Identifying road segments with potential safety hazard

Challenge Provider: PSE

PSE is a company expert in Data Science and Advanced Research, that has been on the market since 1994. Since the beginning of its activity, PSE has been dedicated to Advanced Analytics. Today, this materializes in the implementation of technology solutions, advanced market research, and providing consultancy and data science services.

Context

According to the 2017 UN's Global Mobility Report, "road transport claims the bulk of transport-related fatalities worldwide: it accounts for 97 per cent of the deaths and 93 per cent of the costs. On roads, the fatality risk for motorcyclists is 20 times higher than for car occupants, followed by cycling and walking, with 7 to 9 times higher risk than car travel, respectively. Bus occupants are 10 times safer than car occupants. Rail and air are the safest transport modes."
So, decision agents must know where to "improve the safety of mobility across all modes of transport by avoiding fatalities, injuries, and crashes from transport mishaps".
That is why creating a model that identifies areas or roads with more considerable risks will help where to invest in road safety and which actions are needed.
The goal is not to predict accidents. Instead, it is to identify hazardous roads, so decision agents take measures on where they need to act on their cities to improve road safety.

Goals

Understand which areas of the city of Porto and/or Lisbon are riskier.

Outcome

A heatmap or intensity model that considers traffic intensity data, characteristics of road segments and their speed profile. This heatmap should identify the most conflictive areas or roads.
.

Available Resources

As a reminder, all the data resources can be found here: https://bit.ly/wdl-data. You can also use any open, free and legally available data - Even if it is for another city.

Risk Profiles

A road network level link database with information about its' characteristics, including the average speed and the average daily traffic intensity. It is calculated between April 2019 and March 2020.

Provider: PSE

Submissions

Deadline: 01 - 05 - 2021 @ 14h00 GMT + 1

Don't forget that you will need to deliver the report **using the template provided** (see below) and a 1-minute summary.

Submission template: http://bit.ly/wdl-template

Tips

- Think about what drives people to move around;
- What are the realities that exist in cities that may clash with big traffic intensity roads?
- Do the state-of-the-art research; there might be already a lot of interesting things done before;
- Try to fill in the template from start to finish with a very simple dummy solution first and iterate afterwards;
- You can use other data sources, such as weather, which can be very useful;
- If possible, don't forget to explain the predictions of your model.