

FINAL REPORT //

Cross-border housing markets in Europe

Conceptual framework

Annex No. 1: Technical Annex 1 // July 2022

This Final Report is conducted within the framework of the ESPON 2020 Cooperation Programme, partly financed by the European Regional Development Fund.

The ESPON EGTC is the Single Beneficiary of the ESPON 2020 Cooperation Programme. The Single Operation within the programme is implemented by the ESPON EGTC and co-financed by the European Regional Development Fund, the EU Member States and the Partner States, Iceland, Liechtenstein, Norway and Switzerland.

This delivery does not necessarily reflect the opinions of members of the ESPON 2020 Monitoring Committee.

Authors

Franziska Sielker – University of Cambridge, Department of Land Economy and Cambridge Centre for Planning and Housing Research (UK) and TU Wien, Institute of Spatial Planning (AT)

Brian Longobardi – University of Cambridge, Department of Land Economy (UK)

Efrain Larrea – MCRIT (ES)

Andreu Ulled – MCRIT (ES)

Luis Falcón Martínez de Marañón – inAtlas (ES)

Elisabet Palma Gibert – inAtlas (ES)

Technical Support

Nati Franco – MCRIT (ES)

Silvia Banchini – inAtlas (ES)

Alessandra Cappai – InAtlas (ES)

Mauri Arévalo Amaya – inAtlas (ES)

Advisory group

Marjan Van Herwijnen (project expert, ESPON EGTC)

Caroline Clause (financial expert, ESPON EGTC)

Information on ESPON and its projects can be found at www.espon.eu.

The website provides the possibility to download and examine the most recent documents produced by finalised and ongoing ESPON projects.

ISBN: 978-2-919816-50-7

© **ESPON, 2020**

Published in Luxembourg

Graphic design by BGRAPHIC, Denmark

Printing, reproduction or quotation is authorised provided the source is acknowledged and a copy is forwarded to the ESPON EGTC in Luxembourg.

Contact: info@espon.eu

FINAL REPORT //

**Cross-border housing
markets in Europe**

Conceptual framework

Annex No. 1: Technical Annex 1 // July 2022

Table of Contents

Table of Contents	5
List of Figures	7
List of Tables	8
1 Introduction & Structure of Annex	9
2 Literature Review	10
2.1 Specificities of Housing Markets	10
2.2 Types of Policy Interventions in Housing	10
2.2.1 Statutory Control Over Housing Development	11
2.2.2 Fiscal Control Over Housing Development	12
2.2.3 Public Sector Housing Construction	12
2.2.4 Universal Housing Policy	12
2.3 International Trends in Housing Markets	13
2.4 Housing and the European Union	14
2.5 Border Studies and Housing	14
2.6 Conceptualising and Defining Cross-Border Housing Markets	16
2.7 Conclusion	17
3 Case Study Selection and Methodology	18
3.1 Case Study Selection	18
3.2 Representativeness of a variety of cross-border regions in Europe.	19
3.3 Overview of Methodology of Case Study Implementation	19
4 Indicator Development & Project Update	21
4.1 Obstacles in Developing Indicators for a Cross-Border Housing Market	22
4.2 Logic behind Creating New Indicators for Cross-Border Housing Markets	22
4.2.1 Selection of New Indicators & the Border Effect	24
4.3 Step-by-Step Implementation of Indicators	25
4.3.1 Processing the Big Data file	25
4.3.2 Obtaining complementary data	25
4.3.3 Updating previous indicators	26
4.3.4 New Indicators Accounting for the Border Effect	29
5 Cross-Border Accessibility & Housing	32
5.1 Preparation of the GIS for analysis	32
5.2 Locating the services of general interest (SGI)	34
5.3 Calculating accessibility	35
6 Web scraping of real estate online listings	37
6.1 Generic web scraping methodology	38
6.1.1 Scraping steps	39
6.1.2 Possible scraping problems:	41
6.2 Real Estate websites selection criteria	41
6.2.1 Portal rating	41
6.2.2 Common websites	42
6.2.3 Listings amount	42
6.2.4 Website structure	42
6.2.5 Website content	42

6.2.6	Location data.....	42
6.3	Real Estate websites initially selected	42
6.4	Alternative websites:	43
6.5	Type of data available	43
6.6	Case study specificities.....	44
6.6.1	France-Switzerland Case Study.....	44
6.6.2	Denmark-Sweden Case Study.....	46
6.6.3	Ireland – Northern Ireland Case Study.....	49
6.6.4	Austria – Slovakia Case Study.....	51
6.6.5	Romania – Bulgaria Case Study	53
6.6.6	France – Spain Case Study	55
6.6.7	Summary of scraped portals	57
6.7	Data cleaning, normalization and enrichment.....	57
6.7.1	Methodology.....	58
6.7.1	Scraping results summary.....	62
6.8	Next Steps for Further Studies.....	64
6.8.1	Duplicity of listings.....	64
6.8.2	Length of advertisements.....	65
6.8.3	Scraping monitoring	65
7	Bringing It All Together: Approach to Linking Different Information Sources and Work Processes	67
	References	72

List of Figures

Figure 1: Map of All Case Study Regions	18
Figure 3: Graphic Displaying Three-Step Process for Developing Cross-Border Indicators for Housing Markets	23
Figure 2: Graph displaying differences in PPS per Inhabitant	27
Figure 4: Example of Mean Difference in Income Affordability Indicator	30
Figure 5: Example of navigable graph for the case study of Denmark-Sweden	33
Figure 6: Setting up the Network Analyst tool of ArcGIS	33
Figure 7: Checking the network model results against Google Maps	34
Figure 8: Table of SGI Labels from the Sweden-Denmark Case Study.....	35
Figure 9: Example of calculation of accessibility for the Spanish-French case study	36
Figure 10: Example of calculation of service areas for education centres by car and walking in the Sweden-Denmark case study.....	36
Figure 11: Data Workflow Overview.....	37
Figure 12: Examples of Page in Nestoria.....	40
Figure 13: Example of Content in Nestoria	40
Figure 14: Example of coordinates collection in Top Reality	41
Figure 15: Screenshot of Nestoria France	45
Figure 16: Screenshot of Homegate	45
Figure 17: Screenshot of Boligsiden	46
Figure 18: Screenshot of Boliga	47
Figure 19: Screenshot of Blocket Bostad	48
Figure 20: Screenshot of Hemnet	49
Figure 21: Screenshot of Daft	50
Figure 23: Screenshot of Nestoria UK	51
Figure 24: Screenshot of Top Reality (Austria)	52
Figure 25: Screenshot of Nestoria Austria	52
Figure 26: Screenshot of Top Reality (Slovakia).....	53
Figure 27: Screenshot of Publi24	54
Figure 28: Screenshot of Realestates	55
Figure 29: Screenshot of Nestoria FR.....	56
Figure 30: Screenshot of Nestoria ES.....	57
Figure 31: Histograms of Switzerland dataset (rent and sale listings for multifamily housing typologies) ...	61
Figure 32: Project Overview Graphic	67

List of Tables

Table 1: Case Study Characteristics	19
Table 3: Overview of Reasoning for New Indicators	25
Table 2: GDP per Capita Measured as PPP	26
Table 4: Selected real estate websites information.....	43
Table 5: Alternative websites proposed	43
Table 6: Gathered indicators	44
Table 7: Scraping Summary.....	57
Table 8: Listing Typology Taxonomy	59
Table 9: Multifamily Housing Typologies Taxonomy.....	60
Table 10: Single-family housing typologies taxonomy	60
Table 11: Scraping Results Summary.....	64
Table 12: Spain Fotocasa Results Comparison (sample of 26-03-2022)	65
Table 14: Interview Overviews	69

1 Introduction & Structure of Annex

The technical guidance document is delivered in conjunction with the main report for ESPON 2020 Data and Maps Updates II Lot 4 Updating and Integrating Big Data and Housing Datasets. The Technical Annex presents an extended literature review on cross-border housing markets, the development of indicators for measuring cross-border housing markets, research methodologies, technical processes and project procedures used in preparation for the final delivery of this project.

Specifically, this includes a review of general trends in housing markets and a description of the key characteristics in how a housing market functions as well as an outline of housing policy types and border region particularities. A key task for the project is the analysis of indicators developed from the “ESPON Big Data for Territorial Analysis and Housing Dynamics - Wellbeing of European citizens regarding the affordability of housing” project for the six new cross-border case studies. As such, this report summarises the technical procedures for the update of these indicators, the results of which are presented in the Case Study Annexes 2-7 as well as in the Main Report. The report then outlines the rationalities and reasoning behind the development of new indicators for measuring cross-border housing markets while also highlighting the role in which accessibility can impact decisions on housing.

This ESPON project makes use of innovative web scraping techniques by aiming to make use of the breadth of data available from relevant real estate webpages in order to capture housing market dynamics. As such, the project team found that data available on real estate webpages allows for a good overview of the level of activity in markets across Europe. Additionally, it allows for a good comparison of price ranges across European markets.

However, the results presented here must be read against the background of the COVID-19 pandemic and new developing trends in the housing markets. New desires in housing amenity such as access to green space and home office space have become more important. While the pandemic initially slowed down the real estate market at the advent of the crisis, its long-term progression quickly led to a very active market specifically in the owner-occupied sector. By Spring 2022, the key timeframe of analysis for our web scraping, the results may be skewed by a comparably different selection of offers than in the previous years of the pandemic. This is partly due to the expectation of housing stock which had been sold in 2021 to achieve higher market values due to property speculation. While these unprecedented implications create a somewhat uncertain context for interpreting the residential property market, the use of web scraping techniques has allowed for more detailed analysis along with the identification of price ranges. This is particularly relevant as data about actual transactions and paid prices is complicated or expensive to access for many countries. A shortcoming of the data is that often the most sought about segments of the market are not always publicised online. This can be overcome by analysing the publication timeframe of advertisements tested within this study. However, in order to develop a more robust interpretation, longer time periods would need to be scraped and analysed. Due to the short project timeframe and budget constraints, the scraping was limited to three months. Nevertheless, the project was able to identify best practices for web scraping activities, and around price differences across borders were retrieved as well as a comparison across the cases was done.

This Technical Annex is composed in the following structure: Chapter 2 reports a literature review which contextualises the project within strands of housing and cross-border research. Chapter 3 discusses how the case studies were selected along with the methodology for reviewing the case studies. Chapter 4 informs on how the indicators were developed for a cross-border framework as well as providing an update for the previous project. Chapter 5 examines how the concept of accessibility played a leading role in the conceptualisation and understanding of cross-border housing markets. Chapter 6 consists of the web scraping methodologies and associated technological updates. Chapter 7 then incorporates the web scraping exercise into the production of indicators so as to produce new findings on cross-border housing markets across Europe.

2 Literature Review

In this chapter, we draw on literature stemming from both border studies and housing studies to conceptualise cross-border housing markets as a definitive subject which transects the two literature streams.

Overall, we

- identify international trends affecting housing markets globally
- develop categories for interpreting different types of housing policies
- examine similar studies relating to cross-border spatial dynamics and housing market indicators
- define and conceptualise what a cross-border housing market is.

2.1 Specificities of Housing Markets

Housing and shelter are basic needs for human beings as adopted by the UN in its right to adequate housing (OHCHR, 2009). Yet, the particularity of housing markets in conjunction with the specificities of the land market make financial policies and socio-economic policies rather complex. The specificities of the housing and the subsequent market incorporate the following statements:

- Housing is expensive relative to incomes and the rate new, smaller household formations have available for acquiring independent accommodation. Examples are independent accommodation requirements for e.g. singles or young couples at the beginning of their career.
- Borrowing is influenced by interest rates, which creates uncertainty and differs via creditability.
- Houses are durable and have a value for generations
- Houses are consumer goods and investments
- Housing supply is inelastic: no quick changes in supply is possible due to building time
- Housing stock is immobile
- Land is a scarce resource
- Housing is a necessity (people cannot opt out of consumption)

As in other markets, the housing market is the result of demand and supply. The supply side is dependent on the number of housing completions which is first a result of global markets and supply chains as well as existing labour force, and second, by the planning permissions given or the number of lots cleared for housing developments. Additionally, housing supply includes the existing housing stock that comes onto the market again for resale and excludes any existing stock which changes from housing use into another use, or leaves the market for example for reasons of age.

In view of housing demand, long-term variables such as demographic, social or cultural factors change the demand side. Key variables that impact housing demand in the last decades are the emergence of new household formations. These new household formations are a result of the reduced average size of households, a higher number of family breakdowns, divorces and separations, kids leaving home sooner, humans living longer and living longer independently as well as the unpredictability of net migration. These developments make it challenging to prognose housing needs for the future (Chilla & Sielker, 2016).

2.2 Types of Policy Interventions in Housing

The nature of housing policy tools is diverse and extends across multiple sectors as housing development impacts a multitude of different but interrelated fields (Shucksmith & Sturzaker, 2016). Hence, it is necessary to predefine what this study means when it refers to 'housing policies'. As this research pertains to a number of different regions with varying socio-political and socio-economic conditions, developing a common core understanding to what housing policies are across Europe can better help us analyse and understand the dynamics occurring across cross-border housing markets (Lund, 2011).

The four types of housing policy developed for this research are as follows:

- Statutory Control Over Housing Development
- Fiscal Control Over Housing Development

- Public Sector Housing Construction
- Universal Housing Policy

Within these four categories, we have identified specific tools used to implement such policy in the subsections below.

By categorising housing policies across four different categories, we can encapsulate common policy tools which governments develop to enact and enforce housing policymaking. By categorising housing policy tools within these typologies, we can better interpret and understand the behaviours and dynamics of complex cross-border housing markets (Shucksmith & Sturzaker, 2016). In turn, we can better determine and develop meaningful indicators which can extract critical data central to cross-border markets.

European countries have different forms of housing policies which the European Parliament has placed into four groups back in 1996 that are still relevant today. The first group which is represented for example by the UK or the Netherlands makes use of state interventions to organise housing and are prominent examples for a socially rented sector. The second group of which Austria or France are examples have a large privately rented sector. The third group which Ireland, Belgium and Finland belong to have a relatively small social rented sector and large owner-occupied sectors. The fourth group of which Portugal, Spain and Greece are examples have large owner-occupied sectors with minimal social rented sectors (European Parliament, 1996). While the reference is yet somewhat outdated, it was taken up in this context as it one of the only published sources from the European level. However, since then, much has happened in European Housing markets, and the lines are blurring increasingly between countries. However, until today there does not exist an updated European overview, and comparative research is often limited to specific cities, a research gap that ESPON may fill in the future. Every country displays a different set of these policies, with different competences at individual levels. In the context of this report, this will be explored in more depth in each case studies (see Annexes 2-6).

2.2.1 Statutory Control Over Housing Development

One of the primary forms of government intervention into the housing sector is through legislation which grants the government statutory control over housing development (Shucksmith & Sturzaker, 2016). Through a suite of regulations, states institutionalise property development processes into comprehensive planning systems which allows for the sustainable growth of regions (White & Allmendinger, 2003). States are able to regulate the housing sector via measures such as land-use zoning, planning controls, building regulation and land management (White & Allmendinger, 2003). While different governments choose to implement different instruments for their region's relevant needs and unique requirements, this category of housing policy is characterised through direct government intervention into the property development process (Somerville & Sprigings, 2005).

Across the case studies selected for this research, government interventions range from:

- enforcing zoning codes
- institutionalising the housing development process through granting planning permissions
- applying building controls to new development
- producing a design code which adheres to sustainability principles
- electing an oversight committee to make final decisions

As a consequence of these direct interventions, states build large bureaucratic administrative systems to manage these processes. Hence, by reviewing, understanding and accounting for such direct controls, we can produce a more accurate picture as to how states impact regional housing markets in cross-border regions. In cross-border regions the differences between these legal systems often create a barrier for movement across the border. For example, in order to acquire housing, it may be advantageous to know where zoning regulations may allow for future developments. This kind of information may be more difficult to acquire or observe on the other side of the border. Further building or buying at the other side of the border requires citizens to understand another legal system. This is an additional barrier, but by no means has held back cross-border movements if other incentives make up for the additional burden. Secondary houses or holiday houses are prominent examples of acquisitions in other states. Some places, such as in some local authorities in Austria, are now starting to experiment with regulations to bind new housing stock to a reallocation of one's primary seat of living. This indicates that if the

advantages outweigh the additional burden, these regulatory differences are a complication but not necessarily hindrance.

2.2.2 Fiscal Control Over Housing Development

Another way for the government to exhibit control over the housing sector is through fiscal and monetary policies which impact its housing market (Shucksmith & Sturzaker, 2016). Examples of fiscal policymaking include:

- introducing tax measures to affect market dynamics
- subsidising mortgages or rent through financial schemes
- incentivising new supply through grants
- funding demand-side allowances for certain populations
- controlling macroeconomic monetary policies to stimulate banking activity

Identification of fiscal interventions by the state into the housing sectors will allow us to better grasp housing market dynamics and more closely understand unique characteristics to cross-border housing markets with varied fiscal housing policy tools (Lund, 2011). Understanding fiscal housing policy is particularly insightful into developing a better grasp of socio-demographic and migratory dynamics in the case study regions selected for this study (White & Allmendinger, 2003). To illustrate, if one national government has looser financial regulation, lower mortgage rates and more affordable housing supply than its neighbouring state, the cross-border housing market could impact where people ultimately settle and commute. This study has the potential to illustrate how regions which align and integrate fiscal controls across borders can better manage externalities produced with economic growth (Stamsø, 2010), and what the relation between and dependencies from competence levels is. Investments by the state are one explanation for the size of a social housing sector and are often linked to different mixes of welfare policy regimes in these states. Overall, however, the dependency of local housing markets on fiscal policies by states and financial institutions is not to be underestimated.

2.2.3 Public Sector Housing Construction

Another example of a housing policy tool is through directly supplying housing to inhabitants (Shucksmith & Sturzaker, 2016). The supply of public or social housing varies across countries but is a particularly dominant form of housing tenure in many European countries (Somerville & Sprigings, 2005). Additionally, the supply of public housing has evolved beyond simply the state constructing accommodation to advanced systems which develop public housing through a number of different types of schemes.

Examples of such varied public housing development can range from state-owned rentals to public-private partnerships which stipulates how a certain percentage of each residential development must be reserved for affordable housing units.

Hence, affordable housing bodies have been established across a number also from the case study regions analysed here in order to better support the public sector in supplying housing to its inhabitants. By accounting for regions and cities which have a high percentage of its population living in public housing will provide further depth to our study and understanding of cross-border housing markets (Somerville & Sprigings, 2005).

2.2.4 Universal Housing Policy

Another form of government intervention into housing policy is through policymaking which accounts for more vulnerable population groups (Lund, 2011). Universal housing policies are attempts to focus certain policy toward groups more at risk to housing precarity such as:

- low-income households,
- travelling communities,
- the elderly,
- immigrant communities,
- refugees and migrants,

- at-risk minority communities,
- the homeless

These type of housing policies seek to account for all sectors of the population's housing needs through a range of policy tools. Examples include:

- Affordable housing stipends
- Social housing allocation
- Nursing homes
- Halting sites
- Homeless shelters
- Resettlement strategies
- Housing anti-discrimination legislation

The overall similarities in these types of housing actions relate to ensuring fairness across different socio-economic backgrounds within society and providing universal benefits to those more in need of housing support (Hohmann, 2013). Many of these types of policies are focused on eliminating homelessness and providing housing for all across society (Stamsø, 2010).

2.3 International Trends in Housing Markets

Key international trends affecting housing sectors around the globe are the financialization of housing markets, the commodification of housing stock, the rise of corporate landlordism and growing prices differentials in land value.

The financialization of the housing markets speaks to the transformation of housing as a social good to a commodity. Financialization in its most basic term can be understand as “the process by which something or someone is managed as a fund [... where] the increasing dominance of financial actors, markets, practices, measurements and narratives, at various scales, resulting in a structural transformation of economies, firms (incl. financial institutions), states and households” (Aalbers, 2016). Rolnik (2019) indicates a close relationship between governmental policies and an increase in ownership of real estate stock by construction companies, real estate developers and banks through government-facilitated schemes. Through schemes such as the “right to buy” scheme originating in the UK, the role of the banks and their treatment of potential private clients in allowing citizens to get onto the housing ladder has been criticised widely (Aalbers 2016). Peter Marcuse (2016) has long been a key critical voice outlining the profound challenges of the global housing system. At the same time, the changed relationship between the finance sector and developers has specifically opened more doors for the further commodification of housing stock since the financial crisis of 2007/2008. This is a trend that Marcuse warns will exacerbate social injustices (Marcuse & Madden, 2016).

In treating housing as a vehicle to further wealth and investment, the commodification of housing stock across urban regions has resulted in patterns of systemic exclusion with higher levels of spatial segregation, greater income inequality, growing homelessness and inadequate housing provisions all a consequent of this financialization process (Farha, 2017). In using housing markets as a sink for excess capital in the global flow of cash, housing is perceived as a speculative asset that can be invested in. Hence, luxury housing that sits vacant is built in urban regions over affordable housing requirements as the global financial market dictates development trends instead of meeting local needs (Farha, 2017). This international financialization of housing has been a global phenomenon impacting all European markets through its commodification of housing stock where housing markets in urban and metropolitan areas display substantial profit margins (Christophers, 2021).

As a consequence of the international financialization of housing markets, the commodification of housing has permitted the rise of corporate landlordism (Raymond, et al., 2016) and global corporate landlords. In purchasing large amounts of property and housing (particularly at the downturn of the Global Financial Crisis), global investment funds have become largescale landlords which can control rental prices in local markets through the monopolisation of the private rental sector (Raymond, et al., 2016). Through the process of financialising the housing market and commodifying housing as an asset, these corporate entities have become dominant players across housing markets and are able to manipulate prices and control the building process (Farha, 2017). As a

result, tenants are being forced to pay increasingly higher rents for units which are of lower quality and smaller spaces as a consequent of new accommodation being perceived as an asset instead of an actual home.

Issues generated from the financialization of housing markets are particularly acute in areas considered to be prime land (Farha, 2017). Measures of accessibility and connectivity are consequently having a dramatic impact on land prices as value is linked with location and its mobility (Wittowsky, et al., 2020). Hence, housing in central areas with strong transport connections are increasing in value at a faster rate than housing in peripheral areas with weaker transport connections. This has knock-on effects in terms spatial socio-economic segregation and service provision equity as households with lower incomes are physically forced to live in the periphery of urban regions with less access to public goods and services (Farha, 2017). Even more so, international investors target rundown areas in cities with a “buy it – fix it – sell it” approach in order to make use of the highest rent gaps (Christophers, 2021). These processes are seen as attempts to induce gentrification, a process of transformation where the less affluent are pushed out of an area as more affluent household displace them. Today’s financialization trends are now leading to processes of supergentrification where middle-income families and higher-income families are being displaced by the super-rich in a process also described as the third wave of gentrification in places like London (Butler & Lees, 2006).

Given this research’s focus on border regions though, general international trends across the housing sector need to be contextualised to the unique positions which border regions hold at the periphery of separate states. By analysing border regions, we can focus more closely on the dynamics produced by national boundaries in a context that considers international trends while adequately accounting for unique implications of border differences.

2.4 Housing and the European Union

The European Union does not have a competence over housing policies following the subsidiarity principle that acts as the basis for EU policies. As such the European Union does not have any legislation over housing. Yet, with increasing attention being paid to housing markets as a result of a wide-ranging affordability crisis in Europe (and the world) has put more pressure on politicians to react. In 2021, the European Parliament called on EU Member States to solve the housing crisis by investing in “more decent homes and to ensure everyone has access to affordable housing” (European Parliament, 2021a). The European Parliament adopted a resolution in 2021 which puts the right to adequate housing at its core and suggests enshrining the human right into European law (European Parliament, 2021b). This resolution marks a turning point on attention towards housing at the EU level. As such, it is expected that increasingly projects funded through the European Recovery and Resilience Fund or through the European Regional Development Funds may focus on housing as a topic.

