1.2. What we learned from the data



Checking only non taken orders, we have critical peaks at the lunchtime and most of the non taken order are concentrated in the late-afternoon – night hours.

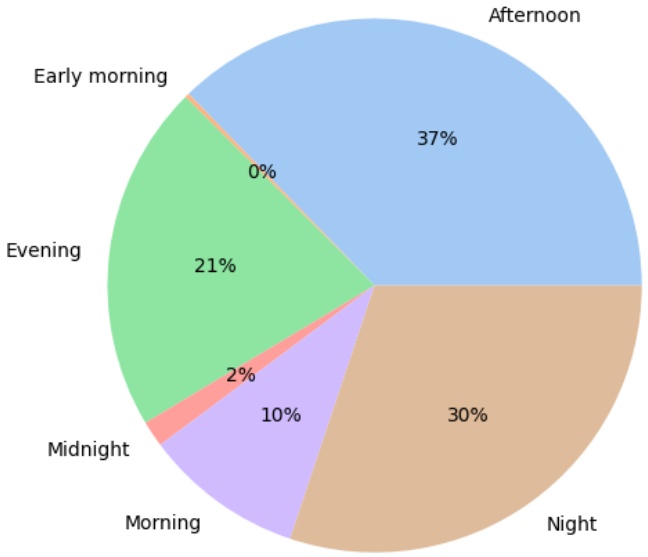
1.2. What we learned from the data

If we categorize the hour of the day when the order was originally created, we can gain some general insights among taken and non taken orders.

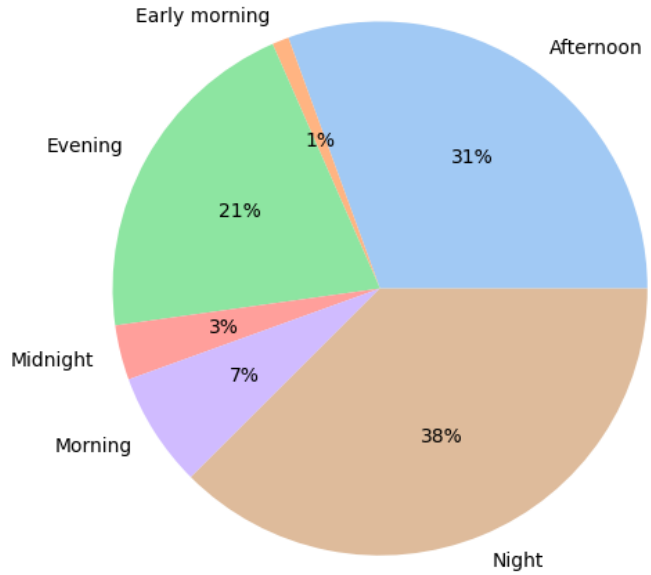
Category	Hours
Early Morning	3 – 6
Morning	7 – 11
Afternoon	12 – 15
Evening	16 – 18
Night	19 – 22
Midnight	23 - 2

1.2. What we learned from the data

Total orders



Non taken orders



1.2. What we learned from the data