Analyzing the Impact of Immigration and Emigration on Local Rent Prices and Homelessness: A Case Study

*Note: This paper studies the relationship between immigration and emigration data and local rent prices and homelessness rates in a specific area.

1st Debmalya Deb
Msc in Data Analytics
National College of Ireland
x21242101@student.ncirl.ie

2nd Giorgia Luzia Pscheidt
Msc in Data Analytics
National College of Ireland
x22184261@student.ncirl.ie

3rdSaheli Dutta
Msc in Data Analytics
National College of Ireland
x21246513@student.ncirl.ie

***Abstract*—This study examines the correlation between immigration and emigration data and local rent prices and homelessness rates in Ireland. The research was conducted by a team of three individuals who gathered data on immigration and emigration, homelessness, and rent prices from online sources. The data was cleaned using the Python programming language, and an ETL (Extract, Transform, Load) process was implemented with Dagster to transfer the data to a PostgreSQL database. The data was merged using SQL queries, and Plotly was used to create visualizations of the data in CSV format. The primary goal of this research was to gain insights into how immigration and emigration patterns influence housing affordability and the cost of renting properties. The results of the study reveal that immigration and emigration can have a significant impact on local rent prices and homelessness rates from the year 2020 to 2021. This project illustrates the significance of utilizing data-driven approaches to address multifaceted social issues like housing affordability and homelessness.**

***Index Terms*—Data cleaning, Python, MongoDB, ETL, PostgreSQL, Data visualization, housing affordability, immigration**

I. Introduction

Housing affordability and homelessness are complex social issues that continue to pose significant challenges for many communities around the world. The ability to access safe, stable, and affordable housing is fundamental to individuals' well-being and to a thriving society. In recent years, there has been growing interest in understanding the impact of immigration and emigration patterns on local rent prices and homelessness rates. However, despite the significant attention given to these issues, there is still a need for more comprehensive and data-driven analyses that can provide insights and solutions.

In this project, we aimed to address this need by conducting a case study to examine the relationship between immigration and emigration data, local rent prices, and homelessness rates in Ireland. Our objectives were to collect and clean data on these variables from web sources, transfer the data to a database using ETL processes, merge the data using SQL queries, and visualize the data using Plotly to gain insights into the relationship between immigration and emigration patterns, housing affordability, and the price of rented properties. The proposed analysis aimed to answer a novel question that would advance our understanding of the factors that contribute to housing affordability and homelessness.

Our results suggest that immigration and emigration can have a significant impact on local rent prices and homelessness rates over the years 2020 to 2021. Overall, this project demonstrates the value of using data-driven approaches to address complex social issues like housing affordability and homelessness. The findings from this study could help inform policy decisions and interventions aimed at improving housing affordability and reducing homelessness rates in communities around the world.

II. Related Work

A. The Impact of Immigration and Emigration on the Housing Crisis: A Look at the Scale of the Problem

Immigration and emigration impact societal issues like housing, schools, jobs, wages, hospitals, and welfare spending. It's crucial to manage immigration levels and integrate immigrants successfully. A report by the Economic and Social Research Institute shows Ireland is doing well in integration, but challenges remain. Immigrants have high employment rates and live in the private rental sector, which affects rental property availability and cost. [1]

B. A Comprehensive Social and Economic Analysis

A report commissioned by Ireland's National Economic and Social Council suggests that immigration has benefited Ireland, but a long-term strategy including integration policies is needed. The report advocates for the government to lead a shared understanding of migration's role in the country's future. Ireland's migration policies were employer-driven, but the report recommends factoring migration into government

policies. Ireland has become an immigration destination, attracting skilled and unskilled workers who have contributed to economic growth. [2]

C. Extending the Analysis: A Further Report Based on Early Research Findings

This report expands on past research on immigration and emigration in Ireland and examines the housing crisis's correlation with homelessness, as well as the influx of immigrants from different regions. To conduct this study, data is collected from three different sources through two semi-structured and one structured dataset. The study utilizes MongoDB, PostgreSQL, and Pg Admin, along with programming languages such as Python and scripts such as SQL and NoSQL, to convert the three datasets into a single source via the Extract Transform and Load method. Finally, the report utilizes various visualization techniques to enhance the reader's understanding and engagement with the findings.

III. Methodology

A. Datasets Overview

As part of the reporting requirements, two semi-structured datasets and one structured dataset are to be collected and merged into a single collection of data through the use of analytical programming. Once this has been accomplished, visualizations will be performed to effectively communicate the findings of the study.

