MSC IN BUSINESS ANALYTICS

Module:

Business Intelligence & Data Visualization - B9BA104

Assignment Title: CA

Data Analysis & Visualization of Dublin's Residential Price Register

Submitted to:

Mr. Clive Gargan

Submitted By:

Lukose Roy Pannapara (10529675)

Table of Contents

1. Introduction & Business Understanding

A series of data analysis & visualization was performed on Dublin's Residential price register for the years 2010 – 2018, with information pertaining to purchase & price of residential properties with attributes like size or type of property.

The **Residential Property Price Register** is produced by the Regulatory Authority for Property Services (PSRA) pursuant to section 86 of the Property Services (Regulation) Act 2011. This contains the date of sale, price and address of all residential properties purchased in Ireland since 1 January 2010, as reported for stamp-duty purposes to the Revenue Commissioners. Unless the property is a new home, the price displayed will be 13.5 percent VAT exempted.

2. Data Understanding

The datasets were collected for from Ireland's Open Data Portal (data.gov.ie) and domain information were segregated from the PSRA official website. Separate datasets were downloaded for each year from 2010-2019 for the Dublin Residential Price Register.

The **Residential Property Register** dataset initially contained 9 columns viz. Date, Postal Code, County, Price, Not Full Market Price, VAT Exclusive, Description Of Property, Property Size Description.

Column Name	Type Of Data	Description
Date	Numerical (Date format)	The date of sale of the property
Address / Postal Code / County, Country	Categorical	Information pertaining the precise location of the property purchased.
Price	Numerical	The Price Of the Residential Property.
Not Full Market Price	Categorical	It indicates if the property was purchased at full market price or not.

VAT Exclusive	Categorical	It denotes information pertaining to property which were purchased excluding the VAT.
Description Of Property	Categorical	It contained information onto whether the property is new or second hand
Property Size Description	Categorical	Description pertaining to the range of size of the property.

3. Data Preparation

- The Residential datasets for the respective years 2010 to 2019 were merged using Power BI transformation. Post merging, the Residential dataset contained 1,23,944 rows and 9 columns.
- The data transformation performed on the data are as follows:
    1. **Merging** datasets.
    2. **Trimming** & **Cleaning** all the columns.
    3. **Removed empty** rows.
    4. Using **Split Function** to separate Date Column. Date column was split into 3 separate columns as Year, Month, Days & setting the data type as a Whole Number.
    5. Setting the **Data type** of each column.
    6. **Standardization** of column names and values.
    7. Identifying and **removing** unwanted columns.
       Columns like Addresses, County and Country were removed.
    8. Filling the **missing values** for each columns with appropriate values.
    9. The 2019 data was only partially available and hence was filtered out during the data transformation process reducing the number of rows in the dataset to 1,14,064.
    10. **Querying and deriving new columns** from existing ones.
        a. VAT Applicable – Column was derived based on the reference that, If the property is a new property, the price shown should be exclusive of VAT.
        b. VAT Incurred – The VAT incurred column was calculated based on reference mentioned in a.
           VAT Incurred = VAT Applicable/100 * Price

c. Total Price - VAT Inclusive – This column was derived by adding up VAT Incurred with existing Price column.
Total Price -VAT Inclusive = VAT Incurred + Price

4. Identifying Key Performance Indicators

(i) Total Revenue & VAT [2010 – 2018]
(ii) Number Of Houses Purchased by Type [2010 – 2018]
(iii) Revenue Growth by Year [2010 – 2018]
(iv) Quarterly & Monthly Revenue [2010 – 2018]
(v) Revenue Growth by Type Of Property [2010 – 2018]
(vi) Percentage Turnover by Property Size [2010 – 2018]

5. Data Analysis & Visualization

(a) Total Revenue & VAT[2010 – 2018]

