Module: Data Storage Solutions for Data Analytics Module

Code: B9DA111

Lecturer: Dr Anesu Nyabadza

Submitted By

Venkaiah Malli 20026676

Table of Contents

1. Introduction

The Airbnb Ireland listings dataset provides rich data on the short-term rental market for Ireland which encompasses properties type and price range, host information, features, availability rates, and customer reviews. Such details documented in the dataset like pricing trends, occupancy rates, customer satisfaction etc, makes it possible to make various kind of analysis by using the set of metrics. In the following roles and usefulness of this dataset: The hosts, guests, investors, and policymakers all derive value from better understanding market dynamics and property performance. In particular, hosts can manage price, investors can find the most profitable zones, and the local government can recognize the tendencies in the share of housing used for short-term lease. This dataset, mainly collected from Inside Airbnb, consists of large volumes of compilations of raw data that can be broken down into formal analyses to help the respective tourism boards. It also helps city planners in adapting their strategies to the current market trends and profile while contributing to data scientists' establishments of predictive models concerning, the prices, demand and customer reviews.

1.1 Reasons for selecting Subject area and DATA

Airbnb Property Listening subject area was selected, it gives comprehensive and detailed information about the short-term rental market in Ireland, covering a wide range of property types, their prices, host information, availability rates, and customer reviews. Such a rich dataset enables multiple analyses on pricing trends, occupancy rates, and customer satisfaction. This is of great value to several stakeholders: hosts can work out the best pricing strategies, investors can determine areas that will be profitable, and policymakers can understand the effects of short-term rentals on housing markets. Offering real-world relevance and analytical potential, the dataset is ideal for research and strategic decision-making in understanding and forecasting market behaviours.

1.2 Business Requirements and Vision

The business requirements used in and vision for the Airbnb Ireland listings dataset are the identification of the short-term rental business requirements and possibilities for gaining new insights into the phenomenon. The vision is to foster a robust data society whereby the property hosts, investors, guests, policy makers and any other stakeholders involved in rental market can be able to use detailed rental information in decision making in a way that will benefit the overall community and at the same time maximize individual revenues.

Core BI needs include performance tracking of property, assessment of pricing solutions, studying occupancy rates, and the measurement of customer satisfaction. For hosts, the dataset must offer solutions for price adaptation according to seasonal conditions and market competition. Leading players within the sector including investors and real estate agencies need more information about areas, properties types that are most in demand to optimally position their investments. Local governments and city planners require data on the incidence of short term lets to enable regulation and planning of cities.

1.3 Business Vision

The main purpose of this use of this data storage solution is to improve the decision-making process for people investing in "Airbnb leases". Based on a reliable data warehouse, the solution offers -

Comprehensive Insights: Allowing hosts, guests, investors, and policymakers to evaluate rental trends, performance and customer behaviours.

Improved Efficiency: Enabling better pricing, property management, and customer satisfaction as a result of more efficient reporting approaches. Scalability: Enabling easy connection to future data sources to capture the expanding short-term renting segment.

1.4 Key Stakeholders

1. Policymakers: To balance the expansion of tourism with the availability of homes to regulate and oversee short-term rentals in Ireland.

Benefit: Make well-informed choices regarding housing laws and tourism rules in important regions.

2. Property owners or hosts: By improving amenities, modifying prices, optimizing listings, and examining visitor comments, you can raise guest satisfaction.

Benefit: more occupancy rates translate into more profits.

3. Visitors and Passengers:

To make well-informed judgments, use reviews to choose homes that fit your criteria for amenities, location, and cost.

Benefit: Improve visitor happiness and their overall vacation experience.

4. **Real estate investors**: To identify attractive short-term rental investments in Ireland, examine market demand and property performance.

Benefit: Improve your investment choices to increase returns in sectors with strong demand.

5. Airbnb and Other Rental Platforms: Enhance offerings, advise hosts on the best prices, and keep up with industry developments.

Benefit: Strengthen the service offerings and competitive positioning.

6. City planners and local governments:

Monitor the effects of short-term rentals on Ireland's local housing market and tourism sector.

Benefit: Make well-informed choices about tourism, zoning, and housing rules.

7. Tourism boards:

Examine traveller patterns, real estate availability, and the impact of short-term rentals on the travel industry.

Benefit: Promote Irish regions strategically to increase tourism.

1.5 Data Source

The data used in this solution utilizes "Inside Airbnb resource" that hosts extensive information about Airbnb listings. The choice of indicators and their frequency is more suitable for assessing the dynamics and shifts in the short-term rental market in the short term in Ireland, particularly the Dublin region.

Source Details Website Link: Inside Airbnb

Complete Dataset Link: Dublin Listings

Dataset Highlights

The dataset provided contains comprehensive data of "Airbnb listings" that includes host data, prices, listing location and dates of availability, and guests' comments. It is accessible to the public and used frequently in order to reflect the current listings in the market.

Column	Data Type	Description
Listing_ID	INT	*"Unique identifier for each property listing.*
Host_ID	INT	*Identifier linking a listing to its host.*
Price	FLOAT	*Rental price per night in USD.*
Availability_365	INT	*Number of days the listing is available in a year.*
Number_of_Reviews	INT	*Total reviews received for the listing.*
Neighborhood_Group	VARCHAR	*High-level area grouping (e.g., Manhattan, Brooklyn).*
Room_Type	VARCHAR	*Type of room offered (e.g., entire home, private room).*
Minimum_Nights	INT	*Minimum nights required for booking.*
Latitude	FLOAT	*Latitude of the listing's location.*
Longitude	FLOAT	*Longitude of the listing's location.*
Column	Data Type	*Description*
Listing_ID	INT	*Unique identifier for each property listing.*
Host_ID	INT	*Identifier linking a listing to its host.*
Price	FLOAT	*Rental price per night in USD.*
Availability_365	INT	*Number of days the listing is available in a year.*
Number_of_Reviews	INT	*Total reviews received for the listing.*
Neighborhood_Group	VARCHAR	*High-level area grouping (e.g., Manhattan, Brooklyn).*
Room_Type	VARCHAR	*Type of room offered (e.g., entire home, private room).*
Minimum_Nights	INT	*Minimum nights required for booking.*
Longitude	FLOAT	*Longitude of the listing's location.*
Latitude	FLOAT	*Latitude of the listing's location."*

Task 2: Data Warehouse Schema Design

Schema Selection

A star schema includes one central fact table, containing measurable data, and a number of dimension tables containing descriptive attributes. This structure will allow for easy analysis and reporting on Airbnb insights like price trends, availability, and reviews. Fact and Dimension Tables

Fact Table: Airbnb Listening Properties

Column	Data Type	Description
Property ID	INT	Primary Key: Unique identifier for each property.
Host ID	INT	Foreign key linked to the Dim Hosts table.
Availability ID	INT	Foreign key linked to the Dim Availability table.
Review ID	INT	Foreign key linked to the Dim Review table.
Location ID	INT	Foreign key linked to the Dim Location Info table.
Price	INT	Price of Airbnb Properties
Minimum Nights	INT	Minimum number of nights to stay
Number Of Reviews	INT	Number of reviews received for Property
Number of Reviews _ltm	INT	Number of reviews received last 12 months

Dim Reviews:

Column	Data Type	Description
Review ID	INT	Unique identifier for each review
Last Review	Date	Last review of the property
Reviews Per month	INT	Reviews per month

Dim_Hosts:

Column	Data Type	Description
Host_ID	INT	Unique identifier for each host.
Host_Name	VARCHAR	Name of the host.
Calculated Host listening Count	INT	Host listening count

Dim Availability

Column	Data Type	Description
Availability ID	INT	Availability unique id
Availability 365	INT	Number of days property available days

Dim Location Info

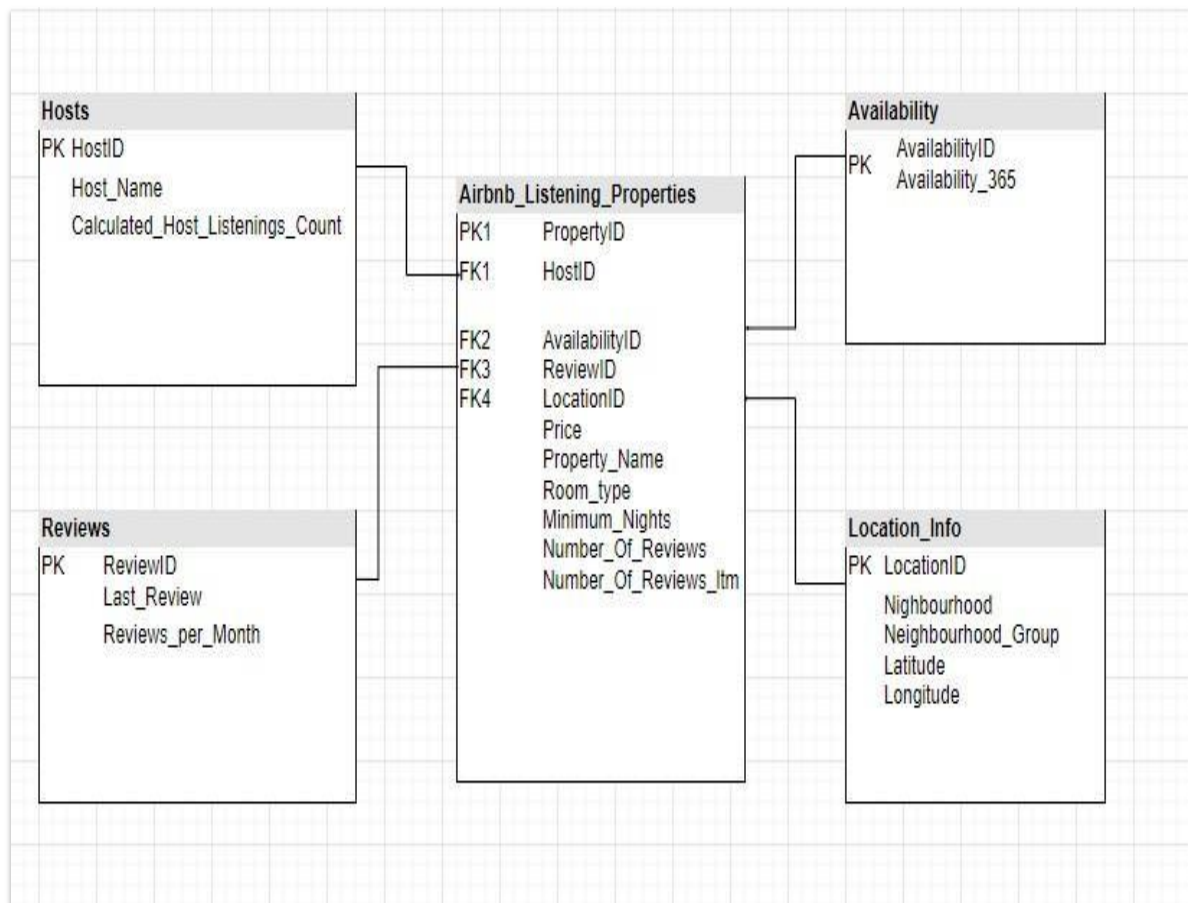| Column	Data Type	Description
Location ID	INT	Location unique Identifier
Neighbourhood	Varchar	Specific area neighbourhood location available
Neighbourhood Group	Varchar	Region of the Property
Latitude	Float	Geographic latitude of the Property
Longitude	Float	Geographic longitude of the Property

