Data Science Telematics Technical Test

- You are given 48 hours to complete this test, but it should take around 4-6 hours to complete, we're not expecting an exhaustive analysis.
- Where you're not sure what approach to take, or you're running out of time, make sensible simplifying assumptions to get to an answer. You won't be penalised for this, but we will want you to be able to justify them.
- You don't need to document every step of your working, but be ready to talk through the method you've used.
- You are free to use whatever tools you like to answer the questions (e.g. Python/R etc.), but you should be able to present your findings and any code/calculations to us during the interview.

Data

`data/data_trips.zip` directory

This folder contains 37 csv files, each of them with acceleration sensors recordings during a trip.

- `timestamp` : `int` unix timestamp of the recorded event
- `ValueX` : `float` value of the smartphone acceleration sensor on the x axis
- `ValueY` : `float` value of the smartphone acceleration sensor on the y axis
- `ValueZ` : `float` value of the smartphone acceleration sensor on the z axis

The name of the files contain the `trip_id` information, which is a unique trip identifier.

`data/trip_classes.csv.zip`

This dataset contains the type of the trips that are in the `data_trips.zip`.

- `trip_id` : `str` unique trip identifier
- `trip_type` : `str` type of trip: car or walk_bicycle_sill
- `trip_data` : `str` name of the csv where the acceleration data of the respective trip is.

Questions

1. How is the noise of the trip `trip_29.csv` ?
2. How would you access the data quality?

Task

Your task is to produce a model able to identify if a coming trip was recorded from a car or not.