A Statistical Analysis of Dublin Rental Properties and Machine Learning Prediction Model for Price

A comprehensive study of properties to rent in Dublin - including insights into their attributes as well as building a machine learning prediction model to identify property feature importance in predicting price.

Colin Hehir 20213371
School of Computing
Dublin City University, Dublin, Ireland
colin.hehir3@mail.dcu.ie

***Abstract* — In today's world, people in Ireland and further afield are becoming more likely to reside in the capital city, Dublin, for educational and professional purposes. As a result, due to the high demand to rent a property in the area, they may struggle financially due to a lack availability and high prices, while also being unsure of what features in a rental property they should be looking for when searching. This study suggests that machine learning techniques can be used to predict a property's rental price per bedroom in Dublin. Using a Random Forest Regressor, it was demonstrated that the average accuracy of the prediction was more than 84.82%, suggesting that it may be useful for individuals who may be interested in renting a property in Dublin, or indeed specific locations or property type in the area, which was influenced by the several statistical tests also carried out in the study.**

CCS CONCEPTS* • *Mathematics of computing* → *Hypothesis testing* • *Computing methodologies* → *Machine Learning

Keywords* —*Cost of Living, Rent Prices, Housing Properties, Dublin, Random Forest Regression, Correlation Coefficient

I. INTRODUCTION AND MOTIVATION

The rise in property prices has been a prominent issue in Ireland over the past few years. With reports that this is to continue to rise, this is a very important issue to gain further insights into [1]. The Covid-19 pandemic has influenced both Dublin's rental market, seeing prices surge and the number of houses available decreasing. The lowest number since records began in 2006 showed that just 2,455 rental homes were available throughout Ireland on 1 August 2021 [2]. Reports on Daft.ie show that this unprecedented scarcity of rental homes caused the average national rents to climb by 5.6% in one year [3]. The persistent imbalance between the housing supply and demand leads to affordability pressures for renters and as a result, this study's objective was to gain a deeper insight into what features influences price, as well as what characteristics in a property are associated with each other to aid individuals in their property search around Ireland's capital.

II. RELATED WORK

In this area of research, a model for predicting rental prices in Amsterdam had an average accuracy of 98.309% by measuring the price by square meter [4]. Elsewhere, JLL, a real-estate advisors and professionals firm based in Ireland, published a residential forecasting model that used 15 economic and real estate market parameters to provide a three-year projection for rentals and rates. The complicated model additionally considered industry knowledge and opinion in order to determine and evaluate the results. Changes to rent price in Dublin were expected to stay stable at 0% in 2021, while rising to 1.5 % in 2022, and 2.5 % in 2023 [5].
Similarly, Statistica, who provides reliable market and business data, performed research showed house prices in the Republic of Ireland are expected to fall after COVID-19, before increasing once more in 2022. This view is more encouraging than it was previously in 2020, when housing market declines of up to 12% were projected [6].

III. DATASET AND EXPLORATORY ANALYSIS

Source of the Data

A dataset from Kaggle of the rent prices in Dublin taken from Daft.ie in September 2020 [7] was chosen where it was found that it contained the required information to conduct the required statistical analysis. It was considered as a 'big data' dataset as the source it was scraped from, Daft.ie, has the following characteristics of big data:

- Volume: Thousand of houses/apartments/properties are advertised for sale or rent.
- Velocity: Hundreds of advertised properties are added daily.
- Variety: Data is coming from different sources such as landlords and with different data types like pictures of the property sand descriptive text.
- Veracity: All users who upload data to Daft.ie must be verified to be a real person before placing an ad.

The attributes in this dataset included ratio type data such as property price and the number of bathrooms and bedrooms as well as nominal data such as its address, whether it was furnished or not, and the specific type of property it was. For every property, its respective latitiude and longitude was included also. Therefore using the Python package gmaps via Google Maps, the following geographical representation was created.

Figure 1: Properties to rent in Dublin in September 2020 created via gmaps.

As anticipated, the majority of the higher rent prices are condensed around the city center. Therefore, locality appears to have a significant impact on rental prices, but it is most certainly not the only influence.

Feature Engineering and Data Processing

Before looking at the other variables in the dataset in greater detail, domain knowledge was used to extract new features such as characteristics from the raw data. This included creating new attributes such as the area district the property was based, as well as it's respective bathroom to bedroom ratio. Please refer to *Appendix A* for more details of the data pre-processing and validation procedures conducted [8].