2.5 Border Studies and Housing

As our understanding and conceptualisation of social spaces have evolved, the expansion of border studies across different disciplines represents our recognition that borders are dynamic entities which transcend simple physical boundaries (Kolossoff & Scott, 2013). Of specific relevance to this project is the work of Christoph Sohn on the “border as a resource” (Sohn, 2013). He suggests that border regions, and specifically for metropolitan regions, may at times use the existence of the border to build on comparative advantages in their formation. He developed his theory around the function of the border as a resource by analysing and reviewing the Luxembourg metropolitan region and accounting for the unique attributes associated with its proximity to several international borders. In conceptualising and understanding the impacts of a border on housing markets, the project consortium will revisit the question of the border as a resource.

While much of the literature within social science border studies relates to understanding, modelling and conceptualising cross-border regions with regard to labour market integration or European policymaking impacts (Sohn, et al., 2009; Fricke, 2015; Decoville & Durand, 2019), few studies delve into how housing specifically impacts the development of a cross-border region. Instead, cross-border literature tends to focus on either specific case study analyses for comparative purposes (Claveres, et al., 2020; Saadi, et al., 2016; Sohn, et al., 2009; Fricke, 2015), measuring cross-border flows and dynamics through the development of cross-border indicators (Chilla & Heugel, 2022; Decoville & Durand, 2019; Heider, 2018; Saadi, et al., 2016; Mitrică, et al., 2016) or defining and understanding different inter-territorial governance structures (Christmann, 2015; Decoville & Durand, 2019; Decoville, et al., 2013; Durand & Decoville, 2020; Kolossoff & Scott, 2013; Sohn, 2013). Some studies do acknowledge how economic asymmetries in property markets can act as an impetus for residential migration within a border region (Sohn, 2013; Christmann, 2015), yet few studies conceptualise cross-border housing markets as

an integrated feature of a border region despite extensive analysis of how labour markets have integrated across these regions due to European integration processes (Claveres, et al., 2020; Dumeignil, 2022).

Likewise, in housing studies, while housing markets in border regions have been understood to have unique cross-border dynamics, few studies incorporate the entire border region into their analysis and instead focus on either specific sub-regions or specific aspects of the market to study (Stevenson, 2004; Micheli, et al., 2019; Fullerton & Villalobos, 2011). Given the difficulties in collecting data across international borders along with researchers seeking to maintain consistency across their work, the conceptualisation and study of cross-border housing markets have not yet been fully realised due to unknown variables and unpredictable characteristics skewing data and challenging final assessments.

As cross-border housing markets are further understood and developed, we found many studies which start to relate how aspects of housing can impact cross-border dynamics. Cross-border studies mentioning housing frequently speak to how it is necessary to develop ways to measure commuting patterns, labour markets and residential migration to better understand how border regions are developing (Chilla & Heugel, 2022; Dumeignil, 2022; Claveres, et al., 2020; Durand & Decoville, 2020; Decoville, et al., 2013). Within this strain of cross-border literature, housing is considered as a topic of discussion with specific regard to how changing housing patterns are impacting mobility for cross-border workers (Decoville, et al., 2013; Chilla & Heugel, 2022; Dumeignil, 2022). Extending this analysis to encompass the housing market itself has not yet been incorporated into many border studies where analysis of price differentials and supply and demand has not been scrutinized as essential aspects to understanding cross-border regions.

Hence, the result is that much of the border literature is case study-led as border theory development is thwarted by unique geographical perspectives across different cross-border case study regions (Kolossoff & Scott, 2013; Decoville, et al., 2013). Likewise, the development of indicators to measure and compare different border regions is often stifled as issues on collecting and harmonising different datasets across different regions makes the process itself difficult. Therefore, comparative studies incorporating quantitative data across multiple border regions are uncommon due to issues in compiling, coordinating and corroborating data from different authorities and spatial levels. This missing gap in the literature illustrates the need to build on comparative studies then by developing cross-border indicators for housing markets in integrated border regions.

Given this difficulty found across border and housing studies, the literature has identified characteristics of a cross-border housing market despite a lack of conceptualisation on the subject. Here are some examples of how cross-border housing markets have been approached across different studies:

- Dumeignil, 2022: Through a study on how an integrating cross-border labour market can impact real estate prices within the Greater Geneva region, Dumeignil suggests that cross-border workers have a unique impact on real estate markets due to their higher skill level and incomes impacting regional differences in income while proximity to the border also plays a role in real estate prices throughout the border region.
- Christmann, 2015: Christmann speaks to the need develop discourse relating all the different factors impacting housing systems in an emerging transnational market developing across the Greater Region of Luxembourg.
- Micheli, et al., 2019: In studying the border region between the Netherlands and Germany, Micheli et al investigates the presence of price discontinuities at a country's boundary and seeks to understand the push factors motivating residential migration when arbitrage allows a cross-border housing market to develop.
- Fullerton & Villalobos, 2015: In performing a hedonic price model on the American sub-market of the El Paso – Ciudad Juarez conurbation at the US-Mexico border, Fullerton & Villalobos found that surface access to an international market (ie proximity to a port of entry) seems to add value to residential real estate in the El Paso sub-market potentially indicating how border crossings can impact housing markets monetarily, and in this case raise the value of the built environment.
- Decoville, et al., 2013: In this expansive study, Decoville et al. speak to how the development of cross-border relations does not imply a reduction in disparity or an increase in territorial cohesion across European borders. They specifically highlight how housing markets play an important role in relation to the mobility of workers as differences between property, land and rental prices constitute powerful factors in encouraging a change in country residence for workers.

Following this overview of border and housing studies, we can thus combine these streams of literature to produce a definite, discrete concept of a cross-border housing market. By conceptualising these integrating markets, we

can approach the development of a cross-border indicator for housing markets with a view to understanding and expanding our knowledge on how housing markets impact cross-border regions.

2.6 Conceptualising and Defining Cross-Border Housing Markets

Our contribution to this growing body of work seeks to recognise that housing markets are impacted by the effects of border crossings and international boundaries. In recognising the role housing markets play in border regions, we define the concept of a cross-border housing market as:

a housing market in which a critical mass of activity has led to a minimum integration where households work and live on different sides of a border. As a consequence, the distribution of jobs and residences across a border forms new functional relations.

Further, the definition of a cross-border housing market incorporates how complex internationally driven market dynamics have evolved as a consequence of cross-border movements driving changes in demand leading to differences in both price levels and household responses. This integration has been made possible by ensuring that a minimum level of accessibility into the territory of the other state is maintained and links foreign households to proximity of job markets near international borders. Reasons for moving include regional differences in income, affordability, cost of living and amenity. Border regions may consequently attract specific migratory patterns that change the dynamics of demand and supply in housing across the border.

Hence, cross-border housing markets are housing markets where price differentials from supply and demand are determined by the dynamics of cross-border movements. This dynamic can develop over a border due to differences in cost of living, in accessing services of general interest and in amenities motivating households to migrate (Chilla & Heugel, 2022). These cross-border households are seeking to take advantage of differences in regional income levels in order to benefit from the economic asymmetries present across the international border (Sohn, 2013). For instance, a cross-border household which earns income in the wealthier region but lives in the poorer region receives a comparative price advantage when accessing certain goods such as housing (Decoville, et al., 2013). When this cross-border flow reaches a critical mass of activity, prices in the poorer region are impacted by the cross-border dynamic, and a cross-border housing market develops as prices are linked to the expanding catchment area of the wealthier incomes.

The overall consequence is the development of a functional cross-border region. The conceptualisation of cross-border housing markets acknowledges that housing prices are impacted by border effects in these functional cross-border regions. We note how cross-border housing markets are more pronounced when agglomerations are close to the border. An agglomeration's proximity to an international border allows strong dynamics to develop and influence economic activity on either side of the border. Yet, housing markets across borders may also materialise when the border region is linked by two major agglomerations and thus allow for inter-municipal commuting patterns as members of the same household travel in different directions.

By conceptualising cross-border housing markets as a term itself, we lay the groundwork for future studies to further theorise, study and understand this concept which finally bridges housing studies with border studies.

In conceptualising cross-border housing markets, we have identified several characteristics associated with these market types. Cross-border housing markets can be typified by having:

- Integrated labour markets
- Cross-regional residential migration
- Varying income levels across borders
- Land value differences based on location
- Shared cross-border services
- Cross-border commuting
- High levels of worker mobility
- Liberalised market conditions

For cross-border housing markets to materialise, certain specificities which encourage cross-border interactions must be present to stimulate the integration of housing markets. First, cross-border housing markets require neighbouring countries to reach border agreements which facilitates unimpeded travel across their borders (Dumeignil, 2022; Decoville, et al., 2013). Border agreements can focus on loosening restrictions on border controls, can focus on economic liberalisation and market integration and can focus on coordinating tax policy in order to encourage further cross-border interactions. Second, cross-border housing markets require an integrated labour market where individuals either commute or migrate across a border to take advantage of economic differences present throughout the border region (Chilla & Heugel, 2022; Dumeignil, 2022). This entails households navigating different economic geographies created by the border to facilitate the establishment of a cross-border market (Sohn, 2013). In this line of argument, we suggest that cross-border housing markets are more likely to occur where the border acts as a resource to acquire a better housing standard with the economic means one has in comparison to what would be possible within the same country. Third, the financial system must reach a level of integration which allows households to develop regular cross-border patterns which consequently impacts housing prices (Christmann, 2015; Decoville & Durand, 2019). Whether through the use of a single currency or development of unique banking arrangements which facilitates currency exchange, a certain degree of economic freedom and security must be present to encourage enough households to cross borders and create a cross-border housing market.

2.7 Conclusion

This broad overarching literature review has sought to identify how international trends in the housing sector are playing a role in the development of cross-border housing markets throughout Europe. The case studies show that these trends are also pertinent in all case study regions, though with varying degrees of intensity. The metropolitan case study regions experience for example a higher influx of international investments, whereas the tourism regions are often experiencing investments from secondary homeowners. The description of the case study regions makes builds on the structure provided here. It has also outlined different types of government interventions in housing policy which will inform our assessment of cross-border markets across 11 different European countries. Critically, it has combined housing and border study literature streams to conceptualise what a cross-border housing market is within the scope of this project. This has also helped to think through the indicator opportunities, and specifically to ensure to take account of the perspectives that each household takes on board in view of the decision-making of moving across the border. For example, the structure of the housing markets across the border may differ due to local, regional or national policy differences. Thus, it is important to understand how a person from one side of the border with their wages and income will be able to live on the other side of the border in comparison to the country they are living in, and what other determinants are. For example, households take decisions for moving mostly in periods of change of household formation or change of job. This helps to understand in which phases of a live households may consider a change of housing or change from rental to buying. Consequently, cross-border indicators for housing markets in the selected case study regions can be developed with regard to how the integrated housing market is impacting border dynamics within each region.

3 Case Study Selection and Methodology

3.1 Case Study Selection

The case study selection process occurred during a joint project meeting where the advantages and disadvantages of different potential border regions were discussed. The ultimate case study selection was agreed upon with the ESPON EGTC.

The six case studies selected are:

- Case Study 1: Switzerland – France: Geneva – Annemasse – Annecy
- Case Study 2: Denmark – Sweden: The Greater Copenhagen Region
- Case Study 3: The Belfast – Dublin Economic Corridor: Newry – Dundalk
- Case Study 4: Romania – Bulgaria: Ruse – Giurgiu
- Case Study 5: Austria – Slovakia: The Vienna-Bratislava Corridor.
- Case Study 6: Spain – France: The San Sebastián – Bayonne Corridor: Irún – Hondarribia – Hendaye.

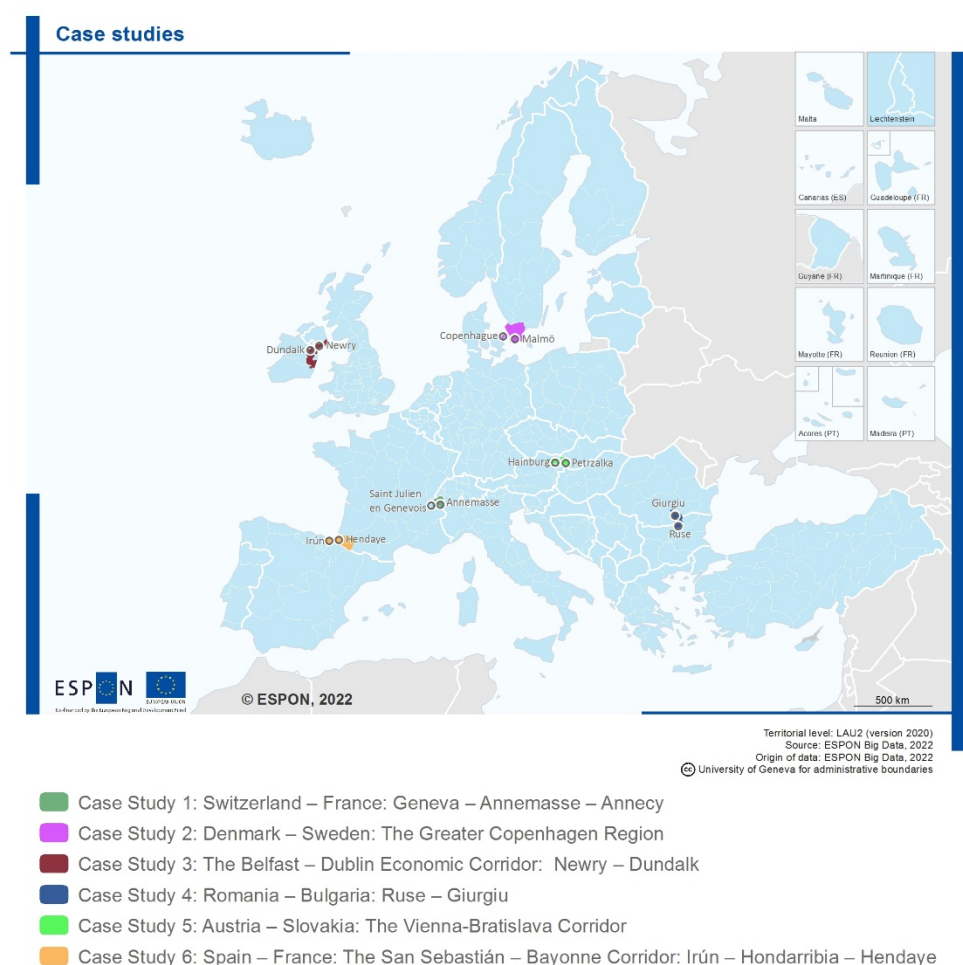


Figure 1: Map of All Case Study Regions

3.2 Representativeness of a variety of cross-border regions in Europe.

The selected case studies aim to reflect on the different dynamics and nature of case studies, which encompass rural border regions, regions with large commuter belts, touristic regions, and so forth. For the purpose of this research, we differentiate between four different types of regions. First, we include areas with large European cities with extensive commuter belts such as Geneva with its commuting belt reaching far into France. The housing market dynamics on one side of the border are driven by one dominant city. Second, the study includes corridors in Europe where large cities in relatively close distance to the border are linked by smaller cities nearer to the border. An example of such a region is the Belfast-Dublin Economic Corridor where the cities of Newry and Dundalk serve as a gateway across the border. The third type of region included are rural border regions such as Bulgaria and Romania where there is a low level of integration, and which are in general displaying a lower population density. Fourth, we include regions that experience tourism, such as for example the border regions of Spain and France alongside the Atlantic coast.

Further we aimed to ensure the

- *Availability of data* and existence of some border cities which experience impacts in their housing markets due proximity to a neighbouring country's developments.
- *Territorial diversity* including old and new EU Member States as well as Southern, Northern and Central European case studies.

Characteristics	Case Studies
Large Cities with Large Commuter Belts	Denmark – Sweden (Greater Copenhagen Region) Ireland – Northern Ireland (Belfast – Dublin Corridor) Geneva
Corridor Development with a Smaller Level of Integration at the Border	Ireland – Northern Ireland (Belfast – Dublin Corridor) Spain – France Slovakia – Austria (Bratislava and Vienna)
Rural Border Regions	Bulgaria – Romania Parts of the Ireland – Northern Ireland Border
Touristic Impacts on Housing Markets	Spain – France Sweden – Denmark (Skåne County)

Table 1: Case Study Characteristics

The case study selection was conducted jointly with the ESPON EGTC upon proposal by the project team.

3.3 Overview of Methodology of Case Study Implementation

The methodology for conducting the policy analysis on each case study research incorporated the following steps:

1. Background Collection on Case Studies

The research commences with extensive desktop research which provided detailed overviews for each of the case study regions. In order to summarise the vast amount of information collected, the team developed comparative fiches for each case study (see Ch 2 of each Case Study Annex) which could contextualise findings through a broad overview of statistical data. For instance, differences in unemployment rates, housing tenure and average mortgage interest rates were collected. This desktop study on the background of each region ultimately allows for better interpretation and understanding of the data.

2. Governance Analysis

The next step in developing a clear picture of each case study region was to understand and analyse the governance structures and legal frameworks operating in each region. We developed institutional schematics for each region's government hierarchy which allowed for quick comparative analysis across the different types of regional governance structures. Furthermore, we completed a document review of local government publications which permit us to understand the spatial planning practices and legal frameworks across the regions. Simultaneous with the document review, we completed a policy compilation in relation to each region's housing sector. This housing policy compilation situated our comprehension of how each government approached their local housing market. As such, the case studies provide an overview of the most relevant political guidance, decisions and strategies to date. Furthermore, the policy analysis incorporated cross-border considerations of housing policies in addition to recognising the impacts of domestic housing policies on border regions.

3. *Literature Review of Housing Market Trends*

Our literature review consisted of two main foci. One focus was to conceptualise cross-border housing market and situate our research within the literature. The second focus was to be able to understand the barriers and opportunities generated from border effects on housing affordability. As such, the literature review provided the team with the knowledge it required to comprehend how different local markets were acting. Further research on grey literature and scoping of local media, allowed the team to then identify the local factors and issues driving potential cross-border housing markets.

4. *Interviews with Regional Stakeholders*

Interviews were conducted to identify the views of key stakeholders from within the regions. The interviewees were asked about key issues and topics relating to housing from their respective positions. In addition, the interviews served as tools to ask stakeholders for further clarification on the proposed territorial coverage of the case studies as well as to ask for advice as to which type of indicators and measurements stakeholders may be interested in.

5. *Update and Development of Indicators*

Following completion of the case study narratives, policy analysis and interviews, the team could proceed with updating the indicators from the previous project while also developing new housing indicators built to measure border effects. Chapter 4 further details some of the updates and novelties incorporated into this research.

6. *Web Scrapping and Data Collection*

With the indicators identified, we could begin the data collection process. This project used a combination of both formal and informal sources to gather data on cross-border housing markets. For the formal data sources, traditional databases such as Eurostat were accessed. For the informal data sources, web scraping was performed on real estate websites to collect statistics on each region's housing market. This annex, and specifically chapters 4,5 and 7 further detail the methodology for web scraping and data collection.

7. *Mapping and Visualisation of Indicators*

Following cleaning and review of the data, the statistics could be mapped and visualised using the indicators as frames of reference to understand the housing dynamics of each region. This form of statistical analysis consequently produced a vast array of thematic maps, regressions, calculations and data tables which could be interpreted and considered against the background contexts for each case study.

8. *Comparative Analysis and Mixed-Method Review*

Finally, we prepared additional statistical analysis targeted at the comparison of case studies, to understand the diversity of cross-border housing markets. We then interpreted and compared the studies with each other by combining the quantitative results with the qualitative findings to produce an answer to our hypothesis on whether a cross-border housing market had developed for each case study region.

4 Indicator Development & Project Update

In this section, we explain how we developed new indicators for measuring cross-border housing markets, and how we updated old indicators from the previous ESPON Big Data and Housing project to reflect the needs of a cross-border region.

The previous project, *ESPON Big Data for Territorial Analysis and Housing Dynamics*, made an exercise of measuring, monitoring and analysing spatial patterns of housing affordability across European FUAs in Poland, Spain, Switzerland, and France. It focused on how housing affordability can affect overall wellbeing through ratios measuring price of housing to income measurements. By mapping the increasing affordability gap, the researchers aimed to understand how social cohesion is being impacted by rising house prices.

The previous project accomplished these tasks by developing a harmonised data collection methodology for big data by using local spatial data to compare affordability measurements within and between different cities. By structuring spatial data through harmonised indicators, unequal spatial patterns of housing affordability were examined at various geographical levels. Ultimately, *ESPON Big Data for Territorial Analysis and Housing Dynamics* provided data, maps and analysis on European wellbeing in relation to housing affordability while also producing a common big data collection methodology, which permits reproducibility for further study of housing market dynamics.

Lot 4 of *ESPON 2020 Data and Maps Update II* seeks to continue this process of analysing European housing markets while adding a cross-border dimension to the research. Hence, this project reproduces the big data methodology used to capture information on housing affordability and applies it to the Geneva case study region in addition to five new case study areas. By reproducing and adapting the big data methodology from the previous project, this work continues to analyse and assess affordability prospects in a similar matter allowing for harmony across data collections processes. Therefore, this project seeks to update the harmonised indicators from the previous project with new data from the new case study areas. The following harmonised indicators have been updated and incorporated into the research methodology for this study:

- Advertised Price of Residential Property for Sale
- Advertised Price of Residential Property for Rent
- PPP (at NUTS3)
- Sales Price Affordability
- Rental Price Affordability
- Number of Offers on Advertised Dwellings per Municipality across the real estate market
- Rental Profitability

Given the cross-border dimension to this research, we have additionally developed new harmonised housing indicators, which measure cross-border dimensions specifically. These indicators have been developed with reproducibility in mind and can be incorporated into future cross-border housing studies as they are developed in accordance with the big data collection methodologies created from the first project. (In the remainder of the report, the methodology for developing the new indicators will be elaborated upon.) Hence, the following are harmonised cross-border housing market indicators developed for this research:

- Mean Difference in Income Affordability to Purchase 1sqm
- Mean Difference in Income Affordability to Rent 1sqm
- Housing Type
- Length of Advertisement Posting
- Accessibility to SGIs
- Accessibility to Public Transport
- Accessibility to Employment Centre

With the seven updated indicators and seven new indicators decided for, the following section describes the methodology used to map and analyse the big data.

This chapter is structured as followed: Chapter 4.1 explains the obstacles faced in developing indicators for cross-border housing markets. Chapter 4.2 details the logic behind our reasoning for creating the new housing indicators for cross-border studies. Chapter 4.3 then provides for a step-by-step explanation of how the indicators were implemented.

4.1 Obstacles in Developing Indicators for a Cross-Border Housing Market

Prior to developing indicators and collecting housing data from the case study areas, we were initially confronted with two primary concerns. First, we recognised a need to cross literature streams between housing studies and border studies in order to build a comprehensive definition of what a cross-border housing market is. Second, we needed to create new cross-border indicators which were capable of aggregating different countries' data into harmonised datasets which would allow for comparative analysis. Hence, clear definitions of concepts and definitive differentiation of data was initially required to be able to adequately develop and then portray the cross-border dynamics revealed through the new indicators.

According to Chapter 2.4, two primary streams of literature were identified as important to this research: literature on housing studies and literature on border studies. When reviewing these streams, we expected to find studies which combined the two streams and which developed cross-border indicators purposefully built for housing markets. While border studies have developed cross-border indicators which can analyse and understand critical dimensions of a cross-border region such as an integrated labour market or commuting flows, many studies fall just short of specifically analysing the housing market itself. Instead, border studies often focus on residential migration within a border region as households relocate to take advantage of economic asymmetries within a region. Conversely, housing studies have extensively developed indicators to measure different types of housing markets, but few studies indeed focused on how international borders would impact the measurement of an integrated housing market in a cross-border region. Hence, we add onto the literature in Chapter 2.5 by defining and conceptualising cross-border housing markets. Following conceptualisation, we could better develop and outline comparative indicators which could be effectively used to measure housing markets in a cross-border region.

