

Estimating missing housing and population data using CSO resources

Seán Healy, Soren Dreano
{healys47,soren.dreano2}@mail.dcu.ie *

December 1, 2021

Abstract

This work attempts to model missing historical and projection data for housing and population on the county level, using available property price, rent price and population data. Analysis was carried out in order to determine statistical correlates. Then linear regression was used for the purpose of determining missing historical property prices, missing population projections, and missing historical rent prices, all on the county level.

Keywords: accommodation, housing, population, rent

1 Introduction

The CSO provide a very extensive collection of data sets at data.cso.ie. These data cover rent and property price, population records, estimates and projections.

Despite the broad array of datasets, it remains difficult to statistically compare variables, since the datasets tend to vary in geographic granularity, interval granularity, and overall timespan.

In ascending order of size, the geographic granularities are **eircode prefix**, **county**, **region** and **state**. Online resources provide a mapping of these granularities to each other, e.g. eircodes are mapped within counties via crowdsourcing [1], as listed in Appendix A. Counties are then mapped to geographic regions by the CSO itself [2]. The remaining **state** level is an aggregate of variables across the entire Republic of Ireland.

Alongside this array of geographic granularities, there are a number of interval granularities used by CSO datasets. Some sets use **monthly** datapoints, others use **quarterly**, **biannual** or **yearly**.

Granularity is not the only incompatibility across data sources from the CSO. The overall timespan for data sets vary greatly. Population records are split across multiple data sets, some dating as far back as 1841. By contrast, county population for 2006 is only released under a standalone dataset. Some rent data is released between the years 2008 and 2021, whereas county-level property price records only begin in 2010.

After aggregation is applied to datasets, rendering them compatible, linear regression can be used to estimate missing data when a strongly correlated variable is present. Then a third issue arises: information overload. Much of the counties near Donegal, for example, follow similar trends for rent. This

makes data difficult to visualise and reason about. For this, “rent profiles” can be created via clustering. Finally, rent, population and property prices can be compared across these few profiles rather than across the overly granular county-levels, or the overly geographic CSO regional level.

The issues that this work attempts to address can be summarised under three headings:

1. Aggregation
2. Estimating missing data
3. Profiling

2 Methodology

2.1 Data sources

Data from the Central Statistics Office and WHEREVERYOUFOUNDYOURPDFIDONTRECALLTHEWEBSITE were used

2.2 Software used

To facilitate collaboration, we used a GitHub repository [6] to host the datasets and the code. As we were both familiar with Python3 and it has a large amount of libraries dedicated for data analysis and exploration, it is the language we went with. In term of libraries, we mainly used Pandas in order to parse and select the data sources, Matplotlib to create graphical representation and Numpy.

3 Results

3.1 Property price

Unsurprisingly, the mean price to buy a property decreased from 2010 to 2012 after the Irish Property Bubble collapsed. The price surged drastically in most counties and most notably in Dublin, Kildare, Wicklow, Cork and Galway. This trend is shown clearly in figure 2.

County	Increase in property price (%)
Dublin	36.53
Kildare	36.49
Wicklow	34.43
Cork	20.37
Galway	11.44

*This work was completed as part of the CA-660 module.

3.2 Rent price

As shown in figure 3, the same pattern applies for the average renting price. Figure 4 presents a tight correlation between the evolution of both property price and rent cost.

4 Factors

4.1 Population

As shown in figure 5, population in Ireland is fairly stable. Indeed, from 2012 to 2016, population in Ireland increased slightly by 3.05% while the mean property price increased by 15.83%.

The increase in inhabitants might drive prices in the housing market up, as there will be more demand for the same number of accommodation. In order to test this theory, we looked at correlations between the total population of a county and the mean property price, from 2010 to 2016, for each county. Figure 6 is a kernel density estimation of the correlation and shows how much this correlation varies from county to county, with the mean being 0.309 and the standard deviation 0.369.

The Donegal, Leitrim, Roscommon, Sligo and Tipperary counties actually have a negative correlation between the evolution of property price and the evolution of the population. Figure 7 shows how these counties are outliers in the general trend.

The property prices decreased while the population continued to increase in all of the aforementioned counties.

County	Pop. 2016	Pop. 2010	Evolution
Donegal	159192	157748	1444
Leitrim	32044	30934	1110
Roscommon	64544	62248	2296
Sligo	65535	64378	1157
Tipperary	159553	155268	4285

5 Discussion

In the later part of the Celtic Tiger, a term referring to the economy of Ireland from the mid 1990s to the late 2000s, the Irish Property Bubble started to appear, which manifested itself as a constant and continuous rise in prices for almost a decade. In September 2008, the Irish government officially acknowledged the country's descent into recession [3], the collapse of the previously mentioned bubble being one of the major contributing factor to the global 2008 financial crisis. Unemployment rate rose from 6.5% in July 2008 to 14.8% 4 years later [4] and house prices fell from 35% between the end of 2007 and the end of 2010 [5].