2. SCHEMA

Figure 1: Warehouse Schema for Airbnb Data Analysis

Reasons for the Design

Star Schema for Simplicity:

The star schema is easy to use in querying and reporting since data is held in a central fact table with dimensions encompassing it.

Ease of Analysis: In addition, in each dimension table, detailed descriptive attributes that enable further analysis are included.

Dim_Listings: This dimension supports the evaluation of rental trends by room type and location. It enables the analysis of average prices and listing patterns per neighborhood or per property type.

Dim_Hosts: Provides data on hosts, such as "superhost performance" and the number of active listings per host. This dimension helps evaluate host activity and reliability.

Dim_Reviews: This facilitates the analysis of guest satisfaction based on reviews per month and the date of the most recent review. It allows for sentiment analysis and service quality assessment.

Dim_Location_Info: Contains geographic information, such as neighborhood and coordinates, enabling location-based insights, like popular areas or pricing trends across different locations.

Fact_Airbnb_Listening_Properties: This is the center fact table that holds information regarding property attributes, availability, prices, room types, and reviews. This table ties all the dimensions together for a thorough analysis.

Support for Historical Trends: By utilizing dimensions such as reviews and availability, the schema enables the tracking of trends across time, such as property or room type popularity over time

Optimized for Reporting: This schema is designed to support various reports, such as: Revenue by neighbourhood. Booking trends by room type. Average price per night by area. Scalability: The schema can be easily extended to cover either additional dimensions or measures in the fact table (such as "Dim Reviews")

2.1 Implementation of Sequential Diagram:
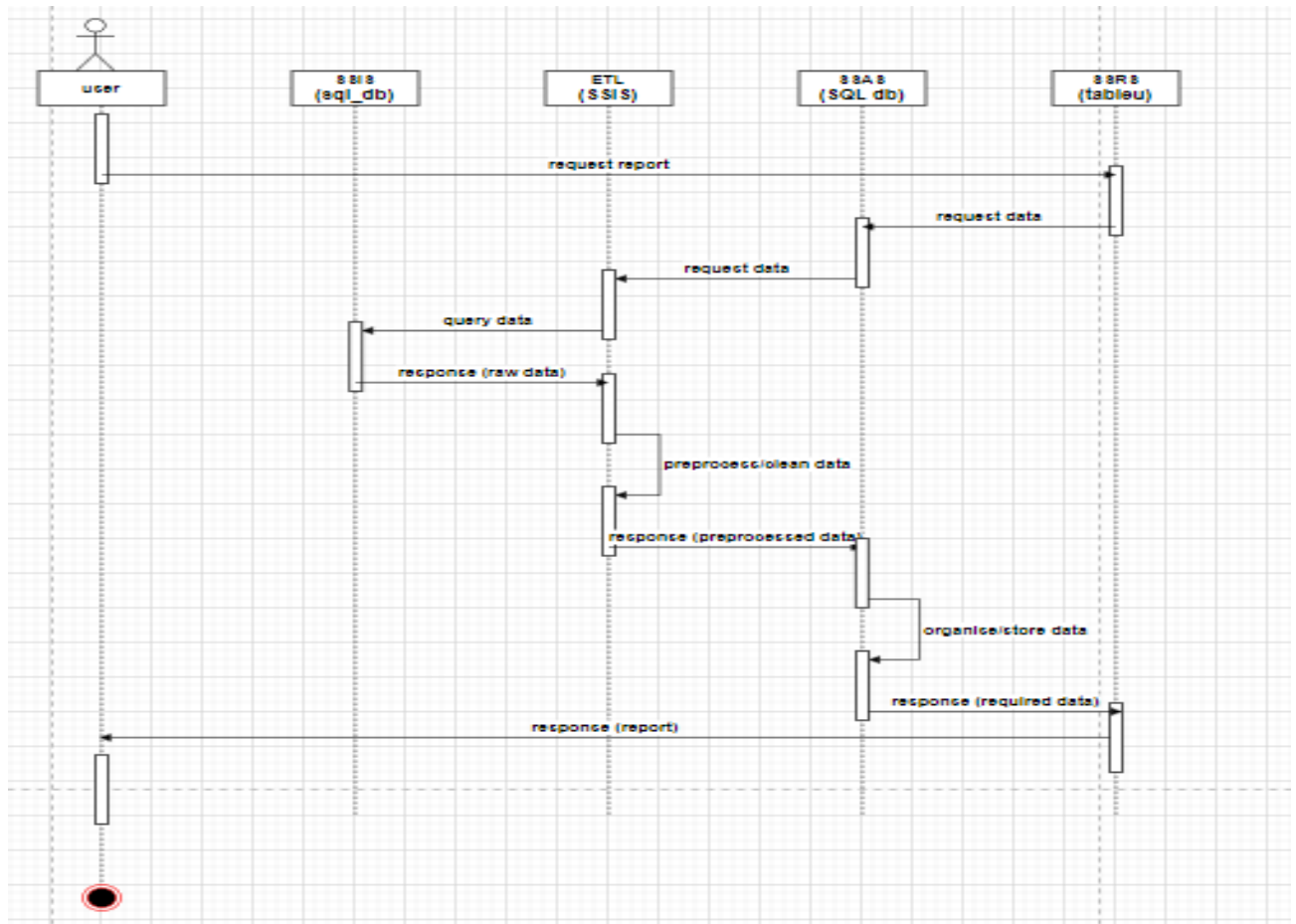
Pipeline Process:

First, the user submits a request SSRS Tableau for a report; this is then processed through SSIS to query raw data from a database. It sends the raw data down to a SSIS ETL (Extract, Transform, Load) process for preprocessing and cleaning. The pre-processed data is forwarded to SSAS (SQL DB), which then organizes and stores the data into a structured format. The final result is that SSRS or Tableau retrieves the needed data, prepares the report, and sends the final response back to the user.

3: ETL Implementation Using SSIS

Creation of Fact and Dimension tables

```sql
/* creating Hosts Dimension Table for Airbnb Listening */

CREATE TABLE Hosts (
    HostID INT PRIMARY KEY,
    Host_Name VARCHAR(255),
    Calculated_Host_Listings_Count INT
);

/* Creating  Availability Dimension Table for Airbnb Listening */

CREATE TABLE Availability (
    AvailabilityID INT PRIMARY KEY AUTO_INCREMENT,
    Availability_365 INT
);
```

```sql
/*Creating  Location_Info Dimension Table for Airbnb listening */

CREATE TABLE Location_Info (
    LocationID INT PRIMARY KEY AUTO_INCREMENT,
    Neighbourhood VARCHAR(255),
    Neighbourhood_Group VARCHAR(255),
    Latitude FLOAT,
    Longitude FLOAT
);
```

```sql
/* Creating  Reviews Dimension Table for Airbnb listening */

CREATE TABLE Reviews (
    ReviewID INT PRIMARY KEY AUTO_INCREMENT,
    Last_Review DATE,
    Reviews_per_Month FLOAT
);
```

11

```
/* Create Fact Table: Airbnb_Listing_Properties */

CREATE TABLE Airbnb_Listing_Properties (
    PropertyID INT PRIMARY KEY,
    HostID INT,
    AvailabilityID INT,
    ReviewID INT,
    LocationID INT,
    Price FLOAT,
    Property_Name VARCHAR(255),
    Minimum_Nights INT,
    Room_Type VARCHAR(255),
    Number_Of_Reviews INT,
    Number_Of_Reviews_Ltm INT,
    FOREIGN KEY (HostID) REFERENCES Hosts(HostID),


    FOREIGN KEY (AvailabilityID) REFERENCES Availability(AvailabilityID),
    FOREIGN KEY (ReviewID) REFERENCES Reviews(ReviewID),
    FOREIGN KEY (LocationID) REFERENCES Location_Info(LocationID)
);
```

Figure 3.1 Creation of Fact and Dimension tables

✔ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0260 seconds.)

```
/* creating Hosts Dimension Table for Airbnb Listening */ CREATE TABLE Hosts ( HostID INT PRIMARY KEY, Host_Name VARCHAR(255),
Calculated_Host_Listings_Count INT );
```

[Edit inline] [Edit] [Create PHP code]

✔ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0222 seconds.)

```
/* Creating Availability Dimension Table for Airbnb Listening */ CREATE TABLE Availability ( AvailabilityID INT PRIMARY KEY
AUTO_INCREMENT, Availability_365 INT );
```

[Edit inline] [Edit] [Create PHP code]

✔ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0265 seconds.)

```
/* Creating Reviews Dimension Table for Airbnb listening */ CREATE TABLE Reviews ( ReviewID INT PRIMARY KEY AUTO_INCREMENT,
Last_Review DATE, Reviews_per_Month FLOAT );
```