Chapter 2.5 of this annex defines and conceptualises cross-border housing markets. With a common understanding of cross-border housing markets developed to frame this study's focus, we could then address potential issues in combining non-harmonised data into cohesive indicators for one integrated market. To accomplish this, the team reviewed previous indicators from the previous ESPON Big Data and Housing project, developed a framework for combining the non-harmonised datasets and approached the task using a comparative perspective. This allowed us to update the old indicators and develop new indicators to suit the needs for measuring and analysing cross-border housing markets.

4.2 Logic behind Creating New Indicators for Cross-Border Housing Markets

To develop the indicators for cross-border housing markets, we first selected six case study regions and then defined the territorial extents for each case study. In selecting suitable regions to study cross-border housing market dynamics, we could then orient the development of the indicators to respond to the dynamics observed across the case study regions.

With the case study regions studied, the researchers were then able to identify potential indicators which could be used to measure the specificity of the cross-border markets. The indicators were suggested on the basis that they could capture information illustrative of cross-border dynamics as identified through the initial narratives developed around each of the case study regions. For example, one proposal was to examine ways to measure how regional differences in income would affect a household's ability to afford housing on either side of the border.

Additionally, selecting the case study regions allowed the researchers to update the old indicators from the previous project to better reflect the dynamics present in a region with internal borders. For instance, previous indicators used to measure metropolitan housing markets such as Sales Price Affordability or Advertised Price of Residential Property for Rent were discussed collectively in order to ensure that data could be aggregated into harmonised datasets which could be easily used for analysis and comparison. As internal borders create non-harmonised data with the case study areas, it was essential to ensure that all indicators could be updated and reconfigured to display information for an integrated cross-border market.

To develop new indicators which could measure and analyse the specificities of a cross-border housing market, the team developed a three-step process to building the new indicators. Box 1 presents an example to illustrate the projects approach.

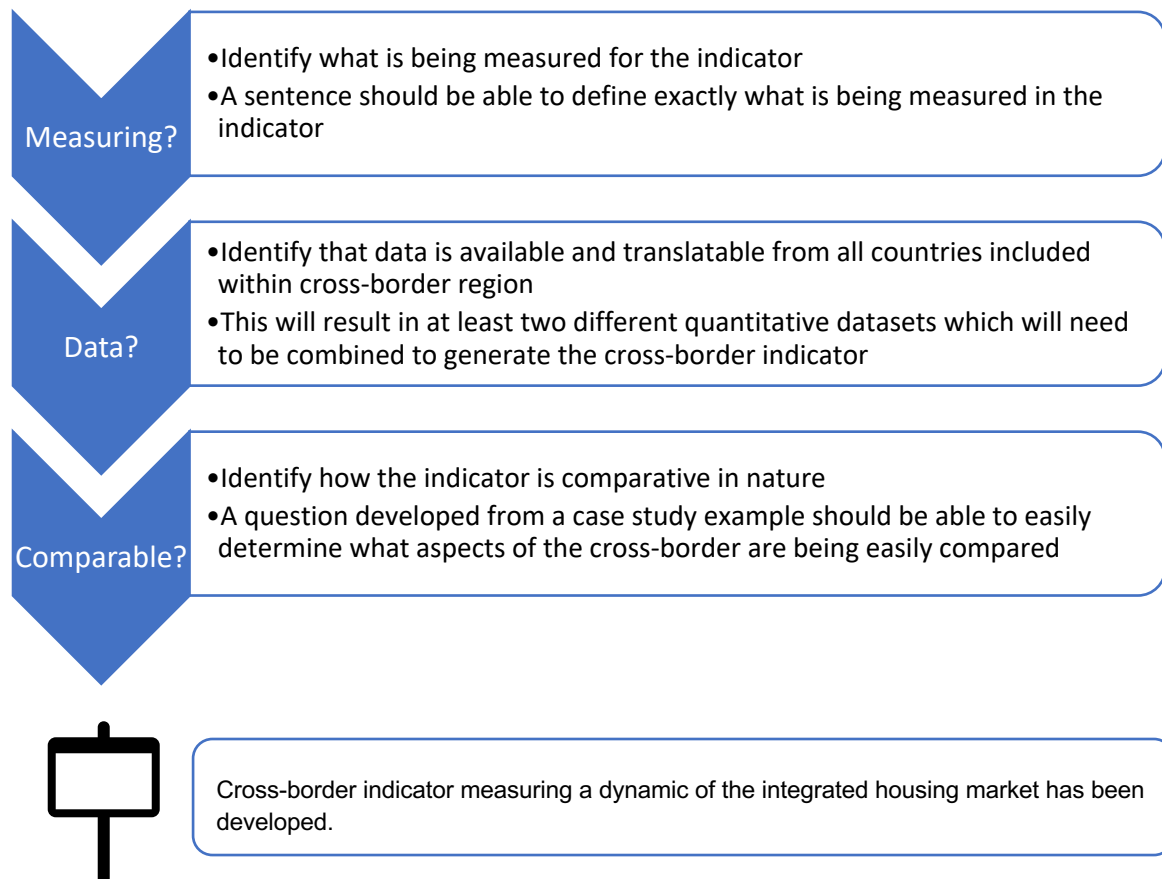


Figure 2: Graphic Displaying Three-Step Process for Developing Cross-Border Indicators for Housing Markets

First, we decided what we would want to measure as with regards to what we defined as a cross-border housing market in Chapter 2.6. We focused primarily on aspects of affordability and accessibility. Second, we identified data sources from both sides of the border which could be combined as a measurement to create a cohesive cross-border indicator. This requires ensuring that suitable data and information is accessible and translatable from each country.

Following identification of the data sources, the available data was made comparable in order to develop an indicator which could analyse a specific aspect of the market. It is essential that the indicator is comparative in nature so that it can address differences across the cross-border region. For instance, by collecting data on regional incomes and overlaying this data with information in relation to house prices, we can comparatively display and analyse how differences in income across a cross-border region can impact affordability of housing on different parts of the region.

Finally, after following these steps, we arrive at a comparative indicator which can be applicable across different cross-border regions. To display such comparative analyses, we visually display side-by-side comparisons of the data and spatial outputs to illustrate the differences across borders. For example, we display across four quadrants how regional differences in income levels can affect how much housing per square metre a household can afford across different parts of a cross-border region.

Box 1: Example of Indicator Development

One of our main interests for measuring cross-border housing markets was understanding how differences in household income within a cross-border region would impact how much a household could afford to buy in different parts of the case study area. To find this, we needed to first identify what exactly needed to be measured. We found that finding the difference in how many days a household would need to work to purchase a property on either side of the border could display this indicator. Hence, we wanted to find the difference in days of average income between both countries that a household would be required to work in order to buy 1sqm.

To build this measurement, we would need four different quantitative datasets:

- Incomes for Country A
- Incomes for Country B
- Advertised Sales Prices for Residential Properties in Country A
- Advertised Sales Prices for Residential Properties in Country B

Having these four datasets would allow us to find how much a household on an income from County A could afford in both Country A and Country B and likewise, how much a household on an income from Country B could afford in both Country A and Country B. This information would allow us to understand how housing market dynamics interplays with cross-border labour market to impact residential mobility in these cross-border regions.

Following acquiring this information and formulating what is being measured, we can ensure that the indicator is suitable for a cross-border comparisons as it can effectively measure how different incomes from different parts of a region can or cannot afford property in different areas of the region. This type of measurement will allow us to then critically analyse how residential migration may occur within a cross-border region and to understand what economic factors may be pushing or pulling people to buy houses in different parts of an integrating cross-border housing market. Hence, we can then ask a question such as: how many days would a household earning its income in Switzerland need to work to buy a house in either Geneva or Annemasse? This style of phrasing is replicable for all parts of our case studies and can also allow for further comparison across different cross-border housing market.

4.2.1 Selection of New Indicators & the Border Effect

In the following table, we have listed the new indicators developed to measure cross-border housing markets using the development process explained above:

<i>Indicator</i>	<i>Justification</i>	<i>Visualisation</i>
Mean Difference in Income Affordability to Purchase 1sqm	Allows us to understand the economic price differential in purchasing a property across the border.	Data table
Mean Difference in Income Affordability to Rent 1sqm	Allows us to understand the economic price differential in renting a property across the border.	Data table
Length of Advertisement Posting	Allows us to understand whether certain parts of the market are more active in terms of new transactions across a border region	Thematic Map
Housing Type	Allows us to see if the type of housing (i.e. single family, apartment, terrace) plays an impact on if households choose to migrate within a cross-border region	Thematic Map
Accessibility to Employment Centres (city centre is used as a proxy)	Allows us to understand how accessibility to job centres near the border can impact housing prices in border regions	Accessibility Maps
Accessibility to SGIs	Allows us to understand how accessibility to services within the catchment area of the border can impact housing prices in border regions	Graphs from Regressions & Accessibility Maps

Table 2: Overview of Reasoning for New Indicators

Given how the six new indicators were developed in consideration of assessing border effects in integrated housing markets, we can use these indicators to further understand how a border might affect price dynamics through the conceptualisation of cross-border housing markets. A key concept that we needed to understand is how differences in income levels can impact household purchasing power across different areas of the functional cross-border region. Therefore, finding the mean difference in income affordability to purchase or rent allows us to directly link income differences to housing prices which in turn provides depth to our comprehension of how cross-border housing markets manifest. Similarly, connecting accessibility to housing prices allows us to decipher how cross-border commuting flows can influence housing prices in a cross-border housing market. Likewise, indicators on the length of advertisement posting and housing type can provide further detail on potential differences observed across a cross-border housing market. Altogether, these indicators build a more realistic picture of how housing price dynamics are impacted by borders in regions in which a cross-border housing market have materialised. Hence, our conceptualisation of cross-border housing markets is substantiated through housing market indicators which were specifically developed to assess border effects.

4.3 Step-by-Step Implementation of Indicators

4.3.1 Processing the Big Data file

The first step for producing the indicators involves processing the raw data obtained and cleaned from the python scripts. This data is available as a big csv file with all necessary fields.

It contains one row for each advertisement collected, including a unique id that allows for tracking the same advertisement across time. This is necessary in order to exclude repeated instances that would produce wrong results.

The table also incorporates a specific field that determines if the collected advertisement is valid, depending on its data's completeness (some adverts do not include information on surface for instance) and whether it is or is not an outlier (detailed methodology of outlier detection is available at 6.7.1.4).

Every advertisement is enriched with information of the geographical unit they are located in; thus for each advert, we assign it a LAU2 code. This will allow for making further aggregations afterwards.

4.3.2 Obtaining complementary data

As an external conventional data source, we needed to include a measure for income in the concerned regions. The most detailed indicator available is the GDP at NUTS3 regions measured in PPP. In this way, results are comparable across different countries.

Data is collected from EUROSTAT and complemented in Northern Ireland with information from the Office of National Statistics.¹

NUTS3	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Ruse	8.700	9.200	9.500	10.500	10.600	11.000	11.400	11.800	12.400	:
Giurgiu	6.900	6.200	7.600	6.900	7.400	7.600	8.500	11.600	10.200	:
Dublin	50.100	51.900	55.600	60.100	64.700	71.200	76.400	82.400	86.500	:
Mid-East	21.700	21.700	22.600	24.100	31.200	34.500	32.100	36.400	36.500	:

¹ <https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/regionalgrossdomesticproductallnutslevelregions>

Calculated as ratio of GDP/cap in pounds * UK GDP PPS / UK GDP pounds

Newry, Mourne and Down	16.555	17.173	17.643	18.492	19.226	19.543	19.970	20.120	20.615	:
Lisburn and Castlereagh	21.167	21.907	22.317	22.483	23.467	23.534	23.844	24.459	24.648	:
Belfast	34.214	34.846	35.364	35.856	37.445	38.399	40.040	40.732	41.025	:
Byen København	46.400	46.200	46.700	49.000	49.800	50.700	54.200	55.100	55.400	55.900
Københavns omegn	47.600	48.500	50.900	51.800	56.400	56.400	59.800	61.600	62.800	64.600
Skåne län	28.800	29.100	29.100	29.600	30.500	30.200	31.300	31.300	32.000	:
Gipuzkoa	30.900	30.700	30.400	30.900	31.900	33.300	35.800	36.200	37.400	:
Pyrénées-Atlantiques	26.200	25.700	26.700	26.600	26.900	27.000	26.900	27.400	29.500	27.700
Bratislavský kraj	49.400	49.000	50.800	50.900	53.000	50.600	49.900	50.300	50.200	48.300
Nordburgenland	24.100	25.900	26.500	26.600	27.500	28.100	28.700	29.500	30.200	:
Wiener Umland/Südteil	37.700	38.900	38.300	39.500	41.500	42.000	43.700	45.900	47.400	:
Wiener Umland/Nordteil	23.800	24.800	25.100	24.900	24.400	24.400	24.800	25.600	26.300	:
Genève	:	:	:	60.100	60.600	59.900	61.500	63.200	:	:
Haute-Savoie	25.400	24.600	25.100	25.200	25.800	26.200	27.000	27.700	28.700	26.600

Table 3: GDP per Capita Measured as PPP

The other important source of data is the calculation of accessibility. This element is explained in detail in Chapter 5.

4.3.3 Updating previous indicators

Following these initial steps, we could bring together information from both the big data sources and conventional sources in order to compute the necessary indicators:

4.3.3.1 Advertised Price of Residential Property for Sale

This indicator is obtained by aggregating the sale price per square meter (or per number of rooms in the Irish-Northern Irish case study) per LAU2 unit and computing the average value.

As a result, we get the advertised average sale price at each LAU2 unit, and this is represented as a choropleth map.

In the case study of Romania-Bulgaria, since the cities of Giurgiu and Ruse are contained within one single LAU2 unit each, we present the results by showing all individual advertisements and performing an IDW calculation².

² IDW or "inverse distance weight" is an algorithm that derives a raster layer based on a set of data points distributed on a plane

4.3.3.2 Advertised Price of Residential Property for Rent

This indicator is obtained by aggregating the rental price per square meter (or per number of rooms in the Irish-Northern Irish case study) per LAU2 unit and computing the average value.

As a result, we get the advertised average rental price at each LAU2 unit, and this is represented as a choropleth map.

In the case study of Romania-Bulgaria, since the cities of Giurgiu and Ruse are contained within one single LAU2 unit each, we present the results by showing all individual advertisements and performing an IDW calculation.

4.3.3.3 PPP (at NUTS3)

This indicator allows for measuring income differences on either side of the border for each case study. Data is available as PPP per capita at NUTS3 level for the 6 case studies with the most recent year of complete information being 2018.

Two sources are used:

- EUROSTAT³
- UK ONS⁴

Results are shown as a chart to show the disparities in each case study but also used for calculating the affordability indicators.

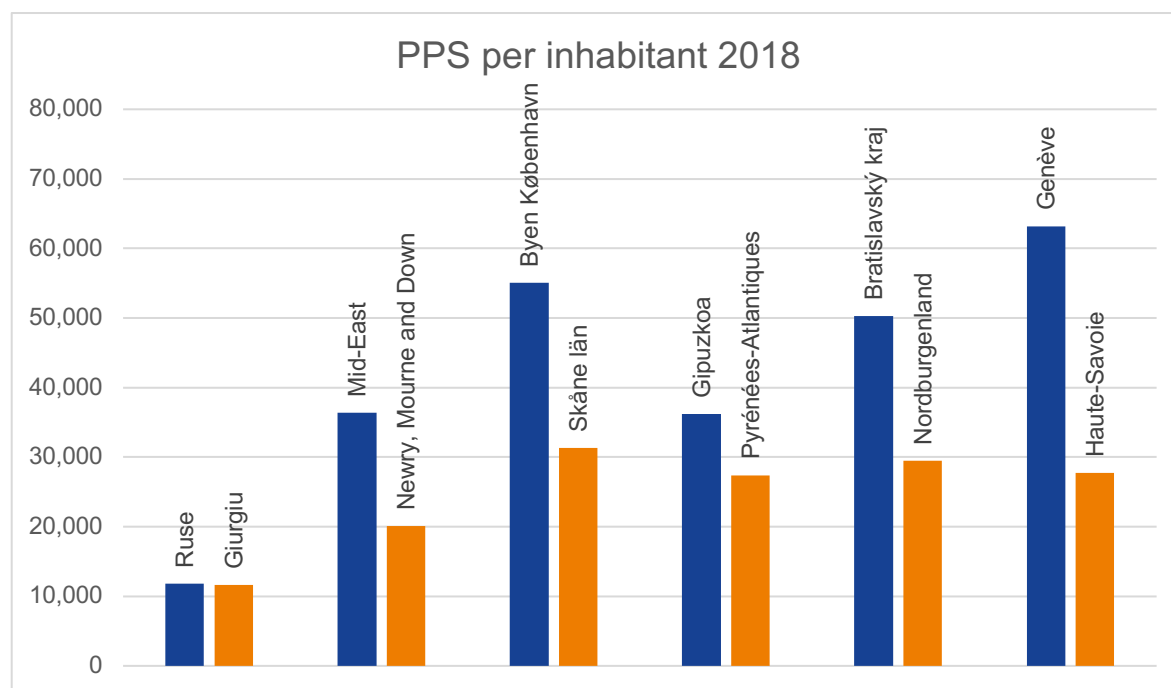


Figure 3: Graph displaying differences in PPS per Inhabitant

³ https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_3gdp&lang=en

⁴ <https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/regionalgrossdomesticproductallnutslevelregions>

4.3.3.4 Sales Price Affordability

Affordability is defined as the cost of buying real estate according to the income level. The calculation is done at LAU2 level and requires information on the average sale price per sqm or room in the LAU2 unit and the income level measured as PPP per inhabitant by NUTS3:

$$SPa_i = \frac{I_j}{\frac{12}{SP_i}}$$

Where:

SPa_i is the sales price affordability of LAU2 i measured in number of months of income needed to buy 1sqm or 1 room

I_j is the income level of NUTS3 j measured in PPP that we divide by 12 to obtain the monthly income

SP_i is the average sales price of LAU2 i measured in €/sqm or €/room

4.3.3.5 Sales Price Affordability – Income Differentiation

Given the cross-border focus in this research, it was necessary to compare regional differences in affordability for potential cross-border households. Hence, we need to cross reference the income levels at either side of the border with the prices of buying at either side of the border.

This was applied in three separate formats:

- Income affordability comparison of the opposite region's income: in this case we measure the affordability of working in one country and living in the other (so dividing price of buying in country A over income of country B and vice versa). The result is one single map per case study showing opposite affordability.
- Income affordability for household income of country A with entire case study area: for this indicator we take all buying prices and divide them by the income level of country A. The resulting map shows how affordable the market is at either side of the border for a person working in one of the sides, in this example country A
- Income affordability for household income of country B with entire case study area: this indicator is the opposite of the previous one. We divide by the income level of country B. This map has to be interpreted alongside its previous companion in order to highlight the differences among them.

4.3.3.6 Rental Price Affordability

Affordability is defined as the cost of renting according to the income level. The calculation is done at LAU2 level and requires information on the average rental price per sqm or room in the LAU2 unit and the income level measured as PPP per inhabitant by NUTS3:

$$SPa_i = \frac{I_j}{\frac{12}{SP_i}}$$

Where:

SPa_i is the rental price affordability of LAU2 i measured in number of months of income needed to rent 1sqm or 1 room

I_j is the income level of NUTS3 j measured in PPP that we divide by 12 to obtain the monthly income

SP_i is the average rental price of LAU2 i measured in €/sqm or €/room

4.3.3.7 Rental Price Affordability – Income Differentiation

Given the cross-border focus in this research, it was necessary to compare regional differences in affordability for potential cross-border households. Hence, we need to cross reference the income levels at either side of the border with the prices of renting at either side of the border.

This was applied in three separate formats:

- Income affordability comparison of the opposite region's income: in this case we measure the affordability of working on one country and living in the other (so dividing price of renting in country A over income of country B and vice versa). The result is one single map per case study showing opposite affordability.
- Income affordability for household income of country A with entire case study area: for this indicator we take all renting prices and divide them by the income level of country A. The resulting map shows how affordable the market is at either side of the border for a person working in one of the sides.
- Income affordability for household income of country B with entire case study area: this indicator is the opposite of the previous one. We divide by the income level of country B. This map has to be interpreted alongside its previous companion in order to highlight the differences among them.

4.3.3.8 Number of Offers on Advertised Dwellings per Municipality across the real estate market

This indicator is a straightforward measure of the number of unique advertisements found on each LAU2 unit. It is important to stress that this dataset requires a cleanup in terms of repeated advertisements across different time periods and across different platforms. This is due to the fact that some advertisements appear repeatedly in different websites and because we sometimes scrap data from aggregators (Nestoria), some ads appear multiple times in the dataset.

The final output for this indicator is the direct number of unique advertisements per LAU2 which we represent as proportional circles located in the centroid of each LAU2 unit.

4.3.3.9 Rental Profitability

Rental profitability is defined as the ratio between sales price and rental price and gives an idea of whether it is more profitable to buy or to rent in a certain area.

For this indicator we use two previous ones:

- Advertised Price of Residential Property for Sale (€ per sqm or room)
- Advertised Price of Residential Property for Rent (€ per sqm or room)

Both indicators are available at LAU2 level; thus, the rental profitability is also computed at LAU2 level. The resulting indicator is a dimensionless index. Given that we put sale price in the numerator and the rental price in the denominator, the bigger the indicator, the lower its rental profitability.

Because not all LAU2 units have both rental and sales advertisements, the number of units without a value is significant, leading to some gaps on the maps. This can be solved by scraping longer time periods, but ultimately it comes down to the real estate offer available.

4.3.4 New Indicators Accounting for the Border Effect

This next sub-chapters outline the methodology used to implement the newly developed housing indicators for cross-border markets.

4.3.4.1 Mean Difference in Income Affordability to Purchase 1sqm

This indicator makes use of a previous one: Sales Price Affordability.

In this case we perform a similar calculation by mixing the sales price in a given LAU2 unit with the income level at the other side of the border. For instance, we have the average sales price in Annecy and instead of calculating the number of months of French income needed to buy 1 sqm, we compute the number of months of Swiss income needed to buy 1 sqm.

Thus, we end up having 4 quadrants per case study:



Figure 4: Example of Mean Difference in Income Affordability Indicator

We represent this indicator with two maps side by side: one with affordability according to the LAU2's corresponding income level and one with affordability according to the LAU2's income for the cross-border country.

On top of that, we also represent the difference in relative value of the two calculations by dividing affordability of corresponding country income over affordability of the neighbouring country income.

4.3.4.2 Mean Difference in Income Affordability to Rent 1sqm

This indicator has the same exact methodology as 4.3.4.1 but uses rental prices instead. The outputs are also equivalent but for rental properties.

4.3.4.3 Housing Type

This indicator displays the two main typologies of housing: apartments (multifamily dwellings) or single houses (single family dwellings). This is obtained from the typologies scraped from the online sources, duly standardized into this binary category, as each country has its own specific typologies.

We represent the indicator by counting the number of unique advertisements per LAU2 unit according to the housing type data column. Thus, we end up having two maps for each case study: one with squares (as the chosen symbols on the map) proportional to the number of multifamily dwellings and one with squares proportional to the number of single-family dwellings.

4.3.4.4 Length of Advertisement Posting

As the scraping of data is carried out in several time periods (on a weekly basis for 2.5 months), we can identify advertisements that appear repeatedly because each publishing platform assigns a unique ID that we also scrap.

