Data Science Challenge

Introduction

The Floow absorbs telematics data from various sources and processes this to enable an understanding to how the vehicle is being driven and any events that occur on the journey. This data can be volatile and can contain anomalies, this test is designed to assess your ability to clean this data and understand what it might imply.

The Challenge

The functional goal of the challenge is to analyse a number of journey files, to understand the data and:

- identify the trips containing an ‘event’ that might be an accident
- identify the time of the ‘event’
- allocate a notional severity index to each ‘event’
- allocate a notional confidence to the ‘event’ being an accident

The technical part of the challenge is to provide a program (ideally in Java, Python or R) that can be run on this or other datasets to achieve the same results.

You should provide a written document describing the methodology that you used along with the results and the source code for your program.

In addition, please prepare a 10min presentation (in PowerPoint or similar), with a target audience of non-technical people to explain their findings.

The challenge will be assessed by comparing your results against known outcomes.

The following bonuses will be taken into consideration if achieved:

- Accuracy & efficiency of the program when run against a similar test set of data
- Any notes considering trade offs between efficiency and accuracy
- Evidence of understanding of the data
- Any advanced data manipulations performed

Data provided:
25 Journey files containing:
Unix Epoch Timestamp, Latitude, Longitude, Altitude, Lateral accuracy, GPS Time, Speed & XYZ accelerometer values.
Accelerometer values are not orientated to a fixed vertical and relate to the axis of the phone that the data was sourced from. The phones are loose in the vehicle so the actual orientation can change over time.
All devices use the same GPS system but some measure the XYZ accelerometer values in G whilst others use m/s2. You will need to identify which are which and make any necessary adjustments.

The Deadline & Submission

Your submission is due as outlined in the email accompanying this challenge. You must provide your results along with the code and all instructions required to successfully run your solution and describe how to view the final results.

Your submission will be assessed internally following which you will be invited along to present and discuss the submission and results.
By accepting this test you also agree to delete all journey data provided no more than two weeks from receipt of the data.