[Edit inline] [Edit] [Create PHP code]

✔ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0213 seconds.)

```
/*Creating Location_Info Dimension Table for Airbnb listening */ CREATE TABLE Location_Info ( LocationID INT PRIMARY KEY
AUTO_INCREMENT, Neighbourhood VARCHAR(255), Neighbourhood_Group VARCHAR(255), Latitude FLOAT, Longitude FLOAT );
```

[Edit inline] [Edit] [Create PHP code]

✔ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0620 seconds.)

```
/* Create Fact Table: Airbnb_Listing_Properties */ CREATE TABLE Airbnb_Listing_Properties ( PropertyID INT PRIMARY KEY, HostID INT,
AvailabilityID INT, ReviewID INT, LocationID INT, Price FLOAT, Property_Name VARCHAR(255), Minimum_Nights INT, Room_Type
VARCHAR(255), Number_Of_Reviews INT, Number_Of_Reviews_Ltm INT, FOREIGN KEY (HostID) REFERENCES Hosts(HostID), FOREIGN KEY
(AvailabilityID) REFERENCES Availability(AvailabilityID), FOREIGN KEY (ReviewID) REFERENCES Reviews(ReviewID), FOREIGN KEY
(LocationID) REFERENCES Location_Info(LocationID) );
```
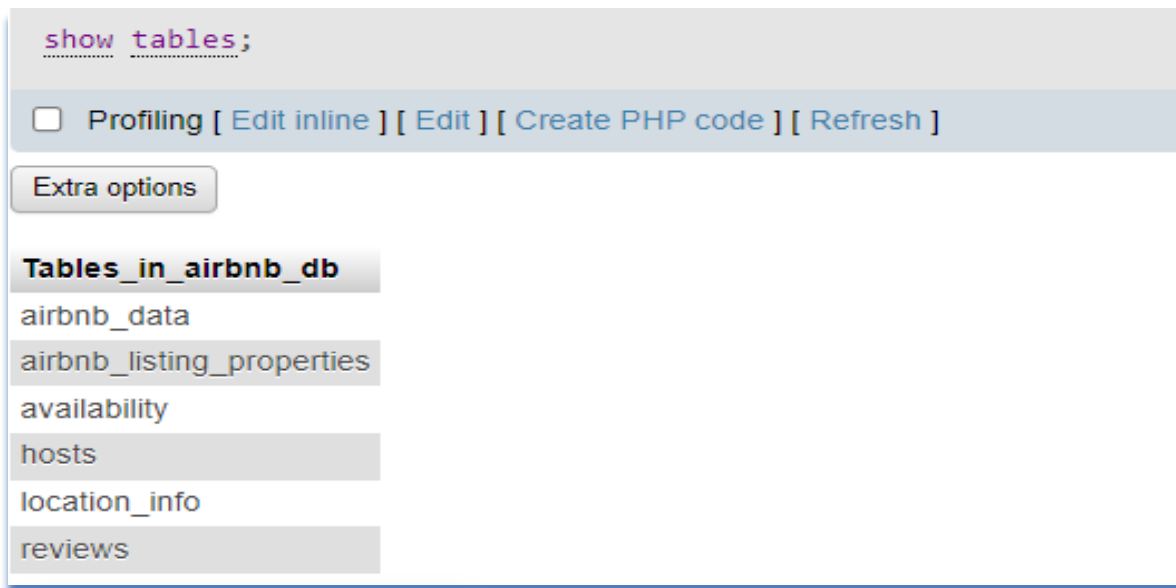
Figure 3.2 creation of Table Outputs

Figure 3.3 All the Tables in Airbnb Database

3.1 ETL Overview

ETL which stands for Extract, Transform, and Load is the process of acquiring data from multiple sources, modifying it to a common format, and then putting that data into a target database usually for analytical purposes (Ciampi *et al.* 2021). In SSIS, ETL prepares source data for analysis and consolidates them into one system to present accurate, standardised data for analysis for BI.

Steps in SSIS

In SSIS, ETL starts from configuring the Data Flow Task, where sources need to be identified, and destinations need to be set. Extracting means setting up source connections like flat files, and databases. Transform utilizes variables such as Data Conversion and Derived Columns. Last, Load transmits the altered data to the destination. Error handling and logging are the procedures that help to achieve a certain level of integrity in the process.

ETL Screenshots

Illustrates the main ETL task in the layout of Control Flow in SSIS.

Figure 6: Excel AirBNB Data Import

This Airbnb dataset import preview in SQL Server displays the initial 50 rows, showcasing attributes like `id`, `name`, `host_id`, `host_name`, `neighbourhood_group`, `latitude`, and `longitude`. Using "Rich Data Type Detection" refines data type precision, automatically detecting field formats and ensuring a closer data fit. The dataset features varied room types and geographic locations across Dublin, aiding comprehensive property and host analysis.
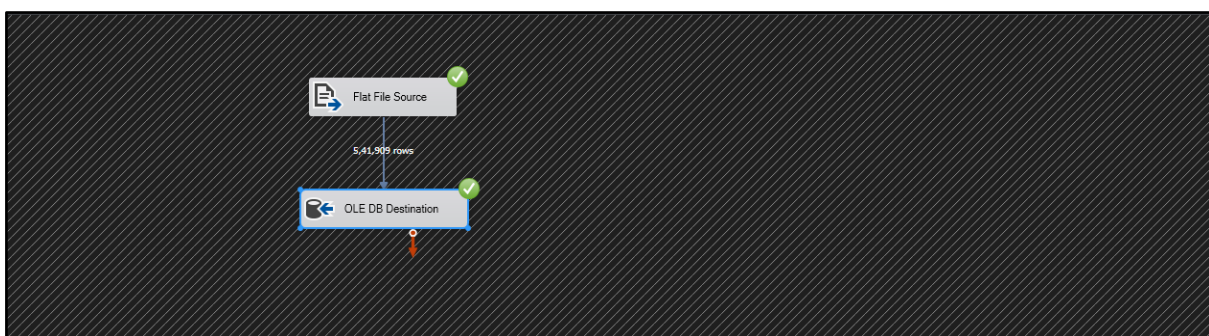

Figure 7: Successfully Transforming the data

Here the data flow transformation is displayed just fine with the row count from the Flat file to the OLE DB.

3.2 Insertion table Values from Airbnb Dataset:

```sql
/* Inserting values for Hosts table from Airbnb dataset */

INSERT INTO hosts (HostID, Host_Name, Calculated_Host_Listings_Count)
SELECT DISTINCT
    host_id AS HostID,
    host_name AS Host_Name,
    calculated_host_listings_count AS Calculated_Host_Listings_Count
FROM airbnb_data;
```

```sql
/* Inserting values availability Dimention table from Airbnb data */
INSERT INTO availability (AvailabilityID, Availability_365)
SELECT DISTINCT
    id AS AvailabilityID,
    availability_365 AS Availability_365
FROM airbnb_data;
```

```sql
/* Inserting values into reviews dimention table from Airbnb data*/
INSERT INTO reviews (ReviewID, Last_Review, Reviews_per_Month)
SELECT DISTINCT
    id AS ReviewID,
    last_review AS Last_Review,
    reviews_per_month AS Reviews_per_Month
FROM airbnb_dataset;
```

```sql
/* Insert values location info dimention table from airbnb dataset */
INSERT INTO location_info (LocationID, Neighbourhood, Neighbourhood_Group, Latitude, Longitude)
SELECT DISTINCT
    id AS LocationID,
    neighbourhood AS Neighbourhood,
    neighbourhood_group AS Neighbourhood_Group,
    latitude,
    longitude
FROM airbnb_data;
```

```sql
/* Inserting values into the Airbnb_Property fact table */
INSERT INTO airbnb_listing_properties(PropertyID, HostID, LocationID, AvailabilityID, ReviewID, Property_Name,
Room_Type, Price, Minimum_Nights, Number_Of_Reviews, Number_Of_Reviews_ltm)
SELECT
    airbnb_data.id AS PropertyID,
    hosts.HostID,
    location_info.LocationID,
    availability.AvailabilityID,
    reviews.ReviewID,
    airbnb_data.name AS Property_Name,
    airbnb_data.room_type AS Room_Type,
    airbnb_data.price AS Price,
    airbnb_data.minimum_nights AS Minimum_Nights,
    airbnb_data.number_of_reviews AS Number_Of_Reviews,
    airbnb_data.number_of_reviews_ltm AS Number_Of_Reviews_ltm
```

15

```
FROM
    airbnb_data
JOIN hosts ON airbnb_data.host_id = hosts.HostID
JOIN location_info ON airbnb_data.id = location_info.LocationID
JOIN availability ON airbnb_data.id = availability.AvailabilityID
JOIN reviews ON airbnb_data.id = reviews.ReviewID;
```

4: Reporting and Visualization

4.1 SSRS Reporting

SQL Server Reporting Services (SSRS) is a comprehensive reporting platform which enables organizations to design, generate, and deliver reports based on their data. The primary purpose of SSRS is to facilitate data analysis and decision-making by providing dynamic, interactive, and customizable reports. It allows users to create detailed, visual reports from diverse data sources, offering features such as drill-downs, data grouping, and parameterized queries, which help to provide in-depth insights into business data. SSRS supports multiple report formats, such as PDF, Excel, and HTML, CSV etc.

