Data Wrangling Project:

I have decided to do wrangling project in NZ transport dataset:

1. First, I decide to check road infrastructure dataset

* Road conditions

<https://catalogue.data.govt.nz/dataset/road-condition>

* Number and length of road built.

<https://catalogue.data.govt.nz/dataset/number-and-length-of-roads-built-reconstructed-or-seal-extended>

* Number of length of bridges built and renewed.

<https://catalogue.data.govt.nz/dataset/number-and-length-of-bridges-built-and-renewed>

1. Investment on Road (expenditure)

* Expenditure on Public Transport services and Public Transport Infrastructure.

<https://catalogue.data.govt.nz/dataset/expenditure-on-road-maintenance-operations-and-renewal-activities>

<https://catalogue.data.govt.nz/dataset/expenditure-on-activity-planning-regional-planning-studies-strategies-and-models>

1. Crash Analysis on the region.

Analysis the data link is already in draft.

Road Conditions: It has region wise condition of the road segregated into urban, rural and all other roads throughout the year.

It has missing values.

No proper variables.

Number and length of road built: It has region wise information of road like the length and build over the last decade.

Investment on Road (expenditure):

Crash Information in each region.

Population of the region

# Team Falcon

We are planning to create a dataset by cleaning and then merging 8 different datasets into a relational data model obtained from 3 NZ government agencies, this would be helpful in finding the correlation between several probable causes of the road crashes in New Zealand.

The probable causes of the road crashes, we are planning to consider are traffic congestion, speed limits, timing (lighting conditions), weather conditions, road condition, age of the driver, state of construction, days (holidays), point of collision (intersections, roundabout, give way, stop sign), alcohol intake among several other reasons.

Data is from 2008 to 2018