

Housing Market Analysis in Ireland

Killian Slater

21378826

Final Year Project – 2025

BSc. Data Science



Department of Computer Science

Maynooth University

Maynooth, Co. Kildare

Ireland

A thesis submitted in partial fulfilment of the requirements for BSc. Data Science

Supervisor: **Dr. Peter Mooney**

Contents

| | |
|---|-----------|
| Declaration | 3 |
| Abstract..... | 4 |
| Chapter 1: Introduction..... | 5 |
| 1.1 Housing Market Problem | 5 |
| 1.2 Previous Research..... | 5 |
| 1.3 Research Questions..... | 6 |
| 1.4 Organisation of thesis | 6 |
| Chapter 2: Background, Related Work and Learning | 7 |
| 2.1 Related Literature | 7 |
| 2.2 Learning Material | 8 |
| Chapter 3: Data Preparation and Methodology | 9 |
| 3.1 Data Cleaning..... | 9 |
| 3.1.1 Annual Population Change | 9 |
| 3.1.2 Property Price Register | 10 |
| 3.1.3 Pobal Deprivation Index..... | 11 |
| 3.2 Research Methodology..... | 12 |
| 3.2.1 Research Questions | 12 |
| Chapter 4: Solution Implementation | 13 |
| 4.1 Data Cleaning..... | 13 |
| 4.1.1 RQ1- “Compare house price trends with net migration trends using line plots” | 13 |
| 4.1.2 RQ2 – “Examine house price disparities on a county and regional level using boxplots” | 14 |
| 4.1.3 RQ3 – “Analyse house price evolution using tabular data” | 14 |
| 4.1.4 RQ4 – “Perform mixed model analysis to determine the influence of socio-economic factors on house prices | 15 |
| 4.2 Statistical Analysis and Modelling | 16 |
| 4.2.1 RQ1 – “Compare house price trends with net migration trends using line plots” | 16 |
| 4.2.2 RQ2 – “Examine house price disparities on a county and regional level using boxplots” | 16 |
| 4.2.3 RQ3 – “Analyse house price evolution using tabular data” | 17 |
| 4.2.4 RQ4 – “Perform mixed model analysis to determine the influence of socio-economic factors on house prices” | 17 |
| Chapter 5 – Evaluation and Research Outcomes | 18 |
| 5.1 Compare house price trends with net migration trends using line plots | 18 |
| 5.2 Examine house disparities on a regional and county level using boxplots..... | 19 |
| 5.3 Analyse house price evolution using tabular data..... | 19 |
| 5.4 Perform mixed model analysis to determine the influence of socio-economic factors on house prices | 20 |
| Chapter 6 – Conclusions and Future Work | 22 |
| 6.1 Conclusion..... | 22 |
| 6.2 Personal Reflection..... | 23 |
| 6.3 Future Work | 24 |
| References and Bibliography | 25 |
| Appendices..... | 27 |

Declaration

I hereby certify that this material which I now submit for assessment on the program of study as part of BSc. Data Science qualification is entirely my own work and has not been taken from the work of others – save and to the extent that such work has been cited and acknowledged within the text of my work.

I hereby acknowledge and accept that this thesis may be distributed to future final year students, as an example of the standard expected for final year projects.

Signed: Killian Slater

Date: 16Th March 2025

Abstract

The Irish Housing market faces major challenges such as a shortage of houses, increased prices, lack of supply and increased population. Utilising datasets such as the Property Price Register, Annual Population Change, and the Pobal Deprivation Index, this study focuses on key aspects affecting the current housing crisis. A meticulous data cleaning process ensured the analysis was precise and focused. Classifying the data into Province, County and Electoral Divisions allowed for a deeper analysis of trends and patterns on a smaller, more problem-focused scale. By using statistical models and interactive visualisations this study aims to explore the relationship between house price increases and net migration trends, examine regional and county-level house price disparities, identify socioeconomic factors influencing house prices, and analyse the decade-long evolution of the Irish Housing Market.

This analysis revealed that net migration and socioeconomic factors, in particular age dependency and education levels, significantly influence house prices. Regional disparities persist, with higher prices in areas of economic growth. The study highlights the ongoing housing crisis, driven by supply shortage and rising demand, emphasising the need for policy intervention.

Chapter 1: Introduction

1.1 Housing Market Problem

The housing market encapsulates the supply and demand dynamics of residential properties, influencing availability, pricing and affordability. In Ireland, the housing market has experienced significant fluctuations over the past decade, with a yearly price increase of around 9% in 2024 according to a report by the Irish Times (L. Slattery, 2025).

A combination of factors such as supply and demand imbalances, a high number of vacant properties and a shortage of social housing have driven house prices upwards. According to social justice Ireland (2024), between the period of 2012-2022, wages rose by only 27%, compared to residential property prices which surged by 75%. This widening gap between income growth and housing costs has placed increased financial pressure on new homebuyers and renters, making homeownership less attainable for many households.

The inspiration for my thesis stems from the fact that these issues are very relevant to my own life. Like many others in my demographic, I face the difficult choice of either remaining in Ireland and struggling to afford a house well into my late twenties or emigrating in search of better opportunities. The housing crisis has made independent living increasingly unattainable for young adults, highlighting the urgency of addressing these challenges.

In this project, I have explored data provided by the Central Statistics Office (CSO), the Pobal Deprivation Index and the Property Price register, focusing on the period from 2010 to 2024. My research investigates two main areas. Firstly, I examined the characteristics of Irish house prices over the last decade. Within this area I compared house prices with the trends in net migration figures, investigated the regional and county level house price disparities, and analysed the evolution of house prices over the past ten years at the county level. The second area of research focuses on the impact of socio-economic factors on house prices using mixed model analysis.

1.2 Previous Research

This project contributes to a growing body of research on the Irish Housing market, with many insightful reports created, such as, **Economic policy issues in the Irish housing market** – Central Bank of Ireland (2025). This report highlights the supply and demand issues within the housing market which are being caused by unmet demand and population growth that has proven faster than expected. It is stated that the number of new housing units added between 2011 and 2022 was far below the growth in population, however the figures from 2023 show that the outturn of houses is getting closer to the expected units of 34,000 to 45,000 units per year.

Another interesting report, **The Daft.ie Rental Price Report** by Ronan Lyons (2024) highlights the per county increases in rental prices in 2024 Q4. Lyons highlights the increase in rental prices for different types of properties, while also breaking down the rental increases per region. It was found that Leinster was recorded to have the lowest year-on-year increase (5%), while Munster had the highest year-on-year increase (9.4%). This report inspired my investigation into the house prices on both a regional and county level.

1.3 Research Questions

The research questions that will inform this thesis are as follows:

RQ.1 – Compare house price trends with net migration trends using line plots: Line plots will be used to compare house price trends with net migration trends over time, identifying potential relationships between migration patterns and house price fluctuations.

RQ.2 – Examine house price disparities on a regional and county level using boxplots: Boxplots will be used to assess regional and county-level disparities between house prices.

RQ.3 – Analyse house price evolution using tabular data: A table that will highlight the decade long evolution of house prices at a county level.

RQ.4 – Perform mixed model analysis to determine the influence of socio-economic factors on house prices: A linear mixed-effects model will be used to evaluate the impact of different socio-economic factors on house prices.

1.4 Organisation of thesis

Chapter 2 explores the background research and foundational work that informed this project. It also discusses the skills acquired throughout the process and the relevant college modules that provided the necessary expertise for this analysis. Chapter 3 outlines the data cleaning process, detailing the steps taken to prepare each dataset for analysis. Chapter 4 focuses on the technical implementation of the research. It provides an in-depth look at the code used for data preprocessing, visualisation and modelling. Chapter 5 presents the results, including the visualisations and models created to address each question. It also includes an interpretation of the findings for each question. Chapter 6 concludes the thesis by summarising key insights from the research and proposing potential next steps to address housing market challenges and guide future research.

