# CASA0006 Research Proposal

Student Number: 22186878

Proposed research question:

|  |  |  |
| --- | --- | --- |
| **Research Topic** | **Pros** | **Cons** |
| Can crash severity be predicted through contextual factors? | * Simple analysis (regression tree) | * Low importance/utility (i.e., would this model have any actual use?) |
| Spatial clustering of traffic accidents in New Zealand to identify ‘black spots’ | * Full lecture devoted to this topic (wk10), so should be acceptable | * Simple methodology (covered entirely within T1 subjects GIS/QM) |
| Causal inference for the effect of transport policy (‘Road to Zero’) on crash frequency | * Complex methodology; applying a novel technique | * Requires significant background research on policy environment, need to identify eligible regions for comparison |
| Predicting crash counts using ML model | * Useful application of predictive modelling, easy to argue | * Would need to aggregate raw data in some way to produce a count per place/time/region? |
| Evaluating implementation of road safety policy using regression discontinuity design | * Complex methodology; novel technique | * Does not build skills in traditional ML model building (essentially just advanced regression) |

Dataset:

Waka Kotahi (NZ Transport Agency) Crash Analysis System (CAS) Data

Contains geographic coordinates, date, severity, and contextual details of road crashes in New Zealand from 1 January 2000.

Method/approach:

Build a random forest (RF) model using a separated testing and training dataset (training will be taken from years 2000 – 2015; testing will use data from 2016-2020).

References: