

CCT College Dublin

## ARC (Academic Research Collection)

---

ICT

---

Summer 2022

### Defining financial risks and market trends through predictive data analysis

Marcelle Louise  
*CCT College Dublin*

Luciana Teixeira  
*CCT College Dublin*

Muhammad Shahbaz  
*CCT College Dublin*

Giovanni Andrade  
*CCT College Dublin*

Follow this and additional works at: <https://arc.cct.ie/ict>



Part of the [Computer Sciences Commons](#)

---

#### Recommended Citation

Louise, Marcelle; Teixeira, Luciana; Shahbaz, Muhammad; and Andrade, Giovanni, "Defining financial risks and market trends through predictive data analysis" (2022). *ICT*. 31.

<https://arc.cct.ie/ict/31>

This Poster is brought to you for free and open access by ARC (Academic Research Collection). It has been accepted for inclusion in ICT by an authorized administrator of ARC (Academic Research Collection). For more information, please contact [debora@cct.ie](mailto:debora@cct.ie).

**MODULE** Problem Solving For Industry  
**SUPERVISOR** Muhammad Iqbal

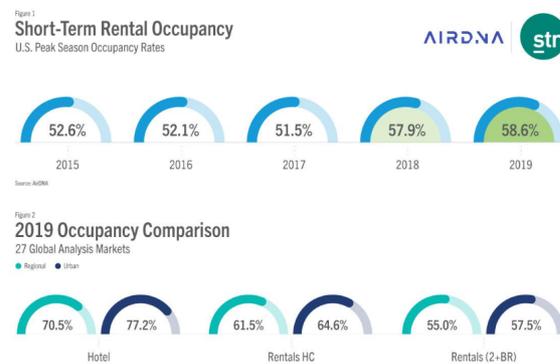
**RESEARCHERS** Marcelle Louise, Luciana Teixeira, Muhammad Shahbaz, Giovanni Andrade

## INTRODUCTION

Despite the challenges posed by the COVID-19 pandemic, a study conducted by STR and AirDNA found that house rentals outperformed hotels; among the most renowned platforms is Airbnb, which has become a symbol of the sharing economy and has changed the way people travel.

This project focuses on Dublin short term rental market opportunities, by developing pricing and rate occupancy prediction models based on machine learning approaches to identify patterns that may impact or aid users in making smarter and cost-effective decisions.

The concept of this research is to show the financial feasibility of data services, as well as how data science can improve business and operational efficiency.



## OBJECTIVE

The aim of this project is to use Machine Learning technologies and online property website datasets to determine financial risks and market trends for real estate investors in Ireland.

## BACKGROUND

### REQUIREMENTS

- Provide an easy and efficient solution for users who are searching for properties.
- Design and deploy a prototype that shows the predictions found through Machine Learning techniques.
- Display information regarding a certain property.
- Provide customized search by adding filters to fit users budget, needs, and general lifestyle.

### MACHINE LEARNING

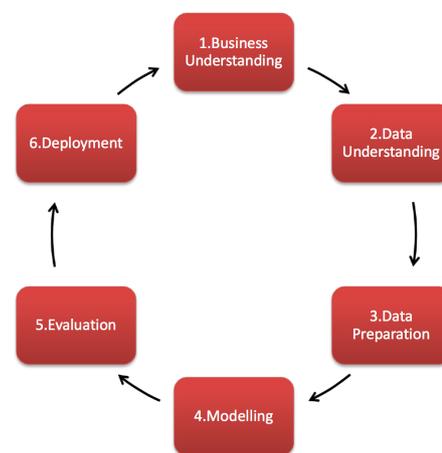
Since this project consists of analysing and combining six datasets, multiple modelling techniques were used to test the accuracy of these predictions:

- Linear Regression
- Multiple Regression
- K-Nearest Neighbour
- Lasso and Ridge Regression
- Sentiment Analysis

## METHODOLOGY

### • CRISP-DM

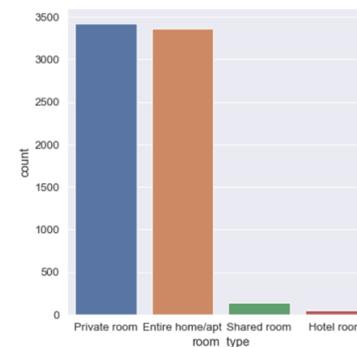
This framework is divided into six phases that correspond to the data science life cycle. Each phase was completed in order to ensure the project's integrity and long term planning.