	Name	Meaning
1	id	The object id or reference id
2	Year	The year of the reported data
3	Month	The month of the reported data
4	Region	County in the region of Ireland
5	Total Adults	N. of total homeless adults
6	Male Adults	N. of total homeless male adults
7	Female Adults	N. of total homeless female adults
8	Adults Aged 18-24	The age group between 18 to 24
9	Adults Aged 25-44	The age group between 25 to 44
10	Adults Aged 45-64	The age group between 45 to 64
11	Adults Aged 65+	The age group on or above 65
12	Number of people who accessed Private Emergency Accommodation	People are being helped with Private Emergency Accommodation
13	Number of people who accessed Supported Temporary Accommodation	People are being helped with Supported Temporary Accommodation
14	Number of people who accessed Temporary Emergency Accommodation	People are being helped with Temporary Emergency Accommodation
15	Number of people who accessed Other Accommodation	People are being helped with Other Accommodation
16	Number of Families	Families are being affected by homeless situation
17	Number of Adults in Families	Count of adults in families
18	Number of Dependants in Families	Count of people who do not earn in a homeless family

Fig. 1. Columns Dataset Homeless

1) Homeless Dataset: The first semi-structured dataset, "homeless.json," is formatted in JSON and has been chosen due to its ability to provide insight into the correlation between rising housing prices and homelessness rates (Fig 1). [3]

2) Immigration and Emigration Dataset: The second dataset used in this study is a structured CSV file containing information on immigration and emigration. This dataset is crucial as it directly supports the analysis of the significant increase in incoming immigration in recent years (Fig. 2). [3]

	Name	Meaning
1	id	The object id or reference id
2	File Name	Case history file name of the migrants
3	File Label	The label name of the migrant's files
4	Year Clone	Year
5	Year	N. of total homeless adults
6	Gender Number	Count of male and female
7	Gender	Gender do the people belong to
8	Nationality Key	Nationality code
9	Nationality	The actual nationality
10	Coin Unit	The unit of the value below
11	Value	Total number of immigrants or emigrants

Fig. 2. Columns Dataset Immigration and Emigration

3) Rent Dataset: The final dataset considered in this study is in JSON format and contains structured data. The significance of this dataset lies in its depiction of the ultimate impact of immigration in the region, namely the increase in rents for different property types and locations (Fig 3). [3]

	Name	Meaning
1	id	The object id or reference id
2	Year	Year when the study was recorded
3	Number of Bedrooms	Count of bedrooms in a property
4	Property Type	The type accomodation
5	Location	The county areas of the region
6	Coin Unit	Currency name
7	Cost	The cost of the property

Fig. 3. Columns Dataset Rent

B. Data cleaning

Data cleaning and validation are crucial steps in the development of any study. They help identify and eliminate errors, inconsistencies, and inaccuracies in the chosen datasets, ensuring that the data used for analysis is accurate and reliable.

- Homeless Dataset, the column names is verified to be loaded correctly, and checks are conducted for duplicate entries or missing data. The data types are analyzed to ensure that values such as integers and float numbers are accurately represented. Fortunately, this dataset has minimal issues to resolve.
- Immigration and Emigration Dataset, data types are analyzed to ensure that numbers are accurately represented, and duplicate entries or missing data are checked. Column names are changed to avoid confusion when working with

the data. After data validation, irrelevant data values are filtered out to facilitate subsequent analysis and minimize the file size.

- Rent Dataset, checks are conducted for duplicate entries or missing data. Unfortunately, 167,444 missing values were identified, and the corresponding rows have to be removed to ensure accurate analysis. The data type of the rent and year columns is changed to ensure that values are accurately represented, and column names are changed. After data validation, irrelevant data values are filtered out to facilitate subsequent analysis.

C. Importing Datasets to MongoDB

MongoDB is a popular NoSQL document-oriented database program that uses JSON-like documents with optional schemas. It is widely used in modern web applications due to its flexibility, scalability, and ease of use. The study uses the MongoDB compass Version 1.35.0 (1.35.0) as it shows in Fig.4 . [4]

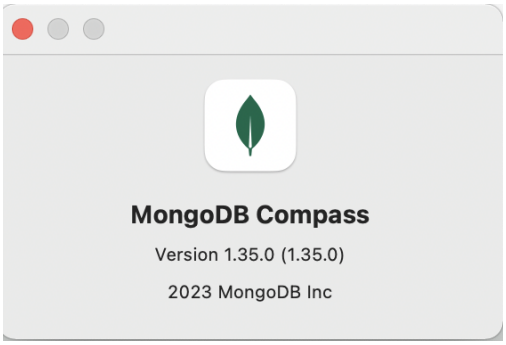

Fig. 4. MongoDB Version

1) Connection Information:

- This case study uses Docker Containers to initiate a database in MongoDB using a Docker Instance. Docker is an open-source platform that isolates applications from infrastructure, allowing for faster software delivery. It manages infrastructure similar to applications, reducing the time between coding and deployment. Docker's techniques can help reduce the time it takes to move code from development to production. The analysis uses Docker Desktop Version 4.16.2 (95914) as it shows in Fig. 5. [5]

Fig. 5. Docker Desktop Version

- In order to establish a container for MongoDB and connect it with specific credentials, the researchers created a docker-compose.yml file. To further initialise a volume in the container, a JavaScript file called mongo-init.js was utilized to create a user. Together, these two files were used to initialise a Docker instance (Fig. 6).

