# About Dataset

Context

A multivariate time series has more than one time-dependent variable. Each variable depends not only on its past values but also has some dependency on other variables.

We have accumulated property sales data for the 2007-2019 period for one specific region. The data contains sales prices for houses and units with 1,2,3,4,5 bedrooms. These are the cross-depended variables. The chart illustrates these variables for houses:

Raw Data:
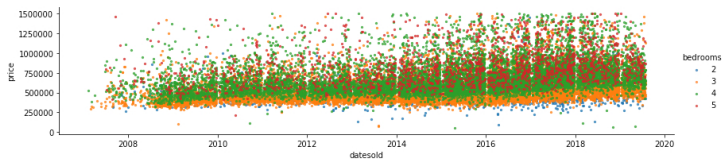
Data re-sampled at quarterly intervals using a median aggregator:
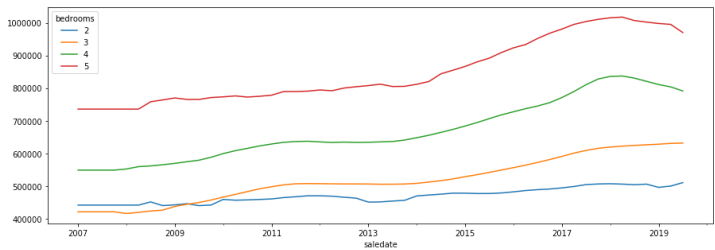
The data can be summarised as:

- date of sale
- price
- property type: unit or house
- number of bedrooms: 1,2,3,4,5 (these are the multi-variables we are targeting)
- 4 digit postcode (this is for reference only, we don't expect that this data point will be useful)

Content

- **raw_sales.csv** contains house sales data ranging from 2007-2019
- **ma_lga_12345.csv** transforms the raw sales data, re-sampling it at quarterly intervals with median price aggregation and step 4 moving average smoothing

Objective:

Forecast MA for 8 future intervals for all # of bedrooms series using a multivariate forecasting model of your choice.