Four SSRS Reporting Queries:

1. Host Performance Based on Number of Listening's and reviews?

```
/* Report 1.Host Perfomance based on Number of Listenings and Reviews?
*/
SELECT
    h.Host_Name AS Host_Name,
    h.Calculated_Host_Listings_Count AS Listings_Count,
    SUM(r.Reviews_per_Month) AS Total_Reviews_Per_Month
FROM
    Hosts AS h
JOIN
    Airbnb_Listing_Properties AS alp ON h.HostID = alp.HostID
JOIN
    Reviews AS r ON alp.ReviewID = r.ReviewID
GROUP BY
    h.Host_Name, h.Calculated_Host_Listings_Count
ORDER BY
    Total_Reviews_Per_Month DESC;
```

Figure 4.1 Host performance based on number of listings and reviews

OUTPUT REPORT:

Host_Name	Listings_Count	Total_Reviews_Per_Month
Mary	1	175.0700003
Diogo	25	144.1300001
GuoQing	7	129.6099997
Martin	19	104.3700008
John	1	103.5700004
Edyta	13	98.1699996
Daniel And G	97	95.76000039
Malcolm	41	94.61000028
Liam	17	92.51000023
Lucas	60	91.22999984
Paul	1	87.55
James	1	85.70000079
Michelle	29	82.48000073
Gregory Joseph	23	77.91999936
John	2	77.81999942
David	1	77.48999934
Sinead	1	73.63999964
Michelle	1	72.21999976
Gillian	12	70.16000032
Deni	9	69.78999877
The Rest	10	68.62000036
Beckett	6	66.17000055
Helen	1	58.60000046
Margaret	1	58.55000004

Report: Host Performance based on Listening Count and Total Review

2. Room Type by Revenue Report?

```
/*2.Estimated Revenue by NEighbourhood Group*/

SELECT
    li.Neighbourhood_Group AS Neighbourhood_Group,
    COUNT(alp.PropertyID) AS Total_Properties,
    SUM(alp.Price * alp.Minimum_Nights) AS Estimated_Revenue
FROM
    Location_Info AS li
JOIN
    Airbnb_Listing_Properties AS alp ON li.LocationID = alp.LocationID
GROUP BY
    li.Neighbourhood_Group
ORDER BY
    Estimated_Revenue DESC;
```

Figure 4.2 Estimated Revenue by Neighbourhood

OUTPUT REPORT:

Neighbourhood_Group	Total_Properties	Estimated_Revenue
Dublin City Council	2804	3625650
Clare County Council	1001	390107
Galway City Council	680	318530
Donegal County Council	732	299410
Wexford County Council	666	297174
Wicklow County Council	487	271830
Fingal County Council	386	221698
Sligo County Council	421	168055
South Dublin County Council	196	130657
Tipperary County Council	340	129996
Cork City Council	79	74817
Offaly County Council	110	31743
Monaghan County Council	97	31006

Revenue by Neighbourhood Group

3. Location wise availability by Room Type and Neighbourhood Group Report?

```
/*Report 3: Average availability by Room Type and Neighbourhood_Group? */

SELECT
    li.Neighbourhood_Group AS Neighbourhood_Group,
    alp.Room_Type AS Room_Type,
    AVG(a.Availability_365) AS Avg_Availability
FROM
    Location_Info AS li
JOIN
    Airbnb_Listing_Properties AS alp ON li.LocationID = alp.LocationID
JOIN
    Availability AS a ON alp.AvailabilityID = a.AvailabilityID
GROUP BY
    li.Neighbourhood_Group, alp.Room_Type
ORDER BY
    Avg_Availability DESC;
```

Figure 4.3 Average availability by room type and Neighbourhood

OUTPUT REPORT:

Neighbourhood_Group	Room_Type	Avg_Availability
Galway City Council	Shared room	296
Wicklow County Council	Hotel room	293
Donegal County Council	Private room	276.2745
Dublin City Council	Hotel room	270.6
Monaghan County Council	Private room	262.1923
Donegal County Council	Entire home/apt	255.7533
Monaghan County Council	Shared room	248
Tipperary County Council	Private room	243.0313
Tipperary County Council	Entire home/apt	239.4109
Galway City Council	Hotel room	237.7143
Monaghan County Council	Entire home/apt	236.7571
Wicklow County Council	Private room	229.9252
Wexford County Council	Private room	229.0316
Offaly County Council	Entire home/apt	227.0519
Sligo County Council	Entire home/apt	223.7903
Offaly County Council	Private room	220.1515
Wicklow County Council	Entire home/apt	219.5119
Clare County Council	Entire home/apt	217.2221
Wexford County Council	Entire home/apt	214.1439

4. Top Rated Properties List Report?

```
/* report 4: Top rated Properties by Price and Room type?*/
SELECT
alp.Property_Name AS Property_Name,
 r.Reviews_per_Month AS Reviews_per_Month,
   li.Neighbourhood AS Neighbourhood,
 alp.Price AS Price, alp.Room_Type AS Room_Type
FROM   Airbnb_Listing_Properties AS alp
JOIN
 Reviews AS r ON alp.ReviewID = r.ReviewID
JOIN  Location_Info AS li ON alp.LocationID = li.LocationID
ORDER BY
 r.Reviews_per_Month DESC
LIMIT 20;
```

4.4 Top rated Properties

OUTPUT REPORT:

Property_Name	Reviews_per_Month	Neighbourhood	Price	Room_Type
Triple Room in Abbey St	50.49	North Inner City LEA-7	161	Private room
Locke Studio at Zanzibar Locke	35.54	North Inner City LEA-7	383	Entire home/apt
Heart of downtown Double-bed Rm+Privacy Bathroom	28.9	North Inner City LEA-7	128	Private room
Charming Triple Room	26.56	South East Inner City LEA-5	162	Private room
Locke Studio at Beckett Locke	26.08	North Inner City LEA-7	314	Entire home/apt
City Studio at Beckett Locke	19.93	North Inner City LEA-7	283	Entire home/apt
Double Room in Abbey St	17.21	North Inner City LEA-7	147	Private room
River Fane Cottage Retreat - Hot Tub~Sauna~Plunge	17.03	Carrickmacross-Castleblayney LEA-6	285	Entire home/apt
Ireland's #1 River Retreat Hot Tub~Sauna~Plunge	16.94	Carrickmacross-Castleblayney LEA-6	259	Entire home/apt
Charming Quadruple Room	16.88	South East Inner City LEA-5	179	Private room
Heart of downtown Triple-bed with Kitchen&Bathroom	16.73	North Inner City LEA-7	152	Private room
Wake to a new dawn in the den	15.54	Cashel-Tipperary LEA-7	83	Entire home/apt
Next stop Cliffs of Moher	15	Ennistimon LEA-4	115	Entire home/apt
2 Bedroom Suite at Beckett Locke	14.76	North Inner City LEA-7	460	Entire home/apt
Dublin Castle Suites â€" 1 Bed Suite	14.28	South East Inner City LEA-5	322	Entire home/apt
Guest Suite in Rosslare Harbour	13.67	Rosslare LEA-5	85	Private room
NEW Eyre Square En Suite Double with Many Extras	13.48	Galway City Central LEA-6	178	Private room
Blue Haven, stop and rest in Doolin	13.33	Ennistimon LEA-4	87	Private room
The Granary	13.21	Wicklow LEA-6	120	Entire home/apt

Top Rated Properties and Price by room type

Tableau Visualisations

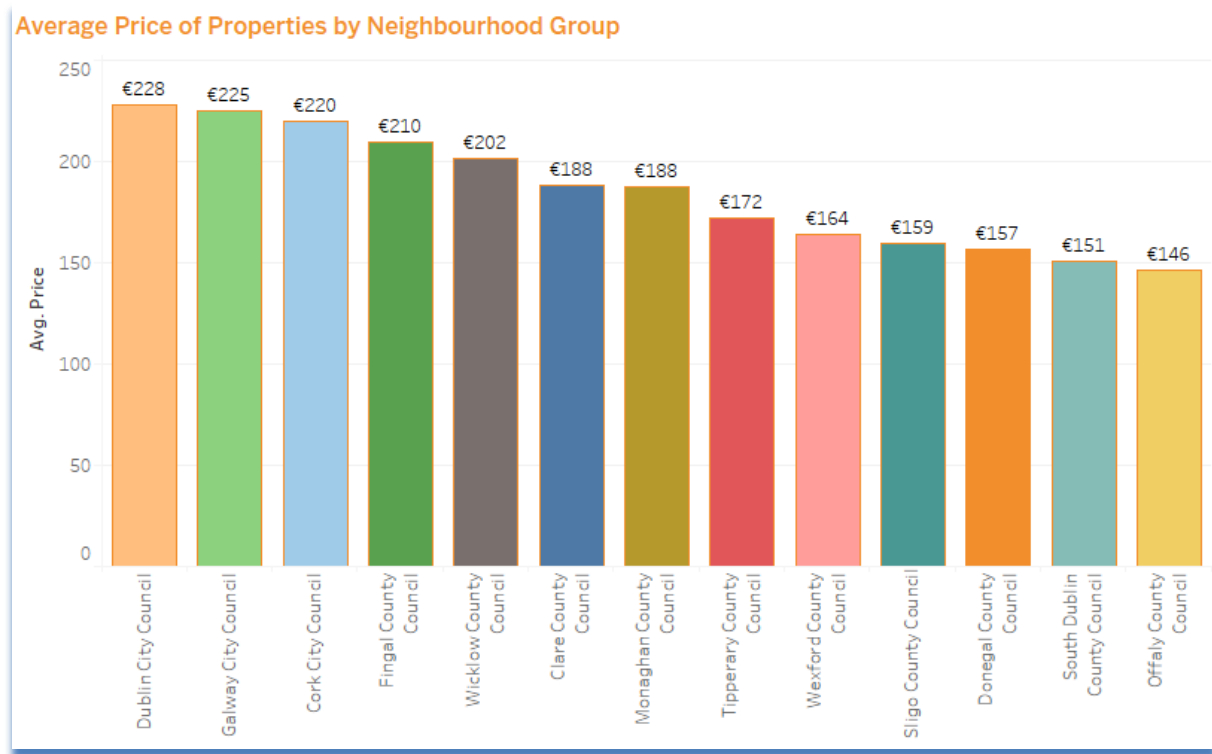1. **Average Price of Properties by Neighbourhood Group**