Chapter 2: Background, Related Work and Learning

2.1 Related Literature

The Irish housing market has been the subject of extensive research, particularly in the areas of price trends, affordability and the impact of socioeconomic factors. To understand the broader context, I first examined reports from the Central Statistics Office (CSO), I found their resource **“The Housing Hub”** particularly useful, especially in the section of Property, Prices and Rents which provided visualisations regarding increases in property prices and rents. This report also provided an insight into the differences between Dublin prices compared to the rest of the country, highlighting the need to take the outlying Dublin prices into account when carrying out my own analysis.

In relation to housing needs and migration figures, Social Justice Ireland (SCI) have carried out considerable research in the area, in their report **“Addressing Irelands housing crisis: urgent policy reforms needed” (2024)**. SCI look into the change in the Irish demographic, highlighting a major increase in our elderly population from 2016-2022. Between 2011 and 2022, Irelands population has grown by 560,887, while the housing stock increased by only 117,276 units. It is also pointed out that Ireland is continually underinvesting in housing, making us the lowest housing investors in the EU despite our strong economy.

It is interesting to look at the impact that the Global Financial Crisis (GFC) had on the Irish housing market, in particular the increase in homelessness. The journal - **“How supply and demand affect national house prices: The case of Ireland (2024)”** looks into the pre, during and post GFC Irish housing market (1995-2019), making comparisons with 21 other countries. Over the three time periods, the Irish housing market experienced the most significant changes. Between 1995 and 2007 house prices rose by an alarming 475 %, whereas during the crash (2007 – 2012), Ireland received the biggest decline in house prices, with a decrease of 51.4%. From the period of 2012-2019 we once again saw an overall increase of 75.4 %, which was significantly higher than any other country. In parallel to rapidly increasing rents, house prices and increasing involvement of real estate investment funds in the Irish private rental market, there has been a significant increase in homelessness from 2013 onwards. In 2012 an average of 8 new families were becoming homeless monthly, while this number became 32 in 2015 according to the study **“Housing financialisation and the creation of homelessness in Ireland (2020)”**. The number of families in emergency accommodation increased by 416% between July 2014 and July 2018, reaching an all time high of 1778 families. The growing homelessness crisis underscores the urgent need to address soaring house prices.

In a research paper titled “**Socio-Economic factors associated with house, prices - Evidence based on key macroeconomic aggregates globally**” (2023) , association between macroeconomic factors and house prices are investigated among 60 countries. For my research I focused on the findings related to unemployment. It was found that high unemployment limits people’s ability to buy houses which in turn causes a decrease in house prices. On the other hand, low unemployment causes the demand for houses to rise which ultimately drives the prices of houses up. While this paper investigated the impact of macroeconomic issues on house prices, for my own research, I intend to investigate socioeconomic factors, including the proportion of the population classified as age dependent, the percentage of individuals with third-level education, and the unemployment rate.

To gain some more insight into socio-economic factors impacting housing prices I looked at Northern Ireland, in particular a journal titled “**Demographic trends and changing housing systems in Northern Ireland (2018)**”. By specifically looking at the period of 1981-2001, it is seen that in Northern Ireland the increase in the number of households over the 20-year period (38%) was almost three times more than the population increase (14%). This generated a much more rapid increase in housing demand than population change alone. The average household size declined faster in the Republic of Ireland and Northern Ireland than in Great Britain. Factors that caused this were growth in single living, a changing mix of household types and lower fertility. As well as this, marriage rates were much lower, and the average age of marriage increased. These issues did not have profound effects on Northern Ireland however due to substantial programmes of redevelopment and new building by the Housing executive, coupled with increased private sector construction in the 1970’s. An environment was created which could accommodate changing demographic changes. These foresighted policies helped Northern Ireland avoid a housing crisis unlike what we are experiencing today in the Republic.

2.2 Learning Material

This project required an in-depth knowledge of R, with a focus on packages such as **ggplot2** for data visualisation, **dplyr** for data manipulation and **lme4** for mixed effects models. The module ST203: R for Statistics and Data Science provided a strong foundation for R programming, while other modules such as ST201: Data Analysis, ST302: Data Visualisation, ST401: Mixed effects models and CS130: Databases aided massively in being able to perform a comprehensive analysis.

Chapter 3: Data Preparation and Methodology

This chapter outlines the data preparation and research methodology used in this study, detailing the steps taken to clean, process and analyse multiple datasets.

3.1 Data Cleaning

This chapter is split into two sections, data preparation and research methodology. The analysis conducted was based off four different datasets, Property Price Register (2010-2024), Annual Population Change (1950- 2024), Deprivation Index (2016) and the Property Price Register (2024 only). Data Cleaning is the process of reformatting data to make it suitable for research purposes.

The data, which was sourced from the CSO, Pobal.ie and the Property Price register included variables and time periods that were not relevant to this analysis. As a result, data cleaning and preprocessing steps were undertaken to filter out unnecessary variables and exclude years outside of the scope of this study. All datasets which were sourced came in the form of CSV files.

The original idea aimed to conduct the research on a regional level by assigning the corresponding Province to each datapoint, allowing for a broad analysis of the disparities within each Province. While conducting this analysis, having only four provinces proved to provide little to no insight into the issues under investigation. As a result of this, the datasets were further disaggregated into county and electoral division levels, allowing for a more detailed and comprehensive analysis.

3.1.1 Annual Population Change

The Annual Population Change dataset was sourced from the CSO. The original dataset included figures for annual births, deaths, immigration, emigration, net migration, population change and population from the years 1950-2024. For this analysis the years from 2010-2024 were selected along with only the net migration for these years. Furthermore, there were a lot of NA's (missing values) in the data which were omitted in the process. See fig 3.1.1.

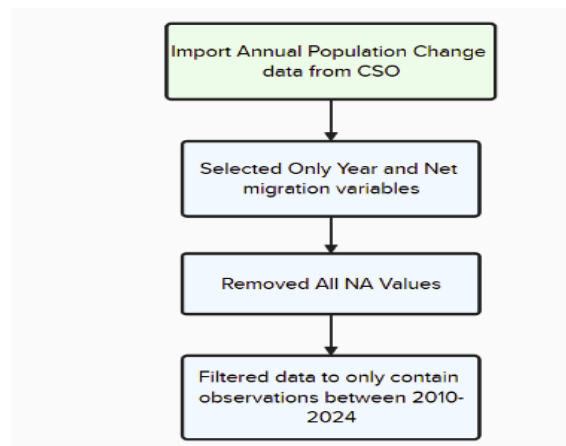


FIG. 3.1.1. THE DATA CLEANING STEPS FOR ANNUAL POPULATION CHANGE DATA

3.1.2 Property Price Register

The Property Price Register data was sourced from the Property Price Register website and the Property Price Register with corresponding electoral divisions was sourced from Shane Lynn (Website). While there were different datasets used, the preparation of the data was the same. The original dataset contained variables such as Date of sale, Address, County, Eircode, Price, VAT, Property type and Property size, with the addition of electoral division in the case of the other dataset. For this analysis Date, County, Price and Electoral division were selected. A new column was created for year, which was extracted from the date column. From here four new data frames were created for each different area of analysis. A data frame for 2024 prices, 2010-2024 prices, 2016 prices and a combined data frame of 2014, 2019 and 2024 prices. See fig 3.1.2

