APPLIED DATA SCIENCE CAPSTONE PROJECT

IDENTIFYING POCKETS OF UNDERVALUED REAL ESTATE IN THE UK

OBJECTIVES

- **Problem:** Brexit has instilled uncertainty into the UK economy and has inevitably weighed on housing prices since the referendum decision in 2016. The market will be looking towards October's deadline. But what can we expect? Well, whether we get a deal or no-deal, one could argue that risk to real estate prices is skewed to the upside from hereon. In the event of an undesirable no-deal scenario, businesses would still gain clarity and can suitably plan for the near-term; moreover, a no-deal scenario would likely keep GBP depressed against other major G-10 currencies, inherently making property relatively cheap for foreign investors.
- **Question:** As a real estate investor who is about to anticipate a turn in property prices, how can we quickly identify pockets of land in Central London that are undervalued in order to make an informed investment decision?

RESOLUTION

With various different data sources, inclusing HM Land Registry, Rightmove and Foursquare Location Data, we can collate data on property prices and determine whether areas (confined to Central London in this study) are over-valued or under-valued.

Using a python script, we scrape data from Rightmove and compare this to HM Land Registry's data to determine whether properties within specified postcodes are undervalued

We use unsupervised machine learning techniques such as K-Means to cluster similar postcodes so that an investor can expand his/her investable universe to other postcodes which are in the same cluster as ones that he is already invested in

SCRIPT INPUT/OUTPUT

Scripting Language	IDE	Inputs	Output
Python	IBM Skills Lab	Open data published by HM Land Registry	List of undervalued postcodes
		Rightmove listings data	Clustered postcodes
		FourSquare Location data	Venues/Facilities close to postcodes

BENEFIT OF THE SCRIPT

- ✓ Real time access to current listings
- ✓ Real time comparison vs HM Land Registry price paid data
- ✓ Visualization aids to determine undervalued vs overvalued postcodes

SCOPE FOR IMPROVEMENT

- ✓ Analyze individual property data
- ✓ Build a linear regression model to value house prices
- ✓ Interactive UI

THANK YOU!