Figure 4.5: Average Price of Properties by Neighbourhood type

The above visualization shows the average price for properties across various neighbourhood groups. It indicates that Dublin City Council has the highest average price, standing at €228, followed closely by Galway City Council (€225) and Cork City Council (€220). At the opposite end of the scale, the average price in Offaly County Council is the lowest, with an average price of €145, which reflects a huge price gap within the regions. This chart shows the property pricing trend, with urban councils like Dublin and Galway maintaining a higher average property price compared to rural councils such as Offaly and South Dublin.
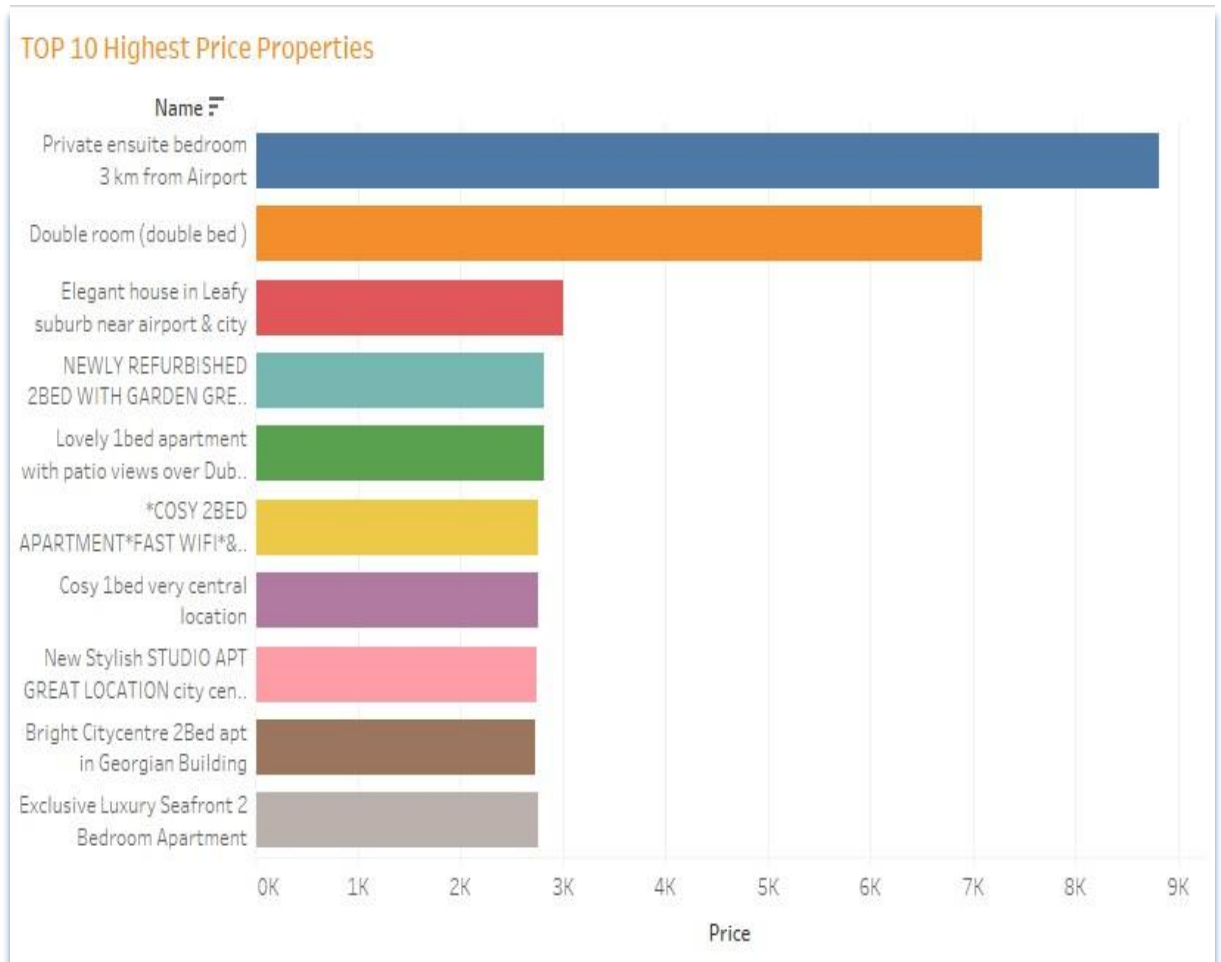
2. **Top 10 Highest Price Properties**

Figure 4.6: Top 10 Highest price Properties

This visualization presents the Top 10 Highest Price Properties, showing the most premium listings in the dataset. The property "Private ensuite bedroom 3 km from Airport" has the highest price, over €8,000, while "Double room (double bed)" is priced at about €7,500. Other properties, like "Elegant house in Leafy suburb near airport & city" and "Exclusive Luxury Seafront 2 Bedroom Apartment", are high-end properties that range between €3,000-€6,000. This analysis gives insights into the trends of the prices of luxury properties, thus helping in the identification of listings for high-budget customers.

3. Distribution Of Property Prices based location

Figure 4.7: Price Distribution by location

This visualization shows the distribution of property listings and prices across neighbourhood groups. It shows that the Dublin City Council has the highest count of prices, which means a high volume in this area, followed by Cork City Council and Galway City Council. The least are Offaly County Council and Tipperary County Council, which have significantly fewer property listings and prices. This distribution reflects regional differences in property availability and demand, with urban areas like Dublin leading in property density and pricing trends.

4. **Revenue Trends for Month**

Figure 4.8: Revenue Trends per month

The above visualization presents the trend of monthly revenues of the Airbnb, showing up and downs of average revenues throughout the year. A September peak is prominent, at an average revenue of €6,483, likely due to seasonal demand. This drops in October to €2,312, while August has the lowest recorded revenue of €457. The early months, especially February (€4,582) and January (€3,545), also exhibit higher revenue compared to mid-year months like May (€622) and June (€1,018). These trends put forth possible seasonality in property demand and revenue generation.

Tableau Dashboard Airbnb Data Analysis
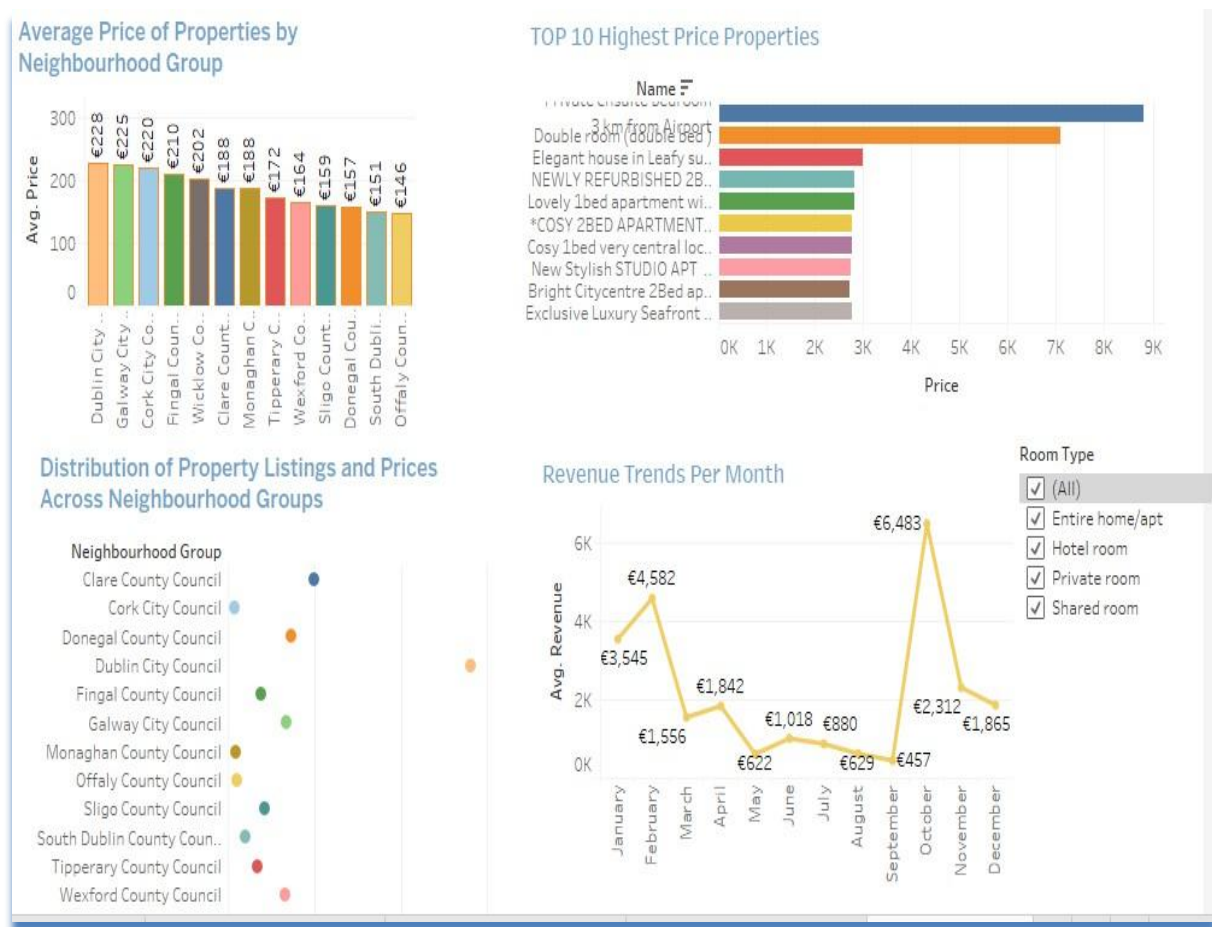
Figure 4.9 : Dashboard of AIRBNB Data Analysis

This dashboard presents, through four visualizations, an analysis of property trends. It portrays the average price of properties across neighbourhoods to show the pricing patterns and differences. The top properties chart shows premium listings for insights into high-end property trends. A distribution chart shows the spread of property counts and prices across neighbourhood groups to depict regional patterns. The monthly revenue trends chart highlights seasonality in property revenue. The dashboard is interactive, enabling the user to filter data by room type for specific insights, hence very useful in property market analysis and decision-making.

