**Finding patterns in crash data**

*Proposed by Richard Arnold*

This project is about finding patterns in publicly available data about road crashes from Waka Kotahi, the New Zealand Transport Agency:

https://opendata-nzta.opendata.arcgis.com/datasets/NZTA::crash-analysis-system-cas-data-1/about

<https://catalogue.data.govt.nz/dataset/crash-analysis-system-cas-data5>

Waka Kotahi have posted a dataset listing crashes on NZ roads since 2000. The first link also includes a link to field definitions for each of the variables listed in the dataset.

Waka Kotahi also provides an anonymised monthly copy of the Motor Vehicle Register, which lists all the vehicles registered in New Zealand.

https://nzta.govt.nz/resources/new-zealand-motor-vehicle-register-statistics/new-zealand-vehicle-fleet-open-data-sets/#data

https://opendata-nzta.opendata.arcgis.com/datasets/NZTA::motor-vehicle-register-api/explore

StatsNZ also have a wide range of datasets describing different aspects of New Zealand, including population numbers, geographical information and business, economic and cultural indicators:

https://www.stats.govt.nz/all-topics/

These datasets can be combined to look for geographical and temporal patterns in crash statistics, or to assess what types of vehicles tend to be involved in crashes.

Students will need to

* Integrate multiple large datasets, some fine-grained and some wide-scale
* Analyse geographical data, count data or time series data
* Carry out an exploratory analysis, choosing a subset of variables and features to report
* Determine which datasets need to be integrated to answer specific questions and/or make specific predictions
* Carry out the analyses required to answer specific questions, and display the chosen model(s) and results
* Consider any ethical aspects when reporting the results