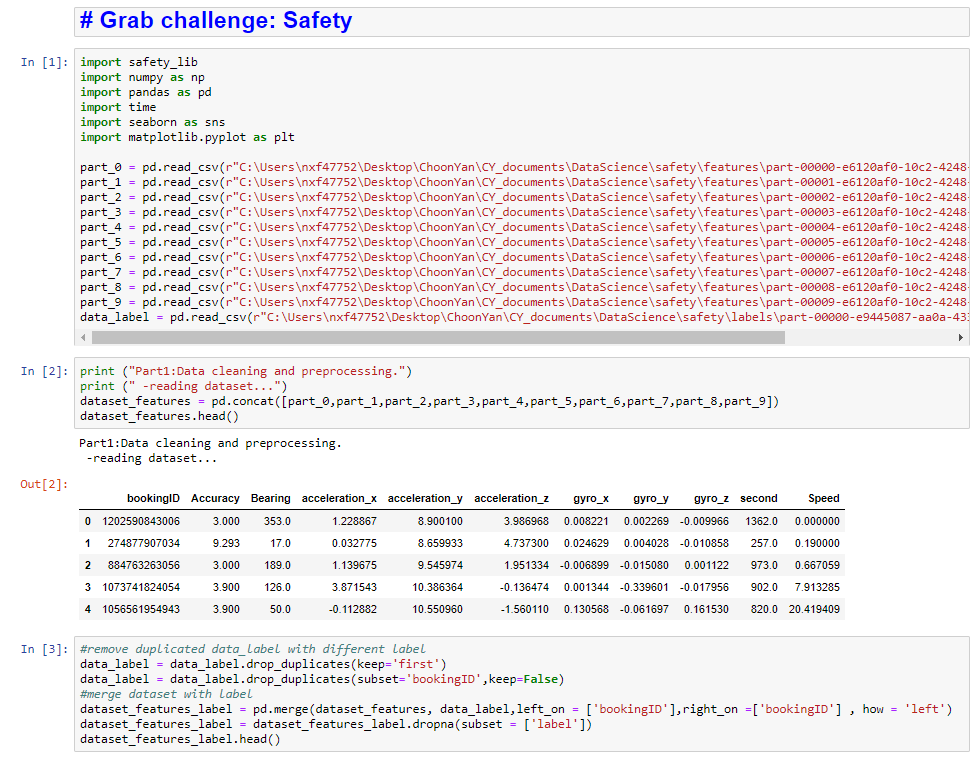
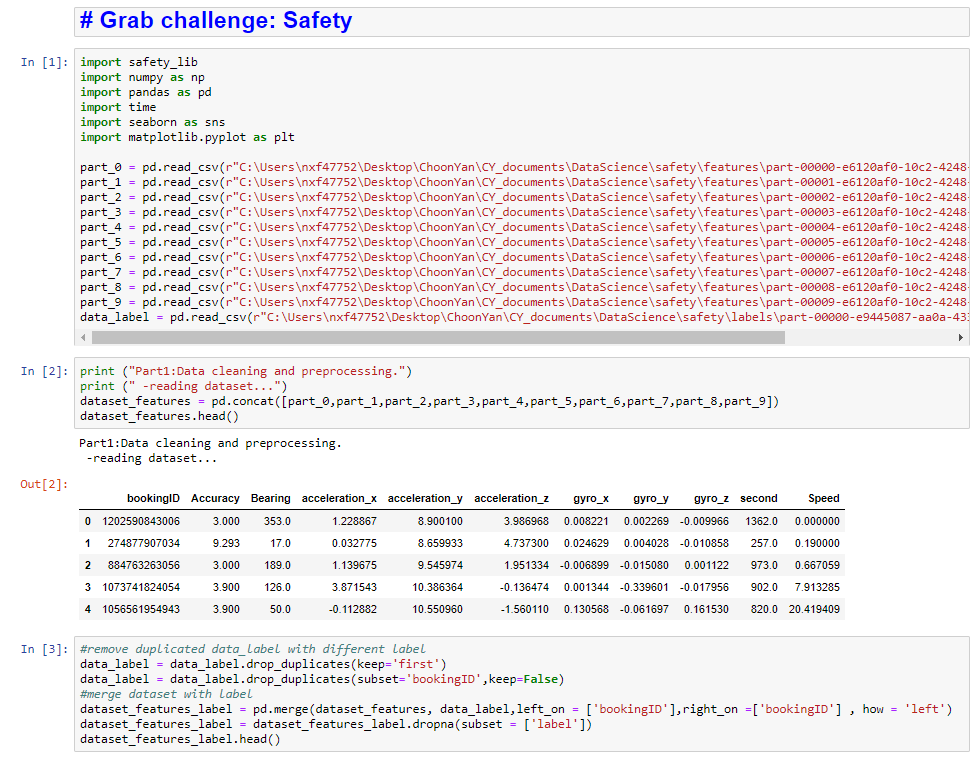
Grab challenge: Safety

