You'll find three CSV files attached with the following data:

data/driver_ids.csv

driver_id - Unique identifier for a driver
driver_onboard_date - Date on which driver was on boarded

data/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 applied on the ride

data/ride_timestamps.csv

ride_id - Unique identifier for a ride
event - event describes the type of event (see below)
timestamp - Time of event

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 XYZ_Taxi 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

After exploring and analyzing the data, please answer the following questions:

● Recommend a Driver's Lifetime Value (i.e. the value a driver to over the entire projected lifetime of a driver). You can assume that XYZ_TAXI takes 20% of each fare.

● What are the main factors that affect a driver's lifetime value?

● What is the average projected lifetime of a driver? That is, once a driver is onboarded, how long does he/she typically continue driving with XYZ_TAXI.

● Explore how drivers churn once they start driving with XYZ_TAXI Are there any predictive indicators for driver churn?

● Do all drivers act alike? Are there specific segments of drivers that generate more value for XYZ_TAXI than the average driver?

● What relevant business uses does this metric have?