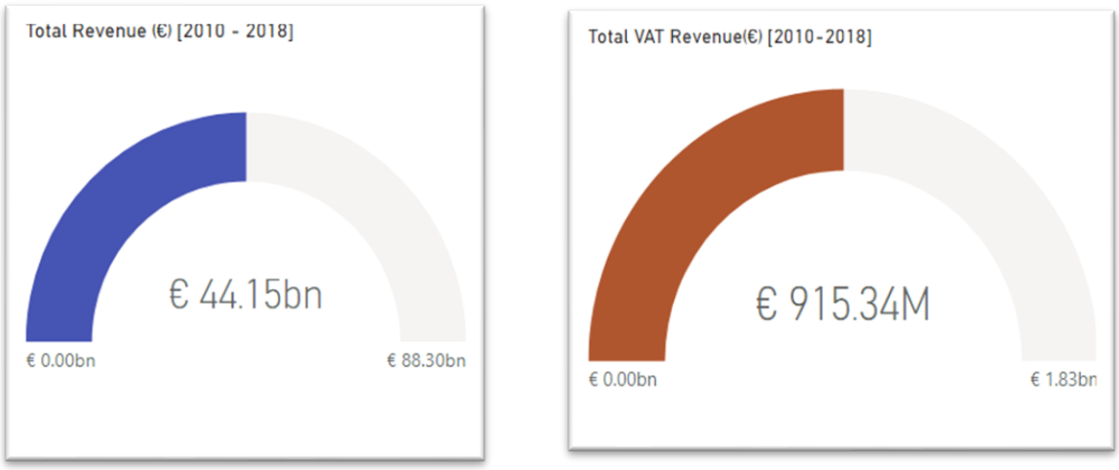

Figure 1: Total Revenue(Left) & VAT(Right)

Gauge Meter is used to visualize metrics with single values like Total Revenue & VAT. The Total Revenue generated over the past 9 years from New Dwelling house/Apartment and Second-Hand house/Apartment inclusive of VAT is 44.15B Euros and the VAT collected is 915.34M euros.

(b) Number Of Houses Purchased by Type [2010 -2018]

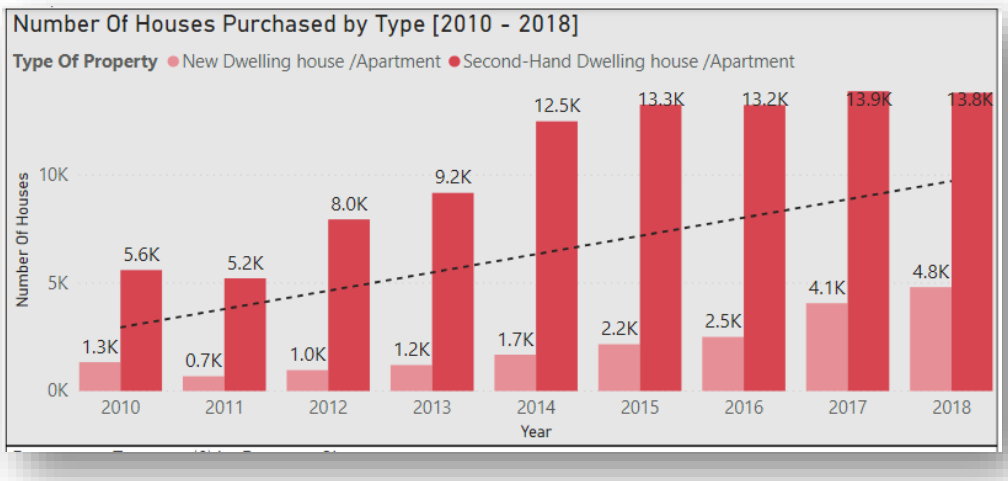

Figure 2: Number Of Houses Purchased by Type

Clustered Column Chart is used to visualize the comparison between number of new and second hand houses. A sudden rise can be seen in the number of houses purchased from 2012 for New and Second Hand houses in Dublin. However, the above graph also suggests that purchase of new dwelling houses grew by almost double in 2017 from the previous year. On the other hand, purchase of second hand houses remained consistent post 2015.

(c) Revenue Growth by Year [2010 – 2018]