This allows for counting how many times a unique advertisement appears on the dataset which we can then represent as information on a map.

The indicator is calculated by representing the exact locations of each unique advertisement as a cloud of points and assigning to each point the length (number of weeks the advertisement appears on the dataset). Then we use the IDW raster creation tool of ArcGIS which interpolates a surface based on a cloud of values (in this case the number of weeks). The resulting raster layer shows the average advertisement duration across the whole area of analysis.

4.3.4.5 Accessibility to SGIs

This is a set of indicators that are explained in detail in chapter 5. It comprises individual calculations per type of Service of General Interest and transport mode:

- Education centres
- Health centres
- Commercial and retail
- Public transportation

- By car
- Walking

Besides from obtaining the so called “service areas” which allow us to know how far away each point of the area of analysis is from a SGI, we also perform a statistical cluster analysis by crossing the exact accessibility value of each scraped location (advertisement) against the price and including different discriminating variables (such as type of dwelling, size...)

4.3.4.6 Accessibility to Employment Centre

This indicator allows for obtaining a measure of how well-located advertisements are in regard to the employment areas.

As there is no exact way of knowing where employment areas are in a given city, we define a centroid in the downtown of each of the 12 cities analysed. Then, we calculate the shortest path from each advertisement to the city centres at either side of the border.

These results are then mapped as service areas with coloured patterns depending on how far away in time each point is.

5 Cross-Border Accessibility & Housing

One of the key metrics that we identified as important to understanding how cross-border housing markets function is in connecting levels of accessibility to the border region with general prices in housing. As cross-border housing markets are a consequence of an integrated labour market where a significant portion of the working population commutes across a border and contributes to the establishment of a cross-border commuting flow, we developed a hypothesis which states that higher levels of accessibility in a cross-border region would correlate with higher housing prices. As accommodation with greater accessibility to the wider labour market would be more impacted by the cross-border flow, we expected these dwellings to receive a boost in value as they are deemed more valuable and accessible to a wider portion of the market.

Hence, we link measures of accessibility with housing data in order to extrapolate to what extent cross-border housing markets are impacted by regional differences in accessibility. In linking accessibility measurements with housing data by creating thematic maps which correlate price to accessibility, we can examine whether these trends are observed in border regions and are able to compare how accessibility impacts housing affordability both within a cross-border functional area and between cross-border regions.

The following section explains the methodology process for developing this specific indicator as it encompasses several steps and many different sources of data and information.

5.1 Preparation of the GIS for analysis

Accessibility is a package of different calculations that use the same theoretical framework to deliver information on how close in terms of time are the services of general interest in relation to the scraped advertisements.

This calculation is similar to other accessibility indicators available on previous ESPON projects, but it has its own specificities. Below, we explain in detail the procedure:

We start by obtaining the necessary dataset for the analysis. This involves downloading subsets of the Openstreetmap planet files (these are backups of the online Openstreetmap transportation network) with one for each case study.

Once we have these raw network files, we prepare them so that they become a navigable graph allowing for network analysis:

- Revision of data associated to each network segment
- Definition of speeds depending on the segment hierarchy
- Definition of senses (distinguishing one way from two ways)
- Definition of non-navigable segments for cars (pedestrian, cycleways, footways) and for pedestrians (motorways)



Figure 5: Example of navigable graph for the case study of Denmark-Sweden

Then, using Network Analyst toolbox from ArcGIS, we define costs (speed and time), restrictions (senses, pedestrian) and the segment hierarchy.

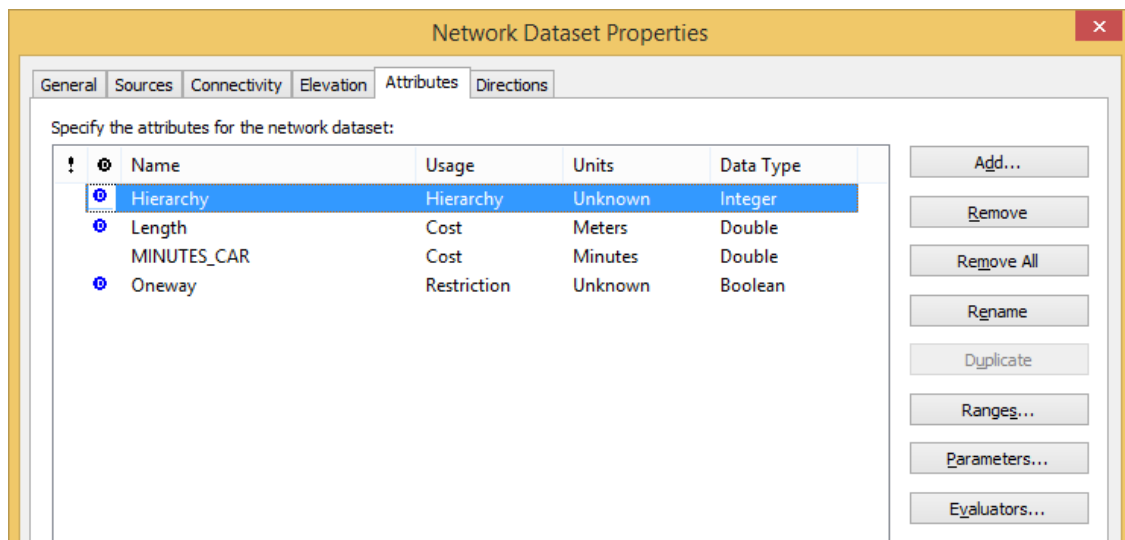


Figure 6: Setting up the Network Analyst tool of ArcGIS

As a final step, we perform a calibration of the network by calculating paths between different random origins and destinations. The resulting paths are then obtained from Google Maps, and we compare the distance and time of the paths between both maps and adjusting the model parameters whenever necessary.

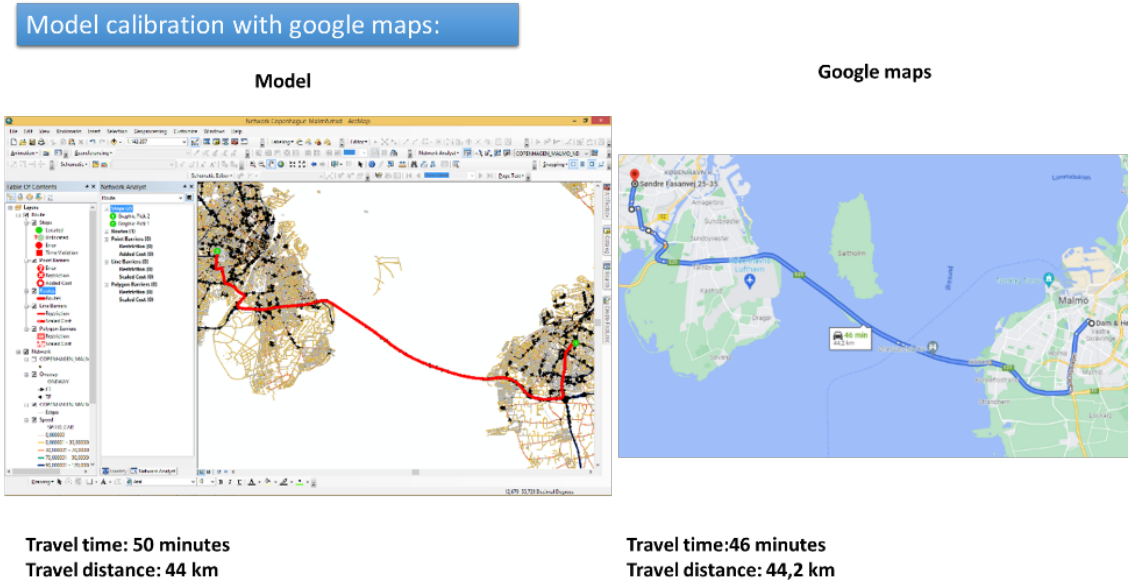


Figure 7: Checking the network model results against Google Maps

5.2 Locating the services of general interest (SGI)

The next step involves determining the points of interest for the accessibility analysis.

We have selected the following four main categories, as being relevant to the price of housing markets:

- Education centres
- Health centres
- Commercial and retail
- Public transportation

The labels selected for each category (as they appear on Openstreetmap) are listed in the following table (example for Sweden-Denmark case study):

TYPE	NAME	POINTS
Retail	department_store	13
Retail	market_place	16
Retail	mall	9
Retail	laundry	100
Retail	jeweller	196
Retail	hairstylist	961
Retail	greengrocer	121
Retail	gift_shop	112
Retail	garden_centre	41
Retail	mobile_phone_shop	71
Retail	florist	238
Retail	kiosk	408
Retail	computer_shop	37
Retail	clothes	920
Retail	butcher	318
Retail	bookshop	119
Retail	bicycle_shop	454

Retail	beverages	348
Retail	beauty_shop	178
Retail	bank	169
Retail	bakery	594
Retail	furniture_shop	138
Retail	supermarket	1.032
Retail	toy_shop	52
Retail	shoe_shop	138
Retail	sports_shop	93
Retail	vending_any	143
Retail	outdoor_shop	36
Retail	post_office	49
Retail	vending_machine	4
Retail	pharmacy	145
Retail	video_shop	12
Education centers	school	563
Health centers	doctors	64
Health centers	pharmacy	137
Health centers	dentist	104
Health centers	optician	156
Health centers	hospital	53
Public transport 1	bus_station	94
Public transport 1	bus_stop	10.972
Public transport 1	taxi	70
Public transport 2	ferry_terminal	117
Public transport 2	Metro	17
Public transport 2	railway_halt	128
Public transport 2	railway_station	264
Public transport 2	tram_stop	96

Figure 8: Table of SGI Labels from the Sweden-Denmark Case Study

Although the bulk of the points of interest are obtained from Openstreetmap, some categories are validated and complemented using Google Maps.

5.3 Calculating accessibility

The *Service Area from the Network Analyst* toolbox from ArcGIS allows for calculating areas with equal access time from each of the generation-attraction points (the services of general interest) towards any point in the case study area.

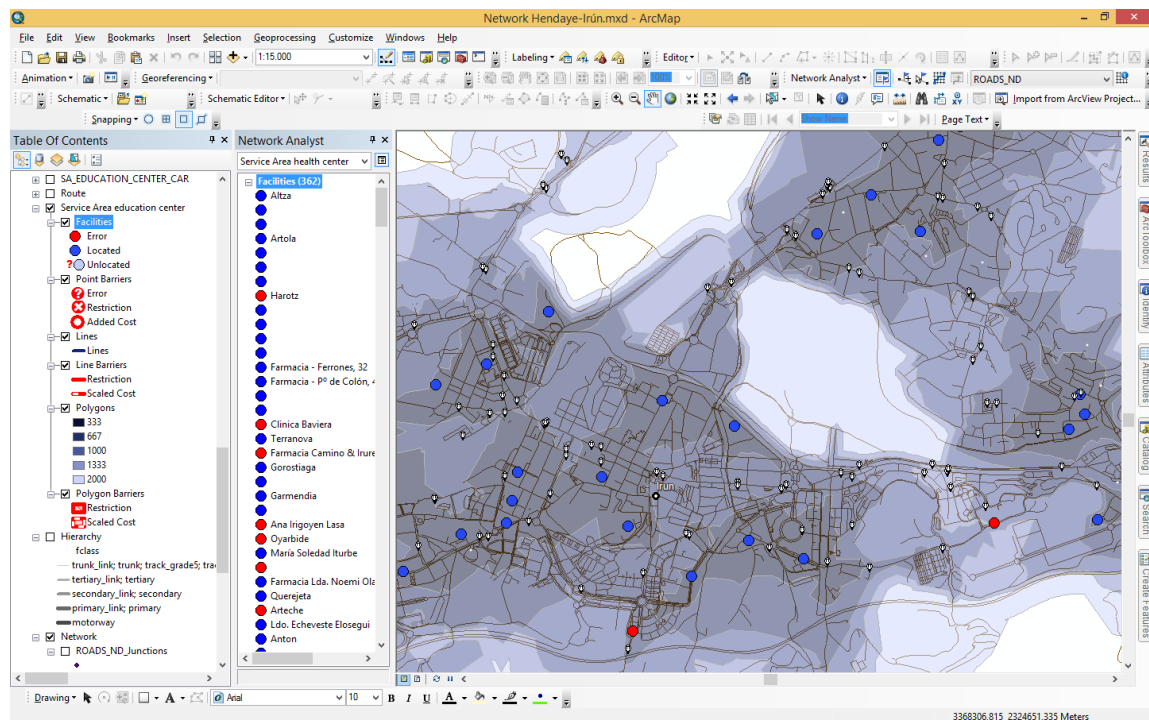


Figure 9: Example of calculation of accessibility for the Spanish-French case study

We perform two separate calculations for each typology of SGI, one by car and another by walking:

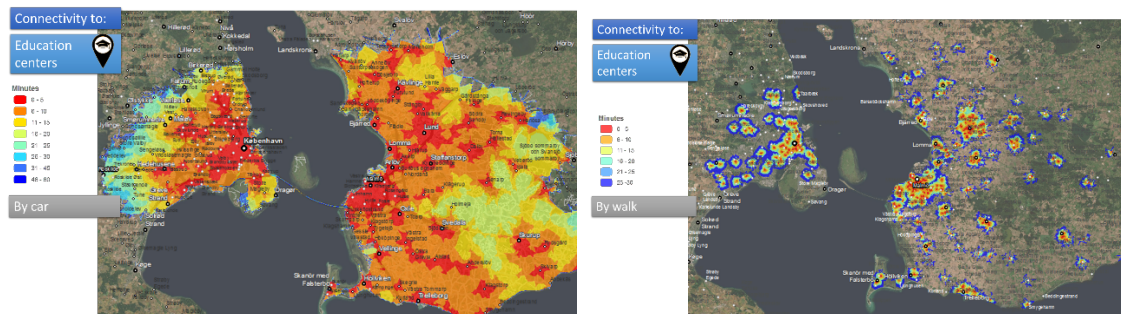


Figure 10: Example of calculation of service areas for education centres by car and walking in the Sweden-Denmark case study

The result are polygon layers for each one of the services areas defined according to the number of minutes of distance.

6 Web scraping of real estate online listings

In this chapter, we explain the methodology carried out to gather unconventional data of real estate portals for this study, as well as the process of data cleaning, harmonization and enrichment. To conceptualize this procedure, the following chart has been designed to show an overview of the whole data workflow.

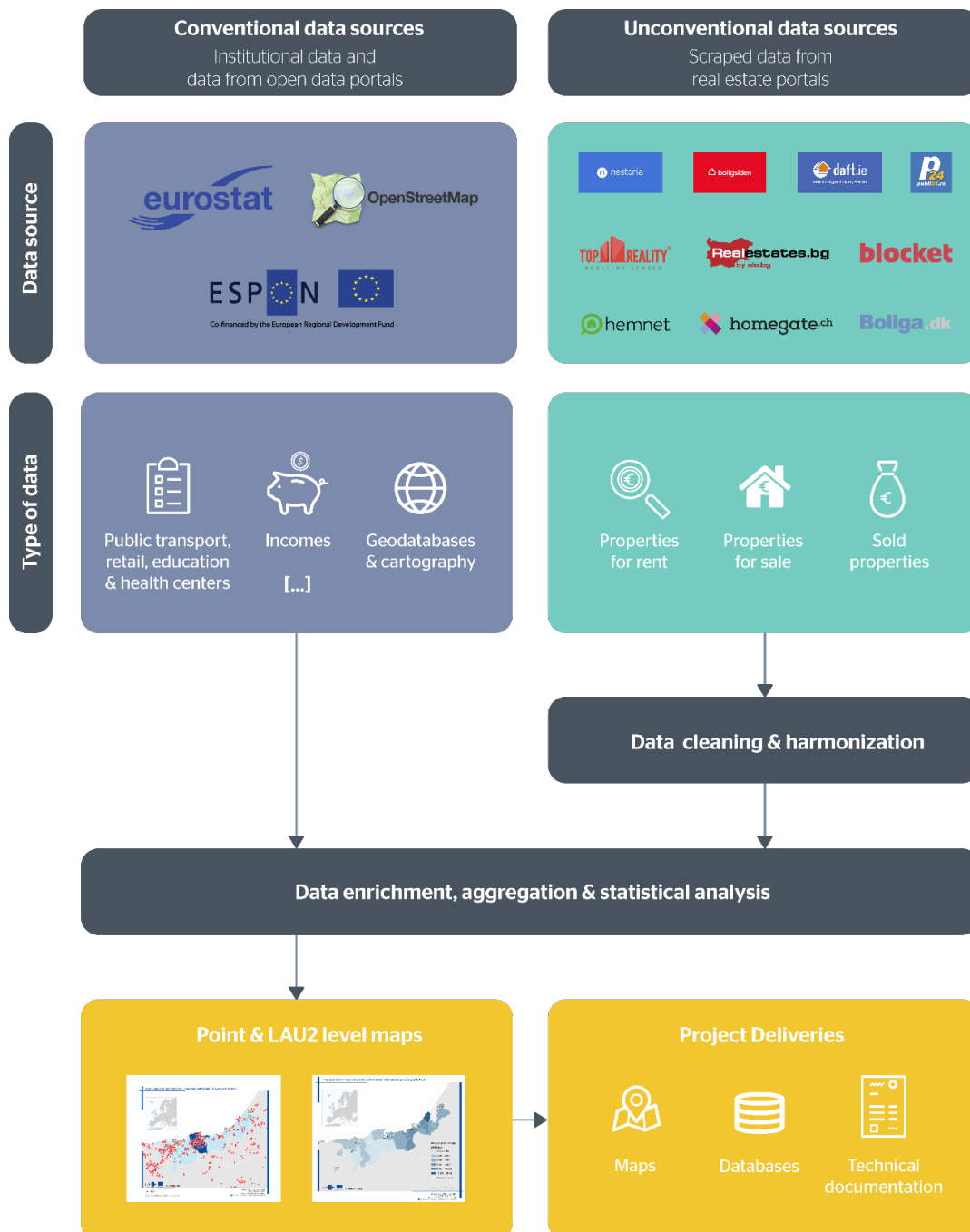


Figure 11: Data Workflow Overview

6.1 Generic web scraping methodology

Data scraping was one of the techniques used to gather the data needed for the analysis of the indicators of this study. This process is conceptually the same as the one that was carried out in the previous project *ESPON Big Data for Territorial Analysis and Housing Dynamics*, but includes some improvements in terms of methodology and performance:

Geospatial information.

- Coordinates (XY). Improvements for coordinate's collection.

In those cases where coordinates were not available directly from 'front page' (html panel), we sent a second call to the API (Application Performance Interface), of the platform web page in order to obtain the coordinates. This process simulates the page's call to its own API in order to receive the coordinates called from the web page itself. Finally, for those addresses with no coordinates, an inverse geocoding (GoogleMap and Nokia) was used to obtain final coordinates.

Data normalization:

- Data cleaning. Improvements on time multiple real-estate sources.

The most common mistakes errors are duplicated ad's (sometimes a real estate ad can be published several times), absence of location coordinates or mistakes when entering the real estate ad (area, price, etc.). Due to build up crawlers on multi-platform source (Nestoria.com) we have built both processes for cleaning up repeating ad, and another one for cleaning up outliers.

- Data aggregation. Once data were cleaned up, web-scraped data were then aggregated in targeted geographical delineation (LAU2, grid). For each and every valid ad we had coordinates (XY).

Scraping languages:

- Python option.

Crawlers were coded in Python in order to facilitate further samples collection from ESPON repository. Once Python is an easier language for any operator this language options is considered the most interoperable talent wise.

Sample methodology and frequency:

- Proxies strategy. Improvements on time collection for one and each sample. Statistical consistency.

In order to collect a more accurate and fit to date sample for each site, we used a multi-proxies (dynamic IP) system. We worked with a system of more than 70Millions proxies (dynamics IP), in order to speed up the number of registers per second to be collected. Average call to web page should run from 0,5 to 3 seconds per register. This frequency is needed to avoid firewalls and robots' identifiers. We use a farm of proxies' servers which can run 1 million call per second.

- Data routines and frequency. Improvements on time patterns for data analytics.

A remarkable improvement of this project was to scrape data on a weekly basis during a period of time in order to analyse the market behaviour and dynamics. Due to valuable indicators concerning time once an offer is online, at least one sample a week from each scrapped data category was needed. Time strip for data sets allowed building up offer patterns in time, and property-level patterns on how long that offer remained online. Ideally this period should be longer, but due to the project duration, it was only collected for a few weeks (from 29/01/2022 to 30/04/2022).

Data for each site was collected on Saturdays in a maximum of 5 hours' time to keep time consistency of offer sample.

Data gathering:

- New indicators:

We gathered data for additional indicators that measure the specificity of cross-border housing markets.

- Multiple real-estate sources. Improvements on unconventional data sources.

We introduced a new wider and homogeneous source of real-estate data (www.nestoria.com).

Nestoria is a vertical search engine which integrate several real-estate platforms in one only repository. It provide a public API in order to collect data from a wide variety of EU countries. To mention some:

France: immobilier.lefigaro.fr

Spain: enalquiler.com; expocasa.com; globaliza.com; habitaclia.com; pisos.com; thinkspain.com; tucasa.com; vivados.es; yaencontre.com; trovimap.com.

UK: thehousesshop.com; citylets.co.uk; country.co.uk; espc.com; scotsproperty.com; london2let.com; smartnewhomes.com; zoopla.co.uk; onthemarket.com; home.co.uk; propertypigeon.co.uk; nethouseprices.com; s1homes.com

Poland: morizon.pl; szybko.pl; rynekpierwotny.pl; domiporta.pl; krm.pl; domy.pl; tabelaofer.pl; nieruchomosci-online.pl; taakidom.pl

Germany: www.immowelt.at

Italy: bakeca.it; wikicasa.it; cambiocasa.it; casa.it; easyavvisi.it; idealista.it; immobiliare.it; annunci.repubblica.it; livellocasa.it; prossimacasa.it

The generic scraping methodology carried out in this study to harvest the data needed from the different housing portals has different steps which are described below.

It needs to be stated that this project is focused on sale and long-term rental data. Short-term rental data was not analysed as it was considered that due to the impact of the COVID-19 pandemic on tourism during the study time period that the short-term rental market would be affected, and results would not be relevant in giving insights on market behaviour.

6.1.1 Scraping steps

6.1.1.1 STEP 1: Defining the geographical scope

The first step is defining the geographical scope for the scraping. The geographical area has to be defined in different ways depending on the website. In some sites, the coordinates of a quadrant containing the area of analysis have to be introduced. In others, the names of the regions containing the geographical area desired have to be inserted.

6.1.1.2 STEP 2: Downloading URL

The second step is obtaining the URLs of all the advertisements on each page of the website. Each website page is going to be one call to the server, and it will get the URLs of the advertisements included in that page. The more pages the website has, the more calls will be needed. The output is a list of URLs, one per each listing that is going to be harvested.

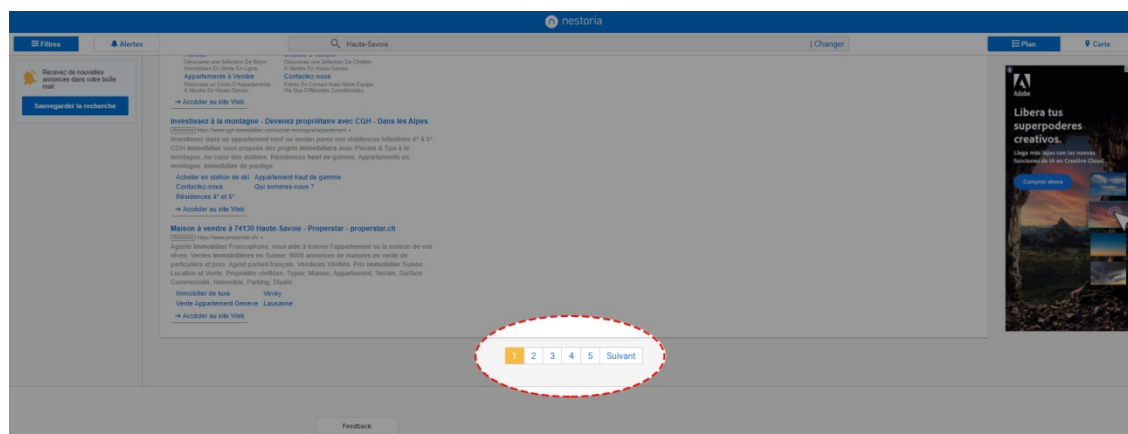


Figure 12: Examples of Page in Nestoria

6.1.1.3 STEP 3: Downloading content

The third step is obtaining the desired content of each advertisement. In this case, another call is made using the URL obtained in step 2 to automatically collect the data. This step is carried out for each advertisement, so it means one call to the server for each listing.