Descriptive Statistics and Visual Exploration

	Min	Max	Mean	Std. Dev
Price Per Month	480	9400	2258.05	965.212

Table 1: Descriptive statistics of the price per month for a rental property.

It was shown that the maximum price per month was €9,400, while the minimum figure in the dataset was €480. The mean of €2,258.05 was consistent with the consensus of the cost of monthly rental in Dublin. However, the extreme maximum price including a standard deviation of 965.21 generated interest, to see if this potentially dependent variable was normally distributed or not.

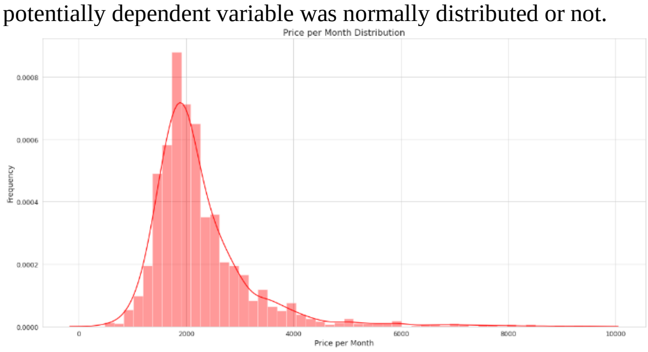

Figure 2: Distribution of the price per month of properties to rent in Dublin.

The distribution of the price per month for a property to rent in Dublin seemed not to be normally distributed as it is slightly skewed to the right in the histogram, and therefore does not fit a bell-shaped curve. To get a clearer picture, a boxplot of the price per month by area was created.

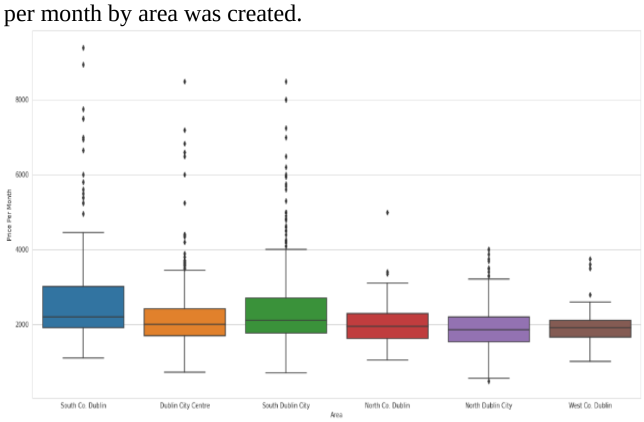

Figure 3: Box-plot of the price per month by the specific area in Dublin.

As can be seen above, there are many extreme outliers on the upper end, predominantly in the South Dublin and City Centre regions. This satisfies the conception that these areas would contain the most expensive properties to rent that are well above the average in Dublin. While some areas in the boxplot demonstrated they are symmetric, the previous ones mentioned do not have the median in the middle of their box and their whiskers are not the same. A similar conclusion can be drawn from the violon plot produced in *Appendix B* where the number of bedrooms was not normally distributed by property type. To definitively and statistically test to see if this attribute was normally distributed, a Shapiro-Wilk's test was conducted, with the output produced below.

```
Statistics=0.775, p=0.000
Sample does not look Gaussian (reject H0)
```

Statistical Result 1: Shapiro-Wilk's test for normality of the price per month variable.

Since the p-value was less than 0.05, the hypothesis of normality was rejected, and it was concluded that the price per month for rent in Dublin was not normally distributed.

IV. HYPOTHESES AND RESEARCH QUESTIONS

After conducting our exploratory analysis, the following questions were identified for further testing and investigation to extract meaningful insights. This was to benefit any student or young professional, for example, that are looking for a room to rent in the Dublin area:

- *Is there a correlation between the number of bedrooms and bathrooms in a property in Dublin?*
- *Is there an association between the price of a property in Dublin and whether it is furnished or unfurnished?*
- *Does the type of a property depend on the area of Dublin it is situated in?*
- *Are there differences in the bathroom to bedroom ratio between the four types of property in Dublin?*
- *What are the characteristics that contribute to a property's price in Dublin?*

V. METHODS USED AND WHY

The following methods were used to conduct our hypothesis testing and help answer the previous research questions identified in the most optimal way [9].