5: Performance Comparison: Relational vs. Graph Databases

5.1 Relational Database SQL Server

Relational databases, commonly known as RDBMS, are designed to store information in a tabular structure. Each table would have rows and columns. A row represents a single record, and a column denotes an attribute of the data. The data in the relational database is usually controlled with a schema that defines the structure of tables and their relationships. These allow relationships between different tables by using primary keys and foreign keys. SQL or Structured Query Language is a standard language used in creating queries, manipulating and managing the data.

```
/* Seven Best Data Analysis Queries based on Airbnb Property Listening*/
/*1. Top 10 Most Expensive Properties ? */
select id,name,price
from airbnb_data
ORDER BY price DESC
LIMIT 10;
```

OUTPUT:

id	name	price ▾ 1
36872270	Private ensuite bedroom 3 km from Airport	8820
42701738	Double room (double bed)	7095
18713449	Elegant house in Leafy suburb near airport & city	3000
4247765	NEWLY REFURBISHED 2BED WITH GARDEN GREAT LOCATION	2823
7576701	Lovely 1bed apartment with patio views over Dublin	2817
15854530	*COSY 2BED APARTMENT*FAST WIFI*&TV*TRENDY RANELAGH	2768
7572709	Cosy 1bed very central location	2766
8776906	New Stylish STUDIO APT GREAT LOCATION city centre	2745
7589060	Central 1bedroom flat with small patio & fast WIFI	2731
7577893	Bright Citycentre 2Bed apt in Georgian Building	2731

Figure 5.1 Top 10 most expensive Properties

Figure 5.1 shows that the top 10 most expensive rental property listings, whose prices range from €2731 to €8820. The most expensive is a private en-suite bedroom near the airport for €8820, while the cheapest, priced at €2731, are apartments in the city center.

```
/*2.Average Price Per Room type?*/
select room_type,AVG(price) as AVG_Price
from airbnb_data
GROUP BY room_type
ORDER BY AVG_Price DESC;
```

OUTPUT:

room_type	AVG_Price ⌄ 1
Hotel room	242.2333
Entire home/apt	224.0314
Private room	128.9169
Shared room	126.2000

Figure 5.2: Average Price by room type

Figure 5.2 summarizes the average prices for different room types. The average price is highest for hotel rooms, at €242.23, followed by entire homes/apartments at €224.03. Private and shared rooms are cheaper, averaging €128.92 and €126.20, respectively.

```
/*3.Top 10 Neighbourhoods with the Most number of Properties ?*/
SELECT neighbourhood,neighbourhood_group,COUNT(*) as Property_Count
from airbnb_data
GROUP BY neighbourhood
ORDER BY Property_Count DESC
LIMIT 10;
```

neighbourhood	neighbourhood_group	Property_Count ⌄ 1
North Inner City LEA-7	Dublin City Council	748
Ennistimon LEA-4	Clare County Council	732
South East Inner City LEA-5	Dublin City Council	646
Glenties LEA-6	Donegal County Council	473
Galway City Central LEA-6	Galway City Council	469
South West Inner City LEA-5	Dublin City Council	345
Pembroke LEA-5	Dublin City Council	292
Carndonagh LEA-4	Donegal County Council	245
Cabra-Glasnevin LEA-7	Dublin City Council	229
Ballymote-Tobercurry LEA-7	Sligo County Council	203

Figure 5.3: Top neighbourhoods with most no of Properties

Figure 5.3 represents the summary of the distribution in the number of properties in different neighbourhoods and their respective local authorities. The North Inner City LEA-7 under Dublin City Council has the highest property count with a total of 748 properties, while the second highest is Ennistimon LEA-4 under Clare County Council, which has 732 properties. Other areas include South East Inner City LEA-5 and Glenties LEA-6, having 646 and 473, respectively. The lowest counts are recorded in areas such as Ballymote-Tubbercurry LEA-7 in Sligo County Council and Cabra-Glasnevin LEA-7 in Dublin City Council, with 203 and 229 properties, respectively.

```
/*4.Total  Available Properties by Room type?*/
SELECT room_type,COUNT(*) as Total_Available_Properties
from airbnb_data
GROUP BY room_type
ORDER BY Total_Available_Properties DESC;
```

OUTPUT:

room_type	Total_Available_Properties ▼ 1
Entire home/apt	5864
Private room	2095
Hotel room	30
Shared room	10

Figure 5.4: Total available properties from Airbnb by Room type

Figure 5.4 summarizes the total available properties on Airbnb by room type. The majority of the available properties are entire homes/apartments, with a total of 5,864 properties. Private rooms come second with 2,095 properties, while hotel rooms and shared rooms are very low in number, standing at 30 and 10 properties, respectively. This shows that entire homes/apartments and private rooms are most preferred among the property types.

```
/*5 what are Top 20 Lowest Price Properties with minimum stay Requirements?
*/
SELECT id AS Property_ID,name AS Property_Name,neighbourhood,price AS Price,minimum_nights AS Minimum_Stay_Nights
FROM airbnb_data
ORDER BY price ASC
LIMIT 10;
```

OUTPUT:

Property_ID	Property_Name	neighbourhood	Price	Minimum_Stay_Nights
150195	Comfy home near Dublin city	Ballyfermot-Drimnagh LEA-5	75	1
724794	Bring a cheeky rascal - pet friendly	Ballyfermot-Drimnagh LEA-5	75	1
2150064	Cozy ensuite room, with mini fridge in Dublin 2.	South East Inner City LEA-5	75	30
3088138	ThornCastle - Small Double Ensuite Room	South East Inner City LEA-5	75	1
4302167	Cosy Double in penthouse	North Inner City LEA-7	75	7
6230130	Safe and comfortable single room with OWN bathroom	North Inner City LEA-7	75	2
6393117	Apartment in converted Georgian house.	North Inner City LEA-7	75	15
6788231	1 Bedroom with King bed in refurbished townhouse	North Inner City LEA-7	75	16
6828409	amazing room in terrace house	North Inner City LEA-7	75	1
10901830	Double Bed in Modern Flat, close to City Center	North Inner City LEA-7	75	2
11110203	Apartment in the City Center	North Inner City LEA-7	75	20
12831874	Christchurch Cathedral Area	South West Inner City LEA-5	75	5

Figure 5.5: Top 12 Lowest Price Properties

Figure 5.5 shows the 12 most affordable listings on Airbnb, which are €75 per night. They are located in Ballyfermot-Drimnagh LEA-5, South East Inner City LEA-5, and North Inner City LEA-7. The minimum nights required range from 1 to 20 nights. Properties like "Comfy home near Dublin city" and "Bring a cheeky rascal - pet friendly" have the lowest minimum stay of 1 night, while others like "Apartment in the City Center" require up to 20 nights.

```
/*6.What are the Properties with Highest Avg reviews per month*/
SELECT id as Property_ID ,name as Property_Name,AVG(reviews_per_month) as Avg_Reviews_per_Month
from airbnb_data
GROUP BY Property_ID,Property_Name
ORDER BY Avg_Reviews_per_Month DESC
LIMIT 10;
```

OUTPUT:

Property_ID	Property_Name	Avg_Reviews_per_Month ▾ 1
5710121	Triple Room in Abbey St	50.4900016784668
8062455	Locke Studio at Zanzibar Locke	35.540000915527344
9730671	Heart of downtown Double-bed Rm+Privacy Bathroom	28.899999618530273
3363852	Charming Triple Room	26.559999465942383
7080298	Locke Studio at Beckett Locke	26.079999923706055
7071316	City Studio at Beckett Locke	19.93000030517578
5558423	Double Room in Abbey St	17.209999084472656
26447755	River Fane Cottage Retreat - Hot Tub~Sauna~Plunge	17.030000686645508
26464386	Ireland's #1 River Retreat Hot Tub~Sauna~Plunge	16.940000534057617
5321592	Charming Quadruple Room	16.8799991607666

Figure 5.6: Properties with Highest AVG Reviews Per Month

Figure 5.6 shows the top properties by average review in a month on Airbnb. The "Triple Room in Abbey St" has the highest, with an average review of about 50.49 in a month, followed by "Locke Studio at Zanzibar Locke" with 35.54, while "Heart of downtown Double-bed Rm+Privacy Bathroom" had 28.89 and "Charming Triple Room" had 26.56. Listings like "Ireland's #1 River Retreat Hot TubSaunaPlunge" and "Charming Quadruple Room" have an average of around 17 reviews a month. This indicates that guests in these listings like the products and services being offered by these listings.

```
/*7.Average Availability by Neighbourhood Group */
SELECT neighbourhood_group,AVG(availability_365) as Average_Availability
from airbnb_data
GROUP BY neighbourhood_group
ORDER BY Average_Availability DESC;
```

OUTPUT:

neighbourhood_group	Average_Availability ▾ 1
Donegal County Council	257.1831
Monaghan County Council	243.6907
Tipperary County Council	239.6441
Offaly County Council	224.9818
Wicklow County Council	222.2526
Sligo County Council	218.6770
Wexford County Council	216.2523
Clare County Council	205.9101
South Dublin County Council	183.0102
Galway City Council	172.9088
Fingal County Council	171.0440
Dublin City Council	144.7322
Cork City Council	143.8987

Figure 5.7: Neighbourhood Group with average Availability

Figure 5.7 shows the average availability of properties in various neighborhood groups. Donegal County Council has the highest average, with 257.18 days, followed by Monaghan County Council with 243.69 days and Tipperary County Council with 239.64 days. Offaly County Council and Wicklow County Council also have high availability, averaging 224.98 and 222.25 days, respectively. At the bottom, Dublin City Council and Cork City Council have the least availability, averaging 144.73 and 143.90 days, respectively. This reflects a very strong variation in the availability of properties across regions.

Graph Database:

Graph databases are specifically designed to manage complex relationships and interconnected data in a very efficient way. Unlike relational databases, which store data in tabular form with predefined rows and columns, graph databases have nodes, edges, and properties as the core model. The nodes can be entities, for example, people, products, or locations, while the edges represent the connections among them, and the properties set details of the nodes and edges. The rationale behind this design is to let graph databases store and traverse relationships directly, making them highly effective in scenarios where the connection of data points is crucial.

