**Project 1 proposal**

**Team number:** 3

**Members:** Franco, Joseph, Catherine, Nasrin

**Project title:** Melbourne Property Market Analysis

**Project description (broad)**

* Brief overview of your project (3 - 4 sentences). Provide an overview of the project scope and broadly define the research questions.
* We all know that Melbourne real estate is booming, regardless of the property type. In this project, we are going to use real world data to analyse what is driving this. For instance, does the convenience of the surroundings (e.g., Schools, supermarket, public transport) affect the house price etc.

**Research question to answer (specific)**

* What do you plan to show with your data analysis?
  + We want to do like for like comparison between the relationship between house prices and each factors below:
    1. Built date
    2. Time (quarterly comparison) ; House Prices vs Sold Date
    3. Suburbs
    4. Number of bedrooms
    5. Type of property, example – House, Townhouse, Unit
    6. Land Size/Building Size
    7. Method of sales: auction, private sales
    8. Agents (does agent make a difference?)
* We would also like to analyse the increase in pricing of houses according to no. of bedroom.
* To confirm the direct relationship between the house prices and distance to CBD)
* More detailed analysis to follow. (deep dive into each factors, with normalised comparison)
* What research questions do you plan to answer? That is, what are your hypotheses?
  1. HYPOTHESIS 1: The house prices jumped every quarter; Suburbs such as east/south-east jump quicker than north/west
  2. HYPOTHESIS 2: The provided amenities also affect the house price positively
  3. HYPOTHESIS 3:
* What do you expect to see as a result of your analysis?
  1. We expect to see quarterly price increases in property, and also different increases depending on the location.
  2. There Is direct correlation between the amount amenities and the value of property in that suburb.

**Datasets required**

* We are going to use 2 datasets:
  + Sales of houses in melbourne dataset (2016-2017)
  + List of all the schools in Melbourne, including the latitude and longitude info.
* Both datasets are downloaded from available website, both include latitude and longitude that will be used in the API analysis later.
* The datasets are both in csv format.
* The possible issue might arise given your choices above could be the age of the data – given it was from 2016-2017
* I believe our data is sufficiently large – first dataset contains more than 13k rows.
* The house price dataset had already removed the houses without prices, so I believe it’s fairly clean.

**Rough breakdown of tasks**

* Even though we have not allocated task, but we concluded that each of us will be doing 2 analysis and do plotting based on that.