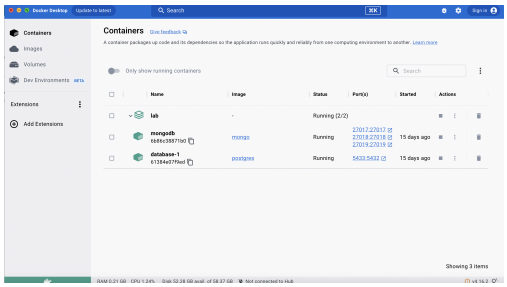

Fig. 6. MongoDB instance in Docker Desktop

- Upon starting the Docker instance, the project initiates a connection with MongoDB Compass using appropriate authentication techniques, and by providing the necessary authentication details. Figure 7 illustrates the MongoDB connection that is established by the Docker instance, which is authenticated by providing the proper credentials.

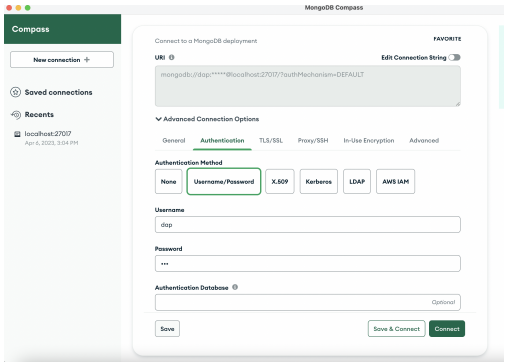

Fig. 7. MongoDB Authentication

2) Programming Language and Code Overview: To import the cleaned datasets into MongoDB, the analysts utilize Python Programming Language and its associated libraries to establish a connection with the MongoDB client.

- The CSV library in Python provides functionality to work with CSV (comma-separated values) files. It allows the reading and writing of CSV files in various formats.
- The JSON library in Python provides functionality to work with JSON (JavaScript Object Notation) data. It allows the encoding and decoding of JSON data in Python objects.
- The pymongo library in Python provides functionality to work with MongoDB databases. It allows for creating connections to MongoDB, querying data, and performing CRUD (create, read, update, delete) operations on data

stored in MongoDB. The MongoClient class is used to create a connection to a MongoDB instance.

- After setting up the necessary Python libraries, the analysts provide the credential details and connect to the MongoDB client user interface where the datasets will be stored. Furthermore, they create a database collection and store the three datasets separately (Fig. 8).

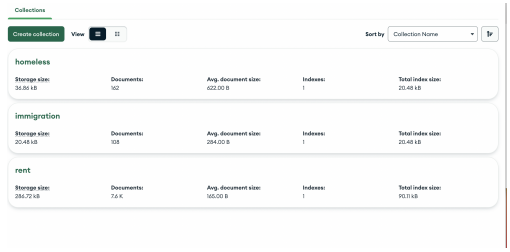

Fig. 8. Datasets present in MongoDB

D. Relational Database

A relational database is structured in tables, where each row represents a record with a unique key. Columns hold data attributes and each record has a value for each attribute. This facilitates the identification of relationships between data points. [6].

1) Software Information:

- PostgreSQL is a powerful open-source object-relational database system that uses SQL language and offers advanced features for secure storage and scalability of complex data workloads. [7]
- In this study, a PostgreSQL instance is utilized which was created through the implementation of the Docker Containerisation technique. To accomplish this, a docker-compose.yml file was composed which contained details such as version, database, image, port, volumes, and connection credentials. Subsequently, through the use of the terminal/command prompt, the analysts were able to successfully create a Docker instance within the Docker Desktop environment.
- PGAdmin is a widely used web-based Graphical User Interface (GUI) management tool for Postgres and related relational databases, enabling communication with both local and remote servers. [8]
- Next, proceed to create a server in pgAdmin 4 using the PostgreSQL docker container that was previously initialized. This process involves using the appropriate user credential that was used during the creation of the docker instance (Fig. 9).
- In addition, this research project creates a new database named ”TeamAHousingCrisisProject”. This database is where all three datasets with mixed structured and semi-structured data will be loaded using the Extract, Transform, and Load (ETL) technique.

