UK Road Traffic Collision

Nonparametric Statistics project

Valeria Iapaolo, Oswaldo Morales, Riccardo Morandi, Abylaikhan Orynbassar,

15 November 2023

POLITECNICO
MILANO 1863

Dataset

Vehicle information of each accident

Geolation of each accident

Road type, weather, lightning situation

Details on the casualties associated with each accident.

2005 → 2022

Goals

Main **stakeholders** and **research questions:**

Car manufacturers

They could be interested in:

- Which parts of the vehicle should be strengthened in future models?
- What type of spare components should be readily available?
- Are there periods during the year when spare parts are in higher demand?

Urban planners

Many road characteristics could contribute both to the number and to the severity of accidents, amongst others:

- types of junctions and roundabouts;
- pedestrian crossings;
- speed limit;
- the manoeuvres that a car takes are influenced by the street layout;
- characteristics of the area;

Which are the most relevant road features that a planner should take into account when designing a new road?

Are there possible measures that urban planners could take to mitigate future accidents?

First responders

Police force and paramedics need to react quickly in the case of an accident.

Some factors could influence the number of crashes, such as:

- time of day and season of the year;
- type of road;
- geographical location of the road;
- weather conditions.

How should the authorities schedule the shifts in hospitals and police stations in order to better cover the crashes?

How should police patrols be placed in the country to maximise effectiveness?

Methods

Next steps:

- Feature selection to understand the relevant features in predicting the number and the severity of the accidents;
- Capture the influence of the selected features using **GAM regression model**
- Identify patterns in the time distribution of crashes modelling using **functional clustering.**
- Spot possible **outliers** in the data (e.g. covid year)
- Using survival analysis to study the time to crash for each observed vehicle
- Since the accidents are geolocated, a **spatial statistical model** could give further insight