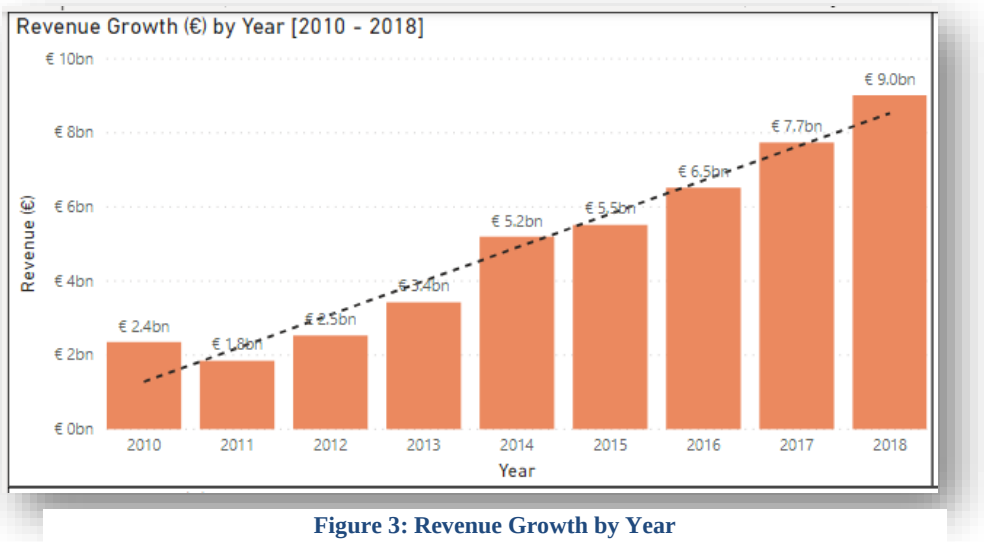

Figure 3: Revenue Growth by Year

A **column chart** is used to draw comparison between Year and Revenue Growth between 2010 – 2018.

The exponential increase in revenue from 2014 suggests that numbers of houses sold (New and Second Dwelling) was on a rise generating a total revenue of 33.99 bn euros until 2018. A positive trend line can be witnessed for the last decade in Dublin which suggests that investment in real estate was on a high.

(d) Quarterly & Monthly Revenue [2010 – 2018]

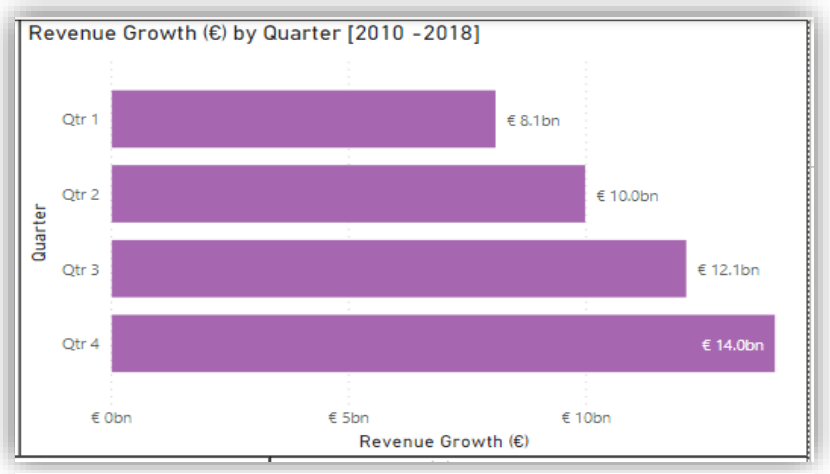

Figure 4: Quarterly Revenue

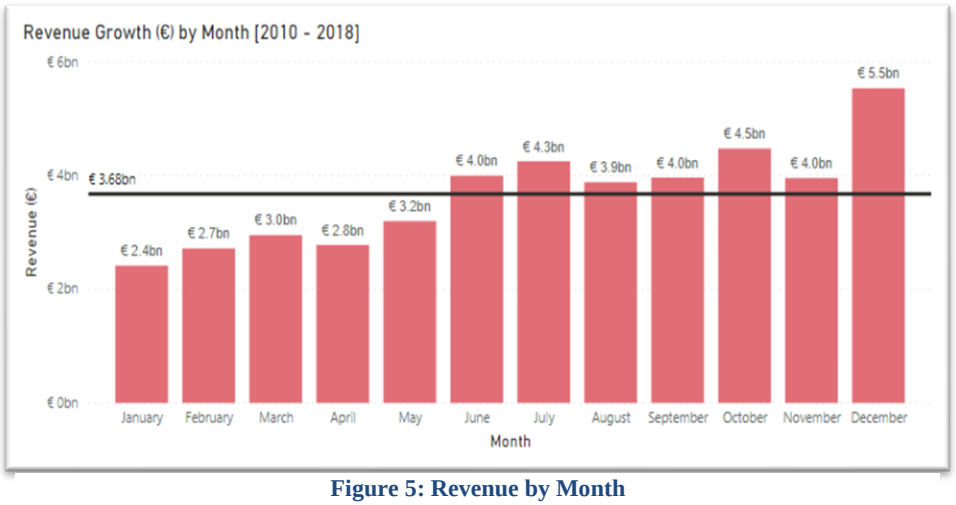

Figure 5: Revenue by Month

The above bar charts suggest that most of the revenue were generated during the second half of the year. An average revenue of 3.68 bn euros was observed each month in the last decade (Figure 5). January to May has witnessed the lowest income whereas the month of December has had the highest sale revenue of all (Figure 5). Quarterly graph suggests that Quarter 3 & 4 has always produced the highest income between 2010 – 2018.