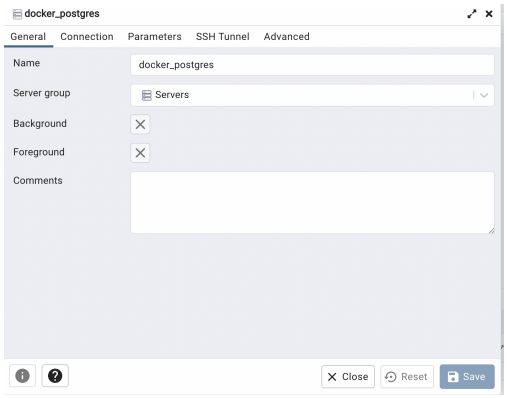

Fig. 9. Server created in PgAdmin

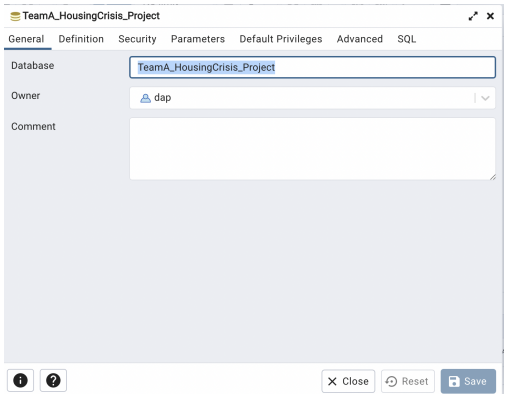

Fig. 10. Database under the dedicated server

- In the final step, the analysts create three tables named ”homeless”, ”immigration”, and ”rent” under the previously created ”TeamAHousingCrisisProject” database. These tables will be used to store the mixed structured and semi-structured datasets extracted from MongoDB after undergoing the ETL process (Fig. 10 and 11).

	table_name name
1	homeless
2	immigration
3	rent

Fig. 11. Created Tables

E. Extract Transform and Load(ETL)

ETL is the process of merging data from various sources into a central repository to prepare datasets for analytics purposes. It helps to clean and structure raw data, leading to more

insightful results, and making it useful for analyzing sales data, forecasting demand, managing inventory, and studying consumer behaviour. [9]

1) Extract: Data extraction involves extracting data from various sources, and data validation is a crucial step to ensure accuracy and conformity to expected values. Rejected data is sent back to the source system for analysis and correction of incorrect records. [10].

- Python libraries are used in the Extract process for data processing and analysis. ”pymongo” communicates with MongoDB. ”Dagster” is a tool for creating pipelines with types like ”Op” and ”Out/In”. The ”PandasColumn” and ”create_dagster_pandas_dataframe_type” are types used for representing Pandas dataframes in Dagster pipelines. Finally, the ”datetime” and ”pandas” libraries are used for working with date/time values and data manipulation, respectively. Overall, these libraries provide a powerful toolkit for working with data in Python.
- Next, the analysts establish a connection to the MongoDB instance using the provided user credentials, with ”dap” as the username and the required password.
- In this project, we have created three Pandas DataFrames - ImmigrationDataFrame, RentDataFrame, and HomelessDataFrame. Each of these dataframes has a specific set of columns that are used to store and manipulate data related to immigration, rent, and homelessness. In this report, we will discuss the columns of each dataframe in detail and explain their significance.
- ImmigrationDataFrame: The ImmigrationDataFrame contains data related to immigration such as immigration_id, file_name, year, gender, nationality, unit, and value. These columns are shown in Fig. 12:

	Name	Meaning
1	Immigration id	A unique identifier for each immigration record
2	File name	The name of the file containing the immigration data
3	Year	Year of the immigration occurred
4	Gender	The gender of the immigrant
5	Nationality	The nationality of the immigrant
6	Unit	The unit of measurement
7	Value	The value of the immigration data

Fig. 12. Columns in ImmigrationDataFrame

- RentDataFrame: The RentDataFrame contains data related to rent, as shown in Fig. 13:

	Name	Meaning
1	Rent id	A unique identifier for each rent record
2	Year	Year in which the rent was paid
3	Number of bedrooms	The number of bedrooms in the rental property
4	Property type	The type of rental property
5	Location	The location of the rental property
6	Cost	The cost of the rental property

Fig. 13. Columns in RentDataFrame

- HomelessDataFrame: The HomelessDataFrame contains data related to homelessness, as shown in Fig. 14:

	Name	Meaning
1	Homeless id	A unique identifier for each homeless record
2	Year	Year in which the homeless count was conducted
3	Month	Month in which the homeless count was conducted
4	Region	Region in which the homeless count was conducted
5	Total Adults	Number of homeless adults in the region
6	Male Adults	Number of homeless male adults in the region
7	Female Adults	Number of homeless female adults in the region

Fig. 14. Columns in HomelessDataFrame

2) Transform:

- The data frames created are extracted from MongoDB and need to be transformed before they can be loaded into the end target. However, after thorough analysis, it has been determined that there are no transformations required for these data frames.
- After conducting a thorough analysis of the data, it has been concluded that there is no need for any transformations to be performed on the three data frames - ImmigrationDataFrame, RentDataFrame, and HomelessDataFrame. This is because all the columns in each of the data frames have been found to contain proper data with no inconsistencies or issues. Therefore, the data is deemed fit for loading into the final database without any further changes or modifications.