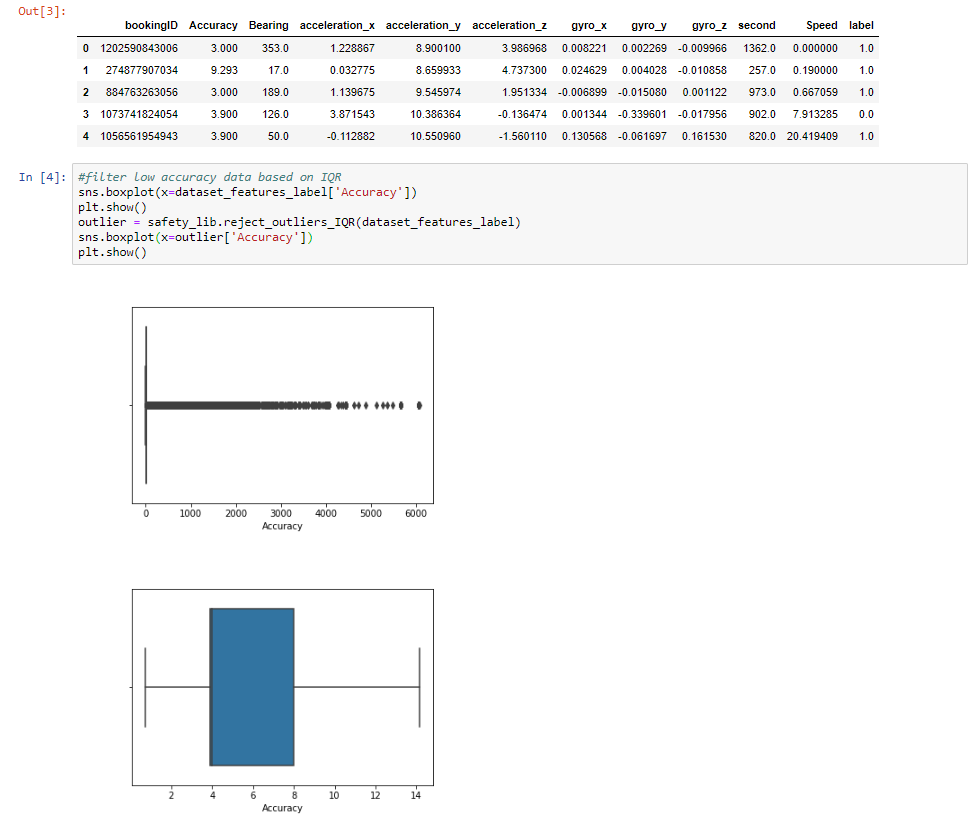
Based on telematics data, how might we detect if the driver is driving dangerously?