Concerns around property prices began to rise a few years later, in mid-2012, as seen in figure 1, the number of queries on the Google search engine for the term "Property price" started surging in this period and costs continue escalating in Ireland [6].

6 Conclusion

The current Tánaiste Leo Varadkar recently said "One of our biggest deficiencies, in housing supply in Ireland, is we're a country of three-bed homes by-and-large and we don't have enough one-bed homes" [7]. We could not find any publicly available data on the matter. Nonetheless, an analysis on the housing supply in Ireland either by size or the number of bedrooms, especially in dense urban areas such as Dublin, Cork or Galway might confirm this deficiency. Further research should focus on the types of housing that sell the most and that have been constructed recently.

Appendixes, if needed, appear before the acknowledgment.

References

- [1] Unknown, "Croudsourced eircode map," 2019, <http://web.archive.org/web/20180109194505/http://eircode.codes/>.
- [2] C. S. Office, "Regional population projections 2017 - 2036," 2016, <https://www.cso.ie/en/releasesandpublications/ep/p-rpp/regionalpopulationprojections2017-2036/appendix2conceptsanddefinitions/>.
- [3] J. Kollwe, "Ireland falls into recession," *The Guardian*, 9 2008, <https://www.theguardian.com/business/2008/sep/25/recession.ireland>.
- [4] C. S. Office, "Seasonally adjusted standardised unemployment rates (sur)," *Central Statistics Office*, 2014, <https://www.cso.ie/en/statistics/labourmarket/principalstatistics/seasonallyadjustedstandardisedunemploymentratesur/>.
- [5] H. Environment and L. Government, "Quarterly house prices bulletin," *Environment, Heritage and Local Government*, 2010, <https://web.archive.org/web/20160112210610/http://www.environ.ie/en/Publications/StatisticsandRegularPublications/HousingStatistics/FileDownload,25191,en.pdf>.
- [6] E. Burke-Kennedy, "House price growth surges to 10.9% as demand outstrips supply," *The Irish Times*, 10 2021, <https://www.irishtimes.com/business/economy/house-price-growth-surges-to-10-9-as-demand-outstrips-supply-1.4700410>.
- [7] D. McGrath, "Young people want one-bedroom apartments, not three-bed homes," *The Irish Times*, 11 2021, <https://www.independent.ie/breaking-news/irish-news/young-people-want-one-bedroom-apartments-not-three-bed-homes-varadkar-41089264.html>.

Appendices

Appendix A Eircode Prefixes

<i>County</i>	<i>Eircode Prefix</i>
Carlow	R21, R93
Cavan	H12, H14, H16
Clare	V14, V15, V95
Cork	P12, P14, P17, P24, P25, P31, P32, P36, P43, P47, P51, P56, P61, P67, P72, P75, P81, P85, T12, T23, T34, T45, T56
Donegal	F92, F93, F94
Dublin	A41, A42, A45, A94, A96, D01, D02, D03, D04, D05, D06, D6W D07, D08, D09, D10, D11, D12, D13, D14, D15, D16, D17, D18, D20, D22, D24, K32, K34, K36, K45, K56, K67, K78
Galway	H53, H54, H62, H65, H71, H91
Kerry	V23, V31, V92, V93
Kildare	R14, R51, R56, W12, W23, W34, W91
Kilkenny	R95
Laois	R32
Leitrim	N41
Limerick	V35, V42, V94
Longford	N39
Louth	A91, A92
Mayo	F12, F23, F26, F28, F31, F35
Meath	A82, A83, A84, A85, A86, C15
Monaghan	A75, A81, H18, H23
Offaly	R35, R42, R45
Roscommon	F42, F45, F52
Sligo	F56, F91
Tipperary	E21, E25, E32, E34, E41, E45, E53, E91
Waterford	X35, X42, X91
Westmeath	N37, N91
Wexford	Y21, Y25, Y34, Y35
Wicklow	A63, A67, A98, Y14

Appendix B CSO Regions

<i>Region</i>	<i>Counties</i>
Border	Donegal, Sligo, Leitrim, Cavan, Monaghan
Dublin	Dublin
Mid-East	Wicklow, Kildare, Meath, Louth
Midlands	Longford, Westmeath, Offaly, Laois
Mid-West	Clare, Tipperary, Limerick
South-East	Waterford, Kilkenny, Carlow, Wexford
South-West	Cork, Kerry
West	Galway, Mayo, Roscommon