3) Load:

- The load.py file contains three operations for loading data from CSV files into a PostgreSQL database. The load_immigration, load_rent, and load_homeless functions each take in a start input and return a boolean output indicating whether the load was successful.
- In each function, the corresponding CSV file is read into a Pandas DataFrame, and then loaded into the database using SQLAlchemy’s to_sql method. Before loading, the target table is truncated to remove any existing data. The number of rows loaded is logged using Dagster’s get_dagster_logger method. In case of any errors, an SQLAlchemyError exception is caught and logged, and the function returns False to indicate that the load was not successful. Overall, this ETL process loads data from CSV files into a PostgreSQL database using Pandas and SQLAlchemy, with error handling and logging for robustness.

4) ETL Workflow:

- The code flow defines an ETL process that extracts data from various sources, transforms it, and loads it into the final target. The code imports the necessary modules and defines an ETL job using the @job decorator. The etl() function calls three functions: extract_immigration(), extract_rent(), and extract_homeless(), which are responsible for extracting data from various sources.

- The extracted data is then passed to the transformation functions: stage_extracted_immigration(), stage_extracted_rent(), and stage_extracted_homeless(), which apply the necessary transformations to the data. The transformed data is then passed to the loading functions: load_immigration(), load_rent(), and load_homeless(), which load the data into the final target (Fig.15).

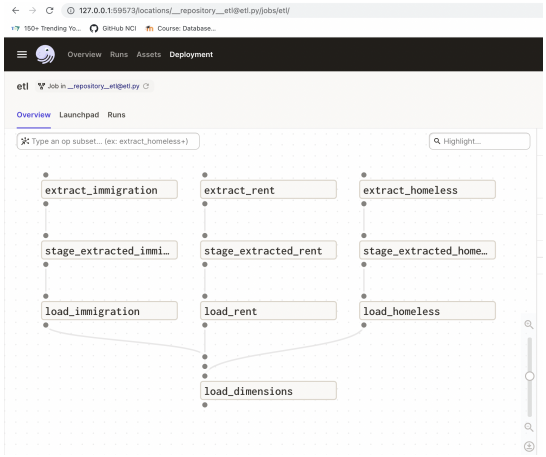

Fig. 15. ETL Overview

- The load_dimensions() function checks if all the data has been successfully loaded into the target. If all the data has been loaded successfully, it returns a boolean value of True.

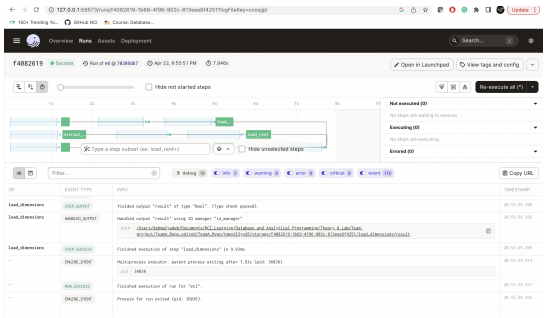

Fig. 16. ETL Launch Pad

F. Transformed Datasets information in pgAdmin 4 after the ETL Process

In this study, the analysts use PgAdmin 4 to create three new tables through an SQL query. These tables are specifically designed to contain only the necessary columns related to the impact of immigration and emigration on the housing market and the resulting homelessness in various counties in Ireland. After extracting the relevant data from these three datasets, the analysts combine them into a single dataset that will be used for the final conclusion and visualization.

1) Homeless table details in pgAdmin 4: This table has been generated by extracting the necessary columns from the ”homeless” dataset stored in MongoDB. The table has been specifically reduced in size, as per the project requirements, to only include columns containing significant information about homeless individuals based on Year, Month, region and Gender.

homeless_id character varying (100)	year integer	month character varying (100)	region character varying (100)	total_adults character varying (100)	male_adults character varying (100)	female_adults character varying (100
643d4ac706c5e7349c739324	2020	January	Dublin	4600	2773	1827
643d4ac706c5e7349c739325	2020	January	Mid-East	301	186	115
643d4ac706c5e7349c739326	2020	January	Midlands	109	58	51
643d4ac706c5e7349c739327	2020	January	Mid-West	341	208	133
643d4ac706c5e7349c739328	2020	January	North-East	128	77	51
643d4ac706c5e7349c739329	2020	January	North-West	67	45	22
643d4ac706c5e7349c73932a	2020	January	South-East	222	160	62
643d4ac706c5e7349c73932b	2020	January	South-West	578	358	220
643d4ac706c5e7349c73932c	2020	January	West	351	202	149
643d4ac706c5e7349c73932d	2020	February	Dublin	4550	2766	1784
643d4ac706c5e7349c73932e	2020	February	Mid-East	312	191	121