1. Concatenate and read out the time series Telematics data:

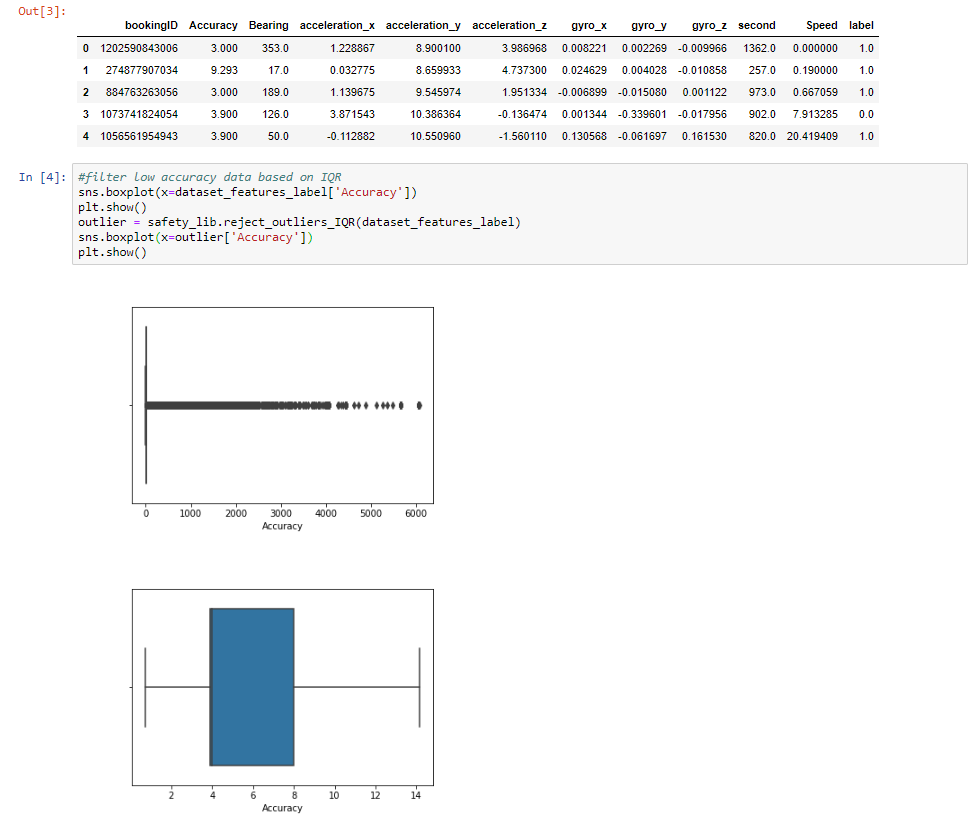


1. Duplicated booking ID with different data label were removed to avoid model accuracy impacted.

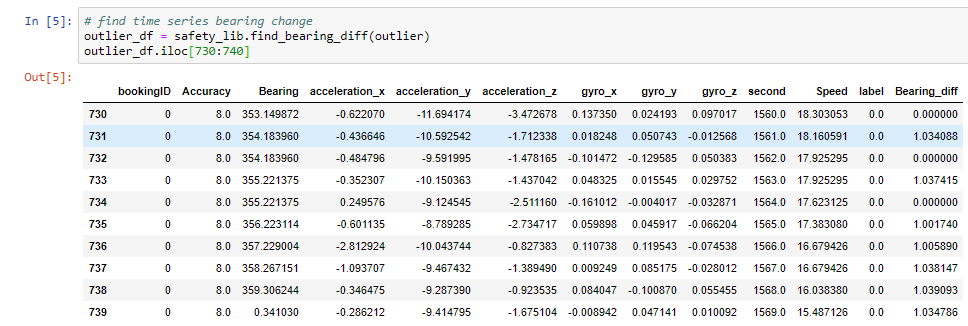




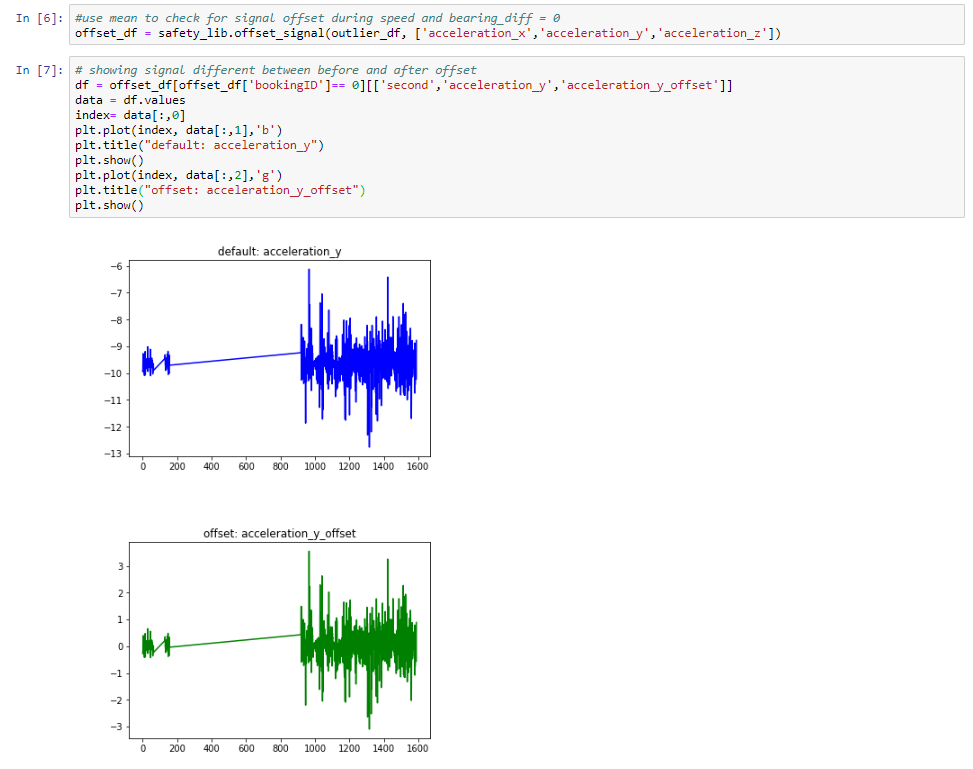
1. Telematics data with low accuracy is removed based on Interquartile range.