A. Pearson Correlation Coefficient

The Pearson correlation coefficient was generated using the Pearson product-moment correlation, denoted as *r*, and it is this coefficient that measures the strength of a direct relationship between two continuous variables. Its value can range from -1 for a perfect negative linear relationship to +1 for a perfect positive linear relationship. A value of zero indicates no relationship between two variables.

B. Point-Biserial Correlation Coefficient

The point-biserial correlation coefficient, r_{pb}, often just called point-biserial correlation, is used to determine the strength of a linear relationship between one continuous variable and one nominal variable with two categories, also known as a dichotomous variable.

Like above, its value can range from -1 to +1. In addition, the coefficient of determination, $r_{pb}{}^2$, can be calculated and used to determine the proportion of variance in one variable that can be described by the other variable.

C. Chi-Squared Test for Independence

The chi-square test can be used to test a variety of sizes of contingency tables, as well as more than one type of null and alternative hypotheses. Contingency tables that are greater than 2 x 2, which are often referred to as r x c contingency tables, tests whether two variables measured at the nominal level are independent.

Most commonly this test is called the chi-square test of independence, but it is also known as the chi-square test for association. Whilst it is also possible to perform the chi-square test of independence on ordinal variables, the ordered nature of the data would be lost by doing so and there would be more suitable tests to conduct. In order to make the correct inferences from a chi-square test of independence, a naturalistic study design will need to have been undertaken.

D. Kruskal-Wallis H Test

The Kruskal-Wallis H test, sometimes also called the one-way ANOVA on ranks, is a rank-based nonparametric test that can be used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable.

The Kruskal-Wallis H test is generally considered the nonparametric alternative to the one-way ANOVA, which can be used when the data fails the assumptions of the one-way ANOVA. This could happen if: (a) the data is non-normally distributed; or (b) there is an ordinal dependent variable, since the one-way ANOVA requires a continuous dependent variable. However, it should be noted that the Kruskal-Wallis H test cannot simply be considered an alternative to the one-way ANOVA. It has its own characteristics that must be considered if the results are to be accurately interpreted.

E. Random Forest Regression

This is a supervised learning technique that uses the ensemble learning approach for prediction. A more precise prediction than one on its own is made by the ensemble learning technique, as it incorporates predictions from several other machine learning models.

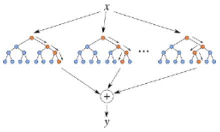

Figure 4: Structure of a Random Forest algorithm.

With no contact, the trees proceed in a straight line, as can be seen in Figure 4. A Random Forest builds many decision trees before outputting the prediction of all the trees from the average of the categories. [10]. The algorithm is as follows:

- From the training set, select *k* data points at random.
- Relating to these *k* data points, construct a decision tree.
- Follow the first two steps again after picking the number *N* of trees that is desired to be constructed.
- Assign the predicted value of *y* for a new data point to the average across all the predicted *y* values for each one of the *N*-tree trees.

It is a robust and precise model. Random Forest Regression generally works well on a wide range of situations, including those with non-linear correlations, such as this scenario for rental prices of properties in Dublin. Average mean-squared error and average accuracy was identified as indicators to determine if the model was reliable.

VI. RESULTS AND FINDINGS

A. Pearson Correlation Coefficient: ***Is there a correlation between the number of bedrooms and bathrooms in a property in Dublin?***

A Pearson's product-moment correlation was run to assess the relationship between the number of bedrooms and the number of bathrooms in in a property based in the Dublin area.

As assessed by Shapiro-Wilk's test ($p < .05$), while preliminary analyses showed the relationship was not to be linear with both variables not normally distributed, because the test is somewhat robust to deviations from normality, the Pearson's correlation was run anyway.

There was a statistically significant, strong positive correlation between the number of bedrooms and the number of bathrooms in a property based in the Dublin area, $r = 0.696$, $p < .0005$, with the number of bedrooms explaining 48% (square of the correlation coefficient, $r^2 = (0.696)^2$) of the variation in the number of bathrooms in a property.

Correlations

		Bedroom	Bathroom
Bedroom	Pearson Correlation	1	.696
	Sig. (2-tailed)		.000

Table 2: Pearson correlation coefficient between bedrooms and bathrooms.