Fig. 17. Homeless table in pgAdmin 4

2) Immigration table details in pgAdmin 4: The ETL process involves creating an immigration table by sorting and filtering datasets imported into MongoDB. Only necessary columns representing housing and homelessness caused by immigration and emigration in Ireland are kept. The resulting table includes immigration ID, file name, year, gender, and nationality. This table is a more concise representation of the original dataset and is used for further analysis. The transformation stage is crucial in formatting relevant data for analysis.

	immigration_id character varying (100)	file_name character varying (100)	year integer	gender character varying (100)	nationality character varying (100)	unit character va
1	643d4ac706c5e7349c7392b7	PEA17	2020	Both sexes	Irish	Thousand
2	643d4ac706c5e7349c7392b8	PEA17	2020	Both sexes	UK (1)	Thousand
3	643d4ac706c5e7349c7392b9	PEA17	2020	Both sexes	EU14 excl Irl (countries in the EU pre 2004 excluding UK & Irelan…	Thousand
4	643d4ac706c5e7349c7392ba	PEA17	2020	Both sexes	EU15 to EU27 (accession countries joined post 2004)	Thousand
5	643d4ac706c5e7349c7392bb	PEA17	2020	Both sexes	Other nationalities(17)	Thousand
6	643d4ac706c5e7349c7392bc	PEA17	2020	Both sexes	All nationalities	Thousand
7	643d4ac706c5e7349c7392bd	PEA17	2020	Male	Irish	Thousand
8	643d4ac706c5e7349c7392be	PEA17	2020	Male	UK (1)	Thousand
9	643d4ac706c5e7349c7392bf	PEA17	2020	Male	EU14 excl Irl (countries in the EU pre 2004 excluding UK & Irelan…	Thousand
10	643d4ac706c5e7349c7392c0	PEA17	2020	Male	EU15 to EU27 (accession countries joined post 2004)	Thousand
11	643d4ac706c5e7349c7392c1	PEA17	2020	Male	Other nationalities(17)	Thousand

Fig. 18. Immigration table in pgAdmin 4

3) Rent table details in pgAdmin 4: This table includes only the relevant columns that provide clear insights into the final outcome, which is the significant increase in rent prices due to the recent immigration inflow in Ireland.

rent_id character varying (100)	year integer	number_of_bedrooms character varying (100)	property_type character varying (100)	location character varying (100)	cost double precision
643d4ac706c5e7349c7393c7	2020	All bedrooms	All property types	Carlow Town	915.37
643d4ac706c5e7349c7393c8	2020	All bedrooms	All property types	Graiguecullen, Carlow	907.25
643d4ac706c5e7349c7393c9	2020	All bedrooms	All property types	Tullow, Carlow	840.54
643d4ac706c5e7349c7393ca	2020	All bedrooms	All property types	Cavan	664.88
643d4ac706c5e7349c7393cb	2020	All bedrooms	All property types	Cavan Town	691.64
643d4ac706c5e7349c7393cc	2020	All bedrooms	All property types	Bailieborough, Cavan	613.4
643d4ac706c5e7349c7393cd	2020	All bedrooms	All property types	Ballyconnell, Cavan	592.51
643d4ac706c5e7349c7393ce	2020	All bedrooms	All property types	Ballyjamesduff, Cavan	667.2
643d4ac706c5e7349c7393cf	2020	All bedrooms	All property types	Belturbet, Cavan	584.85
643d4ac706c5e7349c7393d0	2020	All bedrooms	All property types	Cootehill, Cavan	564.44
643d4ac706c5e7349c7393d1	2020	All bedrooms	All property types	Kingscourt, Cavan	741.9

Fig. 19. Rent table in pgAdmin 4

The rent table contains information such as the rent ID, year of rent, property type, county, and cost. Analysts can use this table to study the patterns and trends in rent prices and draw conclusions about the impact of immigration on the housing market in Ireland.

IV. Results and Evaluation

The three datasets are merged using common columns to create a comprehensive table with information on nationality,

gender, immigration year, and homeless adults. A JOIN operation is performed using the common columns. The resulting table has column names like ”rent location” and ”immigrants gender”. It provides insights into the relationship between immigration, homelessness, and rent prices in Ireland. The Plotly Python graphing library can generate interactive, high-quality visualizations. The table and visualizations can aid analysts in formulating policies to address these issues. [12].