One of the significant benefits when working with graph databases involves their capability to work with highly connected data efficiently.

5.2 Cypher Query Language (CQL) Analysis in Neo4j

```
LOAD CSV WITH HEADERS FROM 'file:///Airbnb_Data.csv' AS row
MERGE (p:Property {id: toInteger(row.id)})
SET p.name = row.name,
    p.room_type = row.room_type,
    p.price = toInteger(row.price),
    p.minimum_nights = toInteger(row.minimum_nights),
    p.number_of_reviews = toInteger(row.number_of_reviews),
    p.reviews_per_month = toFloat(row.reviews_per_month),
    p.availability_365 = toInteger(row.availability_365),
    p.number_of_reviews_ltm = toInteger(row.number_of_reviews_ltm),
    p.neighbourhood = row.neighbourhood,
    p.neighbourhood_group = row.neighbourhood_group;
```

Figure 5.7: Building Listing Nodes for Loading Data

This loading creates Listing nodes from the data in CSV, important attributes such as property ,host ,location ,room_type, price, and availability_365. At this stage of the fundamental step, data is getting organized in this dataset and forming a distinction between distinct Listing entities where further building of relationships becomes possible.
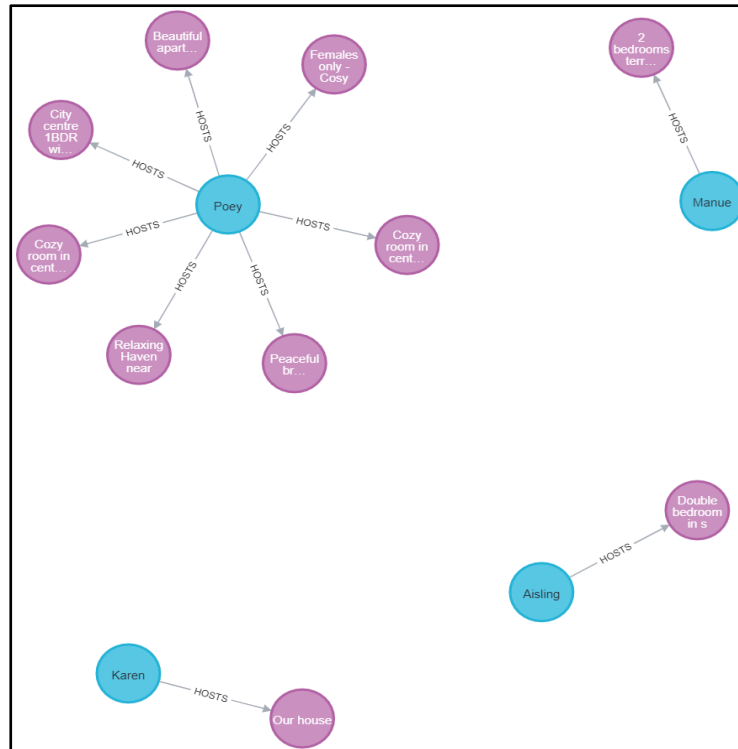

Figure 23: HOSTS Links (Host to Listing)

From this process can observe that the host "Poey" runs multiple listings through having relations with seven different properties that he is handling, making it a professional host account actively managing many different kinds of accommodations. Meanwhile, other hosts, "Manue," "Karen," and "Aisling" handle only one listing, making them more or less casual hosting approaches.
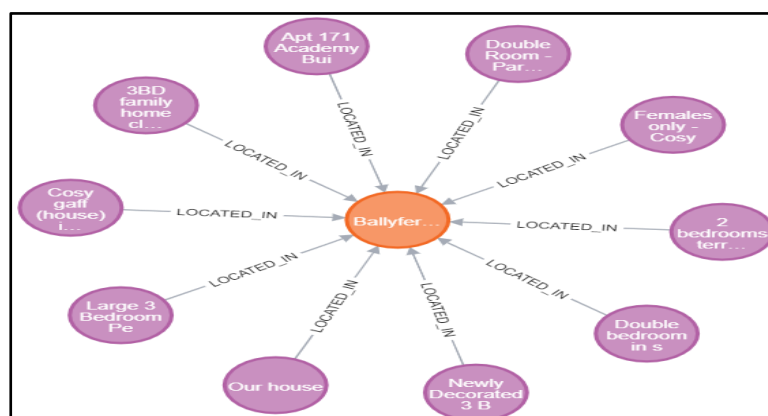

Figure 5.8: LOCATED_IN Relationships (Listing to Neighbourhood)

The second graph is LOCATED_IN, which displays how different listings are geographically concentrated in the neighbourhood "Ballyfermot." This neighbourhood is linked to a high number of listings, suggesting its popularity or suitability for short-term rentals.
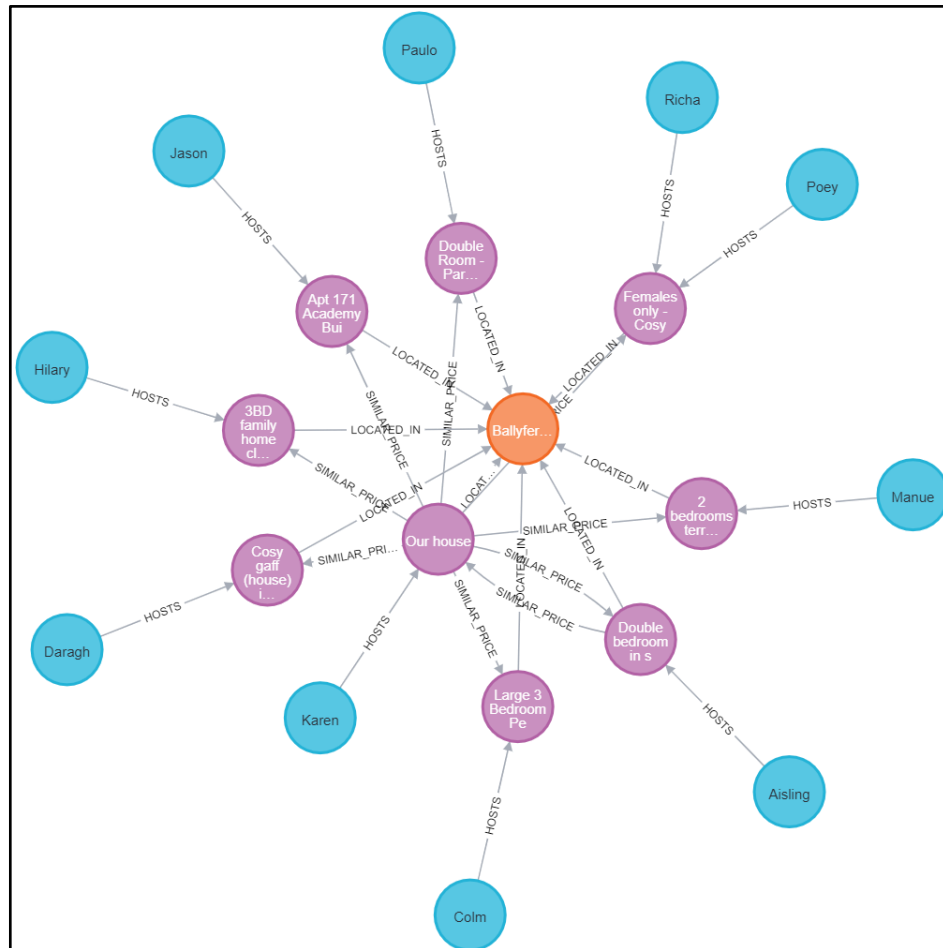

Figure 5.9: Combined View of HOSTS and LOCATED_IN Relationships

The joint view provides a holistic design, showing how hosts lead to neighbourhoods through their listings. This interlinked graph displays a scenario where some hosts may have their listings in the same neighbourhood while others list properties within various regions.

Figure 5.10 : General Relationship Discovery

The final graph is a view of the entire nodes and relationships. It shows that all these nodes connect through the relationship HOSTS and LOCATED_IN. This also gave proof that the system maintains integrity, and all the data proved to be integrated into one system, with the listing system from Airbnb.

Cypher Queries:

```
//1.What are the top 10 most Most expensive Properties?

match(p:Property)
Return p.id as ID,p.name as NAME,p.price as PRICE
Order by p.price DESC
Limit 10;
```

OUTPUT:

ID	NAME	PRICE
36872270	"Private ensuite bedroom \n3 km from Airport"	8820
42701738	"Double room (double bed)"	7095
18713449	"Elegant house in Leafy suburb near airport & city"	3000
4247765	"NEWLY REFURBISHED 2BED WITH GARDEN GREAT LOCATION"	2823
7576701	"Lovely 1bed apartment with patio views over Dublin"	2817
15854530	"*COSY 2BED APARTMENT*FAST WIFI*&TV*TRENDY RANELAGH"	2768
7572709	"Cosy 1bed very central location"	2766
8776906	"New Stylish STUDIO APT GREAT LOCATION city centre"	2745
7589060	"Central 1bedroom flat with small patio & fast WIFI"	2731
7577893	"Bright Citycentre 2Bed apt in Georgian Building"	2731

Figure 5.11: Top 10 Most Expensive Properties

Figure 5.11 shows the properties with the highest and lowest average availability. Donegal County Council has the highest average availability, with 257.18 days, followed by Monaghan County Council at 243.69 days. On the other hand, Dublin City Council has the lowest average availability, with only 144.73 days, closely followed by Cork City Council at 143.90 days. This means that, in general, properties in rural areas are more available than in urban centers such as Dublin and Cork.

```
//2.Average Price per room type?
match(p:Property)
Return p.room_type as Room_Type,avg(p.price) as Avg_Price
Order by Avg_Price desc;
```

```
|Room_Type          |Avg_Price            |

|"Hotel room"       |242.23333333333332   |

|"Entire home/apt"  |224.03137789904477   |

|"Private room"     |128.91694510739882   |

|"Shared room"      |126.2                |
```