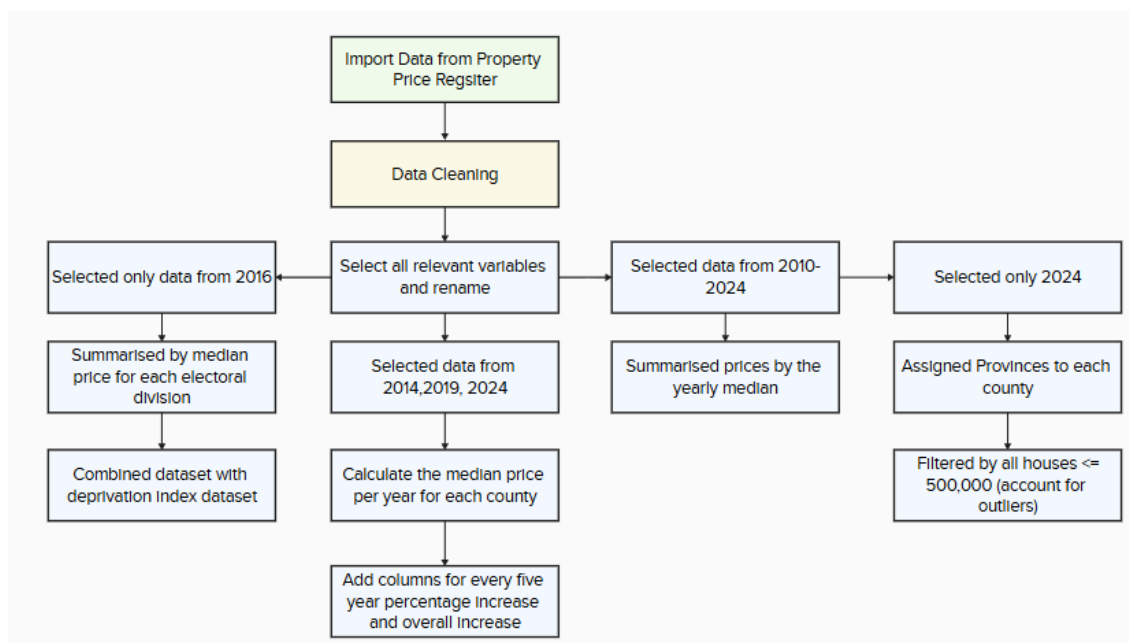


FIG. 3.1.2. THE DATA CLEANING STEPS FOR PROPERTY PRICE REGISTER DATA

3.1.3 Pobal Deprivation Index

The Deprivation Index dataset was sourced from the Pobal website in the format of a CSV file. The initial plan was to use the 2022 deprivation data alongside the 2022 property price register data to analyse the socio-economic factors affecting house prices. However, after extensive searching, no property price register data with corresponding electoral districts was available for use. As a result, the 2016 deprivation index was paired with the 2016 house prices. This methodology can be applied to more recent data if a compatible dataset becomes available.

The original data contained a lot of irrelevant and repeated data. Variables were selected based on the impact that I believed they would have in the modelling. These were Area Name, County, Age Dependency Rate, Third Level Education, Unemployment, Deprivation Category and Deprivation Score. The original columns had long and unclear names, so they were renamed for ease of use. A new column was created to combine the unemployment of males and females for each area into one column. This new data frame was combined with the house prices for 2016 to make one big dataset. See fig 3.1.3

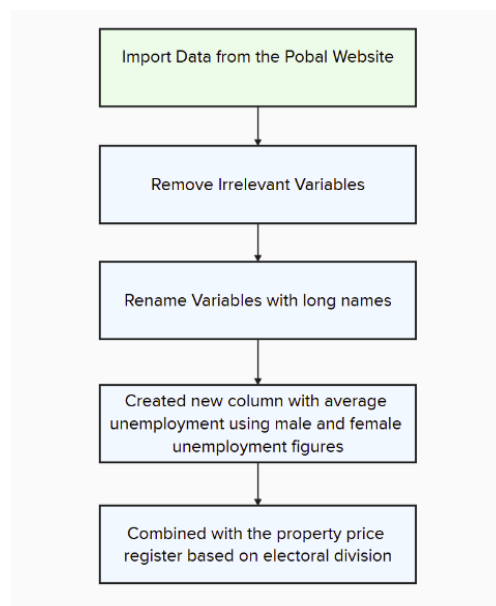


FIG. 3.1.3. THE DATA CLEANING STEPS FOR DEPRIVATION INDEX DATA

3.2 Research Methodology

The datasets used in this analysis provide the foundation for investigating key research questions related to property price trends, regional disparities, and the influence of socio-economic factors on house prices. By leveraging these datasets, various plots, models and statistical analyses will be conducted to explore market patterns and demographic impacts. The research questions outlined below will guide this analysis.

3.2.1 Research Questions

- **RQ1 – “Compare house price trends with net migration trends using line plots”:** These plots aim to explore the trends between net migration figures and median yearly house prices. The analysis here aims to uncover a relationship between the rise of house prices and the increase in migration figures
- **RQ2 – “Examine house price disparities on a regional and county level using boxplots”:** This plot features a box plot with prices from 2024. The plot is split into different counties with a box for each county. Counties are assigned a province and colour-coded accordingly with the data ordered from highest to lowest median price. The visualisation aims to highlight disparities in house prices at both county and regional levels.
- **RQ3 – “Analyse house price evolution using tabular data”:** A table presents percentage price increases at five-year intervals (2014, 2019 and 2024) and the overall increase over ten years.
- **RQ4 – “Perform mixed model analysis to determine the influence of socio-economic factors on house prices”:** This focuses on a mixed model analysis and looks to determine the influence of different socio-economic factors on the house prices. This extends the analysis of house price disparities, looking at factors influencing price variations at the county level.

This analysis examines house price trends, regional disparities and the impact of socioeconomic factors on the Irish house prices. Each plot and model are created to analyse the housing market from a different angle. The findings can help inform future research and policy decisions, with the methodology adaptable for more recent data as it becomes available.

Chapter 4: Solution Implementation

This chapter outlines the technical details behind the solution implementation, providing an in-depth explanation of the code used to conduct the analysis for each research question.

For this project, I chose to use R using RStudio due to its powerful capabilities in statistical analysis and data visualisations. The datasets were available as CSV files, which I downloaded and imported into R. In section 4.1, I will detail the steps taken to clean and preprocess the datasets, while section 4.2 will outline the statistical analysis techniques applied to the data.

4.1 Data Cleaning

4.1.1 RQ1- “Compare house price trends with net migration trends using line plots”

The data for the property price register and annual population change were downloaded as CSV files. The population change data required minimal cleaning – only selecting net migration figures, removing NA values and filtering for 2010-2024.

The house prices data required more preprocessing. Irrelevant variables were removed, keeping only County, Price and Date of Sale. The price column was cleaned using the `gsub()` function to remove the ‘€’ symbol, enabling for calculations. The year was extracted from the Date column, as exact dates were unnecessary. Finally, the data was grouped by year, summarised using median yearly price, and filtered to exclude 2025.

```
```{r}
##Select only relevant columns
Prices_ALL <- PPR_ALL %>%
 select(County, `Price (€)`, `Date of Sale (dd/mm/yyyy)`) %>%
##Rename the columns with difficult to reference names
 rename(Price = `Price (€)`, Date = `Date of Sale (dd/mm/yyyy)`) %>%
 mutate(Price = as.numeric(gsub("[^0-9.]", "", Price)), Year=year(dmy(Date))) %>%
##Group by Year of observation and summarise by yearly median price
 group_by(Year) %>%
 summarize(Median_Price = median(Price, na.rm = TRUE)) %>%
 filter(Year!=2025)

##Filter by Net Migration, Remove NA values and filter data from 2010-2024
annual_net_migration<- AnnualPopulationChange %>%
 filter(Component=="Net migration" & !is.na(VALUE) & Year >=2010)
```
```

FIG. 4.1.1. DATA CLEANING CODE FOR HOUSING PRICE TRENDS IN COMPARISON TO NET MIGRATION FIGURES

4.1.2 RQ2 – “Examine house price disparities on a county and regional level using boxplots”

The data cleaning process for RQ2 was similar to RQ1, as both used the same dataset. After cleaning, observations with prices above €500,000 were filtered out to remove outliers that distorted the plot. Each observation was then assigned a province to enable analysis of both county and provincial level disparities

```
```{r}
##Select only relevant variables and rename for ease of use
PPR_Box<- PPR_ALL %>%
 select(County, `Price (€)`, `Date of Sale (dd/mm/yyyy)`) %>%
 rename(Price = `Price (€)`, Date = `Date of Sale (dd/mm/yyyy)`) %>%
 mutate(Price = as.numeric(gsub("[^0-9.]", "", Price)), Year=year(dmy(Date))) %>%
 ##Filter for houses costing less than 500,000
 filter(Price<500000 & Year==2024) %>%
 ##Assigning a province to each county
 mutate(Province=case_when(
 County %in% Leinster ~ "Leinster",
 County %in% Munster ~ "Munster",
 County %in% Ulster ~ "Ulster",
 County %in% Connacht ~ "Connacht"))
```
```

