# Coursera Capstone Project

## Intro / business problem

Target audience: Local authorities in UK

Problem Statement: predict likelihood of severe accidents on roads on a specific day given the weather conditions. This could also be used to model the reduced risk if temporary measures are put in place (i.e. create a prediction with lower speed limits).

## Data

Data source(s): [UK Department for Transport – Road Safety Data: 2018 dataset](https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data)

I am planning to use the following columns from the “[Road Safety Data – Accidents 2018](http://data.dft.gov.uk/road-accidents-safety-data/dftRoadSafetyData_Accidents_2018.csv)” Dataset:

* Severity of accident
* Police Force
* Road class
* Road Type
* Speed limit
* Day of week
* Time of day \*may be crossover with lighting conditions, however not 1-2-1 as lighting during commute hours in winter will be worse
* Date – extract month \*may be crossover with weather – not sure if needed, visability & weather probs give enough context
* Light Conditions
* Weather Conditions
* Road Surface Conditions (e.g. ice, standing water)
* Urban/Rural area

## Methodology

### Exploratory Data Analysis

### Inferential Statistical testing

### Machine learnings used

## Results

## Discussion

## Conclusion

## Expansion Ideas

Expansion ideas:

* predict who the casualty is - the driver/passenger/pedestrian
* include some of the driver data, so that drivers could look at the personal risk of their journey allowing for the route taken & vehicle type etc

Additional info from vehicle/casualty datasets of interest:

* Junction Detail
* Age of driver
* Vehicle Type
* Age of vehicle
* Engine capacity
* Journey purpose
* Take into account drink driving & speeding
* Casualty age
* Casualty sex
* Casualty category (driver/passenger/pedestrian)
* Casualty severity