B. Point-Biserial Correlation Coefficient: ***Is there an association between the price of a property in Dublin and whether it is furnished or unfurnished?***

On inspection of a boxplot for values greater than 1.5 box-lengths from the edge of the box, there were outliers in the data. As a result, this did not meet the prerequisite assumption of no outliers in the data. At this point, it was envisaged that outliers were included in the analysis anyway. This is because it was not believed the result will be significantly affected as much as removing them instead.

However, other assumptions were not met, as there was no homogeneity of variances for rental prices of furnished and unfurnished properties, as determined by Levene's equality of variances test ($p = .021$). Since the test returns a p-value less than 0.05 ($p < .05$), the population variances were unequal and violated the assumption of homogeneity of variances. Also, rent prices for either furnished or unfurnished properties were not normally distributed, as the p-value in the Shapiro-Wilk's test was less than 0.05.

After checking to see if transforming the continuous variable would change any result above with no success, a change in approach to a non-parametric test was decided via a rank-biserial correlation. Marascuilo and McSweeney [11] demonstrated the use of Kendall's tau-b, τ_b, for this type of scenario, where it was noted that null and alternative hypotheses of this test was not the same as the point-biserial correlation.

Correlations

		Price P/M	(Un)furnished
Price Per Month	Kendall's tau_b corr. coeff.	1.000	-.111
	Sig. (2-tailed)	.	.000

Table 3: Point-Biserial correlation coefficient between price and furnishing.

As previously mentioned, to determine how strong the association is for the different variables is only possible with Pearson's correlation. However, if Kendall's tau-b is close to +1 or -1 there is a strong association whereas it is weak if it is close to zero.

Therefore, it was concluded that there was a weak, negative association between a property's price and whether it was furnished or unfurnished, which was statistically significant, $\tau_b = -0.111$, $p < .05$. Note while the table says otherwise, a p-value of 0 does not mean the significance value is zero, hence it is stated as $p < .05$, which is what it indicates. This conclusion was also verified in the correlation matrix heatmap produced in *Appendix C*.

C. Chi-Squared Test for Independence: ***Does the type of a property depend on the area of Dublin it is situated in?***

A chi-square test of independence was conducted between the property's area and type of property advertised to rent. However, since not all expected cell frequencies were greater than five, it is accepted that the results might not be valid as this assumption was not met.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	162.190^a	18	.000
Likelihood Ratio	182.240	18	.000

Table 4: Chi-Square test between the property's area and type.

There was a statistically significant association conducted between the property's area and type of property advertised to rent, $\chi 2(18) = 162.19$, $p < .0005$. As a result, it can be inferred that these two variables are not independent, and we can reject the null hypothesis at 5% level of significance.

In order to gain information of the strength of the association, Cramer's V was used to measure an estimate of the magnitude of the association between the property's area and type of property advertised to rent. The value of Cramer's V can be found in the Symmetric Measures table below, where it has a value of 0.175. Therefore, it was concluded that the association was between small to moderately strong, according to Cohen's theory [12].

Symmetric Measures

		Value	Approx. Significance
Area * Property Type	Phi	.304	.000
	Cramer's V	.175	.000

Table 5: Cramer's V test between the property's area and type.

Since Cramer's V does not provide further details of this association, an analysis of residuals was used to follow up this statistically significant result. The specification used to determine when a cell provides evidence against the null hypothesis was when the absolute adjusted standardized residuals are greater than 3, also known as standard errors [13] [14]. This shows it deviates significantly from independence.

Upon inspection of the crosstabulation, it was identified that there were seven cells with adjusted standardized residuals larger than three, as presented below. Please refer to *Appendix D* for the full results.

Area * Property Type Crosstabulation

		Property Type			
		Apartment	Flat	House	Studio
Dublin City Centre	Adj. Residual	10.4	-1.5	-10.4	1.8
North Co. Dublin	Adj. Residual	-2.9	-.5	3.3	-.2
North Dublin City	Adj. Residual	-4.9	5.5	2.6	-.5
West Co. Dublin	Adj. Residual	-3.1	-1.6	4.0	-.3

Table 6: Adjusted standardized residuals area and property type.

