

**Edmonton Residential Property Market: Potential ways to Increase Profits**

Executive summary

Our recommendation to increase profit by flipping properties is to find the least expensive properties in the most expensive neighborhood. For example we predicted that Windsor Park is the most expensive neighborhood in Ward 8 (see page,….). We expect that buying the cheapest properties in this Neighborhood will have highest returns on investment (ROI). We also demonstrated it by visualizing the Assessed Value of properties (see page …). We have identified Ward 6,8,9, and 10 (see page…) as the most valuable Wards. We further explore the Neighborhoods in Ward 8.

METHODOLOGY

Data Source: The Property Assessment Data (Historical) was sourced from City of Edmonton Open Data site. Available at: <https://data.edmonton.ca/City-Administration/Property-Assessment-Data-2012-2019-/qi6a-xuwt>. Accessed on: 2022-11-19.

Graphical user interface, application, table

Description automatically generated

Also, we accessed the Lab Assignment2 which contains 384,217 rows of (Account Number, Suite, House Number, Street Name, Garage, Neighbourhood ID, Neighbourhood, Ward, Assessed Value, Latitude and Longitude)

Subsequently, we merged the two files to have a data frame with ‘Ward’ and Neighbourhood ID . We considered the Wards as groups of different Neighborhoods. “Wards are neighbouring groups of residential, commercial, industrial, and other areas that are represented by a City Councillor. Wards are an essential part of Edmonton’s local government system. Voters in each Ward elect a Councillor to deal with local issues and make decisions that impact all residents”. (Ward Boundary Review, City of Edmonton).

Edmonton has 12 wards consisting of different Neighborhoods. We also considered Zoning as a strong variable for our model. Zoning by law is for the purpose of land development, the City of Edmonton is divided into zones. The zone that a particular property is in determines what can be built on that property.

**Exploratory Data Analysis**

We removed all missing values from the entire dataset and used the last assessment date to determine the Assessed Value of property. We considered only ‘RESIDENTIAL’ under “Assessment Class 1” that has:

RESIDENTIAL |323264

COMMERCIAL |12901

OTHER RESIDENTIAL |3168

FARMLAND |10

NONRES MUNICIPAL/RES EDUCATION |3

DESIGNATED IND PROPERTIES |1

We removed all outliers and from the dataset

Account Number

Assessment Year

Latitude

Longitude

Neighbourhood

Actual Year Built

Garage

Zoning

Lot Size

Assessed Value

Age of Property

Assessment Class 1

Ward

Neighbourhood ID

Note: The Age of Property was calculated by subtracting Actual Year Built from current year(2022).

**Predictive Analytics**

Model

Assessed\_Value =β0 +β1Ward + β2Zoning + β3Garage + β4Lot\_Size + β5Age\_of\_Property+α

We imported statsmodels library in pandas for our model

model = smf.ols('Assessed\_Value ~ C(Ward) +C(Zoning) + Lot\_Size +Age\_Property + C(Garage)', data=X)

While ‘X’ is our data.

We also limited our explanatory variable Zoning to most frequent types:

RF1 108355

RSL 41385

RA7 31853

RF5 24650

RF4 18761

RF3 18156

RPL 15258

DC2 9450

DC1 7473

RA9 6083

RMD 5780

RA8 3820

RF6 2858

RF2 2755

HA 1202

HDR 1168

Visualization

We imported the shapefile of Edmonton from the City of Edmonton’s website and plot the coordinates( Longitude and Latitude) using library geo pandas to create GeoPandas data frame.

Result

Inferring from the map, we expect Residential Properties in Ward 6,8,9, and 10 to be highly valued than other Wards. Also, we expect to see a flipping of the blue region of around the green zones of these wards in the nearest future. These are properties that are most likely

Map

Description automatically generatedbelow-market price.

A picture containing table

Description automatically generatedTable

Description automatically generatedText

Description automatically generatedREGRESSION RESULT:

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly

specified.

[2] The condition number is large, 1.31e+04. This might indicate that there are

strong multicollinearity or other numerical problems.

Let’s look at the 95% confidence intervals for the coefficients. Those Wards(Neighborhoods) whose CI are all negative we have considerable evidence that they tend to be priced lower than the reference Ward (Ward1). Those will all positive values tend to be priced higher on average. Also, the F-statistics tells us if we have sufficient evidence that the mean Assessed Value differs based on the Ward (as that is proxy explanatory variable Neighborhood). The small p-value here tells us to reject this null hypothesis here. Apparently, the location of property does matter.

We also looked at the boxplot:

Chart, box and whisker chart

Description automatically generated

Zoning:

Chart, box and whisker chart

Description automatically generated

The frequent Zoning bylaw in Edmonton from our research shows that only

RF1, RF2, RF3, RSL, HA, are highly priced than DC1 except RF4, RF5, RF6, RMD RPL, DC2, RA9, RA8, HDR, and RA7.

We plot the Assessed\_Value against the Age of property. From the regression output, we are 95% confident that newer properties tend to be priced higher than old houses.

Chart, line chart

Description automatically generated

We further constructed a model for Neighborhoods in Ward 8 using:

Assessed\_Value =β0 + β1Neighborhood + β2Garage + β3Lot\_Size + β4Age\_of\_Property+α

We imported statsmodels library in pandas for our model

model = smf.ols('Assessed\_Value ~ C(Neighborhood)+ Lot\_Size +Age\_Property + C(Garage)', data=X1)

While ‘X1’ is our data.

1

Table

Description automatically generated with low confidenceTable

Description automatically generatedText, letter

Description automatically generated

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly

specified.

[2] The condition number is large, 7.04e+04. This might indicate that there are

strong multicollinearity or other numerical problems.

Windsor Park seems to be the most priced Neighborhood in Ward 8, we are 95% confident that this Neighborhood is highly priced than Belgravia which is our reference. We are also confident that Belgravia is the second most priced Neighborhood in Ward 8.