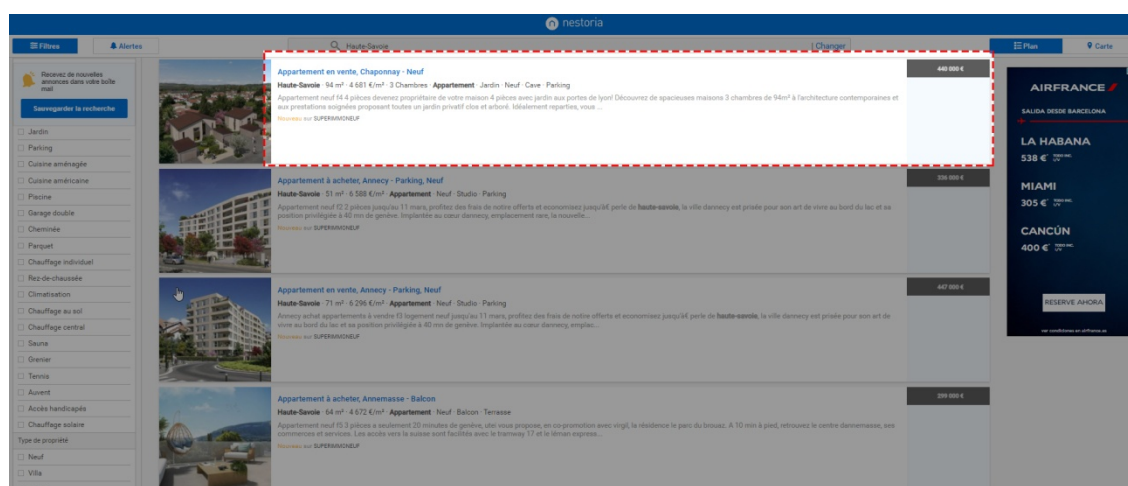


Figure 13: Example of Content in Nestoria

6.1.1.4 STEP 4: Downloading latitude and longitude

The fourth step is not needed in all websites, but some of them require this additional step in order to obtain the coordinates (latitude and longitude). In these cases, another call to the server is carried out for each advertisement.

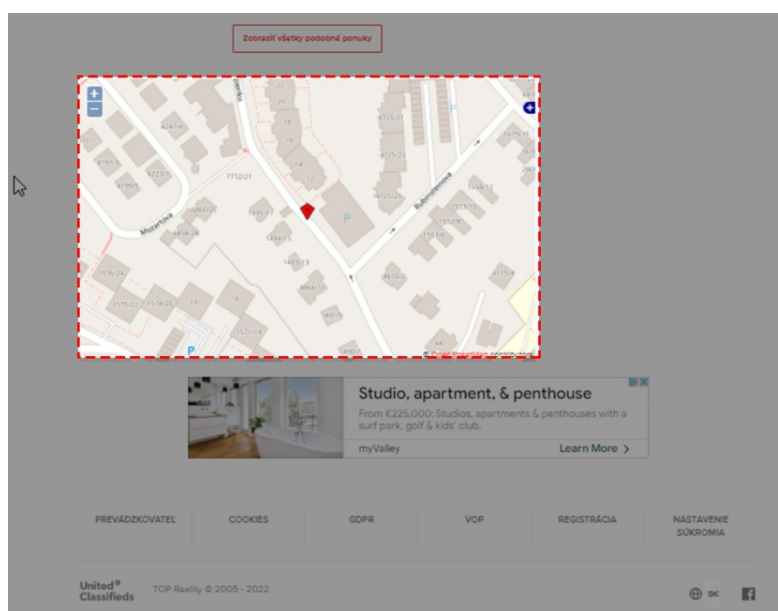


Figure 14: Example of coordinates collection in Top Reality

6.1.2 Possible scraping problems:

Possible problems which may arise during the scraping period are the following:

- **Server crash:** if the website server crashes due to any troubleshoot external to the process of scraping, the website will not be available for data scraping.
- **No advertisements:** the possibility of not getting any advertisements exists at some locations on a specific date during the period of data scraping, especially in areas with a low accommodation offering.
- **Language:** problems regarding language may be encountered in some countries if the website cannot be translated.

For the previous reasons, a proposal of alternative websites for each case study was made (section 6.4) in case any of the mentioned problems arose.

6.2 Real Estate websites selection criteria

Taking into account the previous methodology, the criteria to select the real estate websites which were going to be used for the data harvesting in each case study were defined:

6.2.1 Portal rating

Housing sale and long-term rental data was needed for the study. In order to select the websites to obtain the data, the best real estate websites in each country were analysed using the real estate ranking sites listed below.

Real Estate Ranking Websites:

- <https://foreignbuyerswatch.com>
- <https://www.allyoucanread.com>
- <https://www.similarweb.com/top-websites/>
- <https://www.expatica.com/ch/living/household/swiss-apps-191780/>
- <https://www.bystored.com/blog/best-property-websites>

6.2.2 Common websites

It needs to be taken into consideration that having homogeneous data between case studies is important for comparison, so if a common website among the different case studies existed, it was also considered, despite of not being in the ranking of best websites analysed during the previous step.

6.2.3 Listings amount

It is important to understand that the cross-border case studies contain small cities, and due to that, some of the best websites might not have any advertisements in the selected areas of study. In those cases, other sites at a more local scale were also considered.

6.2.4 Website structure

Apart from the previous criteria, the structure of the website needed to be analysed in order to know if the portal data was possible to harvest.

6.2.5 Website content

The content of each site was taken into account to see if the data desired for the analysis of all the case studies was available (such as surface, price, rooms,...) and if their offer included both rental and sale advertisements. Language was also considered as an important issue when gathering the data, to ensure the viability of the analysis.

6.2.6 Location data

A really important point was the availability of the map with the location of the advertisements, in order to be able to geolocate each advertisement. If the location was not available, then the site was not considered among the available possibilities of each case study.

6.3 Real Estate websites initially selected

The websites initially selected for each case study are listed below. Some characteristics of each site are detailed in order to provide an idea of the appropriateness of each site.

CASE STUDY	COUNTRY	WEBSITE	PORTAL RATING	COMMON WEBSITES	LISTINGS AMOUNT	WEBSITE STRUCTURE	WEBSITE CONTENT	LOCATION DATA
Geneva-Annecy	France	https://www.nestoria.fr/	High	France/Spain/UK	High	Easy harvest to	Complete/needs translation	Available
Geneva-Annecy	Switzerland	https://www.homegate.ch	High	Switzerland	High	Easy harvest to	Complete/needs translation	Available
North Ireland – Ireland	North Ireland	https://www.nestoria.co.uk/	High	France/Spain/UK	Medium	Easy harvest to	Complete/English	Available
North Ireland – Ireland	Ireland	https://www.daft.ie/	High	Ireland	High	Easy harvest to	Complete/English	Available
Denmark – Sweden	Denmark	https://www.boligsiden.dk	High	Denmark	High	Easy harvest to	Complete/needs translation	Available
Denmark – Sweden	Sweden	https://www.blocket.se	High	Sweden	High	Easy harvest to	Complete/needs translation	Available
Slovakia – Austria	Slovakia	https://www.topreality.sk	High	Slovakia/Austria	High	Easy harvest to	Complete/needs translation	Available

Slovakia – Austria	Austria	https://www.topreality.sk	Medium	Slovakia/Austria	Medium	Easy harvest	to	Complete/needs translation	Available
Romania – Bulgaria	Romania	https://www.publi24.ro	High	Romania	High	Easy harvest	to	Complete/needs translation	Available
Romania – Bulgaria	Bulgaria	https://www.holprop.com	Medium	Bulgaria	Medium	Easy harvest	to	Complete/needs translation	Available
France-Spain	France	https://www.nestoria.fr/	High	France/Spain /UK	High	Easy harvest	to	Complete/needs translation	Available
France-Spain	Spain	https://www.nestoria.es/	High	France/Spain /UK	High	Easy harvest	to	Complete/needs translation	Available

Table 4: Selected real estate websites information

6.4 Alternative websites:

In case any problems arose when harvesting the data from the previous websites, the following ones were proposed as alternative sites for scraping:

CASE STUDY	COUNTRY	ALTERNATIVE WEBSITES
Geneva-Annecy	France	https://www.seloger.com
Geneva-Annecy	Switzerland	https://www.immobilier.ch
North Ireland – Ireland	North Ireland	https://www.zoopla.co.uk https://www.rightmove.co.uk/
North Ireland – Ireland	Ireland	https://www.myhome.ie
Denmark – Sweden	Denmark	https://www.nybolig.dk
Denmark – Sweden	Sweden	https://www.hemnet.se https://www.booli.se
Slovakia – Austria	Slovakia	https://www.real-estate-slovakia.com https://www.areality.sk https://ringo.topky.sk
Slovakia – Austria	Austria	http://www.nestoria.at https://www.properstar.co.uk https://www.immowelt.at http://www.bit-immobilien.at
Romania-Bulgaria	Romania	https://world-estate.com
Romania-Bulgaria	Bulgaria	https://www.bulgarianproperties.com https://www.property.bg https://www.luximmo.com
France-Spain	France	https://www.seloger.com
France-Spain	Spain	https://www.fotocasa.es

Table 5: Alternative websites proposed

6.5 Type of data available

A list of indicators that were gathered from the different sites is listed below.

DATA	DESCRIPTION	DATA	DESCRIPTION
id	ID of the dataset	listing_sold_price*	Sold price
listing_url	URL of the listing	listing_build_year*	Construction year
listing_title	Title of the listing	listing_sold_time*	Date of sale
advertiser	Name of the advertiser	listing_area_price*	Price/Sqm
advertiser_code	Code of the advertiser	listing_percentage*	% sold price over advertised price
advertisement_code	Code of the listing	listing_bathrooms*	Number of bathrooms

lng	Longitude	listing_address*	Adress of the listing
lat	Latitude	listing_floors*	Number of floors
listing_type	Typology (Rent/Sale)	listing_floor_number*	Floor number of the listing
listing_id	ID of the listing	listing_garden*	Garden availability (True/False)
listing_prop_type	Type of property (House/Apartment)	listing_balcony*	Balcony availability (True/False)
listing_area	Surface	listing_elevator*	Elevator availability (True/False)
listing_price	Price	listing_terrace*	Terrace availability (True/False)
listing_rooms	Number of rooms	listing_swimming_pool	Swimming pool availability (True/False)
search_code	Date of scraping	listing_communal_pool*	Communal pool availability (True/False)
listing_crawled_platform	Name of the website of the listing	listing_private_pool*	Private pool availability (True/False)

Table 6: Gathered indicators
*(*Data not available in all websites)*

6.6 Case study specificities

In this sub-chapter, some specific characteristics on the scraping of each case study are mentioned, as well as scraping limitations, results, problems encountered during the process and solutions adopted.

6.6.1 France-Switzerland Case Study

For this case study, there was no common website with listings in both countries. Two different portals were harvested which meant having less homogeneous data for the analysis. However, in both countries the websites had complete information for rent and sale.

6.6.1.1 France

- **WEBSITE:** The final website used to collect data for the French part of the case study was <https://www.nestoria.fr/> as it was common to other case studies.
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope was defined using the coordinates of a quadrant containing the area of analysis.
- **OBSERVATIONS:** Attribute data obtained for rental and sale listings was complete, but some indicators needed translation and homogenization (such as housing typologies or transaction type) in order to compare the data with the rest of the case studies.

Furthermore, large groups of advertisements were found located at the same coordinates, as sometimes the location of the listings was a centroid of the postal code/district or the location of the advertising agency. This type of inaccuracy is always a limitation when conducting spatial analysis.

It needs to be taken into account that Nestoria is an aggregator platform, meaning it collects advertisements from different portals. Therefore, adverts can appear more than once. In order to be able to compare the number of listings in the same way it was done for other case studies (where only one platform was being scraped), the portal with the biggest amount of adverts in each LAU and for each typology (rent and sale) was chosen.

Another peculiarity of this case study that was detected during the scraping period was that the dataset of the first sample had around 30.000 records, after that it got down to 8.000 for a few weeks and then

it rose again to 30.000 listings. After conducting some analysis on the results, it appeared that there were some issues in the platform during the scraping period, and listings of some advertisers inside Nestoria were not being collected properly for some time.

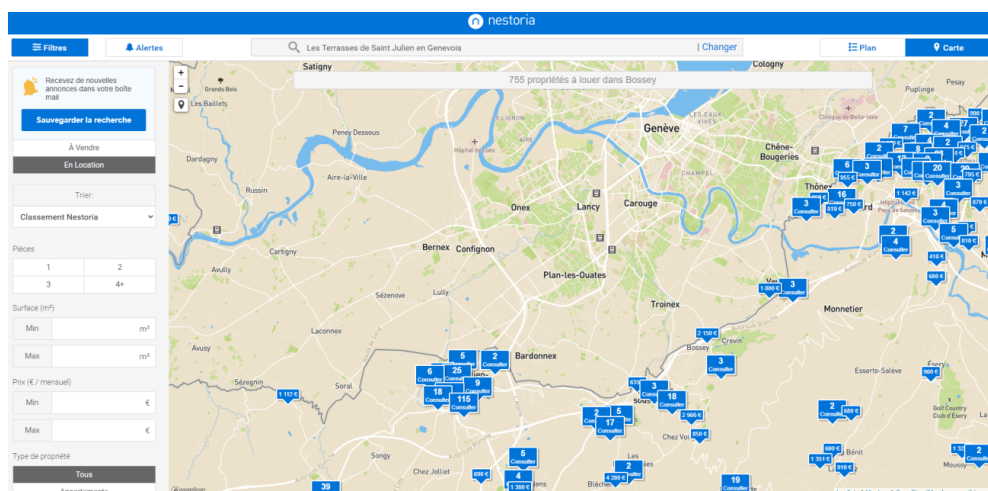


Figure 15: Screenshot of Nestoria France

6.6.1.2 Switzerland

- **WEBSITE:** In the case of Switzerland, the site scraped was <https://www.homegate.ch/de>
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope was defined using the names of the two regions containing the area of study (Canon de Genève and District de Nyon).
- **OBSERVATIONS:** Data gathered for rent and sale was complete, however some indicators needed translation and currency needed conversion from CHF to EUR for dataset homogenization and analysis.

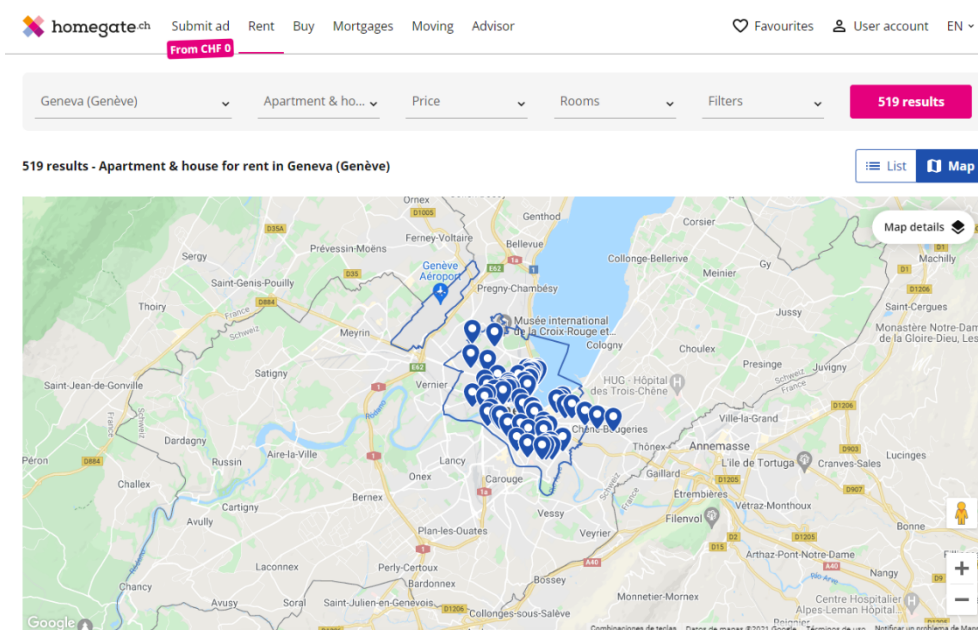


Figure 16: Screenshot of Homegate

6.6.2 Denmark-Sweden Case Study

For this case study, there was not a common portal with listings from both countries, so two websites were used. Furthermore, in the case of Sweden, the selected portal turned out to not have enough listings for sale typology, so one of the alternative websites proposed was additionally harvested in order to complement the results.

6.6.2.1 Denmark

- **WEBSITE:** The website used to collect data for Denmark was <https://www.boligsiden.dk/>. However, as in Denmark historical data of sold properties was available, another website <https://www.boliga.dk> was scraped in order to obtain it.
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope was defined using a quadrant which contained the area of analysis. In the case of historical data, the whole country was considered.
- **OBSERVATIONS:** The results turned out having a huge dataset of listings for sale, but a smaller sample for rent (comparing to alternative websites which only offered rent). However, even though the sample was small, it was considered enough for the analysis taking into account the timings of the project \, and the efforts needed to prepare a new script in order to include an additional website.

Once more, some indicators needed translation in order to homogenize the dataset and currency needed conversion from DKK to EUR for the analysis.

A specific issue that arose with Boligsiden during the scraping period was that, at some point, the taxonomy of housing typologies which initially came in Danish changed to English, so the process of data homogenization had to be adapted to this change.

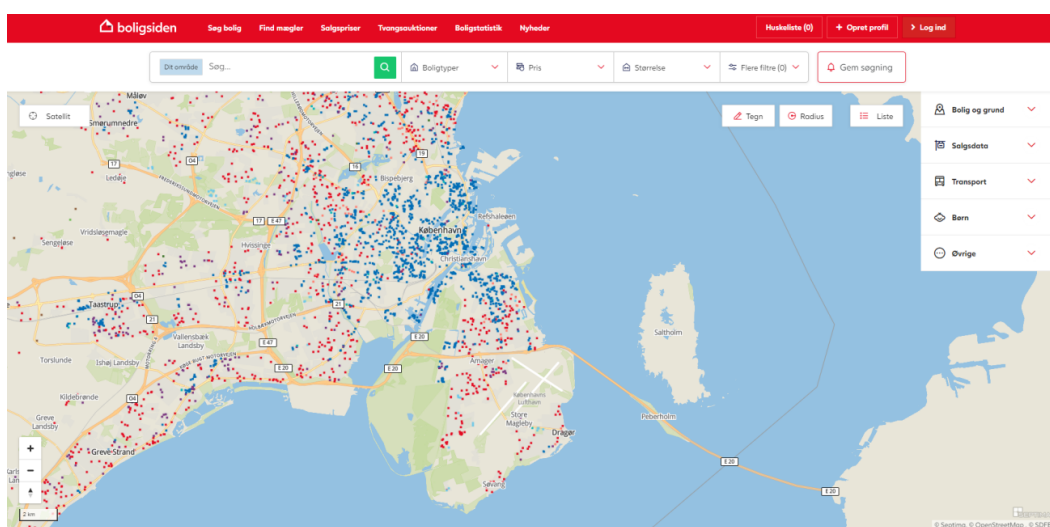


Figure 17: Screenshot of Boligsiden

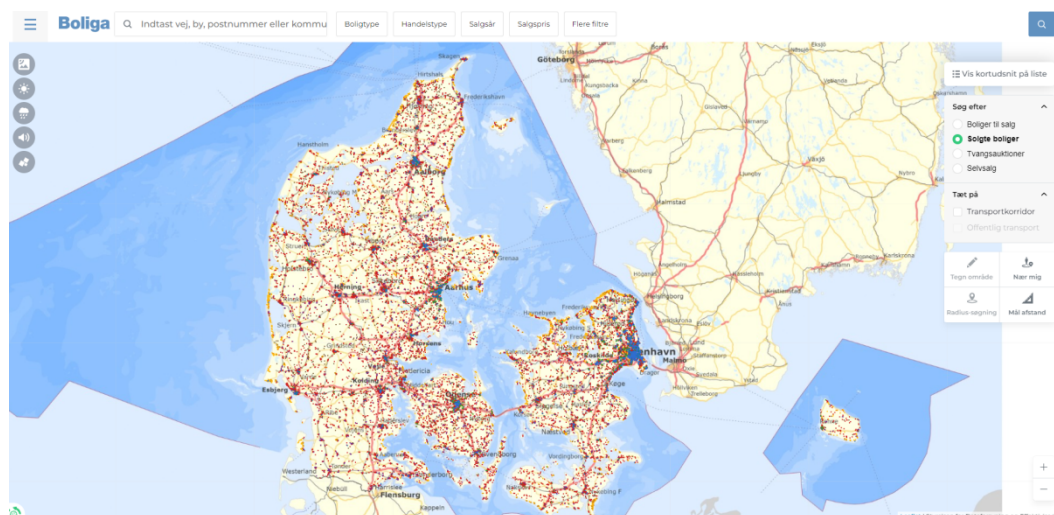


Figure 18: Screenshot of Boliga

6.6.2.2 Sweden

- **WEBSITE:** In Sweden, one of the alternative websites needed to be added to the selected one, as it turned out that, despite appearing to be one of the best websites for renting in Sweden, it did not have any sale listings in the area of analysis. Therefore, to select the best alternative website, some research on property purchasing process in Sweden was carried out:
 - The sale is usually conducted through a bidding process where the Real Estate agent has exclusive rights of the property of the seller (so properties are usually not advertised for more than one agent).
 - The agent is the one in charge of the whole purchasing process (no lawyers or solicitors needed) and manages the bidding process⁵.
 - The advertisements are mainly published in two websites called Hemnet and Booli, which have the advertised prices of listings and historical data of sales which can be used as references for bidders.

Considering that information and keeping in mind it was necessary to complement our data on sale listings, it was decided to add Hemnet site to our scraping. This would not only allow us to collect data on advertised prices but also historical data of sold properties too.

For this reason, two websites were used for data gathering in Sweden, <https://www.blocket.se/> for rent and <https://www.hemnet.se/> for sale and sold properties.

- **GEOGRAPHICAL SCOPE:** For sale and rental data, the geographical scope was defined using the name of the County which comprised the area of analysis (Skåne). In the case of historical data, a limited number of records could be scraped in each call, so instead of launching one single call for the region, 10 calls were conducted (one for each municipality inside the study area) in order to gather more data.

⁵ Swedbank Fastighetsbyrå AB, 'The road home - buying a residence in Sweden', *Fastighetsbyrån* [website], <https://www.fastighetsbyran.com/sv/sverige/artiklar/engelska-sprakstod/>, (accessed 25 April 2022)

- **OBSERVATIONS:** Rental data and historical data of sold properties were complete; however, in the case of sale data, coordinates could not be gathered. For this reason, a geocoding process was launched using the address of the listings to obtain the coordinates. In case the address was not accurate, the centroid of the postal code was used to assign a location, although some listings could still not be geolocated.