FIG. 4.1.2. DATA CLEANING CODE FOR 2024 PROPERTY PRICE DISPARITIES

4.1.3 RQ3 – “Analyse house price evolution using tabular data”

Using the Property Price Register dataset, I created three data frames for prices from 2014, 2019 and 2024. Within each, prices were grouped by county and summarised by the median price. After cleaning each dataset, they were combined using the `left_join()` function from the `dplyr` package in R.

```
```{r}
##Creating individual data frames for each year
Prices_2014<- PPR %>% filter(Year==2014) %>%
 select(County,Price) %>% group_by(County) %>% summarise(Price_2014=median(price))
Prices_2019<- PPR %>% filter(Year==2019) %>%
 select(County,Price) %>% group_by(County) %>% summarise(Price_2019=median(price))
Prices_2024<- PPR %>% filter(Year=2024) %>%
 select(County,Price) %>% group_by(County) %>% summarise(Price_2024=median(price))
##Combining 2014 and 2019 data frame
PPR2014_2019<- Prices_2014 %>%
 left_join(Prices_2019,by="County")
##Combining all data frames together
Price_Comparison<-PPR2014_2019 %>%
 left_join(Prices_2024,by="County")
```
```

FIG. 4.1.3. DATA CLEANING CODE FOR THE EVOLUTION OF HOUSE PRICES

4.1.4 RQ4 – “Perform mixed model analysis to determine the influence of socio-economic factors on house prices

The Pobal Deprivation Index variables were renamed for clarity. The geocoded property price register data (2012-2017) included electoral districts and was filtered to contain only prices from 2016. Keeping only price and electoral district, the data was grouped by distinct electoral district and summarised by the median price, creating a simplified dataset. The deprivation data was then merged with the property price data using `left_join()`, and a new column was added to represent the average unemployment rate for males and females in each district.

```
```{r}
##Selecting relevant variables for mixed models analysis
electoral_all<-electoral %>%
 rename(electoral_district="ed_name",
 Age_Dependency_Rate="agedependencyratio",
 Proportion_Third_Ed="propthirdlevel",
 Unemployed_Male="unemploymentratemale",
 Unemployed_Female="unemploymentratefemale",
 Pobal_Status="pobalhpdescription",
 Pobal_Score="pobalhpindex2016",
 County="county_1")
##Simplifying geocoded data to combine with deprivation data
PPR_Geo_summary<-PPR_Geo %>%
 filter(year==2016) %>%
 select(price,electoral_district) %>%
 group_by(electoral_district) %>%
 summarize(Price=median(price,na.rm=T))
##Combining deprivation and geocoded PPR by electoral division
electoral_data<- electoral_all %>%
 left_join(PPR_Geo_summary,by="electoral_district") %>%
 filter(!is.na(Price)) %>% mutate(Unemployment=(Unemployed_Female+Unemployed_Male)/2)
```
```

FIG. 4.1.4. DATA CLEANING CODE FOR MIXED MODEL ANALYSIS

4.2 Statistical Analysis and Modelling

4.2.1 RQ1 – “Compare house price trends with net migration trends using line plots”

Utilising functions from the **ggplot2** and **gridExtra** packages in R, the side-by-side plots were created to compare the net migration trends with house price trends. In both plots the year was plotted on the x-axis. The net migration plot used VALUE on the y-axis, representing net migration figures divided by 1000 for easier interpretation. The house price plot used median house prices on the y-axis, also divided by 1000. Data points were plotted for each year, with a smoothed line added to highlight trends more clearly.

```
```{r}
##Plot the line plot for annual net migration
Annual_net <- ggplot(annual_net_migration, aes(x = Year, y = VALUE)) +
 geom_point(size = 3, color = "#0072B2", alpha = 0.7) +
 geom_smooth(se = FALSE, colour = "#D55E00", linewidth = 1.2, method = "loess") +
 labs(title = "Annual Net Migration in Ireland",
 subtitle = "Net migration trends over time", x = "Year", y = "Number of People (in 1000s)") +
 theme_minimal()

##Plot the line plot for Yearly median house prices
plot_median <- ggplot(Prices_ALL, aes(x = Year, y = Median_Price)) +
 geom_point(size = 3, color = "#0072B2", alpha = 0.7) +
 geom_smooth(se = FALSE, colour = "#009E73", linewidth = 1.2, method = "loess") +
 labs(title = "Median Property Price Trends", subtitle = "Median house prices over time",
 x = "Year", y = "Price (in 1000s €)") +
 theme_minimal()
##Arrange the plots in a side by side grid
grid.arrange(plot_median, Annual_net, ncol=2)
```
```

FIG. 4.2.1. CREATING SIDE-BY-SIDE LINE PLOTS FOR NET MIGRATION AND MEDIAN HOUSE PRICES

4.2.2 RQ2 – “Examine house price disparities on a county and regional level using boxplots”

Utilizing the **ggplot2** package for its powerful data visualisation functionality again, I have created a boxplot for the 2024 property price register data. House Prices are plotted on the x-axis, while Counties are plotted on the y-axis. The boxes for each county are coloured based on the province they belong to and reordered from highest to lowest median house prices for interpretability. By using colouring and ordering, the regional and county price disparities become much clearer.

```
```{r}
##Create box plot, reordered based on median house price
ggplot(PPR_Box, aes(x=Price/1000, y=reorder(County, Price, FUN=median), fill=Province)) +
 geom_boxplot() +
 labs(title="Comparison of the property prices across each county",
 y="County", x="Price (Thousands)")
```
```

FIG. 4.2.2. CREATING A BOX PLOT SHOWING REGIONAL AND COUNTY HOUSE PRICE DISPARITIES

4.2.3 RQ3 – “Analyse house price evolution using tabular data”

To analyse house price evolution from 2014 to 2024, calculations were made to show the increases between each interval:

- First Increase = $((\text{Price}_{2019} - \text{Price}_{2014}) / \text{Price}_{2014}) * 100$
- Second Increase = $((\text{Price}_{2024} - \text{Price}_{2019}) / \text{Price}_{2019}) * 100$
- Overall Increase = $((\text{Price}_{2024} - \text{Price}_{2014}) / \text{Price}_{2014}) * 100$

These calculations were added as new columns to the dataset and reordered for clarity. The table was then converted from an R file to a CSV file using the **write_xlsx** function from the **openxlsx** package in R.

```
##{r}
##Adding new column for each five year increase, and overall increase
Percentage_Table <- Price_Comparison %>%
  mutate(First_Increase = round(((Price_2019 - Price_2014) / Price_2014) * 100,digits=2),
         Second_Increase=round(((Price_2024-Price_2019)/Price_2019)*100,digits=2),
         Overall_Increase=round(((Price_2024-Price_2014)/Price_2014)*100,digits=2))
##Reordering the columns for easier interpretation
Percentage_Table<-Percentage_Table[,c("County","Price_2014","First_Increase","Price_2019","Second_Increase",
                                     "Price_2024","Overall_Increase")]
##Creating a csv file with the outputted table
write_xlsx(Percentage_Table,"House_Price_Comparison.xlsx")
##}
```

FIG. 4.2.3. CALCULATING THE FIVE YEAR AND OVERALL INTERVALS AND ADDING TO A NEW DATA FRAME

4.2.4 RQ4 – “Perform mixed model analysis to determine the influence of socio-economic factors on house prices”

I aimed to examine the underlying factors affecting house prices by constructing a linear mixed-effects model using the **lmer** function from the **lme4** package in R. The purpose of this model is to determine the influence of Pobal Score, Age Dependency, Third Level Education and Unemployment on the house prices. The random effect in the model is the county. Initially I had intended to use electoral area as the random effect, but due to the way the datasets were joined in 4.1.4, there was only one observation per area. Using county as the random effects provided more meaningful insights.