(e) Revenue Growth by Type Of Property [2010 – 2018]

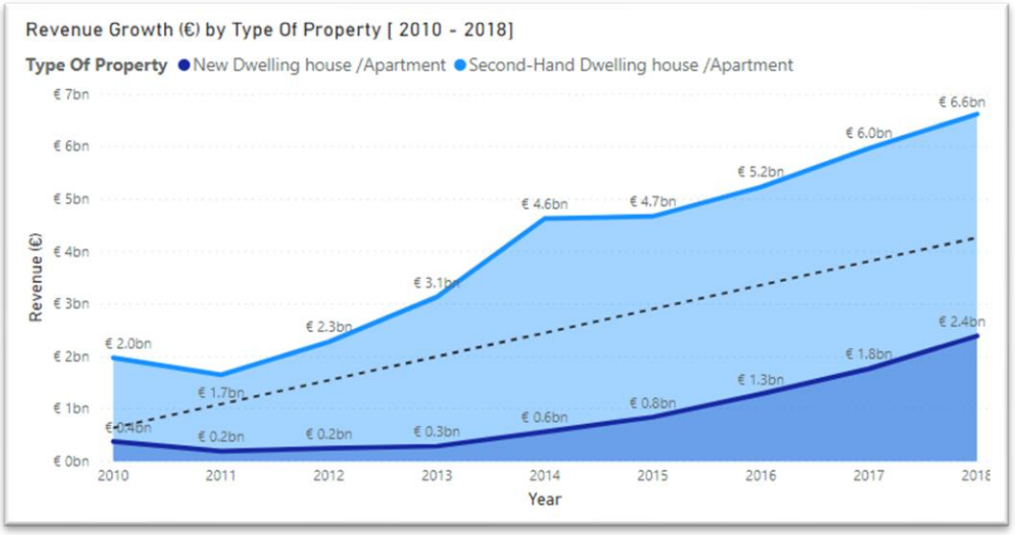

Figure 6: Yearly Income by Type Of Property

Area Chart (Figure 6) is used to graphically display the comparison between quantitative data. **Pie Chart** (Figure 7) is used to reprensent proportional data for various categories. The area chart shows that Second hand dwelling houses or apartments are more preferred and purchased as compared to New houses in Dublin. A pie chart visualization was developed to understand the percentage of property type purchased and the data suggests that 82.98% of the overall houses purchased were Second-Hand Dwelling Houses /Apartments (Figure 7).

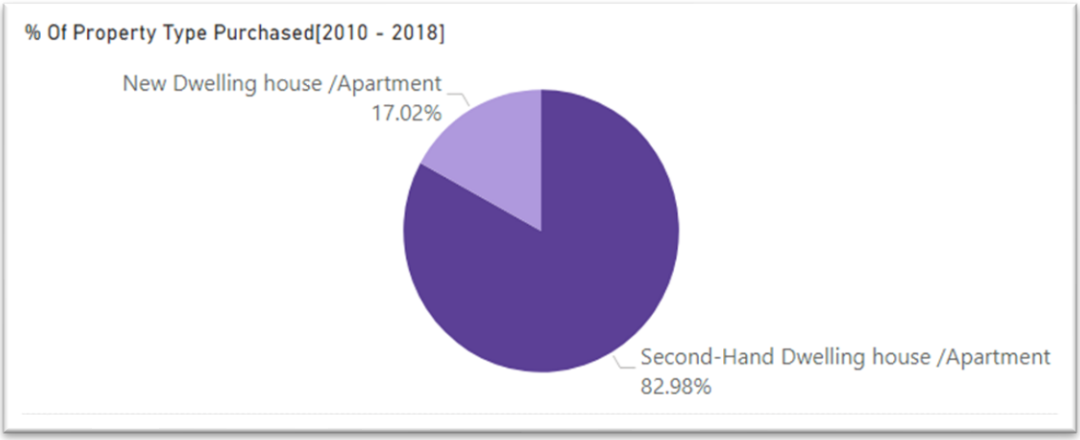

Figure 7: Percentage Of Property Type Purchased

(f) Percentage Turnover by Property Size [2010 – 2018]