Figure 5.12: Average Price Per Room type

Figure summarizes the average prices of different room types on Airbnb. The "Hotel rooms" room type is the most expensive, with an average price of €242.23, while the second most expensive are "Entire homes/apartments," at €224.03. "Private rooms" are cheaper, with an average of €128.92, and "Shared rooms" have the lowest average price of €126.20.

```
//3.Top 10 Neighbourhood with Most number of Properties
MATCH(p:Property)
Return p.neighbourhood as Neighbourhood,p.neighbourhood_group as Neighbourhood_Group,COUNT(*) AS Property_Count

order by Property_Count DESC
limit 10;
```

OUTPUT:

Neighbourhood	Neighbourhood_Group	Property_Count
"North Inner City LEA-7"	"Dublin City Council"	748
"Ennistimon LEA-4"	"Clare County Council"	732
"South East Inner City LEA-5"	"Dublin City Council"	646
"Glenties LEA-6"	"Donegal County Council"	473
"Galway City Central LEA-6"	"Galway City Council"	469
"South West Inner City LEA-5"	"Dublin City Council"	345
"Pembroke LEA-5"	"Dublin City Council"	292
"Carndonagh LEA-4"	"Donegal County Council"	245
"Cabra-Glasnevin LEA-7"	"Dublin City Council"	229
"Ballymote-Tobercurry LEA-7"	"Sligo County Council"	203

Figure 5.13: Neighbourhood with most number of Properties

Figure 5.13 shows the count of properties per neighborhood and respective local authority. The property count for "North Inner City LEA-7" under Dublin City Council is the highest, with a total of 748 properties, while the lowest is "Ballymote-Tubbercurry LEA-7" under Sligo County Council, with just 203 properties.

```
//4.Total Available Properties by Room Type?
MATCH(p:Property)
Return p.room_type as Room_Type,count(*) as Total_Available_Properties
ORDER BY Total_Available_Properties DESC;
```

OUTPUT:

Room_Type	Total_Available_Properties
"Entire home/apt"	5864
"Private room"	2095
"Hotel room"	30
"Shared room"	10

Figure 5.14: Room Type Total Available Properties

Figure 5.14 provides an overview of the total available properties on Airbnb by room type. Most of the listings fall under the category "Entire home/apt," with 5,864 properties available; this makes it the most common room type available. On the contrary, "Shared rooms" have the least number of listings, totaling only 10 properties. This simply means that entire homes/apartments are preferred among property types.

```
//5.What are the Top 12 Lowest Price Properties along with stay requirements?
MATCH(p:Property)
RETURN p.id as Property_ID,p.name as Property_Name,p.neighbourhood,p.price as Price,p.minimum_nights
order by Price ASC
LIMIT 12;
```

OUTPUT:

Property_ID	Property_Name	Nighbourhood	Price	Minimum_Nights
11110203	"Apartment in the City Center"	"North Inner City LEA-7"	75	20
6788231	"1 Bedroom with King bed in refurbished townhouse"	"North Inner City LEA-7"	75	16
6828409	"amazing room in terrace house"	"North Inner City LEA-7"	75	1
10901830	"Double Bed in Modern Flat, close to City Center"	"North Inner City LEA-7"	75	2
4302167	"Cosy Double in penthouse"	"North Inner City LEA-7"	75	7
2150064	"Cozy ensuite room, with mini fridge in Dublin 2."	"South East Inner City LEA-5"	75	30
6230130	"Safe and comfortable single room with OWN bathroom"	"North Inner City LEA-7"	75	2
3088138	"ThornCastle - Small Double Ensuite Room"	"South East Inner City LEA-5"	75	1
6393117	"Apartment in converted Georgian house."	"North Inner City LEA-7"	75	15

Figure 5.15: Top 12 Lowest Price Properties

Figure 5.15 Low-priced properties are listed that have consistent pricing, at €75 per night. Starting with the "Apartment in the City Center," located in "North Inner City LEA-7," is an option for visitors who will stay in Dublin longer because it calls for a minimum stay of 20 nights. While the "Apartment in converted Georgian house" is also situated in "North Inner City LEA-7," but requires only a minimum stay of 15 nights, making this home suitable for medium-term visitors.

```
//6.What are the Properties with Highest Average Reviews per month?
MATCH (p:Property)
WITH p.id AS Property_ID, p.name AS Property_Name, AVG(p.reviews_per_month) AS Avg_Reviews_Per_Month
RETURN Property_ID, Property_Name, Avg_Reviews_Per_Month
ORDER BY Avg_Reviews_Per_Month DESC
LIMIT 10;
```

OUTPUT:

```
|Property_ID| Property_Name                                       |Avg_Reviews_Per_Month|
|5710121    |"Triple Room in Abbey St"                            |50.49                |
|8062455    |"Locke Studio at Zanzibar Locke"                     |35.54                |
|9730671    |"Heart of downtown Double-bed Rm+Privacy Bathroom"   |28.9                 |
|3363852    |"Charming Triple Room"                               |26.56                |
|7080298    |"Locke Studio at Beckett Locke"                      |26.08                |
|7071316    |"City Studio at Beckett Locke"                       |19.93                |
|5558423    |"Double Room in Abbey St"                            |17.21                |
|26447755   |"River Fane Cottage Retreat - Hot Tub~Sauna~Plunge"  |17.03                |
|26464386   |"Ireland's #1 River Retreat Hot Tub~Sauna~Plunge"    |16.94                |
```

Figure 5.16: Property Average Reviews per Month

Figure 24 shows the properties which have the highest average review per month. The first property, "Triple Room in Abbey St", leads the chart with as many as 50.49 reviews per month, which indicates its popularity and customer satisfaction. The last property, "Ireland's #1 River Retreat Hot TubSaunaPlunge", has an average of 16.94 reviews per month, indicating a consistent level of engagement.

```
//7.Average Availability by Neighbourhood Group
MATCH(p:Property)
RETURN p.neighbourhood_group as Neighbourhood_Group ,AVG(p.availability_365) as Average_Availability
ORDER BY Average_Availability DESC;
```

```
Neighbourhood_Group             |Average_Availability|

"Donegal County Council"        |257.18306010928944  |

"Monaghan County Council"       |243.69072164948446  |

"Tipperary County Council"      |239.64411764705864  |

"Offaly County Council"         |224.98181818181817  |

"Wicklow County Council"        |222.25256673511274  |

"Sligo County Council"          |218.6769596199525   |

"Wexford County Council"        |216.25225225225228  |

"Clare County Council"          |205.91008991009025  |

"South Dublin County Council"   |183.0102040816327   |

"Galway City Council"           |172.9088235294117   |
```

Figure 5.17: Average Availability by Neighbourhood Group

Figure 5.17 gives the average availability of properties across different neighborhood groups. It can be observed that "Donegal County Council" has the highest average of 257.18 days, thus reflecting a very strong supply in terms of available properties in this region. In contrast, the "Galway City Council" has the lowest average of 172.91 days, which could reflect poor availability of properties within the city's urban environment.

Comparison between Relational vs. Graph Databases

The Neo4j graph database and the SQL Server relational database allow for some very unique strengths in analysing Airbnb data. Neo4j is really good at visualizing complex relationships, like HOSTS and LOCATED_IN, which enables an intuitive understanding of the connections between hosts, listings, and neighbourhoods. Ideal for network-based queries or quick identification of patterns-for instance, multi-listing hosts or geographic clusters. SQL Server is of very great use as the tables are well structured, where we can aggregate and summarize data

in terms of average price or count of reviews. Neo4j does relationship-based analysis while SQL Server does statistical or tabular insights.

5. Security Authentication:

For protecting the database against unauthorized access and data modification, it's a good practice to create users and regulate their access. To do that, users can be created through commands like CREATE USER, using the username, password, and host. Give permissions using GRANT to provide levels of access such as reading with SELECT, modifying with INSERT, UPDATE, DELETE, or ALL PRIVILEGES. Permissions can also be revoked using the REVOKE command to restrict access when necessary. Best practices include granting minimal permissions required for a user's role, regular audits of privileges, and enforcing strong password policies. This helps maintain data security and integrity.

20025600@dbs.ie	localhost	Yes	ALL PRIVILEGES	Yes	Edit privileges	Export	Lock
20035725@mydbs.ie	localhost	Yes	ALL PRIVILEGES	Yes	Edit privileges	Export	Lock
20036237@mydbs.ie	localhost	Yes	ALL PRIVILEGES	Yes	Edit privileges	Export	Lock
20036676@mydbs.ie	localhost	Yes	ALL PRIVILEGES	Yes	Edit privileges	Export	Lock
20037113@mydbs.ie	localhost	Yes	ALL PRIVILEGES	Yes	Edit privileges	Export	Lock

6. Conclusion

In this particular assignment, learned about the Data Warehouse and Relational Databases, and then compared Relational Databases with Graph Databases. In the case of two databases, the disparities in facets of relationship queries were brought out by executing the two programs created. The project highlighted the need to recommend the most adequate database technology given certain usage characteristics and hence make it easier to store, access and analyse data in a way that fits business needs.

7. Reference List

Akter, S., Michael, K., Uddin, M.R., McCarthy, G. and Rahman, M., 2022. Transforming business using digital innovations: The application of AI, blockchain, cloud and data analytics. *Annals of Operations Research*, pp.1-33.

Ciampi, F., Demi, S., Magrini, A., Marzi, G. and Papa, A., 2021. Exploring the impact of big data analytics capabilities on business model innovation: The mediating role of entrepreneurial orientation. *Journal of Business Research*, *123*, pp.1-13.

Guo, C. and Chen, J., 2023. Big data analytics in healthcare. In *Knowledge technology and systems: Toward establishing knowledge systems science* (pp. 27-70). Singapore: Springer Nature Singapore.

Maddireddy, B.R. and Maddireddy, B.R., 2022. Real-Time Data Analytics with AI: Improving Security Event Monitoring and Management. *Unique Endeavor in Business & Social Sciences*, *1*(2), pp.47-62.