To improve the interpretation of the coefficients, I transformed the house prices onto a log scale. This helps in reducing the skewness of the data, making the model more interpretable and ensuring that the effects are proportional.

```
##{r}
##Performing mixed model analysis
model1<- lmer(log(Price)~ Pobal_Score + Age_Dependency_Rate +
              Proportion_Third_Ed+ Unemployment+(1|County),data = electoral_data)
summary(model1)
##}
```

FIG. 4.2.4. DEFINING A MIXED-EFFECT MODEL AND PRINTING OUT THE SUMMARY OF THE MODEL

Having laid the groundwork for the analysis, the following chapter will focus on the results and their interpretation in the context of the housing market dynamics and socioeconomic trends in Ireland.

Chapter 5 – Evaluation and Research Outcomes

This chapter evaluates the findings from the statistical analyses and modelling presented in chapter 4. The structure of this chapter follows the research questions outlined earlier, presenting key insights and findings for each.

5.1 Compare house price trends with net migration trends using line plots

I used a variety of plots to investigate different variables throughout my analysis. In my first area of analysis, I sought to compare the increase in house prices to trends in net migration between 2010-2024. Initially looking at the median yearly property prices it is clear to see a significant increase in property prices, recorded at a median of €200,000 in 2010 and then skyrocketing to €350,000 in 2024, an increase of 75%.

Similarly, annual net migration figures show a steady increase over time. In 2010 net migration stood at -30,000 people, meaning that more people were leaving Ireland than entering. This is linked to the global financial crash which happened in 2008, which plunged Ireland into recession. Over the next fourteen years net migration rose significantly, reaching an all-time high of 85,000. Net migration dropped sharply in 2021 due to COVID-19 travel restrictions which limited entry to the country. Despite this the surge resumed once it was possible to enter Ireland again.

Both house price and net migration graphs follow a similar trend, initially house prices dropped from €200,000 in 2010 to around €130,000 in 2013, before rising rapidly. Similarly, the net migration figures slowly increased over the same period, but still showing more people emigrating than immigrating. From 2015 onward both house price and net migration graphs show a rapid increase. As house prices increase so do net migration figures. The rising population led to increased housing demand. As noted in section 2.1, housing stock has not kept pace with population growth, as confirmed by this analysis.

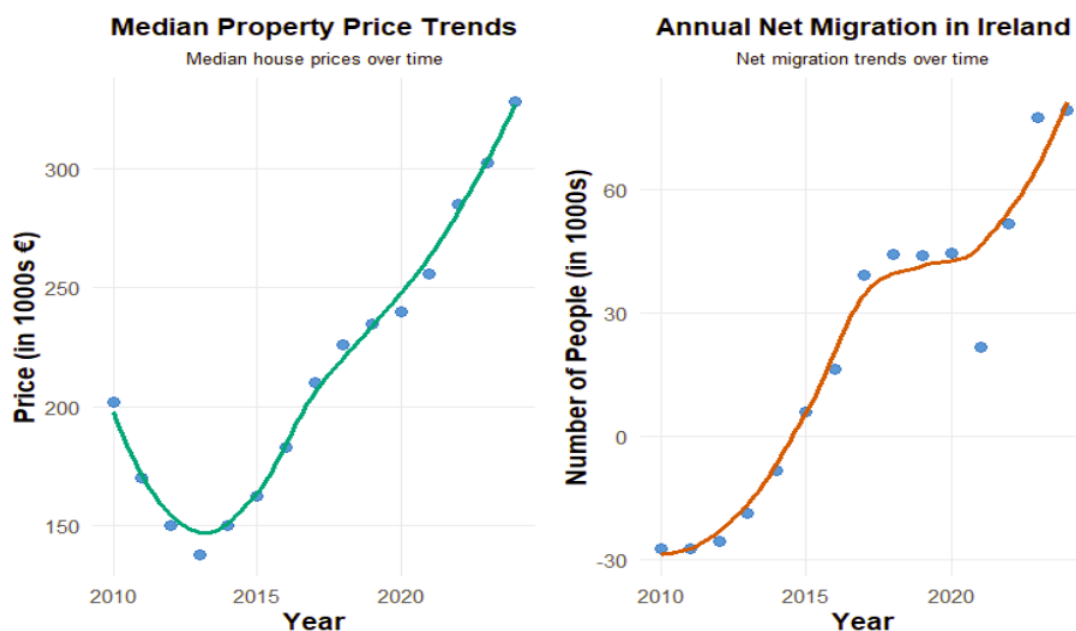


FIG. 5.1.1. LINE PLOT COMPARISON OF NET MIGRATION AND HOUSE PRICE TRENDS (2010-2024)

5.2 Examine house disparities on a regional and county level using boxplots

Next, I examined the house price disparities on a regional and county level. The house prices were plotted for each county and ordered based on the highest median house price. Each box in the box plot was coloured based on region. The plot clearly shows that Dublin and Wicklow have the highest median house prices (€350,000), while Leitrim, Donegal, and Longford fall on the lower end at around (€175,000). The disparity box plot also reveals that counties in Leinster and Munster tend to have higher house prices than counties in Connacht and Ulster. Proximity to Dublin appears to be a key factor in house prices with counties such as Meath, Kildare and Wicklow, all of which are a commutable distance from Dublin, rank among the top three most expensive counties (excluding Dublin). This analysis has highlighted the clear house price disparities across regions and counties. These disparities can be attributed to factors such as infrastructure, employment opportunities and economic activity.

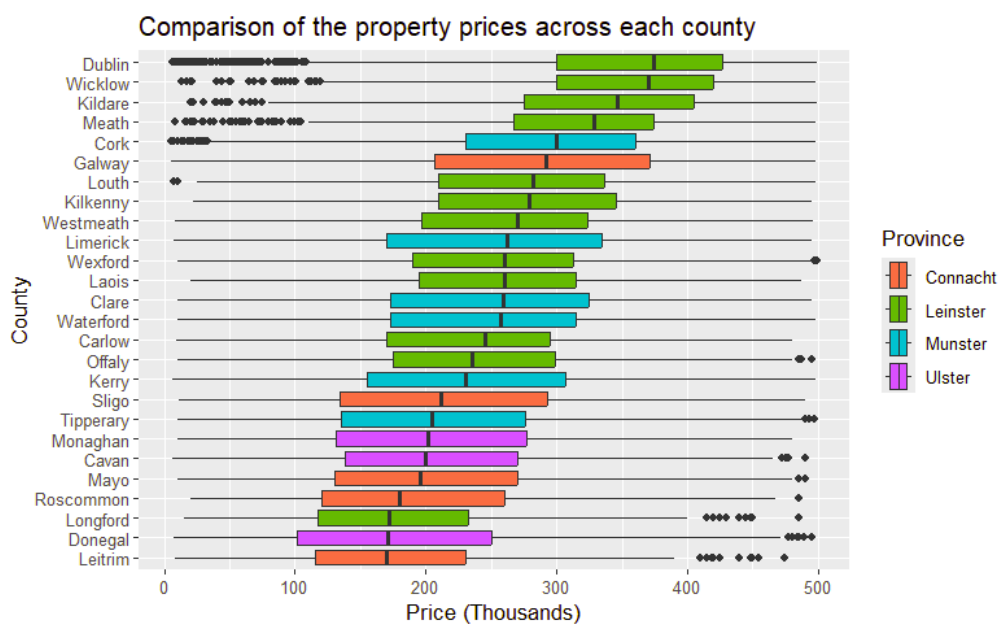


FIG. 5.2.1. BOX PLOT TO HIGHLIGHT THE COUNTY AND REGIONAL DISPARITIES IN HOUSE PRICES (2024)

5.3 Analyse house price evolution using tabular data

In my next study, I wanted to analyse the evolution of house prices from 2014-2024. Dublin has been highlighted as the county with the lowest overall percentage increase of 74.55% (In green), while Longford was identified as the county with the highest overall percentage increase of 225.71% (In red). On a nationwide scale house prices rose an average of 66.82 % between 2014 and 2019, 49.53% between 2019 and 2024, and 149.79% across the decade. The average absolute increase in house prices was also €150,109.10.