## FINDINGS

Through Python visualizations we discovered patterns and trends that had a higher likelihood of impacting the final prediction models by analysing each dataset.

- **Most Popular Types Of Property**  
Private Rooms and Entire Houses
- **Average Price Per Neighbourhood**  
From €88.72 up to €359.31
- **Price VS Location**  
Dublin South East Area:  
Over €130.00 per night



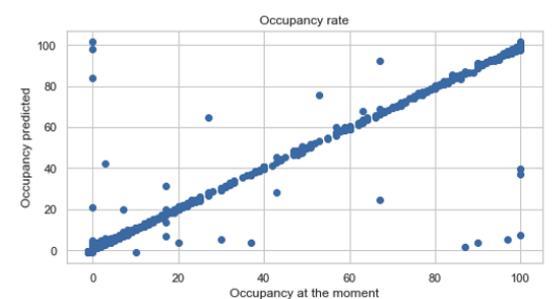
## DEPLOYMENT TOOLS



## RESULTS

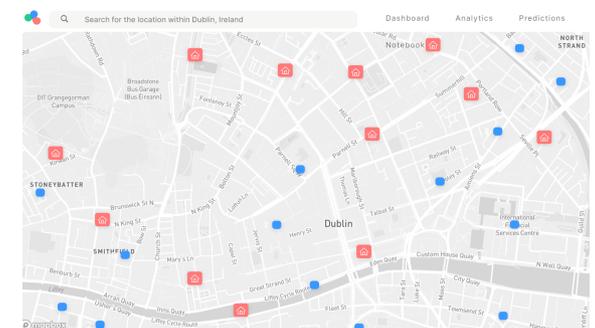
Between the number of modelling methods used, the following options have achieved a greater accuracy:

- **Price Prediction**  
K-Nearest Neighbour: 85%
- **Comments Review**  
Sentiment Analysis: 98%
- **Occupancy Rate**  
Lasso Regression: 93%



## DASHBOARD

### • View Prototype



## CONCLUSION

As previously discussed, the real estate business has witnessed a major development in the use of data in recent years, and purchases must be thoroughly analyzed, with comprehensive analysis and evaluation of the best investment according to the investor's needs or preferences. By the end of this project, we felt positive that the outcomes outlined have been achieved. The application deployed shows a solution for real estate investments that employs modern technologies to support investors in general in making data-driven decisions.

## REFERENCES

- Sanford, W. and DuBois, D., 2022. COVID-19 impact on hotels and short-term rentals. [online] STR. Available at: <https://str.com/whitepaper/covid-19-impact-on-hotels-and-short-term-rentals-airdna> [Accessed 15 May 2022].
- Luna, Z. (2021). Understanding CRISP-DM and its importance in Data Science projects. [online] Analytics Vidhya. Available at: https://medium.com/analytics-vidhya/understanding-crisp-dm-and-its-importance-in-data-science-projects-91c8742c9f9b. [Accessed 01 May 2022].
- Insideairbnb.com. 2022. [online] Available at: <http://insideairbnb.com/dublin> [Accessed 15 May 2022].
- Bevans, R. (2020). Simple Linear Regression | An Easy Introduction & Examples. [online] Scribbr. Available at: https://www.scribbr.com/statistics/simple-linear-regression/. [Accessed 02 May 2022].
- Goel, A. (2022). 10 Best Web Development Frameworks. Hackr.io. Available at: https://hackr.io/blog/web-development-frameworks [Accessed 14 May 2022].

## ACKNOWLEDGMENTS

We wish to express our sincere gratitude to the professors in CCT, and particularly to our supervisor, Dr Muhammad Iqbal. His guidance and mentorship made this experience extremely memorable and rewarding to all of us.

The completion of this undertaking could not have been possible without the participation and assistance of the members of this group. Each one's contributions are sincerely appreciated and gratefully acknowledged.