Figure 20 is a scatter plot showing the correlation between rental costs, immigrant population, and homelessness in Ireland. The abscissa represents rental costs in Euros, and the ordinate denotes the number of immigrants from a particular nationality. Each data point corresponds to a specific region in Ireland, with bubble size representing the number of homeless people and colour denoting the homeless region. The chart visually demonstrates the connection between the three variables.

Fig. 20. Distribution of immigrants' nationality by rent cost per region in Ireland, along with the number of homeless adults in each region

The rental costs tend to increase as the concentration of immigrants increases, as evidenced by the visual representation of the data. Specifically, Irish and EU15-EU27 citizens have a higher number of homeless adults compared to other nationalities. The scatter plot revealed a significant concentration of homeless adults among these two groups. The data also showed that the number of homeless adults among other nationalities was relatively low. Additionally, the chart indicates that the concentration of homeless adults is higher in regions where the number of immigrants is more pronounced.

Figure 21 is a line chart that displays the trend of average rent costs over time in different locations in Ireland. The chart shows the average rent costs for each region in Ireland during the years 2020 and 2021, with the abscissa indicating the year and the ordinate displaying the average rent cost in Euros. Each line on the chart corresponds to a specific region, with the color of the line indicating the region.

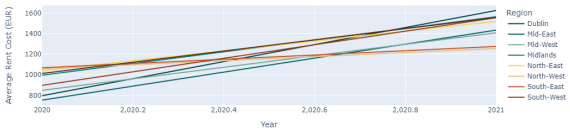

Fig. 21. Average rent costs by location over time in Ireland

The visual representation indicates that certain regions, such as Dublin and South-West, consistently exhibit higher rent costs compared to other regions. Furthermore, the chart underscores the impact of the COVID-19 pandemic on the rental market in Ireland. Notably, there has been a discernible increase in rent costs across several regions since the outbreak of the pandemic in 2020, thereby suggesting a significant impact of the pandemic on the housing market in Ireland.

Figure 22 depicts the distribution of rent costs across distinct regions within Ireland, with the abscissa representing the regions and the ordinate displaying the rent cost in Euros. Each box portrayed on the chart corresponds to a specific region, with the box plot serving to visualize the distribution of rent costs for that region.

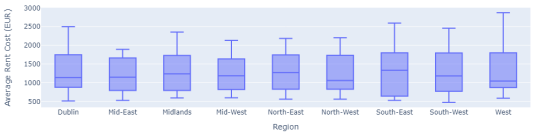

Fig. 22. Distribution of rent costs across Ireland

This box plot showcases that certain regions exhibit a higher median rent cost in comparison to others. Notably, The box plot for rent costs in Dublin gives us an idea about the distribution of rent costs in the region. The range of rent costs in Dublin, as shown by the box plot, is from 513 Euros to 2496 Euros. The median rent cost, which is the value separating the higher half from the lower half of the data set, is 1138 Euros, indicating that the distribution is somewhat skewed towards the higher end of the range. The interquartile range, which is the range between the first and third quartiles, is 870 Euros, suggesting that there is some variability in rent costs within Dublin.

Similarly, we can also see that the distribution of rent costs in the West region appears to be somewhat skewed to the right. This means that there are likely more rental properties with lower rent costs than higher rent costs in this region, Overall, the box plot provides a useful summary of the range and distribution of rent costs by region, and can help renters and landlords alike to make more informed decisions about rental pricing.

Figure 23 presents a histogram that illustrates the frequency distribution of rent costs in different homeless regions. The chart demonstrates that the prevalence of rent costs varies among the homeless regions.

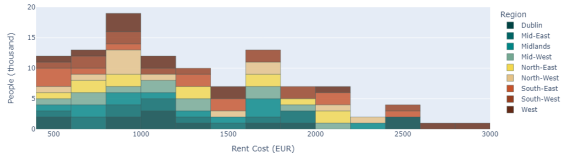

Fig. 23. Histogram of frequency of different rent costs

Specifically, the histogram for Dublin depicts that rent costs within the range of 1000-1500 EUR are the most frequent, whereas for West, the most common rent costs are within the range of 500-1000 EUR. Moreover, the chart reveals that the

distribution of rent costs across regions is not uniform. For instance, the histogram for Dublin has a broader range of rent costs compared to that of the South-West, signifying that the distribution of rent costs in Dublin is more diverse than that in the South-West.

Figure 24 presents a grouped bar chart presenting the distribution of homeless adults by gender and region. Each region has two bars, one for the number of homeless male adults and the other for the number of homeless female adults. According to the grouped bar chart, Dublin has the highest number of homeless adults, irrespective of gender, followed by the South-West region. In contrast, the North-West region has the lowest number of homeless male and female adults.