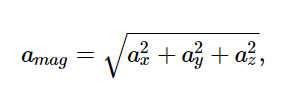
1. Checkout the bearing changes/differences by second by comparing current Bearing vs previous Bearing.

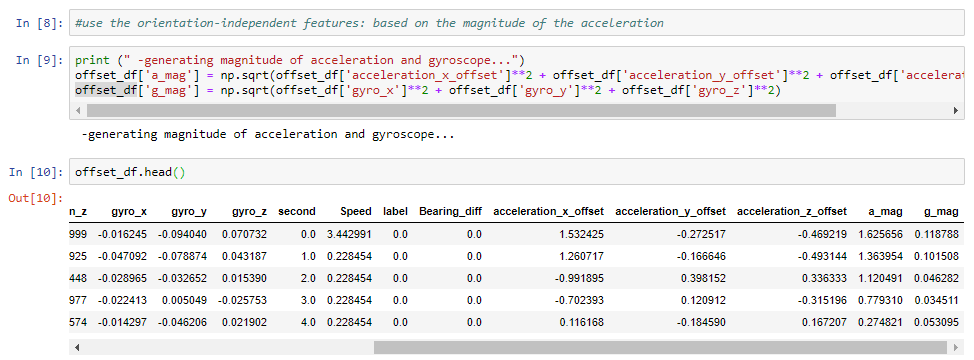


1. This step will offset the x,y,z by taking mean of data during speed and bearing different are 0. This method can normalized the data regardless of the phone angle/orientation placed in the vehicle.



