

## Data Analysis Exercise

*We expect this exercise to take approximately 2-3 hours.*

May Mobility use Lexus vehicles for our operations that are run in shifts. A shift can be thought of as a cycle from ignition-on to ignition-off. A variety of data is produced during a shift. One of them is time-sampled data for horizontal vehicle velocity and acceleration that we will be using for this exercise. Provided with the exercise are two datasets.

The dataset file `vehicle_kinematics.csv` consists of the following columns:

- **vehicle\_id** (no units)
- **time** (seconds since shift started)
- **velocity** (m/s)
- **acceleration** (m/s<sup>2</sup>)

The dataset file `vehicle_gps.csv` consists of the following columns:

- **vehicle\_id** (no units)
- **time** (seconds since shift started)
- **latitude** (degrees)
- **longitude** (degrees)

**Please return an output file with the answers to the following questions. Also please provide the code with comments in a script or notebook that was used to generate the output file.**

### Question 1

Please load the dataset and provide a summary of the data provided. Does the dataset match your expectations about the data?

### Question 2

For each vehicle in the dataset, calculate and print out the total (absolute) distance traveled in meters, rounded to the nearest meter. Assume that the velocity is constant between time samples. All values are floating point values, but the output should be an integer. Provide an analysis of your solution including algorithmic complexity and failure cases.

### Question 3

Calculate and print out the total distance travelled by the fleet (all vehicles) rounded to the nearest meter. The output should be an integer.

#### Question 4

Are there any data insights that you can provide that would be of interest to stakeholders of the company? Feel free to add a plot for clarity if you think it is needed.

**Note:** State any assumptions you make about the data in your solution that was not already given in the question.