Initially, I had thought that Dublin would have the highest percentage increase overall with counties on the lower scale such as Leitrim and Donegal having the lowest percentage increase, based off my analysis in section 5.2. These results urged me to add a new column to my table, presenting the overall price increase over the ten-year period (in €'s). This provided insight into why the results did not agree with my initial predictions. The highest overall

house price increase (disregarding percentages) is Wicklow with an increase of €199,060, while the lowest overall increase being Donegal, with an increase of €94,000.

This analysis shows that while Dublin had the lowest percentage increase in house prices, it still saw significant price increases. Meanwhile, lower-priced counties like Longford experienced the highest percentage increase. My initial expectations that Dublin would experience the highest growth was likely due to its already high prices. By looking into the absolute change in prices and not on a percentage basis, initial expectations were satisfied. This highlights the importance of looking at both percentage and absolute change when analysing house prices.

| County | 2014 | Increase (2014-2019) | 2019 | Increase (2019-2024) | 2024 | Increase Overall (2014-2024) | Overall (€) Increase |
|-----------|--------|----------------------|-----------|----------------------|------------|------------------------------|----------------------------|
| Carlow | 110000 | 53.67% | 169033.9 | 44.94% | 245000 | 122.73% | 135000 |
| Cavan | 60000 | 126.67% | 136000 | 41.91% | 193000 | 221.67% | 133000 |
| Clare | 98750 | 72.15% | 170000 | 54.99% | 263480 | 166.82% | 164730 |
| Cork | 145000 | 58.62% | 230000 | 34.78% | 310000 | 113.79% | 165000 |
| Donegal | 78500 | 43.63% | 112750 | 53.44% | 173000 | 120.38% | 94500 |
| Dublin | 252077 | 34.88% | 340000 | 29.41% | 440000 | 74.55% | 187923 |
| Galway | 134000 | 60.45% | 215000 | 46.51% | 315000 | 135.07% | 181000 |
| Kerry | 105000 | 57.14% | 165000 | 42.42% | 235000 | 123.81% | 130000 |
| Kildare | 205000 | 32.05% | 270704.34 | 34.21% | 363300 | 77.22% | 158300 |
| Kilkenny | 112750 | 61.42% | 182000 | 59.34% | 290000 | 157.21% | 177250 |
| Laois | 83000 | 104.82% | 170000 | 52.94% | 260000 | 213.25% | 177000 |
| Leitrim | 60000 | 66.67% | 100000 | 71.40% | 171403.085 | 185.67% | 111403.15 |
| Limerick | 95000 | 89.47% | 180000 | 50% | 270000 | 184.21% | 175000 |
| Longford | 52500 | 109.52% | 110000 | 55.45% | 171000 | 225.71% | 118500 |
| Louth | 110131 | 84.33% | 203000 | 41.06% | 286343.61 | 160% | 176212.6 |
| Mayo | 81250 | 57.23% | 127753 | 52.64% | 195000 | 140% | 113750 |
| Meath | 158237 | 66.05% | 262750 | 27.50% | 335000 | 111.71% | 176763 |
| Monaghan | 80000 | 81.25% | 145000 | 37.93% | 200000 | 150% | 120000 |
| Offaly | 100000 | 50% | 150000 | 63.33% | 245000 | 145% | 145000 |
| Roscommon | 60000 | 80% | 108000 | 66.67% | 180000 | 200% | 120000 |
| Sligo | 94000 | 37.27% | 129036 | 68.56% | 217500 | 131.38% | 123500 |
| Tipperary | 90000 | 55.56% | 140000 | 47.14% | 206000 | 128.89% | 116000 |
| Waterford | 95000 | 76.84% | 168000 | 57.31% | 264281.07 | 178.19% | 169281.1 |
| Westmeath | 94500 | 74.60% | 165000 | 65% | 272250 | 188.10% | 177750 |
| Wexford | 107000 | 63.55% | 175000 | 50.79% | 263876.65 | 146.61% | 156876.7 |
| Wicklow | 215000 | 39.53% | 300000 | 38.03% | 414096 | 92.60% | 199060 |

FIG. 5.3.1. TABLE REPRESENTING THE EVOLUTION OF HOUSE PRICES (2014-2024)

5.4 Perform mixed model analysis to determine the influence of socio-economic factors on house prices

Lastly, I wanted to examine the underlying socioeconomic factors impacting house prices. I used the Pobal Deprivation index data (2016) along with house prices (2016) in a combined dataset. I constructed a linear mixed-effects model (Fig 5.4.1), incorporating county as a random effect. This allowed me to account for variability across counties. The purpose of this model is to determine the impact of factors such as, Pobal Score, proportion of age

dependency, proportion of individuals with third-level education, and the unemployment rate on the property prices across all the counties. Each of these factors is hypothesised to have an impact on the house prices. The null hypothesis (H0) of this model states that there is no effect of Pobal score, age dependency, third level education or unemployment on house prices. Conversely, the alternative hypothesis (H1) suggests that these factors do have an effect on house prices. By evaluating the p-values of the fixed effects model, we can test the hypothesis.

From the output in Fig 5.4.1, the model suggests that Age dependency rate and Third level education positively influence house prices, with t-values of 4.547 and 7.502 respectively. We can reject the null hypothesis and conclude that these factors do have a significant impact on house prices. Counties with higher levels of education tend to have higher house prices. Areas with a higher proportion of dependents (elderly and young) tend to have higher house prices. This could be due to increased demand for housing in areas with high dependency and also due to policies put in place to cater to these group. Unemployment negatively affects house prices, with a t-value of -2.51, indicating that areas of high unemployment having lower house prices, which can be put down to lack of demand. Pobal score has no effect on house prices, as reflected in its t-value of 0.023.

Looking into the correlation between the fixed effects, it is evident that there is strong positive correlation between Pobal score and unemployment as more deprived areas tend to have higher unemployment rates, as well as this, Pobal score and education are negatively correlated, indicating that deprived areas have less education. This analysis provided valuable insight into the main socioeconomic factors impacting house prices, which were age dependency and education, while factors such as unemployment negatively impacted prices and deprivation score had no effect at all.

```
Linear mixed model fit by REML ['lmerMod']
Formula:
log(Price) ~ Pobal_Score + Age_Dependency_Rate + Proportion_Third_Ed +
  Unemployment + (1 | County)
Data: electoral_data

REML criterion at convergence: 3888.2

Scaled residuals:
    Min       1Q   Median       3Q      Max
-5.3591 -0.4961  0.0898  0.6075  5.2056

Random effects:
 Groups   Name                Variance Std.Dev.
 County   (Intercept)          0.07312  0.2704
 Residual                    0.25475  0.5047
Number of obs: 2561, groups: County, 26

Fixed effects:
              Estimate Std. Error t value
(Intercept)    10.8362200  0.1413075  76.685
Pobal_Score      0.0001499  0.0064134   0.023
Age_Dependency_Rate 0.0124452  0.0027369   4.547
Proportion_Third_Ed 0.0189924  0.0025318  7.502
Unemployment    -0.0103298  0.0041150  -2.510

Correlation of Fixed Effects:
      (Intr) Pbl_Sc Ag_D_R Pr_T_E
Pobal_Score -0.050
Ag_Dpndnc_R -0.796  0.395
Prprtn_Th_E -0.291 -0.866 -0.143
Unemploymnt -0.365  0.845  0.522 -0.574
```

FIG. 5.4.1. MIXED MODEL OUTPUT OF SOCIO-ECONOMIC IMPACTS ON HOUSE PRICES (2016)

Chapter 6 – Conclusions and Future Work

6.1 Conclusion

This project provided some interesting insights into the current housing crisis in Ireland, in particular in the area of regional house price disparities, the effects of demographic changes on house price fluctuations and the impact of different socioeconomic factors on house prices. This project also contributes to a growing body of research into this issue, highlighting the need for policy changes and housing reforms if we are to ever make home ownership attainable. Reflecting on the specific questions framed in Chapter 3, here are the conclusions for the analysis in Chapter 5.