The largest adjusted standardized residual was for apartments that were in Dublin City. For this case, less than x1.3 the number of apartments were in Dublin City Centre compared to what would be expected if the null hypothesis was true, with an adjusted standardized residual of 10.4. This and other adjusted residuals greater than three can partly explain the rejection of the null hypothesis of independence.

For reference, a cell of the crosstabulation that does satisfy the null hypothesis of independence, is a flat based in North Co. Dublin. The standardized adjusted residual is -0.5, and therefore well within the threshold of three, where the number of flats in North Co. Dublin is nearly identical compared to what would be expected if the null hypothesis was true.

D. Kruskal-Wallis H Test: ***Are there differences in the bathroom to bedroom ratio between the four types of property in Dublin***

A Kruskal-Wallis H test was run to determine if there were differences in bathroom to bedroom ratio between four types of property: apartment, house, flat and a studio.

The distribution of the Bathroom to Bedroom Ratio was not the same across categories of Property Type. Since the dependent variable did not have similarly shaped distributions for all groups of the independent variable, inferences could not be made about differences in medians between groups.

Independent-Samples Kruskal-Wallis Test Summary

Test Statistic	159.540
Degree Of Freedom	3
Asymptotic Sig. (2-sided test)	.000

Table 7: Kruskal-Wallis H Test of bathroom to beds across property type.

Instead, the differences in distributions, lower/higher ratios and mean ranks were investigated, where distributions and mean ranks of the bathroom to bedroom ratio between the types of property were statistically significantly different between groups, $\chi^2(3) = 159.540$, $p < .0005$.

As a result of this, to discover which groups are different to which other groups, a post hoc test was run to interpret all pairwise comparisons using Dunn's procedure [15] with a Bonferroni adjustment.

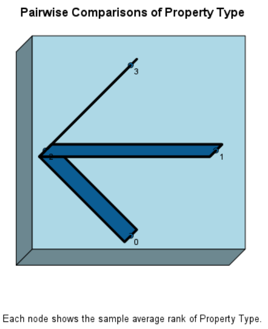

Figure 5: Graphic of the pairwise comparision of each property type via their node.

Table of Pairwise Comparisons of Property Type

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.
House-Apartment	316.608	25.513	12.410	.000	.000
House-Flat	358.588	57.545	6.231	.000	.000
House-Studio	-607.750	438.451	-1.386	.166	.994
Apartment-Flat	-41.980	54.581	-.769	.442	1.000
Apartment-Studio	-291.142	438.072	-.665	.506	1.000
Flat-Studio	-249.162	441.098	-.565	.572	1.000

Table 8: Kruskal-Wallis H Test of bathroom to beds across property type.

Adjusted p-values are presented where the post hoc analysis revealed statistically significant differences in the bathroom to bedroom ratio between a house and an apartment ($p < .0005$), and a house and a flat ($p < .0005$), but not a studio or any other group combination.

E. Random Forest Regression: ***What are the characteristics that contribute to a property's price in Dublin?***

Using the K-Fold (with K = 10) Cross Validation approach with the Random Forest Regression via the Python package sklearn, the data was sliced into 9 training parts and tested against the 10th part A different part of the ten sections of the data was used as the testing data in each iteration, in order to maximize the size of our dataset and obtain better results. Please refer to *Appendix E* for the full results of this experiment and *Appendix F* for the respective decision tree produced. It resulted in an average accuracy of 84.82% in predicting the price per month per bed for a rental property in Dublin.

Random Forest Regression K-Fold Cross Validation Results (K=10)	
Average Mean Absolute Error	170.81
Average Accuracy	84.82%

Analyzing the feature importance graph produced, the number of bedrooms was by far the most important factor (37.44%) in determining rental prices for a bed per month in Dublin. This was following by the specific location features in Longitude and Latitude, with nearly identical relative importance of over 24%. It was also noted, all else being equal, that a rental property that was a flat or based in either Dublin City Centre or South Dublin City influences the price by around 1%.

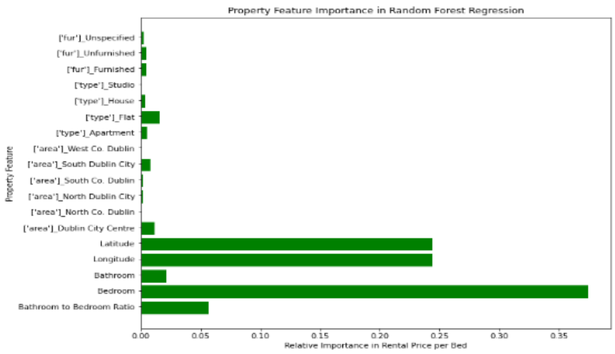

Figure 7: Property feature importance in Random Forest Regression.

