# Meeting Minute - Week 7 (2)

**Meeting Time**: Monday 5th September

**Members Attended**: Yihan Bao, Chaoran Xu, Shunran Wang, Xiangyi Huang, Hairong Xu

What We Have Done Last Week (Week 6)

* **Property datasets**: Finished scraping all properties, read into dataframe, test for missing values.
* **ABS District Boundaries:** Found shp file for district boundaries, able to draw map using folium.
* **External datasets:** Population, population prediction, income, school location, gym location, health data(on district)
* **Public transport :** annual customer each train stop, location of train stop

TODO List This Week(Week 7):

**Looking For More Datasets:**

* Learn how to use scraping API: school, shopping centre, entertainment, clinic/hospital, park.
* Crime rate. (by Wednesday)
* Climate. (by Wed)

**Preprocessing:**

* Missing values
* Outliers
* Merge datasets based of features decided

Derive more features from existing datasets: distance to nearest train, distance to CBD

**Decide Feature(respect to research questions):**

1. Each property should have: attribute from property datasets, number of gym,school,shopping centre in 2km, number of trains 5km, number of bus/tram 5km, distance to nearest train, annual customer of nearest train, population density of the suburb, income of suburb.
2. (use data 2014-2018) Property median, avg population, annual customer of train stop, income.
3. Each suburb should have: population density, climate, average rental price, average income, number of gym,shopping,school in the suburb, distance to CBD, public transport convenience, clinic/hospital, number of factory, crime rate.