- **RQ1 - Compare house price trends with net migration trends using line plots:** The house price comparison with net migration figures study demonstrated the increase in the Irish population, which consequently caused a surge in house prices. This revealed a supply and demand issue within the housing market, where the government has struggled to keep up with housing demands. Addressing this challenge will require urgent policy reforms, including increased investment in housing supply and an increase in social housing to tackle the high costs.
- **RQ2 - Examine house price disparities on a regional and county level using boxplots:** The analysis of regional and county house price disparities revealed that Leinster and Munster have significantly higher house prices than Connacht and Ulster. It also showed that proximity to Dublin causes much higher house prices. Commutable counties such as Meath, Kildare and Wicklow ranked among the most expensive counties. This can be attributed to factors such as improved infrastructure and employment opportunities. Potential next steps to tackle these house price disparities are to improve transport links between lower priced counties and major job hubs to attract more home buyers. Also, the government could provide incentives to encourage businesses to set up in these counties, improving infrastructure.
- **RQ3 - Analyse house price evolution using tabular data:** This study revealed that despite its high house prices, Dublin had the lowest percentage increase in prices over the period of 2014 to 2024, while counties with lower house prices such as Longford and Donegal had the highest percentage increases of house prices over the same period. It was also found that overall, Ireland has seen an increase of 149.79 % over the decade or, in absolute terms, an increase of €150,109. While Longford had the highest percentage increase, Wicklow had the highest absolute increase (€199,060). This study showed that it is important to look at percentage increase as well as absolute change when analysing house prices.
- **RQ4 - Perform mixed model analysis to determine the influence of socio-economic factors on house prices:** The mixed model results revealed that age dependency and education were positively associated with house prices, while unemployment had a negative impact. In contrast, the deprivation score showed no significant effect on house prices. These findings highlight the strong influence of socioeconomic factors on house prices, suggesting that areas with higher education levels and higher age dependency tend to have higher house prices, while areas of low unemployment have lower house prices.

6.2 Personal Reflection

In undertaking this project, I was able to bring together all of the skills and knowledge that I have gained from my data science and computer science modules. I have also gained a lot of knowledge through my own study. From ST201: Data Analysis and ST203: R for statistics and data science, I learned techniques for data manipulation using functions from the **dplyr** package in R. This project required a significant amount of data manipulation as it included four different datasets, all of which contained variables and time frames that were not relevant to my data. The names of variables in these datasets also contained symbols and white spaces (e.g. **Age Dependency Rate** instead of **Age_Dependency_Rate**) which made referencing within R quite difficult. I was able to rename these variables easily using the **rename** function. In these modules I also learned how to analyse outputs such as graphs and tables, to extract valuable insights on the data. In the module ST302: Data Visualisation, I was taught how to effectively use visualisation techniques to get a certain point across. Within this project I used colouring and ordering in my price disparity comparison, which allows the reader to clearly see the regional disparities between house prices. I also ordered the data from highest to lowest prices for clarity. I used grids in RQ1 to allow for side-by-side plots in the output. Lastly, in ST401: Mixed effects models, I learned how to apply mixed model analysis to a given dataset, to extract valuable information about the impact of some predictor variables on an output variable. I also learned how to infer output from the model summaries.

Throughout this project I ran into many issues with gathering data and not getting the outputs that I had expected from the data. One of the major issues that I faced was in the case of my study of socioeconomic factors on house prices. The original intention was to compare the 2022 property price register with the 2022 Pobal deprivation index. I was unable to find a property price register from 2022 that included the electoral area for each house sale. This caused an issue when modelling as, I needed to use the electoral divisions to combine both datasets together in order to use this in my mixed model. While looking for this data, I ended up finding a property price register dataset (2012-2017) which included the electoral areas. I decided to use this dataset in conjunction with the 2016 deprivation index in order to answer the research question at hand. The idea behind this was to allow future researchers to use the techniques that I have applied on more recent data should it become available.

I also spent a lot of time going down avenues of research that ultimately led to dead ends or conclusions that I did not deem useful within this project. I had originally intended to investigate new house registrations in the major cities in Ireland, creating a forecast for registrations for the next three years. In creating a forecast, I found that the insights were not very useful and didn't successfully answer any research questions. Another area I looked into was to investigate the impact of socioeconomic factors on house prices, on a much smaller scale. I had applied the same model as used in this analysis to house prices and deprivation only in Limerick. I had chosen Limerick as it was the area with the lowest deprivation score in Ireland. By using only one county in the analysis, I found that there was not enough information within the data to get any reasonable answers to my research question. This led to the choice to look at the whole country as this allowed for a lot more data to be used, improving the model significantly.

6.3 Future Work

There are a number of potential areas of future research stemming from this research project. As stated in section 6.2, it could be of use to apply the same techniques that I used on more recent data to investigate the current socioeconomic factors impacting house prices, this could lead to the implementation of new policies to improve disadvantaged areas, improving infrastructure and home ownership attainment. With more time, it may be beneficial to analyse the evolution of house prices from a much earlier stage such as the start of the twentieth century, to see if the trends that have been seen in this research are different from the trends seen from a century ago. If there is a change in trends, it would be interesting to find reasons for this and see if there are any possible solutions to slow down the rapid increases in house prices. Another possible area of research that could be interesting is to use the property price register data to create a forecast of house prices for the next decade. This research would give the government an idea of which way the market is heading and will help to inform policies which will positively impact house prices.

In relation to a point that I had made earlier in my introduction, I believe that this research showed that for myself and others in my demographic, there is no evidence to suggest that homeownership will be attainable any time in the foreseeable future. With house prices at an all-time high and an increasing demand for new houses, without drastic measures put in place, it does not seem that this issue will be fixed any time soon.

References and Bibliography

Addressing Ireland's housing crisis: Urgent policy reforms needed (2024) *Social Justice Ireland*. Available at: <https://www.socialjustice.ie/article/addressing-irelands-housing-crisis-urgent-policy-reforms-needed> (Accessed: 05 March 2025).

Central Bank's Working Group on Housing and the Macroeconomy (2024) *Economic policy issues in the Irish housing market*, [centralbank.ie](https://www.centralbank.ie/docs/default-source/publications/quarterly-bulletins/quarterly-bulletin-signed-articles/economic-policy-issues-in-the-irish-housing-market.pdf). Available at: <https://www.centralbank.ie/docs/default-source/publications/quarterly-bulletins/quarterly-bulletin-signed-articles/economic-policy-issues-in-the-irish-housing-market.pdf> (Accessed: 05 March 2025).

Egan, P., McQuinn, K. and O'Toole, C. (2024) 'How supply and demand affect national house prices: The case of Ireland', *Journal of Housing Economics*, 65, p. 102006. doi:10.1016/j.jhe.2024.102006.

Lewandowska, G., Taracha, M. and Maciuk, K. (2023) (PDF) *socio-economic factors associated with house prices. evidence based on key macroeconomic aggregates globally*, *researchgate.net*. Available at: https://www.researchgate.net/publication/374406819_Socio-economic_factors_associated_with_house_prices_Evidence_based_on_key_macerconomic_aggregates_globally (Accessed: 06 March 2025).

Lima, V., Hearne, R. and Murphy, M.P. (2022) 'Housing financialisation and the creation of homelessness in Ireland', *Housing Studies*, 38(9), pp. 1695–1718. doi:10.1080/02673037.2022.2042493.

Lyons, R. (2024) *An analysis of recent trends in the Irish residential market 2024 Q4 - Year in Review*, *daft.ie*. Available at: https://ww1.daft.ie/report/2024-Q2-rentalprice-daftreport.pdf?d_rd=1 (Accessed: 05 March 2025).

Paris, C. and Frey, J. (2018) 'Demographic trends and changing housing systems in Northern Ireland', *Housing Studies*, 33(8), pp. 1264–1285. doi:10.1080/02673037.2018.1424805.