As in previous case studies, some indicators needed to be translated and currency converted from SEK to EUR for data homogenization and analysis.

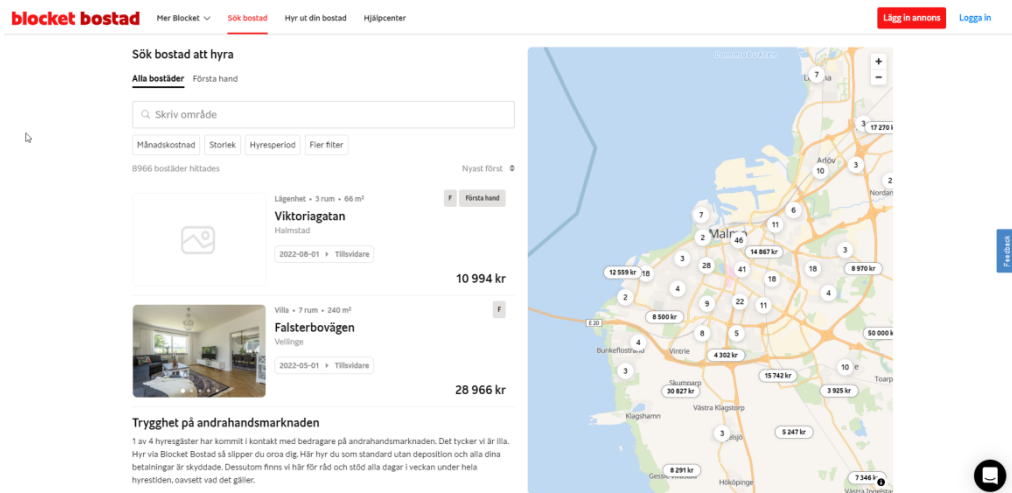


Figure 19: Screenshot of Blocket Bostad

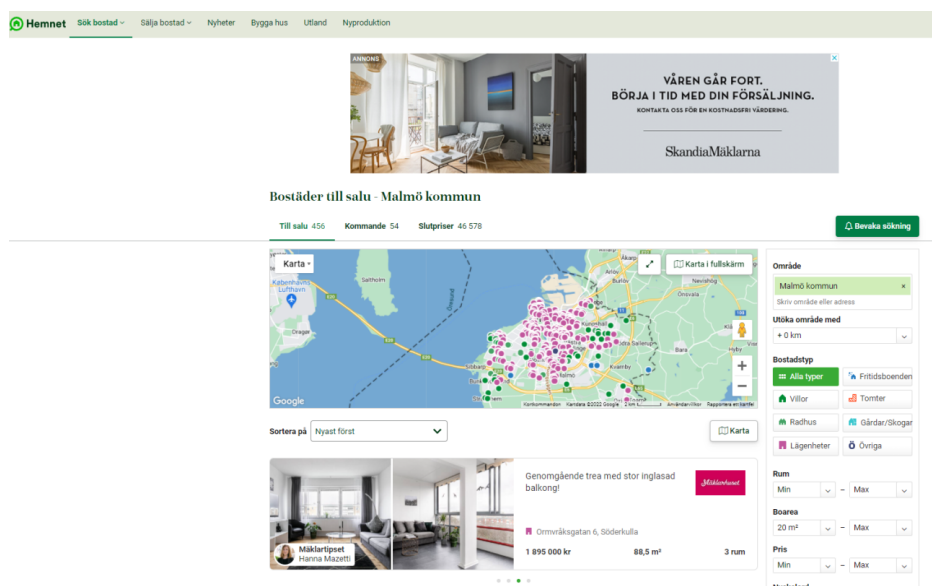


Figure 20: Screenshot of Hemnet

6.6.3 Ireland – Northern Ireland Case Study

In this case study, there was no unified portal with listings in both countries, so two different portals were harvested.

A particularity of this case study was that the Northern Ireland portals did not publish the listings surface area. Therefore, some research was conducted in order to understand that behavior. It was detected that *“When properties are being marketed, the norm in most European countries is to define dwelling size by floor area, whereas in the UK it is by the number of bedrooms.”*⁶.

Keeping that in mind, it was mandatory to obtain data on the number of rooms and that the website selected for Ireland also had that indicator available, so the two countries could be compared using the same metrics (price/number of bedrooms instead of price/sqm) for this case study.

6.6.3.1 Ireland

- **WEBSITE:** The selected website used to collect data for Ireland was <https://www.daft.ie/>
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope for the scraping was the whole country.
- **OBSERVATIONS:** Attribute data obtained for both, rental and sale listings, was complete, containing data on the number of bedrooms.

⁶ Greater London Authority, 'Housing space standards', London, Greater London Authority, August 2006, https://www.london.gov.uk/sites/default/files/hatc_housing_space_standards_report_for_gla_2006.pdf, (accessed 25 April 2022)

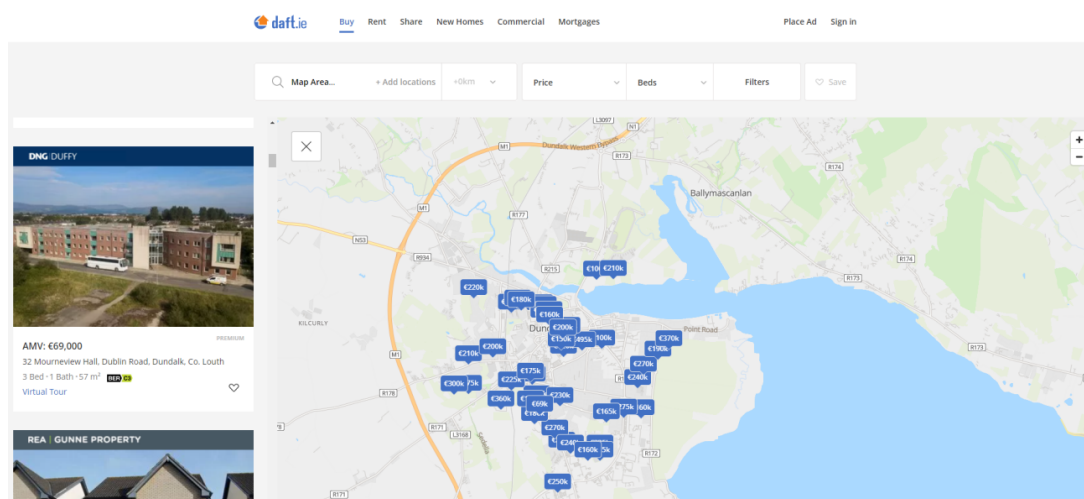


Figure 21: Screenshot of Daft

6.6.3.2 Northern Ireland

- **WEBSITE:** The website used for Northern Ireland was <https://www.nestoria.co.uk/> as it was common to other case studies.
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope for the scraping was a quadrant containing the area of analysis.
- **OBSERVATIONS:** The records in the area of analysis were complete; however, most listings were mainly for sale, and there were few advertisements for rent near the border. With these results in mind, alternative websites proposed were analyzed in order to figure out the amount of listings that were advertised on other platforms in that same zone. Alternative websites appear to have the same few listings for rent; therefore, the scraped results were considered correct.

In this case, prices came in GBP, so currency needed conversion to EUR for the analysis.

Moreover, as mentioned before, Nestoria is an aggregator platform and advertisements can appear more than once. So, in order to be able to compare the data in the same way it was being done in other case studies, again the advertiser with the highest amount of adverts in each LAU and for each typology was chosen.

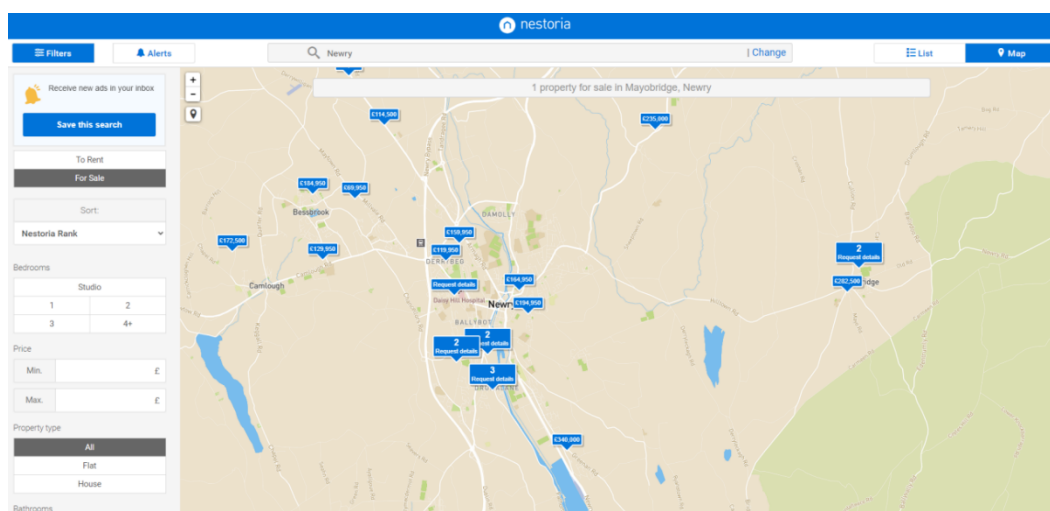


Figure 22: Screenshot of Nestoria UK

6.6.4 Austria – Slovakia Case Study

In this case study, there was a unified portal with listings for both countries, so that one was selected for the scraping. However, for Austria, an alternative website was added to the scraping to complement the results, as it turned out that listings were concentrated in the core area, but the rest of the corridor had no listings.

6.6.4.1 Austria

- **WEBSITE:** In the case of Austria, most listings of the initially selected website appeared to be near the border. That was an advantage compared to other portals analyzed in Austria, as for the cross-border case study this was relevant, but the rest of the corridor appeared to have almost no listings. For that reason, an alternative website was added to the scraping in order to complement the rest of the corridor. The supplementary site selected was Nestoria, as it was also common to other case studies.

Therefore, two websites were used for Austria: <https://www.topreality.sk/> (with listings near the border), and <https://www.nestoria.at/> (with listings in the rest of the corridor).

- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope for the scraping was a quadrant containing the area of analysis -for the case of Nestoria- and the name of 3 regions (Wein, Niederösterreich and Burgenland) containing the area of study -for the case of Top Reality.
- **OBSERVATIONS:** Records collected were complete in both cases although some indicators needed to be translated for data homogenization and analysis.

Once again, as Nestoria platform was used to gather the data for Austria; to avoid duplicities the advertiser with the highest amount of listings for each LAU and typology was chosen.

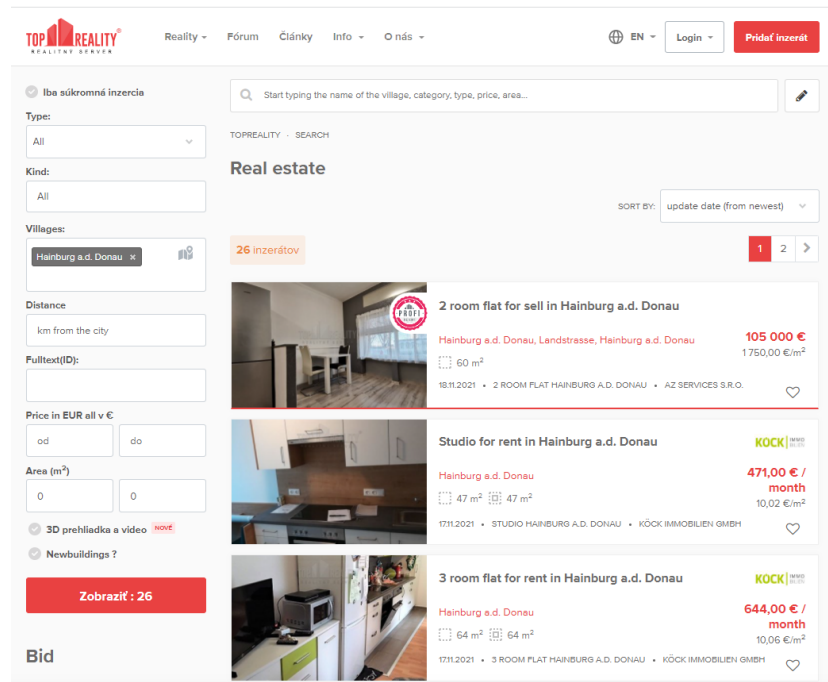


Figure 23: Screenshot of Top Reality (Austria)

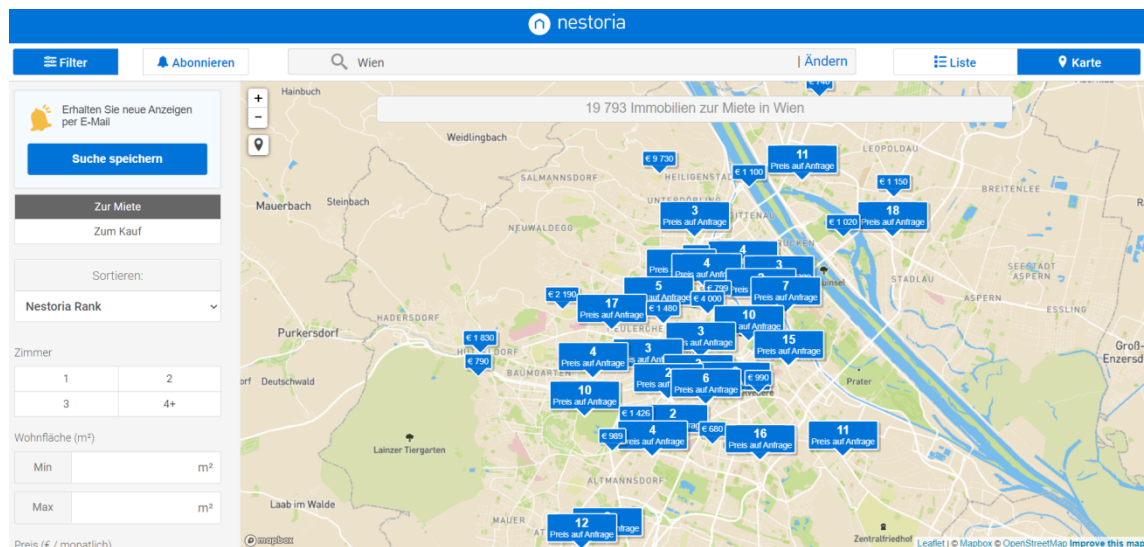


Figure 24: Screenshot of Nestoria Austria

6.6.4.2 Slovakia

- **WEBSITE:** The website used for Slovakia was <https://www.topreality.sk/> as it had complete information, and it was common to Austria.
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope for the scraping was Bratislava Region which contained the area of study.

- **OBSERVATIONS:** Some records collected did not have coordinates nor an accurate address to launch the geocoder, so they could not be geolocated. The rest of the information was complete, although some indicators needed to be translated for data homogenization and analysis.

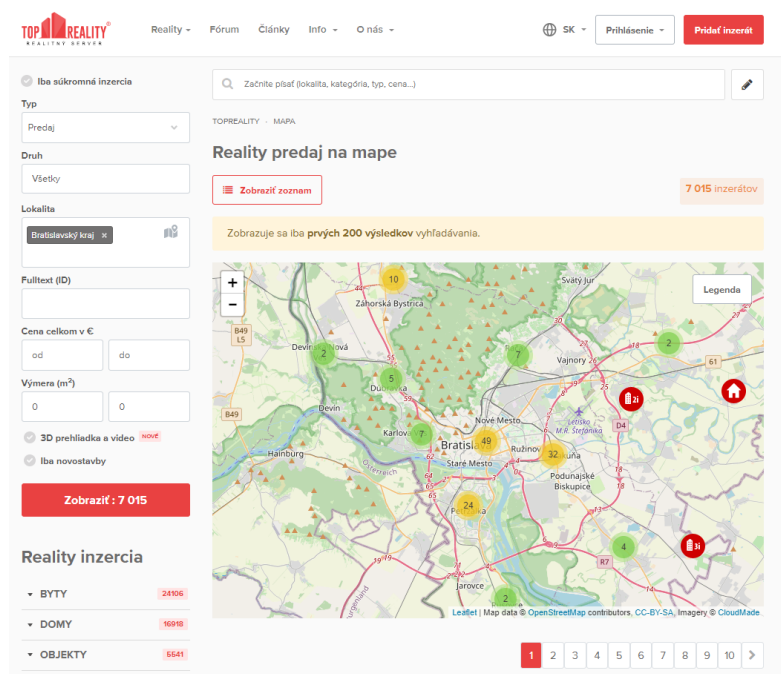


Figure 25: Screenshot of Top Reality (Slovakia)

6.6.5 Romania – Bulgaria Case Study

In this case study, there was not a unified portal with listings for both countries, so two different portals were used for the scraping. A peculiarity of this case study was that in Bulgaria the chosen website turned out to not have rental data on the housing typologies within the area of study nor in the alternative portals suggested. Due to that, some research on different portals was conducted in order to find a platform with listings of both types in the area of analysis.

6.6.5.1 Romania

- **WEBSITE:** The website used for Romania was <https://www.publi24.ro/>
- **GEOGRAPHICAL SCOPE:** The geographical scope was defined using Giurgiu Region, which was the area of study.
- **OBSERVATIONS:** Data gathered was complete, although there were few records for rental data. Again, as in other case studies, some indicators needed to be translated for data homogenization and analysis.

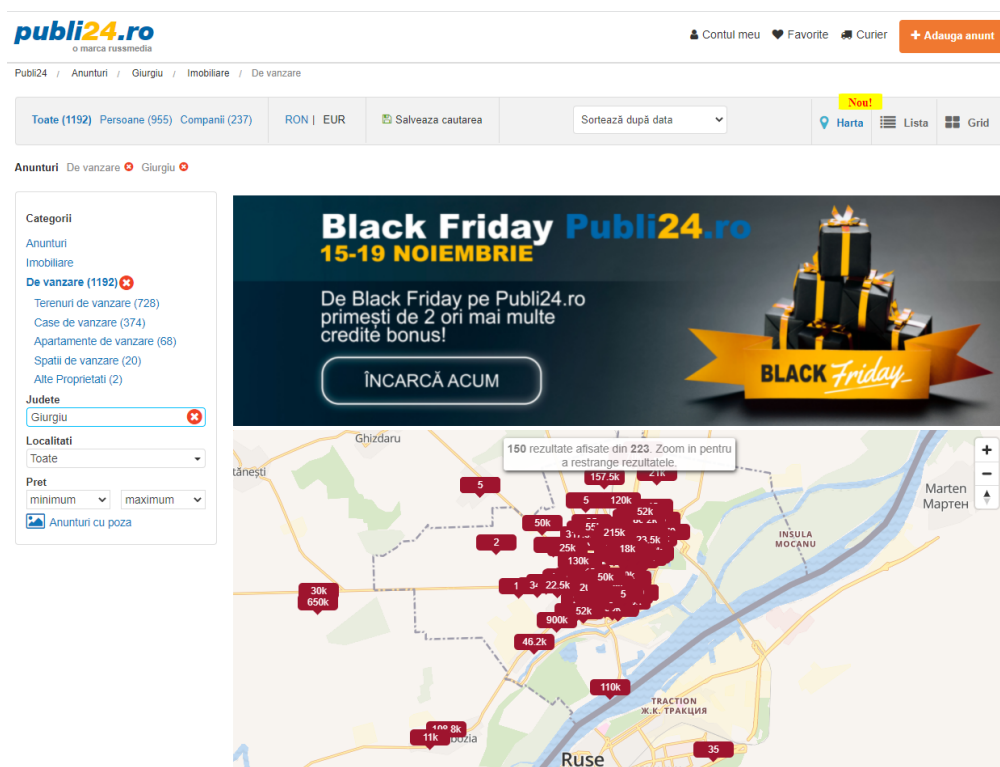


Figure 26: Screenshot of Publi24

6.6.5.2 Bulgaria

- **WEBSITE:** As mentioned before, the selected website for Bulgaria turned out to not have housing rental data in the area of study nor in the alternative ones. So, after looking for other real estate websites, two options were proposed: <https://www.indomio.bg/> (which was not able to be scraped) and <https://en.realestates.bg/> which finally was the website used to gather the data.
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope for the scraping was defined using the regions containing the area of study (Ruse, Veliko Tarnovo, Targovishte and Razgrad).
- **OBSERVATIONS:** Data collected was complete for both rent and sale even though there were few records in the area of analysis.

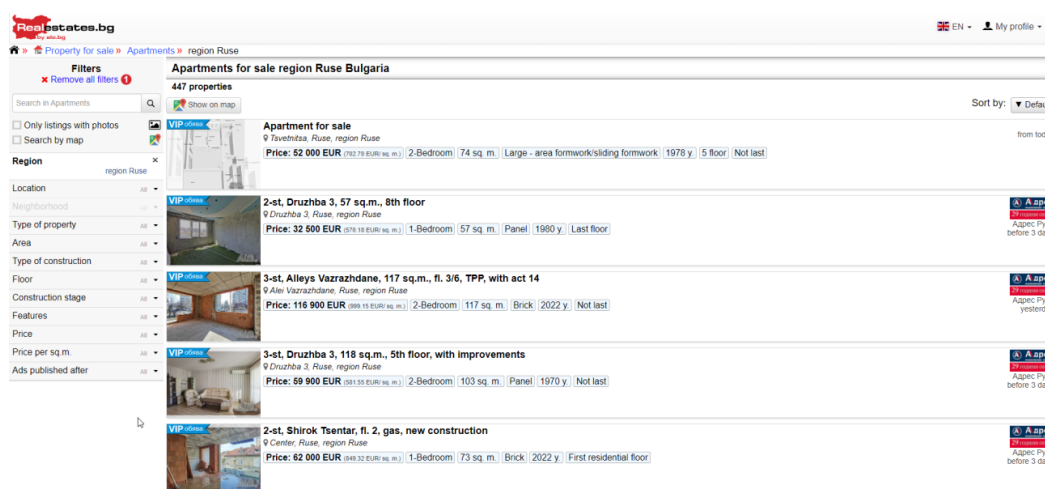


Figure 27: Screenshot of Realestates

6.6.6 France – Spain Case Study

In the Spanish-French case study, there was a common website for both countries, so it was selected for the scraping.

6.6.6.1 France

- **WEBSITE:** The website used to collect data for the French part of the case study was <https://www.nestoria.fr/> as it was common to other case studies.
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope was defined using the coordinates of a quadrant containing the area of analysis.
- **OBSERVATIONS:** Attribute data obtained for rental and sale listings was complete, but some indicators needed translation, such as housing typologies or transaction type, in order to normalize the data with the rest of the case studies.

In this case, large groups of advertisements were found located at the same coordinates.

Moreover, the first sample gathered around 6.000 records, after that it got down to 1.200 for a few weeks and then it rose again to 6.000 records. As mentioned before, it seemed that there was some issue in the platform during the scraping period.

As in the case study of France-Switzerland, Nestoria was used again to gather the data for France, the advertiser with the biggest amount of adverts for each LAU and typology was chosen to avoid duplicities.

