**Project Overview - Project 1**

Project Title:

An Exploration Into the Food Industry in the Perth Metropolitan Area.

Group Members:

* Josh Cullen
* Hideaki Kaneko
* Matt Vella

Project Description:

Considering restaurants and cafes, is there a link between certain characteristics of eateries relative to typical socio-economic factors such as age, gender, income. Similarly, is a trend or representation which can be made between these eateries, socio-economic factors and other indicators such as postcodes.

Questions to Answer:

* Is there a link between food outlets and public transport access? E.g. number of restaurants, cuisine type
* Geographic primary ethnicities vs. popular cuisine types
* Age and Gender by Postcode / Council Area vs. Number of and Average Rating of Eateries and Cuisine Popularity
* Average Household Income by Postcode / Council Area vs Number of and Average Rating of Eateries and Cuisine Popularity

Datasets:

* Zomato API
  + [How to use the Zomato API in 3 Easy Steps [Tutorial] (rapidapi.com)](https://rapidapi.com/blog/how-to-use-zomato-api/)
  + [API Policy (zomato.com)](https://www.zomato.com/api_policy)
* Census Data
  + [Find Census data | Australian Bureau of Statistics (abs.gov.au)](https://www.abs.gov.au/census/find-census-data)
* General ABS Data
  + <https://www.abs.gov.au/>
* Kaggle
  + [Find Open Datasets and Machine Learning Projects | Kaggle](https://www.kaggle.com/datasets?search=perth)
  + E.g. Perth House Prices
    - Link between house prices and certain restaurant characteristics
* Public Transport Information
  + Availability of public transport to restaurant popularity
  + [Datasets - data.wa.gov.au](https://catalogue.data.wa.gov.au/dataset?theme=Transport)

Potential Limitations of Data sources:

* Zomato API
  + Live Data
  + Limited Call capacity per day (free account)
* Australia Government Data
  + May be lagging
  + Preliminary research would suggest 2016 would be the standardised year data source
  + Granularity of data

Breakdown of Tasks and Timeframes:

* Hideaki and Matt unavailable on Tuesday 27th of April
* All of group to conduct cursory review of data sources by Thursday 22nd of April
* Commence writing of code in class on 22nd of April.
* Division of work subsequent to this in preparation for weekend and following week.