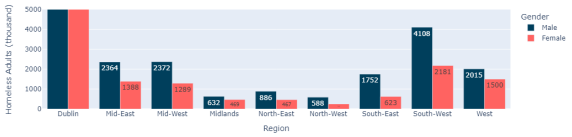

Fig. 24. Breakdown of homeless adults by gender and region

This information highlights the significant differences in the distribution of homeless adults across regions, with some regions experiencing higher levels of homelessness than others. The chart helps in identifying significant gender disparities in the number of homeless adults across regions and whether the distribution of homeless adults by gender varies significantly between regions.

Figure 25 utilizes the OpenCageGeocode API to extract the latitude and longitude coordinates of various rent regions in Ireland. The chart employs the Plotly Express library to generate a scatter map box visualization, which displays the rent cost on a map. The map is color-coded based on the

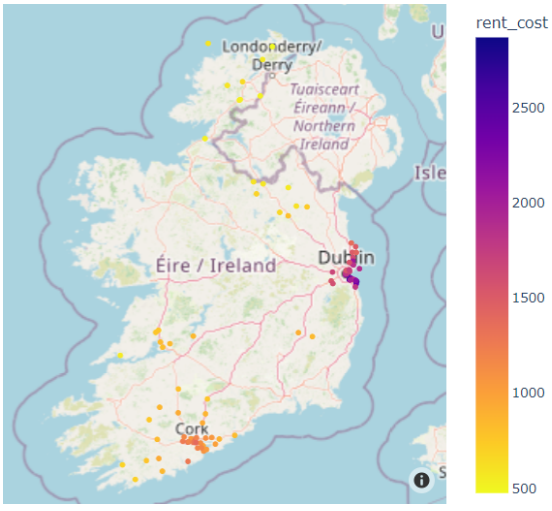

Fig. 25. Spatial Distribution of Rent Costs in Ireland

rent cost, thereby facilitating an understanding of the spatial distribution of rent costs across Ireland. From the visual, we can conclude that most rent regions in Ireland are clustered in Dublin and Cork(South West region). Through this chart, it is possible to identify areas with higher or lower rent costs and explore potential factors contributing to these differences. This map-based visualization thus provides valuable insights into the geographic variations in rent costs in Ireland [13].

V. Conclusions and Future Work

The case study conducted to examine the relationship between immigration and emigration data, local rent prices, and homelessness rates in Ireland provided valuable insights. The charts analyzed in this study highlighted significant spatial and temporal variations in rent costs and homelessness rates across different regions in Ireland. The study also found a positive correlation between rent costs and the number of immigrants in a region, indicating that housing affordability could be a significant issue for immigrants in Ireland. Additionally, the study identified gender disparities in the number of homeless adults across different regions, emphasizing the need for targeted interventions to address the needs of homeless populations.

Based on the insights found in the analysis, several areas for future research can be identified. One potential is to investigate the specific factors that contribute to regional variations in rent costs and homelessness rates. Additionally, further research could be conducted to examine the effectiveness of existing policies and programs aimed at addressing homelessness and housing insecurity among immigrant and vulnerable populations in Ireland.

References

[1] David Quinn, Independent.ie, WEBSITE © INDEPENDENT.IE, Huge scale of immigration is making our housing crisis worse. [Online] Available: https://www.independent.ie/opinion/columnists/david-quinn/huge-scale-of-immigration-is-making-our-housing-crisis-worse-35498057.html/

[2] ION UN migration, Publications platform, Managing Migration in Ireland: A Social and Economic Analysis. [Online] Available: https://publications.iom.int/books/managing-migration-ireland-social-and-economic-analysis: :text=The

[3] Government of Ireland. [Online] Available: https://data.gov.ie/

[4] Getting Started, MongoDb Documentation [Online] Available: https://www.mongodb.com/docs/manual/tutorial/getting-started/

[5] Docker Docs [Online] Available: https://docs.docker.com/get-started/overview/

[6] Oracle Ireland [Online] Available: https://www.oracle.com/ie/database/what-is-a-relational-database/: :text=A

[7] The PostgreSQL Global Development Group [Online] Available: https://www.postgresql.org/about/

[8] Adservio [Online] Available: https://www.adservio.fr/post/what-is-pgadmin/

[9] AWS [Online] Available: https://aws.amazon.com/what-is/etl/: :text=Extract

[10] Extract, Wikipedia [Online] Available: https://en.wikipedia.org/wiki/Extract,$transform$,$load$/

[11] Transform, Wikipedia [Online] Available: https://en.wikipedia.org/wiki/Extract,$transform$,$load$/

[12] Plotly Open Source Graphing Library for Python, Python [Online] Available: https://plotly.com/python//

[13] OpenCage Geocoding API Documentation, OpenCageData [Online] Available: https://opencagedata.com/api/