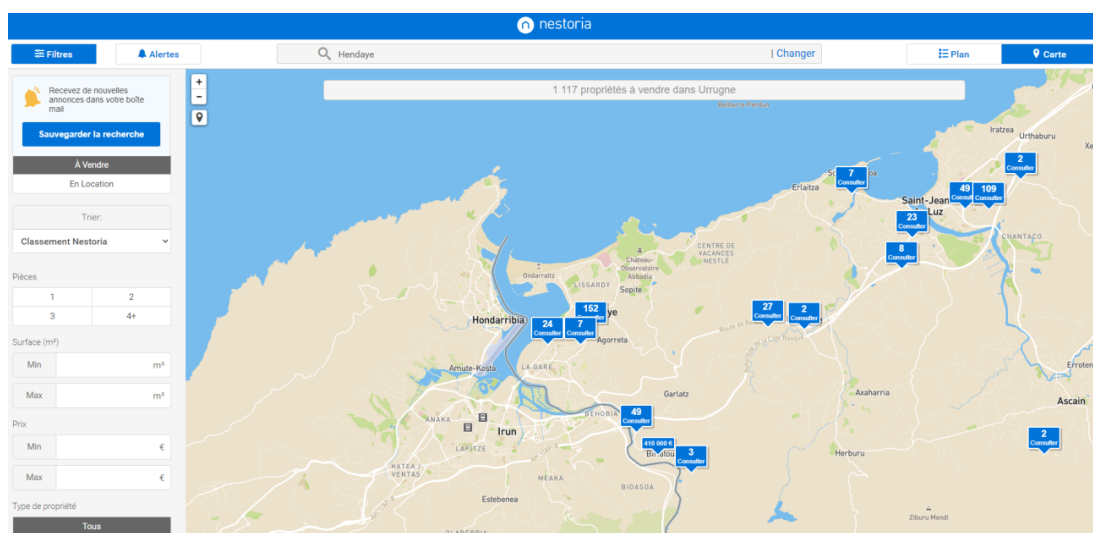


Figure 28: Screenshot of Nestoria FR

6.6.6.2 Spain

- **WEBSITE:** The website used to collect data for Spain <https://www.nestoria.es/> as it was common to other case studies.
- **GEOGRAPHICAL SCOPE:** In this case, the geographical scope was defined using the coordinates of a quadrant containing the area of analysis
- **OBSERVATIONS:** Attribute data obtained for rent and sale listings was complete, but some indicators needed translation, in order to homogenize the data with the rest of the case studies.

As in other case studies where Nestoria was used for the data gathering, it seemed that there was some issue in the platform during the scraping period, and in this case, the first sample had around 4.000 records; after that, it got down to 1.400 for a few weeks and then rose again to 4.000 records.

In order to avoid Nestoria duplicities, again the advertiser with the biggest amount of listings for each LAU and typology (rent and sale) was chosen.

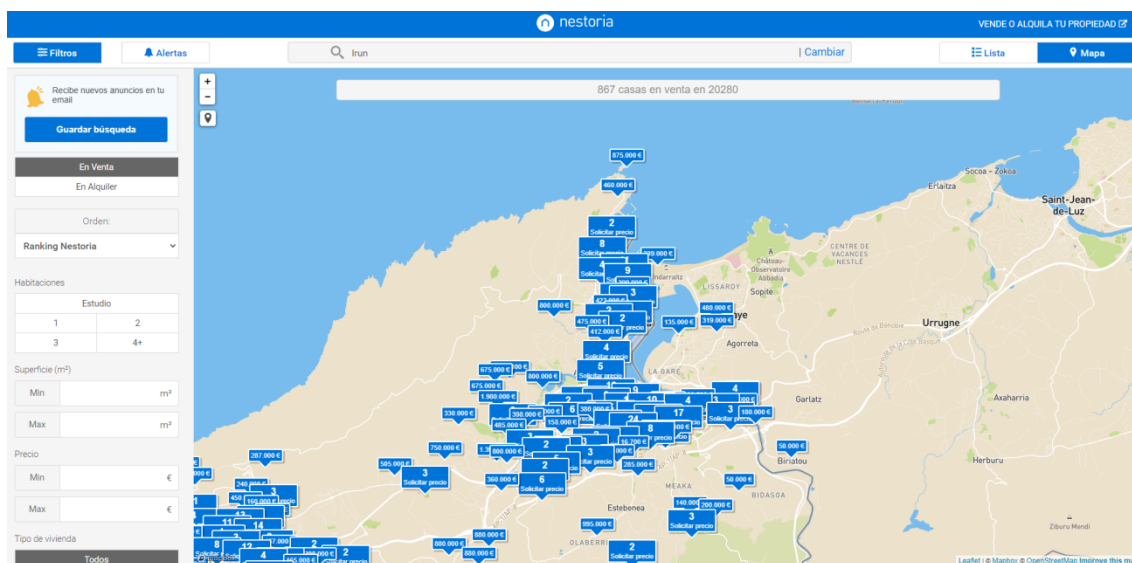


Figure 29: Screenshot of Nestoria ES

6.6.7 Summary of scraped portals

CASE STUDY	COUNTRY	WEBSITE	GEOGRAPHICAL SCOPE	AVERAGE WEEKLY LISTINGS OBTAINED
Geneva-Annecy	France	https://www.nestoria.fr/	Quadrant	8.000-30.000
Geneva-Annecy	Switzerland	https://www.homegate.ch/	Regions	2.400-2.600
Northern Ireland	Northern Ireland	https://www.nestoria.co.uk/	Quadrant	1.700-2.200
Northern Ireland	Ireland	https://www.daft.ie/	Country	14.000-15.000
Denmark – Sweden	Denmark	https://www.boligsiden.dk/ https://www.boliga.dk/	Quadrant (Rent/Sale) Country (Historical data sold properties)	9.000-10.000 (Rent/Sale) 370.000 (historical data sold properties)
Denmark – Sweden	Sweden	https://www.blocket.se/ https://www.hemnet.se/	Regions (Rent) Regions (Sale) / Municipalities (historical data sold properties)	1.000 (Rent) 2.000 (Sale) / 21.700 historical data sold properties
Slovakia – Austria	Slovakia	https://www.topreality.sk/	Regions	13.000-14.000
Slovakia – Austria	Austria	https://www.topreality.sk/ https://www.nestoria.at/	Regions (Cross-border area) Quadrant	30.000-50.000
Romania-Bulgaria	Romania	https://www.publi24.ro/	Regions	1.400
Romania-Bulgaria	Bulgaria	https://en.realestates.bg/	Regions	700-800
France-Spain	France	https://www.nestoria.fr/	Quadrant	1.200-6.000
France-Spain	Spain	https://www.nestoria.es/	Quadrant	1.500-4.000

Table 7: Scraping Summary

6.7 Data cleaning, normalization and enrichment

This section explains the methodology followed to clean the data coming from the scraping along with the way in which it was standardized and enriched for the posterior analysis.

6.7.1 Methodology

Results obtained with scraping always need to be cleaned as listings can come with incomplete information or wrong content (such as coordinates, area, price, etc...). So, the following methodology was carried out in order to clean the raw data collected and prepare and enrich the final dataset for the data analysis and map creation.

6.7.1.1 Geographical scope delimitation

- **Coordinates homogenization:**

The first step of the data cleaning was homogenizing the coordinates of each case study as depending on the platform they come as latitude/longitude or as longitude/latitude.

- **Geocoding process:**

For those records where coordinates could not be gathered during the scraping (such as the case of properties for sale in Sweden), a geocoding process was launched using the address of each property, and in cases where the address was not accurate, the postal code was used to assign a coordinate.

The geocoding process returns coordinates from a given address by connecting to network data sets of each of two suppliers: Google and HERE. The process designed allows choosing the most appropriate match from each one. It needs to be remarked that the geocoder process does not return the coordinates if the given address does not match with the network data set provided by each API's suppliers.

- **Data enrichment:**

The following step of the analysis was to enrich the data with the LAU2 geometries for the delimitation and aggregation of the data. A geospatial process was conducted using the digital cartography of LAU2 in order to assign the geometry that belonged to each record.

This enrichment was also used for the delimitation of the dataset in the scraping as the records inside the geometries of the corridor can be identified and selected.

6.7.1.2 Complete records identification

The following step of the data cleaning process was to identify those records with incomplete information. For this reason, all listings with no price or no surface were taken out from the dataset.

In the case of Northern Ireland, as mentioned before, surface data was not published in the advertisements, so number of rooms was used for the analysis. Therefore, in Northern Ireland those listings without surface data were kept in the dataset.

6.7.1.3 Housing typology definition

- **Taxonomy of listing typology:**

As data came in different languages depending on the country and the platform, translation and normalization was needed in order to classify the listings in two typologies: Rent and Sale. For this reason, the following table was created with the taxonomy for all countries to facilitate the subsequent analysis:

RENT	SALE
MIETEN	KAUFEN
ALQUILER	COMPRAR
RENT	SALE
LOCATION	VENTE
RESIDENTIAL-TO-RENT	BUY
DE-INCHIRIAT	RESIDENTIAL-FOR-SALE
	DE-VENZARE
	TILL SALU

Table 8: Listing Typology Taxonomy

It needs to be remarked that those listings with no classification into a typology of any language were taken out from the dataset.

- Taxonomy housing typology:**

In order to standardize the data of all countries for a correct comparison, a taxonomy for housing typologies was created with a classification into single-family (houses) and multifamily (apartments) as it was suggested on the previous report in the Technical Guidance Document: *“Further studies should systematically differentiate apartments and houses for the analysis.”*⁷

Therefore, for each case study, housing typologies were identified and non-housing typologies (such as land, office, etc.) were taken out from the dataset. In some countries, translation was needed for the normalization process. In the case of Denmark, as mentioned before, during the scraping period the platform changed the typology language from Danish to English, so the homogenization process established had to be adapted to this change.

The following tables show the taxonomy of typologies considered as housing, and its classification into single-family and multifamily:

MULTIFAMILY HOUSING TYPLOGIES			
APPARTEMENT	APARTMENT;ROOF_FLAT	APARTMÃN / PREDAJ	EJERLEJLIGHED
APARTMENT	APARTMENT;SINGLE_ROOM	APARTMÃN / PRENÃJOM	WOHNUNG
APARTMENTS	APARTMENT;STUDIO	BYTOVÃ DOM / PREDAJ	2-BEDROOM
STUDIO	APARTMENT;TERRACE_FLAT	BYTOVÃ DOM / PRENÃJOM	1-BEDROOM
FLAT	1 IZBOVÃ BYT / PREDAJ	DVOJGARSÃNKA / PREDAJ	ATTIC
PISO	1 IZBOVÃ BYT / PRENÃJOM	DVOJGARSÃNKA / PRENÃJOM	ATELIER / STUDIO
DUPLEX	2 IZBOVÃ BYT / PREDAJ	GARSÃNKA / PREDAJ	LARGE
APARTMENT;ATTIC_FLAT	2 IZBOVÃ BYT / PRENÃJOM	GARSÃNKA / PRENÃJOM	APARTAMENTE DE VANZARE
APARTMENT;BACHELOR_FLAT	3 IZBOVÃ BYT / PREDAJ	INÃ BYT / PREDAJ	APARTAMENTE DE INCHIRIAT

⁷ ESPON EGTC, 'Technical Guidance Document. ESPON Big Data for Territorial Analysis and Housing Dynamics. Wellbeing of European citizens regarding the affordability of housing'. ESPON EGTC, 12 November 2019. https://www.espon.eu/sites/default/files/attachments/ESPON%20Big%20Data-Guidance_Document.pdf (accessed 25/04/2022).

APARTMENT;DUPLEX	3 IZBOVĚ BYT / PRENĚJOM	INĚ BYT / PRENĚJOM	LĚ,GENHET
APARTMENT;DUPLEX;MAISONETTE	4 IZBOVĚ BYT / PREDĚJ	LOFT / PREDĚJ	VILLA APARTMENT
APARTMENT;FLAT	4 IZBOVĚ BYT / PRENĚJOM	LOFT / PRENĚJOM	CONDO
APARTMENT;FURNISHED_FLAT	5 IZBOVĚ BYT A VIAC / PREDĚJ	MEZONET / PREDĚJ	
APARTMENT;LOFT	5 IZBOVĚ BYT A VIAC / PRENĚJOM	MEZONET / PRENĚJOM	

Table 9: Multifamily Housing Typologies Taxonomy

SINGLE-FAMILY HOUSING TYPLOGIES			
MAISON	HOUSE;TERRACE_HOUSE	COTTAGE	FRITIDSBOLIG
DETACHED	HOUSE;VILLA	TERRACE_HOUSE	LANDEJENDOM / HELĚ...RSGRUND
END OF TERRACE	CHALUPA, REKREAČENĚ DOMĚČEK / PREDĚJ	VILLA / HELĚ...RSGRUND	LANDEJENDOM / LYSTJEJENDOM
HOUSE	CHALUPA, REKREAČENĚ DOMĚČEK / PRENĚJOM	LANDEJENDOM / VILLA	HAUS
HOUSES	CHATA, DREVENICA, ZRUB / PREDĚJ	VILLA / LYSTJEJENDOM	MAISONETTE
SEMI-D	CHATA, DREVENICA, ZRUB / PRENĚJOM	VILLA	CASE DE VANZARE
TERRACE	NĚJOMNĚ DOM / PREDĚJ	RĚKKEHUS	CASE DE INCHIRIAT
TOWNHOUSE	NĚJOMNĚ DOM / PRENĚJOM	LANDEJENDOM	PAR-/KEDJE-/RADHUS
HOUSE;BIFAMILIAR_HOUSE	RODINNĚ VILA / PREDĚJ	FRITIDSBOLIG / FRITIDSGRUND	RADHUS
HOUSE;CASTLE	RODINNĚ VILA / PRENĚJOM	RĚKKEHUS / FRITIDSBOLIG	FRITIDSHUS
HOUSE;CHALET	RODINNĚ DOM / PREDĚJ	VILLA / LANDEJENDOM	VINTERBONAT FRITIDSHUS
HOUSE;FARM_HOUSE	RODINNĚ DOM / PRENĚJOM	LYSTEJENDOM / FRITIDSBOLIG	ADOSADO O CHALET
HOUSE;MULTIPLE_DWELLING	VIDIECKY DOM / PREDĚJ	LYSTEJENDOM / VILLA	HOLIDAY HOUSE
HOUSE;ROW_HOUSE	VIDIECKY DOM / PRENĚJOM	FRITIDSBOLIG / VILLA	TERRACED HOUSE
HOUSE;SINGLE_HOUSE	ZĚHRADNĚ CHATA / PREDĚJ	VILLA / FRITIDSBOLIG	

Table 10: Single-family housing typologies taxonomy

- Minimum surface:**

When cleaning the data for the price/sqm analysis, a minimum surface was defined to take out invalid listings such as listings advertised as housing despite being a single room, garage, etc.

In order to establish the minimum surface, some research was conducted.

According to Apolloni et D'Alessandro (2021)⁸, a comparison of dimensional parameters in housing and living space examined standards of nine different European countries and showed that Spain was the country with the least apartment area at 20sqm minimum.

Therefore, all records with less than 20sqm were taken out from the dataset except in the case of Northern Ireland where there is no surface data available as mentioned before. For instance, in the case of Switzerland, approximately 2% of the data was removed.

6.7.1.4 Outlier detection

The last step of the raw data treatment was the statistical data cleaning. In order to properly analyze the price/sqm data, outliers needed to be detected to yield accurate results.

Some visualization of the rent and sale prices for single-family and multifamily typologies was conducted in order to see how the data was distributed.

The following histograms show an example of the data distribution for single-family typologies in Switzerland for rent and for sale:

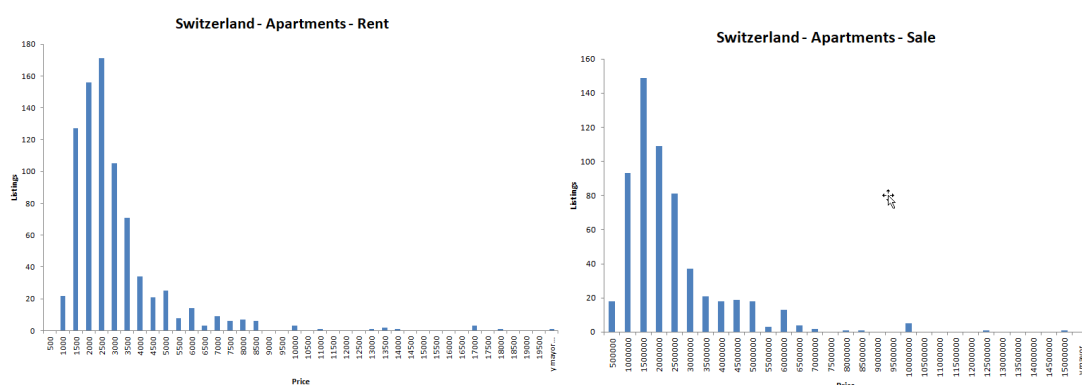


Figure 30: Histograms of Switzerland dataset (rent and sale listings for multifamily housing typologies)

As it can be observed in the histograms of Switzerland, the datasets show skewed distributions, a pattern which was similar to all case studies for each type of listing (rent/sale) and housing typology (single-family/multifamily). All datasets showed long tails on the right side. Therefore, the Interquartile Range methodology for outlier detection (IQR) was chosen and applied disaggregating the data for each case study, country, typology (rent or sale) and housing type (single-family or multifamily).

However, as the data was segregated into various groups, and histograms showed really long tails, in order to not to lose many records, the extreme outlier detection criterion (3xIQR) was applied instead of the general potential outlier detection rule (1.5xIQR)⁹. For instance, in the case of Switzerland, approximately 3% of the data was removed after applying the methodology.

6.7.1 Scraping results summary

Approximately 130.000 listings were collected from scraping in a weekly basis, with a total amount of 1.320.382 listings gathered during the whole period. From those, 463.124 were unique listings and 102.238 of those were identified as valid records to be used in the study after the data cleaning process.

The following table shows a summary of the results obtained by case study for each type (rent/sale) and each housing typology (single-family or multifamily):

CASE STUDY	COUNTRY	WEBSITE	TYPE	HOUSING TYPE	TOTAL ¹⁰	INSIDE CORRIDOR ¹¹	COMPLETE ¹²	HOUSING TYPE ¹³	VALID UNIQUE LISTINGS ¹⁴
FR-CH	France	Nestoria			266.155	266.155	216.943	210.681	14.498

			Rent						3.361
				SINGLE-FAMILY					201
				MULTIFAMILY					3.160
			Sale						11.137
				SINGLE-FAMILY					2.985
				MULTIFAMILY					8.152
FR-CH	Switzerland	Homegate			28.292	28.241	20.704	20.684	6.457
			Rent						3.121
				SINGLE-FAMILY					309
				MULTIFAMILY					2.812
			Sale						3.336
				SINGLE-FAMILY					1.411
				MULTIFAMILY					1.925
DK-SE	Denmark	Boligsiden			104.451	26.099	26.049	25.321	8.796
			Rent						1.412
				SINGLE-FAMILY					45
				MULTIFAMILY					1.367
			Sale						7.384
				SINGLE-FAMILY					2.712
				MULTIFAMILY					4.672
DK-SE	Sweden	Blocket + Hemnet¹⁵			33.876	14.548	13.921	13.335	4.365
		Blocket	Rent						1.113
				SINGLE-FAMILY					224
				MULTIFAMILY					889

⁸ Appolloni, L. & D'Alessandro, D. (2021). Housing Spaces in Nine European Countries: A Comparison of Dimensional Requirements. *International journal of environmental research and public health*, 18(8), 4278. <https://doi.org/10.3390/ijerph18084278>

⁹ University of Florida, 'Biostatistics Open Learning Textbook', University of Florida [website], <https://bolt.mph.ufl.edu/6050-6052/unit-1/one-quantitative-variable-introduction/understanding-outliers/>, (accessed 25 April 2022).

¹⁰ Total listings gathered from scraping.

¹¹ Total listings inside the area of study.

¹² Total listings with complete information.

¹³ Total listings considering only housing typologies.

¹⁴ Total unique listings remaining after data cleaning, normalization and enrichment. In the case of Nestoria, the values show the result after choosing the advertiser with the biggest amount of listings per LAU and typology to avoid duplicates.

¹⁵ In this case study, Blocket platform was used for rent and Hemnet for sale.

		Hemnet	Sale						3.252
				SINGLE-FAMILY					960
				MULTIFAMILY					2.292
IE-UK	Ireland	Daft			182.820	42.289	41.631	38.584	9.502
			Rent						2.603
				SINGLE-FAMILY					788
				MULTIFAMILY					1.815
			Sale						6.899
				SINGLE-FAMILY					4.857
				MULTIFAMILY					2.042
IE-UK	Northern Ireland	Nestoria			25.195	15.663	15.553	14.867	3.048
			Rent						1.641
				SINGLE-FAMILY					934
				MULTIFAMILY					707
			Sale						1.407
				SINGLE-FAMILY					1.180
				MULTIFAMILY					227
AT-SK	Austria	Topreality + Nestoria			318.369	303.777	237.335	216.129	28.957
			Rent						18.691
				SINGLE-FAMILY					314
				MULTIFAMILY					18.377
			Sale						10.266
				SINGLE-FAMILY					1.016
				MULTIFAMILY					9.250
AT-SK	Slovakia	Topreality			126.721	126.536	106.919	70.439	18.555
			Rent						8.202
				SINGLE-FAMILY					512
				MULTIFAMILY					7.690
			Sale						10.353
				SINGLE-FAMILY					2.688
				MULTIFAMILY					7.665
RO-BG	Romania	Publi24			11.147	1.908	1.402	902	169
			Rent						32
				SINGLE-FAMILY					6
				MULTIFAMILY					26
			Sale						137
				SINGLE-FAMILY					66
				MULTIFAMILY					71
RO-BG	Bulgaria	Realstates			6.057	6.057	6.013	6.013	1.272
			Rent						576
				SINGLE-FAMILY					0
				MULTIFAMILY					576
			Sale						696

				SINGLE-FAMILY					30
				MULTIFAMILY					666
SP-FR	Spain	Nestoria			42.776	33.662	31.425	28.246	3.728
			Rent						578
				SINGLE-FAMILY					7
				MULTIFAMILY					571
			Sale						3.150
				SINGLE-FAMILY					392
				MULTIFAMILY					2.758
SP-FR	France	Nestoria			62.181	62.181	47.564	46.330	2.891
			Rent						140
				SINGLE-FAMILY					30
				MULTIFAMILY					110
			Sale						2.751
				SINGLE-FAMILY					598
				MULTIFAMILY					2.153

Table 11: Scraping Results Summary

6.8 Next Steps for Further Studies

In this chapter we propose future steps, ideas and opportunities regarding scraping methodological aspects that could be improved in further studies.

6.8.1 Duplicity of listings

When carrying out scraping of real estate offers, it needs to be taken into account that some adverts can be published more than once inside the same website or in different ones.

It is a complex task to identify which advertisements are published various times in the same platform. It needs to be remarked that for the scope of this study these kinds of duplicities were not identified.

Moreover, when scraping a portal like Nestoria, which is an aggregator of data from different portals, adverts can appear more than once because they could be published in more than one website. In order to avoid these kinds of duplicities, in this study, the name of the advertiser was gathered as an indicator during the crawling, and a decision was made to avoid these kinds of duplicities by selecting the advertiser with the biggest amount of adverts for each LAU and each typology (rent and sale). However, it needs to be remarked that the Nestoria portal was chosen because it is a platform that is common to various case studies, so it facilitated the homogeneity of data among them and also speeded up the crawling process.

In addition, for the case of Spain, a comparison of the number of listings coming directly from Fotocasa website and those coming from Fotocasa advertiser inside Nestoria was carried out as inAtlas has Fotocasa data available for Spain. That comparison allowed us to analyse the behavior of Nestoria as an aggregator platform. The results showed that Nestoria does not gather all the data published in Fotocasa portal:

SOURCE	LISTING TYPE	CORRIDOR LISTINGS	CORE CORRIDOR LISTINGS
NESTORIA	TOTAL	424	122
	RENT	0	0
	SALE	424	122

FOTOCASA	TOTAL	1.660	362
	RENT	1.408	234
	SALE	341	14

Table 12: Spain Fotocasa Results Comparison (sample of 26-03-2022)

A recommendation for future research to improve the analysis would be to design an algorithm to identify duplicated listings inside the same platform but also inside aggregator platforms such as Nestoria. Aggregators are a good choice in order to keep the biggest amount of offers possible coming from different portals in order to have a more accurate picture of the market offer. In order to design this algorithm, more listing indicators should be gathered during the scraping process in order to be able to identify these duplicated adverts.

