1. Use the dataset (**DfTRoadSafety\_Accidents\_2014.csv**) or download it from the <https://data.gov.uk/dataset/road-accidents-safety-data>
2. Build a model that predicts if a police officer is likely to attend an accident or not. The variable’s name is *Did\_Police\_Officer\_Attend\_Scene\_of\_Accident*. There is a dictionary that explains all the features included in the dataset called Road-**Accident-Safety-Data-Guide.xls.**
3. Produce a report that summarises your findings. Things to Include in the report:
   1. Describe the features and the target included in the dataset. What insight did you gain from working with the data?
   2. Explain how you transformed the features
   3. The validation approach taken
   4. Which performance metrics did you use and why?
   5. Which algorithms did you use?
   6. Which were the most important features?
   7. How useful is the model from a practical point of view?
   8. What might you do differently if you had more time/resource?
4. This is not meant to be an exercise to develop a ‘final solution’ but more a first pass, to understand how you would approach this kind of exercise: Suggested time breakdown:
   1. 3-4 hours getting data and building model
   2. 1 hour creating the report/presentation
5. Submit the report and any source code (these may be part of the same document if you are working with e.g. jupyter notebook).