

**IBM Professional Certificate in Data Science – Capstone Project (Coursera)**

Pictures of the public housing in Singapore.

*Source: istockphoto.com*

**Predicting Resale Prices of HDB Flats in Singapore**



Contents

[**1.** **Introduction** 2](#_Toc57120841)

[**2.** **The Problem** 2](#_Toc57120842)

[**3.** **The Data Set** 2](#_Toc57120843)

[**4.** **Using API to link Data to One Map** 2](#_Toc57120844)

[**5.** **Exploratory Data Analysis** 2](#_Toc57120845)

[**6.** **Linear Regression** 2](#_Toc57120846)

[**7.** **The Analysis** 2](#_Toc57120847)

[**8.** **Conclusion** 2](#_Toc57120848)

[**9.** **Appendices** 2](#_Toc57120849)

# **Introduction**

This is a project on predicting the resale prices of HDB flats.

The Housing and Development Board (“HDB”) is Singapore’s public housing authority. It plans and develops Singapore’s housing estates, building homes and transforming towns to create a quality living environment for the citizens. Within each town, HDB also provides various commercial, recreational and social amenities in the towns for the convenience of the residents.

HDB flats are home to over 80% of Singapore’s resident population, with about 90% of these resident households proudly owning their homes.

# **The Problem**

Buying an HDB flat is one of the most talked about subject in Singapore among young couple starting their own families. Housing, by its very nature, is possibly one of the largest financial commitment for a young adult. Given a certain budget, where and what time of resale flat should one buy?

# **The Data Set**

The data sources I use are publicly available sources and are easily found online. For HDB historical prices, I retrieved the resale transactions from Aug 2017 to date, retrieved from <https://data.gov.sg/dataset/resale-flat-prices>.

To extract the location data of the transaction, I extracted the coordinates (latitude and longitude) from onemap.sg using API and Python.

From the link, I will merge the information into a single file and used this merged file for analysis. The reason for separating the task is because of the long run time for the API data merge to happen.

Using predominantly linear regression, which has been proven to be a better predictor of property prices in other research paper, I shall develop a simple model to predict the resale prices of HDB flats.

# **Using API to link Data to One Map**

[Note: This will eventually include some description of the source code]

# **Using API to link Data to Four Square**

Notwithstanding that Singapore is a rather small city state and that the amenities such as schools and eateries are easily accessible, given that the HDB has a standard township plan, I shall utilise the API to Four Square to plot out the key venues for a typical HDB Town.

# **Exploratory Data Analysis**

[Note: Using the date, I shall plot the lat and long onto Folium and leverage on Foursquare location data to solve and execute]

# **Linear Regression**

# **The Analysis**

# **Conclusion**

# **Appendices**