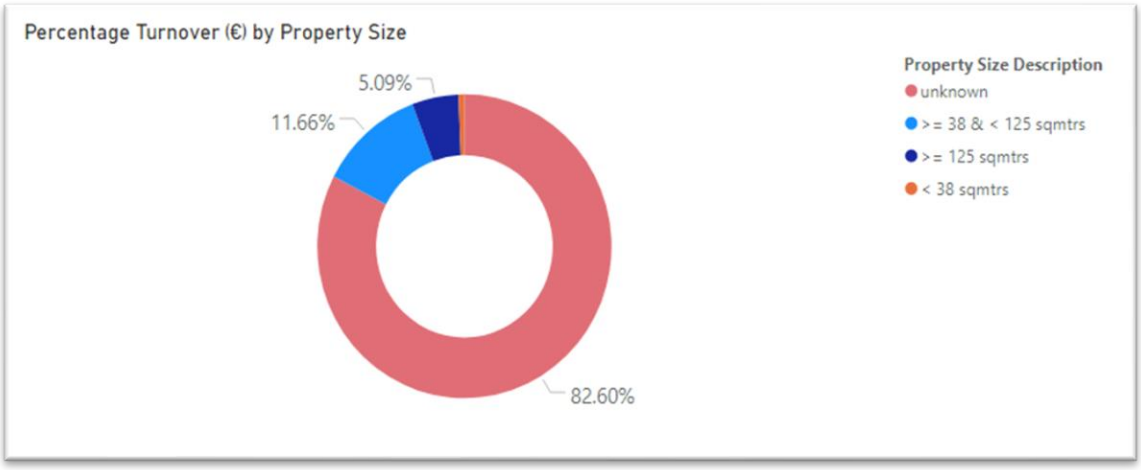

Figure 8: Percentage Turnover by Property Size – New & Second Hand Residencies

The **donut graph** from figure 8 represents the percentage turnover on the basis of property size. Upon visualizing the data, it can be seen that 82.60% of the rows had no information of their respective property sizes. Majority of this unknown data belong to second hand dwelling houses or apartments.

The below graph (Figure 9) suggests that, out of all the New houses purchased, 64.59% percentage accounted to a property size of >=38 and <=125 square meters and 28.21% of the new residential properties purchased had a size greater than or equal to 125 square meters. The notable feature from the analysis is that the month of December 4th Quarter of the year sold the most number of new houses and the from the above mentioned categories.

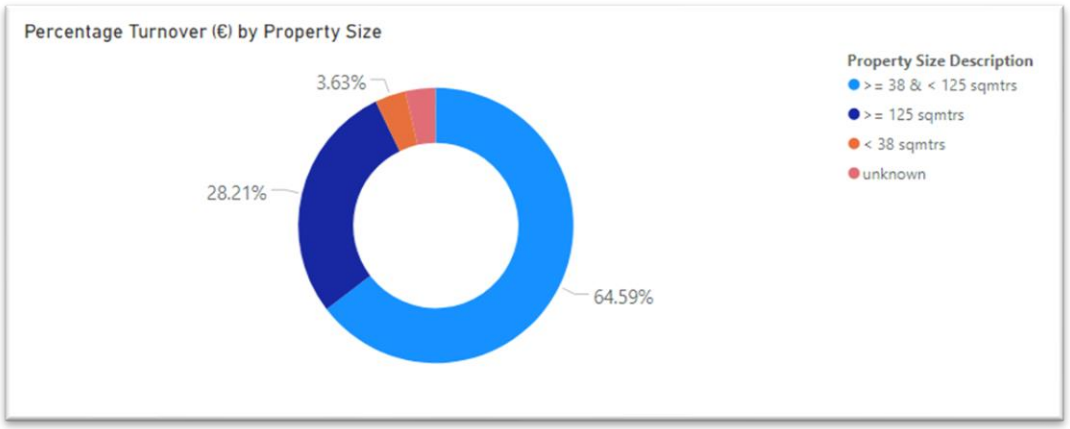

Figure 9: Percentage Turnover by Property Size – New Dwelling

6. Conclusion

After carefully analyzing and visualizing Dublin's Residential Register, it can be concluded that, an exponential increase was observed in purchase of new and second hand properties post 2013. The month of December played a significant role for the sale of New Dwelling Apartments/Houses whereas approximately 80% of the overall houses purchased were second hand. Also, 4th Quarter of the year proved to be the best selling time as most number of houses were purchased during this period. It is also worth noticing that 64% of the new houses purchased had a property size greater than equal to 38 and less than 125 square meters.

7. References

- "PSRA - Dublin Residential Property Price Register - Data.Gov.Ie." *Data.Gov.Ie*, data.gov.ie/dataset/property-price-register. Accessed 16 Apr. 2020.
- Property Services Regulatory Authority. "Residential Property Price Register - Home Page." *Propertypriceregister.Ie*, 2011, propertypriceregister.ie/Website/NPSRA/pprweb.nsf/page/ppr-home-en, /Website/NPSRA/pprweb.nsf/page/ppr-home.

8. Challenges

- The dataset obtained for year 2019 contained partial data and had to be filtered out.
- Too many empty rows within the dataset. Multiple columns with empty rows were filled with values like 'unknown' or were dropped if not required (Addresses, County, Country, Postal Code).