

Jenny Huang <jh7717@nyu.edu>

DS project 3 - Jenny

1 message

Rebecca yu Su <rebecca.y.su520@gmail.com> To: Jenny Huang <jh7717@nyu.edu> Fri, Jun 2, 2023 at 10:33 PM



Hi Jenny,

Please check if you have access to the data for this project first, and please let me know if you have any issue accessing the three .csv files ASAP, thank you.

Here's the details for your 3rd project:

INTRODUCTION

In this assignment, we're asking you to use the given data to complete the following presentation. Your presentation will be evaluated based on your technical analysis and interpretation of the data using the .csv files and information provided.

INSTRUCTIONS

Deliverable: Submit a 20-30 minute presentation that analyzes a driver's Lifetime Value (LTV) using the data provided, and identifies trends and insights of the factors that impact LTV.

Assumptions to keep in mind when creating your presentation:

- Turn in a presentation that can be read and interpreted without a voiceover. Please include all working materials, assumptions, and supporting analyses in an appendix.
- Assume this presentation will be delivered to a cross-functional audience.
- The data provided is a snapshot in time and does not encompass all rides data. Drivers may have continued to drive after the snapshot was taken.
- Be sure to state any and all assumptions you make in your assignment.

Questions to consider when building your presentation:

- Recommend a Driver's Lifetime Value (LTV) (i.e. the value of a driver over the entire projected lifetime of a driver). State the components of your LTV equation and the assumptions made to calculate your final LTV.
- What is the average projected lifetime of a driver? That is, once a driver is onboarded, how long do they typically continue driving with the company.
- Number of Rides and Active Drivers over time (in weekly buckets)
- Explore how drivers churn once they start driving with the company.
- Are there any predictive indicators for driver churn?
- o % of Active Drivers over time (in weekly buckets, cohorted by onboard week)
- Segment the driver population to identify driving behavior that may lead to churn
- What are the main factors that affect a driver's lifetime value? Please support your answers using your data analysis.
- How does the demand for rides impact the number of drivers?
- What actionable recommendations are there for the business?

DATA DETAILS

Attached are three CSV files containing the data you will need for this assignment. More information below.

The three CSV files attached contain the following data:

driver ids.csv

driver id: Unique identifier for a driver

driver_onboard_date: Date the driver was approved to drive with the company

ride_ids.csv

driver_id: Unique identifier for a driver

ride_id: Unique identifier for a ride that was completed by the driver

ride_distance: Ride distance in meters ride_duration: Ride durations in seconds

ride_prime_time: PrimeTime multiplier (%) applied on the ride

ride timestamps.csv

ride id: Unique identifier for a ride

event: event describes the type of event (see below)

timestamp: Time of event You can assume that:

- All rides in the data set occurred in San Francisco.
- All timestamps in the data set are in UTC.
- This is a snapshot of onboarding and ride history data for a 3 month period. You may assume the data is complete for these drivers during the given time period; however, additional rides may have occurred before and after the time period included in the data.

Here's an overview of the event types:

requested_at: passenger requested a ride

accepted_at: driver accepted a passenger request

arrived_at: driver arrived at pickup point

picked_up_at: driver picked up the passenger

dropped_off_at: driver dropped off a passenger at destination

You can make the following assumptions about the rate card:

Base Fare \$2.00

Cost per Mile \$1.15

Cost per Minute \$0.22

Service Fee \$1.75

Minimum Fare \$5.00

Maximum Fare \$400.00