Introduction

Cruise is building the world's most advanced self-driving vehicles to safely connect people with the places, things, and experiences they care about. One of the things that make Cruise unique is that we are building the world's largest fleet of all-electric self-driving cars, which also means we're investing in electric vehicle (EV) infrastructure in at a large scale. Cruise already owns nearly 40 percent of all EV fast chargers in San Francisco, and now we are building the largest EV fast charger station in the country right here in San Francisco.

The purpose of this assignment is to learn how comfortable you are with data and how you approach real-world questions under time constraints. **Please clearly articulate your assumptions, framework, methods of analysis, and justification for your recommendations. You can choose to work in any language or tool (Python, SQL, R) that is familiar to you as long as your code is adequately documented. Please provide a summary write-up of your results, methods, and recommendations, and a file containing your code.**

Part 1: SQL

Question 1

Our team has identified 10 potential locations in San Francisco to build the future EV fast charger station. You have been given some basic information on these locations and a table with ride service receipts. We would like to choose a location that is within the proximity of the most pickup and drop-off activities. Please develop a SQL query to identify the best location for charger station out of these ten. Feel free to write it in any SQL dialect you prefer.

Table 1: locations

Column name	Data type	Description	Example
id	string	Unique identifier of each location	d41d8cd98f00b204e9800
name	string	Name of location	Dolores Park
lat	FLOAT64	Latitude of location	37.759086
long	FLOAT64	Longitudinal of location	-122.426987

Table 2: cruise_ride_receipts

Column name	Data type	Description	Example
id	string	Unique identifier of each trip	5a1b935e8a4377883b3a7
request_time	timestamp	Timestamp of trip request in UTC	2019-07-01 02:39:33
pickup_time	timestamp	Timestamp of pick up event in UTC	2019-07-01 02:41:03
dropoff_time	timestamp	Timestamp of dropoff event in UTC	2019-07-01 04:03:56
receipt_sent_time	timestamp	Timestamp when email receipt was sent	2019-07-01 04:04:30
pickup_city	string	City of pick up location	San Francisco
pickup_state	string	State of pick up location	CA
pickup_address	string	Address of pick up location	1201 Bryant Street
dropoff_city	string	City of dropoff location	San Francisco

dropoff_state	string	State of dropoff location	CA
dropoff_address	string	Address of drop off location	3380 21st Street
pickup_zipcode	INT64	Zip code of pickup location	94103
dropoff_zipcode	INT64	Zip code of dropoff location	94110
order_total	FLOAT64	Subtotal of trip	18.12
taxi_ride_distance	FLOAT64	Total distance of ride	2.3
vin	string	VIN number of autonomous vehicle	5G21A6P0XL4100014
car_name	string	Autonomous vehicle name	Poppy
trip_request_lat	FLOAT64	Latitude of trip request location	37.769886
trip_request_long	FLOAT64	Longitude of trip request location	-122.409705
pickup_lat	FLOAT64	Latitude of pick up location	37.769950
pickup_long	FLOAT64	Longitude of pick up location	-122.410363
dropoff_lat	FLOAT64	Latitude of dropoff location	37.756929
dropoff_long	FLOAT64	Longitude of dropoff location	-122.422802
user_id	string	Unique identifier of the user	bd60635614a36d22c8ef6
trip_type	string	Trip type: could be one of "ridesharing", "doordash", "grocery", "testing", and "demo"	ridesharing
status	string	Trip status: could be one of "completed", "cancelled", and "aborted"	completed
Trip rating	INT64	Rating of the trip	5

Question 2

Cruise offers its employees unlimited free autonomous vehicle ridesharing service. The team is now curious about how many trips are generated by Cruise employees for commuting to/from our 1201 Bryant Street office. You can assume that the trips that requested by Cruise employees must start or end within 0.5 miles from Cruise HQ (the "name" field in Table 1) from Monday to Friday between 7am to 10am (morning) and 4pm to 10pm (evening).

Please design a query that returns the number of pickups and drop-offs requested by Cruise employees broken down by morning/evening, and by date. The output should look like this:

Date	Number of Cruise Employee Pickup (7am - 10am)	Number of Cruise Employee drop-offs (7am - 10am)	Number of Cruise Employee Pickup (4pm - 10pm)	Number of Cruise Employee drop-offs (4pm - 10pm)
2019-07-01				
2019-07-02				
2019-07-03				
...				

Part 2: Analytics

The second part of this exercise involves analyzing a publicly available Travel Decision survey data from residents of San Francisco and surrounding areas to make critical future product decisions. This dataset can be accessed here: https://data.sfgov.org/Transportation/Travel-Decision-Survey-Data-2017/cxi3-57f8

While this data is extensive, the goal of this exercise is to articulate a data driven case for our chief product officer and write a report with the following set of recommendations:

1. Whether or not we should launch a ridesharing service for residents of San Francisco that will operate within the boundaries of the city, or for residents outside San Francisco to commute into the city?
2. Are there any specific groups of customers (age group, gender, income group, or any combination thereof) that we should target as our first set of customers and why.

Please clearly articulate your assumptions, framework, methods of analysis, and justification for your recommendations. You can choose to work in any language or tool (Python, SQL, R) that is familiar to you as long as your code is adequately documented. Please provide a summary write-up of your results, methods, and recommendations, and a file containing your code.