6.8.2 Length of advertisements

The scraping process was launched on a weekly basis in order to be able to analyze the length of the advertisements over time.

To calculate how long an advert had been published, the listing ID was used to identify the same advert in different samples. After that, a calculation of the number of samples (weeks) each listing appeared gave us the length value.

It needs to be taken into account that the scraping was conducted for only few weeks, so the length of advertisements which appeared and disappeared during that period of time could be easily calculated. However, the length of those which were gathered in the first sample of scraping (that means they were already in the market) and those gathered in the last samples (which still remain in the market) could not be calculated. Nevertheless, the results are still relevant, as they show if adverts remain for a long or short period of time.

A recommendation for future research would be to carry out the scraping for a longer period of time in order to have more detailed results.

6.8.3 Scraping monitoring

This project prepared crawling scripts for 15 different platforms which were launched automatically on a weekly basis. The results obtained were periodically checked in order to see if the crawling was correctly working. However, for the week of 5 May 2022 some of the scrapings failed and only data from two platforms could be gathered.

A recommendation for future studies is to set an automatic way to monitor the crawling processes in order to be able to immediately detect errors and correct them if necessary so as to not lose any data.

7 Bringing It All Together: Approach to Linking Different Information Sources and Work Processes

In this chapter, we outline the process of combining the work packages in order to deliver a final product.



Figure 31: Project Overview Graphic

In commencing this research, the project team held several meetings and engaged in extensive discussions in order to select the case study areas, to delineate the territorial scope and to identify the indicators which will be updated and developed for the project. When the five case study regions were selected in addition to the Geneva Metropolitan Region, it was essential to select what parts of the region to include in the case study. Further explanation and narratives on how and why the territorial scopes were selected is included within each case study's annex. Furthermore, the team collaborated at an early stage as to what indicators would be possible to update and analyse along with discussing what further information could be extrapolated from the data and developed into new indicators used specifically to measure cross-border housing markets.

In parallel with selecting the territorial scopes and identifying the harmonised indicators, the Cambridge team began research for each case study's policy background. As there was a vast array of housing policies

documented across all of the case study regions, we developed overviews which listed and explained pertinent policies while also connecting the policy contexts to the housing dynamics of each cross-border market. Furthermore, semi-structured interviews were carried out with key individuals from different parts of each case study region. Despite initial difficulties receiving responses for interview, eight interviews were carried out covering three of the case study regions. In carrying out these interviews, we developed a stronger understanding of the local dynamics within each case study region while we were also able to add further depth of understanding to policy contexts. Below is a table which provides an overview of the interviews conducted.

Representative	Region	Key Points
Representative from Oresund Institute	Greater Copenhagen	<ul style="list-style-type: none"> - Difficulties crossing the border since 2015 have discouraged some cross-border migration due to longer commuting times, border controls and more regulation - Different estate systems exist between Denmark and Sweden which makes integration and coordination more difficult - Lots of new supply from new apartment developments has helped normalise rising prices – predominance for private rental sector though - Political developments across the region are not followed cohesively – establishment of an Oresund Region news agency to interrelate both sides - Price is not the most important motivating factor in people migrating – nature, amenities and differences in law all play an important role
Representative from Finance Administration of City of Copenhagen	Denmark	<ul style="list-style-type: none"> - 2015 Legislation stipulating that 25% of new residential development must be reserved for the non-profit affordable housing sector - Issues with supply: new supply is mostly private rental sector (due to government tax subsidy) which is pushing up owner occupied prices - 24% difference in prices between the same apartment from the private rental sector and the non-profit affordable housing sector - Issues regarding the lack of regulation and coordination of new housing output at the regional level – there is no instrument to unify new development proposals across the region - City housing policies: mostly focused on creating new land for development along with densifying city through renewal
Representative from Architecture Department at City of Malmo	Sweden	<ul style="list-style-type: none"> - Reasons for price differences dependent on which sector: prices within the owner-occupied sector would be more impacted by accessibility to transport links to Denmark while prices within the private rental sector are more impacted by the age of the building/renovations/amenities/etc - There was a survey completed on incomes in relation to the cost of flats (In Swedish though) - Swedish market is more 'accessible' with cheaper flats and easier access to getting a flat – navigating the Swedish market is less difficult - Housing policies focused on achieving more equitable balance across housing tenures – currently a preponderance for owner occupied - Very little collaboration exists on housing across the strait, previous coordination of comprehensive municipal plans have been resultant of political leaders sharing same political parties – ie politics drives a lot of work in this
Representatives with the Northern Ireland Executive	Northern Ireland	<ul style="list-style-type: none"> - A lot of focus in Northern Ireland is on the provision of public socialised housing - Currently publishing a new Housing Supply Strategy which approaches the housing sector with a whole system approach across all tenures

		<ul style="list-style-type: none"> - Intermediate Rent Sector – a new & secured form of private rental sector - Lack of a lot of cross-border data available - Strategic Housing Market Accessibilities currently being updated for new Housing Supply Strategy - Newry stands out as an anomaly across different tenures within the housing system – issues with private market stress, growth pressures, high demand within the private rental sector, high demand for social housing – also existence of geographic constraints for new development in Newry (town built into granite rock on a hill)
Representative from Louth County Council Services Directorate	Ireland	<ul style="list-style-type: none"> - 4,000 social housing units are managed by LouthCoCo - Issues relating to SGIs specifically in relation to medical services in the North (access to NHS medical care) - Acknowledgement of households living in North for better cost of living but working in the South - Border is not seen as a 'border' as communities are strongly interlinked but different housing systems are still strong and separate the two markets - Housing policy directed by national government in Ireland – centralised decision making limits LA abilities
Representatives from Annemasse	France (Geneva)	<ul style="list-style-type: none"> - Local housing market: new development is small, expensive, built to Swiss incomes - French salaries are unable to compete in Annemasse requiring affordable housing help from the government - More construction and development on the French side of the border due to less restraints on building new supply - State imposed quota of 25% social housing on new development with fines and annual observation if the local authority does not meet quota - High income inequality (fourth largest in France) - Within private market, 20% price increase in Annemasse over last 6 years - Challenges to implement local housing policies as centralisation of policymaking does not meet needs of region - New innovations within social/affordable housing sector
Representatives from Grand Geneva	Grand Geneva	<ul style="list-style-type: none"> - Disagreement with map perimeters, should be cohesive with one of their boundaries - INSEE – data access to cross border workers - Collaboration amongst institutions is sought for - Grand Geneva has competency over interregional mobility and ecological issues but no competency over housing except to push for higher densification of the region
Representatives from Canton of Geneva	Switzerland	<ul style="list-style-type: none"> - Missing information of cross-border dynamics due to households refusing to declare to authorities - Flat ownership (PPE) is a new concept to Geneva commencing in the 1970s - Discussed the municipal preview in planning decisions - Benefits from tax allocations contingent on cross-border cooperation which promotes policy coordination - Strong control of housing pricing from the State - New legislation seeks to promote new housing development while also seeking to protect against any further environmental degradation

Table 13: Interview Overviews

While research collection continued for the housing policy contexts of each region, once the territorial parameters and harmonised indicators were selected, the inAtlas team could commence the data scrapping process. Over the course of several weeks, we ran the data scrapping software in order to ascertain a complete picture of each case study's cross-border housing market. Chapter 6 delves further into the

methodologies used the scrapping of big data and harvesting of information from the different cross-border markets.

After the data collected from the web scrapping processes were cleaned and harmonised, the data was then sent to MCRIT to be incorporated into the selected indicators chosen at the beginning of the project. The methodologies on how we turned the data into spatial outputs have been included in Chapters 3 and 5. To summarise the process, we incorporated data from the web scrapping with data from official sources of data (ie Eurostat) in order to produce indicators on housing affordability in relation to regional differences in income across the different case study regions.

Following the mapping and visualisation exercises, we could analyse and assess the data through a series of different maps and graphs which visualise the indicators and allow for comparison within and between the chosen cross-border functional areas. Then using the policy context from the desktop research study, we could interconnect policy objectives with on-the-ground information. This allows us to consider how housing policies designed to encourage the development of affordable housing and an equitable market are impacting housing prices in regions with substantial cross-border flows. Moreover, it allows us to assess how effective policymakers have been in their attempts to legislate for affordable housing practices across different housing markets where unique forms of tenure and differences in homeownership can impact how housing is delivered. Thus, by incorporating information from the housing policy overview into our assessment of the measured indicators, we draw conclusions on the dynamics of housing prices in cross-border housing markets. Such conclusions can illustrate how further policy coordination and cross-border collaboration on housing policies in cross-border regions can strengthen social cohesion and ensure equity within a cross-border housing market.

References

- Aalbers, M. B., 2016. *The Financialization of Housing*. 1st Edition ed. New York: Routledge.
- Antalovsky, E. & Löw, J., 2019. *Why Vienna Gets High Marks?*, Luxembourg: EIB.
- Aris, C. & Panayotis, C., 2019. *JRC Publications Repository - Measuring cross-border road accessibility in the European Union*. [Online]
Available at: <https://publications.jrc.ec.europa.eu/repository/handle/JRC117094>
[Accessed 25 11 2021].
- Aris, C. & Panayotis, C., 2019. Measuring cross-border road accessibility in the European Union. *Sustainability*, 11(15), p. 4000.
- Blair, N., Corrigan, J., Magennis, E. & Broin, D. Ó., 2021. *Dublin-Belfast Economic Corridor: Current Profile, Potential for Recovery & Opportunities for Collaboration*, Belfast: DBEC.
- Bradley, S., 2018. Geneva's property shortage continues to fuel exodus. *Swissinfo.ch*, 5 Sept, p. Online.
- Butler, T. & Lees, L., 2006. Supergentrification in Barnsbury, London: globalisation and gentrifying global elites at the neighbourhood level.. *Transactions of the Institute of British Geographers*, Volume 31, pp. 467-487.
- Chilla, T. & Heugel, A., 2022. Cross-border Commuting Dynamics: Patterns and Driving Forces in the Alpine Macro-region. *Journal of borderlands studies*, 37(1), pp. 17-35.
- Chilla, T. & Sielker, F., 2016. *Leitfaden Wohnraumprognose: Lehrforschung: Regionalentwicklung - Konzepte der Wohnraumentwicklung*, Erlangen: FAU.
- Christmann, N., 2015. *Dwelling trends in border regions - towards an inter-urban discourse analysis*. Lisbon, European Network for Housing Research.
- Christophers, B., 2021. Mind the rent gap: Blackstone, housing investment and the reordering of urban rent surfaces. *Urban Studies*, 59(4), pp. 698-716.
- City of Vienna, 2021. *Vienna - a regional authority*, Vienna: City of Vienna.
- Clarke, S., Levitt, D. & Gutiérrez, P., 2019. This is the Irish Border, Monday 7am. *The Guardian*, 3 Sept, p. Online.
- Claveres, G. et al., 2020. *Housing and inequality: The case of Luxembourg and its cross-border workers*, Paris: OECD.
- Dascher, K. & Haupt, A., 2011. The political economy of regional integration projects at borders where poor and rich meet: The role of cross-border shopping and community sorting. *Journal of urban economics*, 69(1), pp. 148-164.
- Decoville, A. & Durand, F., 2019. Exploring cross-border integration in Europe: How do populations cross borders and perceive their neighbours?. *European urban and regional studies*, 26(2), pp. 134-157.
- Decoville, A., Durand, F., Sohn, C. & Walther, O., 2013. Comparing Cross-border Metropolitan Integration in Europe: Towards a Functional Typology. *Journal of Borderland Studies*, 28(2), pp. 221-237.
- Dillinger, T., 2004. Facing suburbanisation in the area of Bratislava: The border between Austria and the Slovak Republic. *Urbani Izziv*, 152(2), pp. 121-124.
- Doval Tedin, M. J. & Faubert, V., 2020. *Housing Affordability in Ireland Economic Brief 061*, Brussels: European Commission.
- Dumeignil, C., 2022. The Impact of Cross-border Labor Mobility on Real Estate Price Trends: A Natural Experiment. *International Regional Science Review*, 45(1), pp. 108-132.
- Durand, F. & Decoville, A., 2020. A multidimensional measurement of the integration between European border regions. *Journal of European Integration*, 42(2), pp. 163-178.

- European Commission, 2017. *COMMUNICATION FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT Boosting growth and cohesion in EU border regions {SWD(2017) 307 final}*, Luxembourg: s.n.
- European Parliament, 1996. *Overview of Housing Policies*, Strasbourg: European Union.
- European Parliament, 2021a. *Parliament calls for action to solve housing crisis*, Strasbourg: European Union.
- European Parliament, 2021b. *Decent and affordable housing for all*, Strasbourg: European Union.
- Farha, L., 2017. *Report of the Special Rapporteur on adequate housing as a component of the right to an adequate standard of living, and on the right to non-discrimination in this context*, New York: United Nations Human Rights Council.
- Federal Ministry of Digital and Economic Affairs AT, 2021. *Housing policy*, Vienna: Government of Austria.
- Fricke, C., 2015. Spatial Governance across Borders Revisited: Organizational Forms and Spatial Planning in Metropolitan Cross-border Regions. *European planning studies*, 23(5), pp. 849-870.
- Fullerton, T. M. & Villalobos, E., 2011. Street Widths, International Ports of Entry and Border Region Housing Values. *Journal of Economic Issues*, 45(2), pp. 493-510.
- Gkartzios, M. & Scott, M., 2009. Planning for Rural Housing in the Republic of Ireland: From National Spatial Strategies to Development Plans. *European Planning Studies*, 17(12), pp. 1751-1780.
- Government Offices of Sweden, 2021. *The Swedish model of government administration*, Stockholm: Government of Sweden.
- Govt of France, 2017. *The Government's housing strategy*, Paris: Government of France.
- Grand Geneve, 2021. *Historical Timeline of Grand Geneve*, Geneva: Grand Geneve.
- Greater Copenhagen, 2021. *About Greater Copenhagen*, Copenhagen: Greater Copenhagen Secretariat.
- Hanzl-Weiss, D., Holzner, M. & Römisch, R., 2018. *BRATISLAVA and VIENNA: Twin Cities with big Potentials*, Vienna: The Vienna Institute for International Economic Studies.
- Heider, B., 2018. The impact of EU Eastern enlargement on urban growth and decline: New insights from Germany's Eastern border. *Papers in Regional Science*, 98(3), pp. 1443-1468.
- Heinrich, R., 2019. *Vienna calling — An assessment of the Austrian housing market*, Vienna: Ren & Heinrich.
- Hilber, C. & Schöni, O., 2016. Housing policies in Switzerland, the United Kingdom, and the United States. In: N. Yoshino & M. Helble, eds. *The Housing Challenge in Emerging Asia : Options and Solutions*. Tokyo: Asian Development Bank Institute, pp. 210-259.
- Hohmann, J., 2013. *The Right to Housing: Law, Concepts, Possibilities*. London: Bloomsbury Publishing.
- Housing 2030, 2022. *Slovakian State Housing Policy*, Brussels: Housing Europe.
- IIBW, 2022. *Housing Policy*, Vienna: IIBW.
- Jaššo, M., Leibenath, M., Korcelli-Olejniczak, E. K. & Knippschild, R., 2008. Cross-border cooperation challenges: Positioning the Vienna-Bratislava region. In: *Cross-Border Governance and Sustainable Development*. New York City: Springer, pp. 87-100.
- Kolossov, V. & Scott, J., 2013. Selected conceptual issues in border studies. *Belgian Journal of Geography*, Volume 1, p. Online.
- Kuben Management & Copenhagen Economics, 2018. *HOUSING MARKET ANALYSIS OF GREATER COPENHAGEN*, Copenhagen: Danish Ministry of Transport, Buildings & Housing.
- Lennon, M. & Waldron, R., 2019. De-democratising the Irish planning system. *European planning studies*, 27(8), pp. 1607-1625.
- Lund, B., 2011. *Understanding Housing Policy*. Bristol: Policy Press.

- Marcuse, P. & Madden, D., 2016. *In Defense of Housing: The Politics of Crisis*. New York: Verso.
- Micheli, M., Rouwendal, J. & Dekkers, J., 2019. Border Effects in House Prices. *Real estate economics*, 47(3), pp. 757-783.
- migration.gv.at, 2021. *The political, administrative and legal systems*, Austria: Government of Austria.
- Ministry of the Interior & Housing, 2021. *About municipalities & regions*, Copenhagen: Government of Denmark.
- Mitrică, B., Mocanu, I., Dumitrașcu, M. & Grigorescu, I., 2016. Socio-Economic Disparities in the Development of the Romania's Border Areas. *Social Indicators Research*, 134(3), pp. 899-916.
- Møller, I. H. & Larsen, J. E., 2015. The Socioeconomic and Ethnic Segregation of Living Conditions in Copenhagen. *Centro de Estudos Sociais da Universidade de Coimbra Revista crítica de ciencias sociais*, Volume 108, pp. 7-30.
- Municipality of Giurgiu, 2021. *Municipality of Giurgiu Departments*, Giurgiu: Giurgiu Office of Mayor.
- Municipality of Ruse, 2021. *Municipality of Ruse Government Structure*, Ruse: Ruse City Government .
- NI Direct, 2021. *Property & Housing*, Belfast: Northern Ireland Executive.
- O'Dea, C., 2020. Switzerland is putting its housing crisis to a referendum. *The Irish Times*, 29 Jan, p. Online edition.
- OECD, 2017. *The Governance of Land Use Country Fact Sheet Austria*, Paris: OECD.
- OECD, 2017. *The Governance of Land Use Country Fact Sheet Slovak Republic*, Paris: OECD.
- OHCHR, 2009. *The Right to Adequate Housing*, Geneva: United Nations.
- Presence Switzerland, 2021. *Political system*, Bern: Swiss Confederation.
- Raymond, E. et al., 2016. *Corporate Landlords, Institutional Investors, and Displacement: Eviction Rates in Single Family Rents*, Atlanta: Federal Reserve Bank of Atlanta.
- Rolnik, R., 2019. *Urban Warfare: Housing Under the Empire of Finance*. New York: Verso.
- RSM Consulting, 2019. *Brexit and the Housing Market*, Belfast: Northern Ireland Housing Executive.
- Saadi, I., Boussauw, K., Teller, J. & Cools, M., 2016. Trends in regional jobs-housing proximity based on the minimum commute: The case of Belgium. *Journal of transport geography*, Volume 57, pp. 171-183.
- Schaerer, C. & Baranzini, A., 2009. Where and How Do Swiss and Foreigners Live? Segregation in the Geneva and Zurich Housing Markets. *Swiss Journal of Sociology*, 35(3), pp. 571-592.
- Scharmann, L., 2020. *Introduction to the Austrian Spatial Planning System*, Leipzig: Interreg.
- Schwartz, H. & Seabrooke, L., 2009. *The Politics of Housing Booms and Busts*. London: Palgrave Macmillan.
- Sharma, R., 2021. The rise of eastern Europe is a forgotten economic success story. *Financial Times*, 13 Sept , p. Online edition.
- Shucksmith, M. & Sturzaker, J., 2016. Rural Housing. In: S. J. Smith, ed. *International Encyclopedia of Housing and Home*. Amsterdam: Elsevier Ltd, pp. 204-209.
- Sohn, C., 2013. The Border as a Resource in the Global Urban Space: A Contribution to the Cross-Border Metropolis Hypothesis. *International Journal of Urban and Regional Research*, 38(5), pp. 1697-1711.
- Sohn, C., Reitel, B. & Walther, O., 2009. Cross-Border Metropolitan Integration in Europe: The Case of Luxembourg, Basel, and Geneva. *Environment and planning. C, Government & policy*, 27(5), pp. 922-939.
- Somerville, P. & Sprigings, N., 2005. *Housing and Social Policy: Contemporary Themes and Critical Perspectives*. Milton Park: Routledge.
- Stadt Wien, 2016. *Housing in Vienna Annual Report 2016*, Vienna: Geschäftsgruppe Wohnen, Wohnbau und Stadterneuerung.

Stamsø, M.-A., 2010. Housing and Welfare Policy - Changing Relations? A Cross-National Comparison. *Housing, Theory & Society*, 27(1), pp. 64-75.

Staré Mesto Bratislava, 2021. *Statutes, Powers and Authorities*, Bratislava: The Local Authority of the City District of Bratislava-Staré Mesto.

Statistical Office of the SR, 2022. *Bratislavský kraj - Characteristic of the region*, Bratislava: Statistical Office of the Slovak Republic.

Stevenson, S., 2004. House price diffusion and inter-regional and cross-border house price dynamics. *Journal of property research*, 21(4), pp. 301-320.

Stirling, T., 2020. *A glimpse of Copenhagen and Malmö*, Cardiff: Welsh Housing Quarterly.

Stoychev, K., Klisarova, M. & Zarkov, V., 2016. *Crossing the Borders: Studies of Cross-Border Cooperation within the Danube Region*, Sofia: Central European Service for Cross-Border Initiatives.

Thinggaard Hetland, S., Pedersen, J., Juul Hviid, S. & Schmith, A. M., 2021. *Housing Market Robustness Should Be Strengthened*, Copenhagen: Danmarks Nationalbank.

White, M. & Allmendinger, P., 2003. Land-use Planning and the Housing Market: A Comparative Review of the UK and the USA. *Urban Studies*, 40(5/6), pp. 953-972.

Wittowsky, D., Hoekveld, J., Welsch, J. & Steier, M., 2020. Residential housing prices: impact of housing characteristics, accessibility and neighbouring apartments – a case study of Dortmund, Germany. *Urban, Planning and Transport Research*, 8(1), pp. 44-70.

Articles from Webscrapping :

Appolloni, L. & D'Alessandro, D. (2021). Housing Spaces in Nine European Countries: A Comparison of Dimensional Requirements. *International journal of environmental research and public health*, 18(8), 4278. <https://doi.org/10.3390/ijerph18084278>

Kaliyaperumal, Senthamarai & Kuppusamy, Manoj & Arumugam, S.. (2015). Labeling Methods for Identifying Outliers. *International Journal of Statistics and Systems*. 10. 231-238.

Norling, M. & Hjulfors, S. (2017), Bidding strategies in the Swedish housing market, DiVA. Accessible here: <http://www.diva-portal.org/smash/get/diva2:1093030/FULLTEXT01.pdf>.

Wienand, D., Paulheim, H. (2014). Detecting Incorrect Numerical Data in DBpedia. In: Presutti, V., d'Amato, C., Gandon, F., d'Aquin, M., Staab, S., Tordai, A. (eds) *The Semantic Web: Trends and Challenges*. ESWC 2014. Lecture Notes in Computer Science, vol 8465. Springer, Cham. https://doi.org/10.1007/978-3-319-07443-6_34

Websites from Webscrapping:

A Medium Corporation, 'Housing rent prices and venues data analysis of London - A geographical clustering approach for house seekers', *A Medium Corporation* [website], Analytics Vidhya, 9 December 2019, <https://medium.com/analytics-vidhya/housing-rent-prices-and-venues-data-analysis-of-london-a-geographical-clustering-approach-for-3abbf999b02>, (accessed 25 April 2022).

Greater London Authority, 'Housing space standards', London, Greater London Authority, August 2006,

ESPON 2020

ESPON EGTC
4 rue Erasme, L-1468 Luxembourg
Grand Duchy of Luxembourg
Phone: +352 20 600 280
Email: info@espon.eu
www.espon.eu

The Single Operation within the programme is implemented by the ESPON EGTC and co-financed by the European Regional Development Fund, the EU Member States, the United Kingdom and the Partner States, Iceland, Liechtenstein, Norway and Switzerland.

Disclaimer

This delivery does not necessarily reflect the opinion of the members of the ESPON 2020 Monitoring Committee.