1. Create feature for magnitude of acceleration using the orientation independent method.



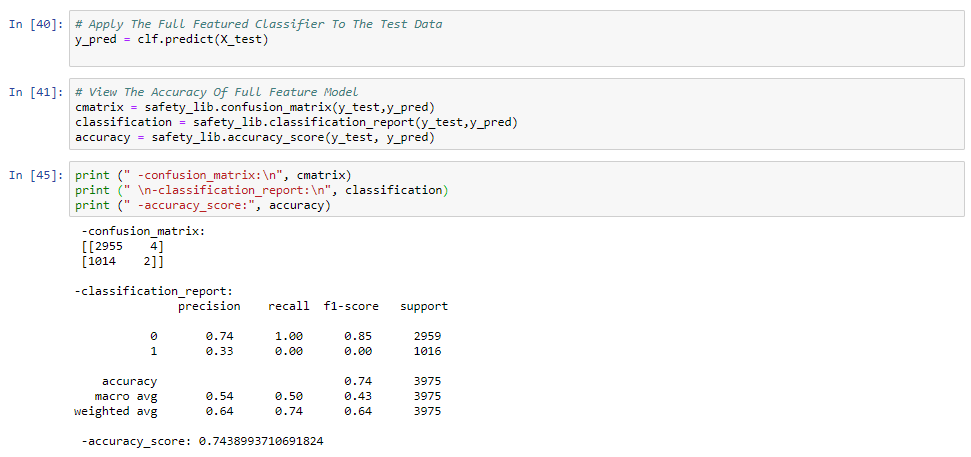


1. Modelling:

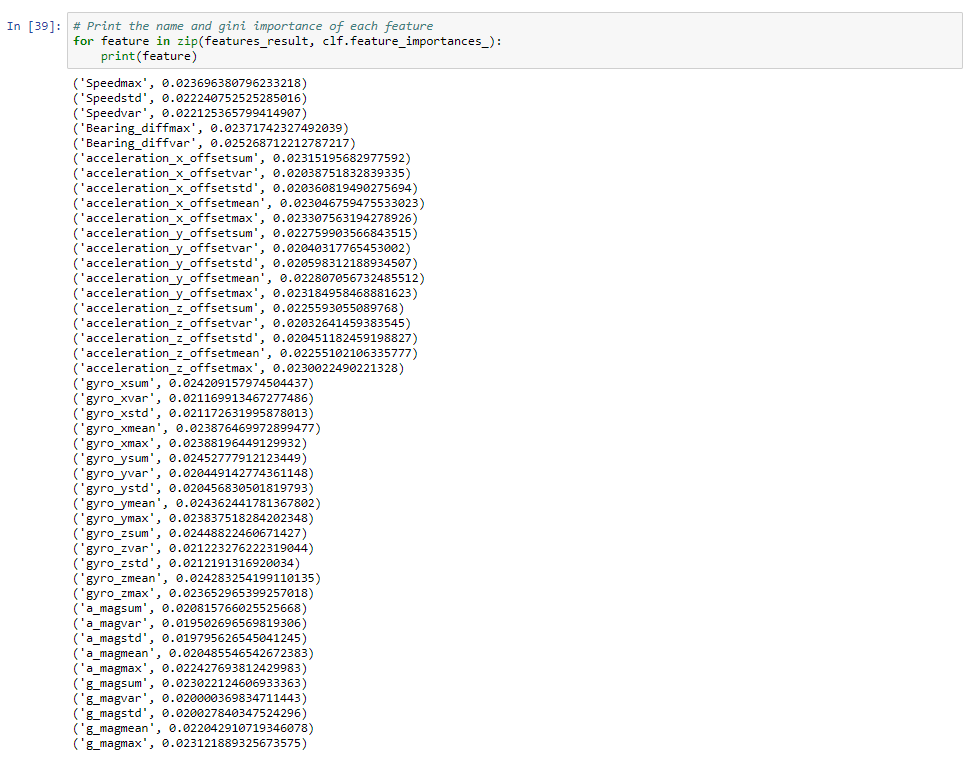
* Two common classification algorithms (Random Forest and Naïve Bayes) are used in this challenge.
* Based on the accuracy result, Random Forest (RF) is 74.39% while Naïve Bayes (NB) is slightly lower with accuracy 73.26%.
* However, in term of computation time in running the model, RF took ~15mins to run while NB only 0.15s.

1. Running with RF:





* Checking the features importance



1. Running with NB with accuracy 73.26%. Computation time is quick: 0.15s only.