Privalko, I. and Maître, B. (2022) *Housing supplements and deprivation in Ireland*, <https://www.esri.ie/>. Available at: <https://www.esr.ie/article/view/1965/735> (Accessed: 05 March 2025).

Slattery, L. (2025) *Irish house prices climb at 'still very high' rate of 9.4%*, *The Irish Times*. Available at: <https://www.irishtimes.com/business/2025/01/22/irish-house-prices-climb-at-still-very-high-rate-of-94/> (Accessed: 05 March 2025).

Social Justice Ireland (2024) *Addressing Ireland's housing crisis: Urgent policy reforms needed*, *Social Justice Ireland*. Available at: <https://www.socialjustice.ie/article/addressing-irelands-housing-crisis-urgent-policy-reforms-needed> (Accessed: 06 March 2025).

CSO Housing Hub - <https://www.cso.ie/en/releasesandpublications/hubs/p-hh/housinghub/>

Annual Population Change data - <https://data.cso.ie/table/PEA15>

Geocoded Property Price Register Data - <https://www.shanelynn.ie/the-irish-property-price-register-geocoded-to-small-areas/>

Pobal Deprivation Index Data -

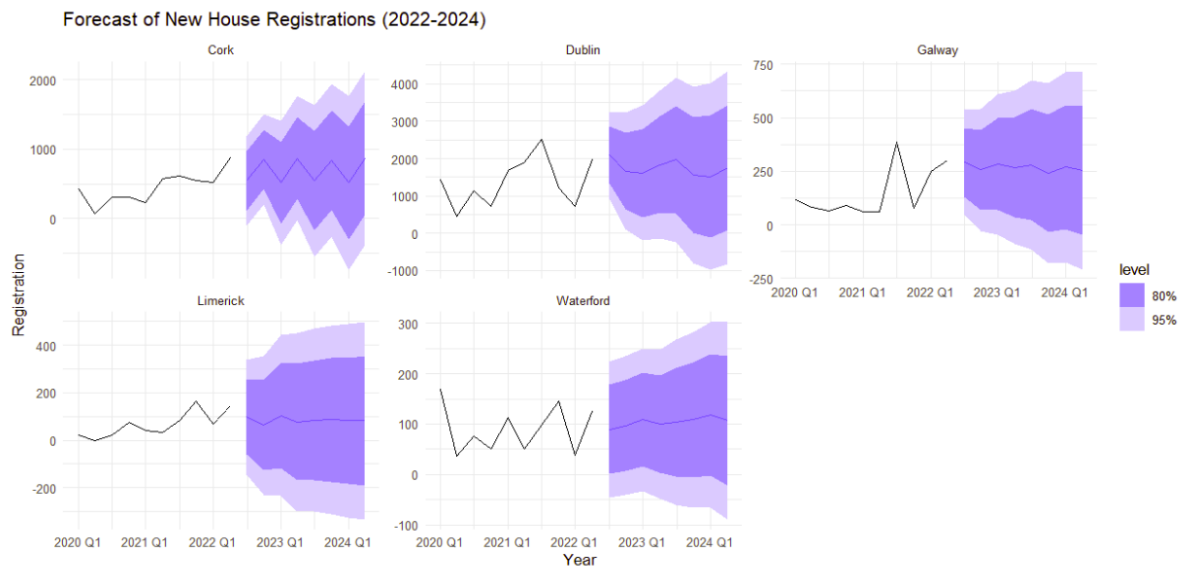
<https://data.pobal.ie/portal/apps/experiencebuilder/experience/?id=a184b6f21e9e499482340573c5d945b6>

Property Price Register Data -

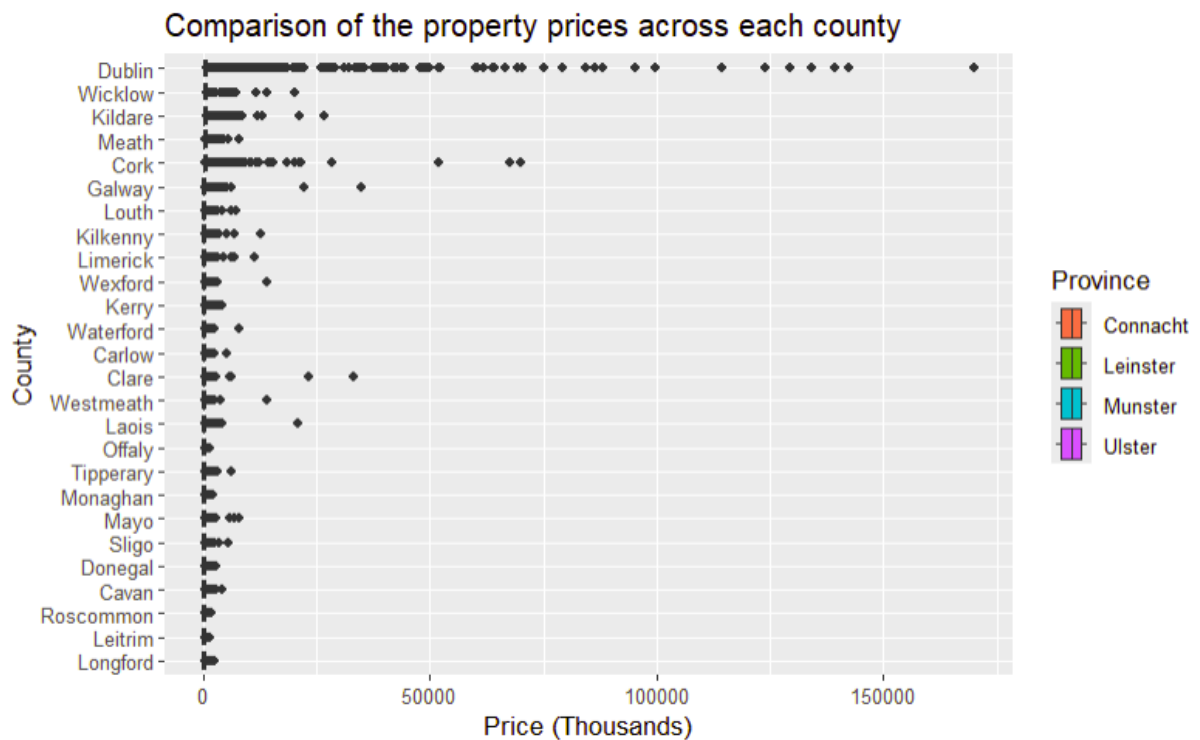
<https://www.propertypriceregister.ie/Website/NPSRA/pprweb.nsf/page/ppr-home-en>

Appendices

Appendix A: Forecast for new house registrations (2022-2024) for the five major Irish cities.



Appendix B: Regional and County disparities plot when not filtered to show prices less than €500,000.



Appendix C: Code to assign regions to each county

```
##{r}
Leinster<- c("Carlow", "Dublin", "Kildare", "Kilkenny", "Laois", "Longford", "Louth", "Meath", "Offaly", "Westmeath", "Wexford", "Wicklow")
Munster<-c("Clare", "Cork", "Kerry", "Limerick", "Tipperary", "Waterford")
Ulster<-c("Cavan", "Donegal", "Monaghan")
Connacht<-c("Galway", "Leitrim", "Mayo", "Roscommon", "Sligo")
Leinster_No_Dublin<- c("Carlow", "Kildare", "Kilkenny", "Laois", "Longford", "Louth", "Meath", "Offaly", "Westmeath", "Wexford", "Wicklow")
Province<- c("Leinster", "Ulster", "Connacht", "Munster")
```

Appendix D: Creating two new datasets for analysis of socioeconomic factors impacting house prices in Limerick

```
##{r}
Deprivation_Limerick<- DepIndexFixed %>%
  filter(County=="Limerick")
Prices_Limerick<- PropertyPriceRegister %>% filter(COUNTY=="Limerick") %>%
  mutate(Year=year(dmy(SALE_DATE))) %>%
  filter(Year==2021)
```