The feature importance table of the results also reinforced our understanding that, whether a property was furnished or not, it essentially had no influence on the rental price per month. Surprisingly being based in South County Dublin, for example, had in-essence no effect on price, as well as whether it was a studio or a house.

Property Feature	**Relative Importance in Price Per Bed Per Month**	**Percentage (100% Total)**
Bathroom to Bedroom Ratio	0.05695480070	5.70%
Bedroom	0.37435746000	37.44%
Bathroom	0.02139702620	2.14%
Longitude	0.24389974800	24.39%
Latitude	0.24379700800	24.38%
Area - Dublin City Centre	0.01169084190	1.17%
Area – North Co. Dublin	0.00021229453	0.02%
Area – North Dublin City	0.00153809684	0.15%
Area – South Co. Dublin	0.00145973764	0.15%
Area – South Dublin City	0.00774265798	0.77%
Area – West Co. Dublin	0.00030166382	0.03%
Type - Apartment	0.00546052973	0.55%
Type - Flat	0.01562635030	1.56%
Type – House	0.00354634855	0.35%
Type - Studio	0.00058223047	0.06%
State - Furnished	0.00419909295	0.42%
State - Unfurnished	0.00455187576	0.46%
State - Unspecified	0.00268223688	0.27%

Table 9: Property feature relative importance in price per bed per month.

VII. CONCLUDING DISCUSSION AND OUTLOOK

This paper described a random forest regression method that may be able to predict the price of a rental property in Dublin, by using the most relevant important features of a property, such as the type of house or its location. One of this paper's critical contribution was to have statistically demonstrated that there were significant differences in the bathroom to bedroom ratio between a house and an apartment, as well as a house and a flat.

It was also found that there was a statistically significant association between a property's area and type of property advertised to rent, where an apartment was likely to be advertised in Dublin City Centre, while a flat was more prominent in North Dublin City. Simultaneously, it was concluded that there was a weak, negative association between a property's price and whether it was furnished or unfurnished, which may be against the general perception that a furnished property would be more expensive to rent.

Further work in this area should aim to increase this model's accuracy using other models, such as pre-trained Convolutional Neural Networks, with the final objective of an ensemble model. Also, it can be safely argued that the hyperparameter tuning of the Random Forest Regression model will improve these results even further. Other areas to explore include extending the model to include rental properties across Ireland, as well as ones that are on the market to purchase fully also. Removing features, like the coordinates and the number of bedrooms in the model, would result in a better understanding of how much the other features influences the price per month per bed.

VIII. REFERENCES

[1] Pope, C., 2021. House prices could rise by over 12% by year-end, survey shows. [online] The Irish Times. Available at: https://www.irishtimes.com/news/ireland/irish-news/house-prices-could-rise-by-over-12-by-ye ar-end-and-survey-shows-1.4665462 [Accessed 20 November 2021].

[2] Curran, S., 2021. 'Chronic and worsening' supply crisis heaps pressure on renters as prices climb sharply. [online] TheJournal.ie. Available at: https://www.thejournal.ie/daft-rental-report-summer-5518739-Aug2021/ [Accessed 20 November 2021].

[3] Lyons, R., 2021. [online] ww1.daft.ie. Available at: https://ww1.daft.ie/report/2021-Q3-houseprice-daftreport.pdf?d_rd=1 [Accessed 20 November 2021].

[4] Torino, B., 2020. [online] Machine Learning and Real State: Predicting Rental Prices in Amsterdam. Available at: https://towardsdatascience.com/ai-and-real-state-renting-in-amsterdam-part-1-5fce18238dbc [Accessed 21 November 2021].

[5] Dwyer, H., 2021. JLL Bi-Annual Dublin Residential Forecasts Report for H2 2021. [online] https://www.jll.ie/en/trends-and-insights/research/jll-dublin-residential-forecasts (Accessed: 21 November 2021).

[6] Statista Research Department., 2021. Forecast of the percentage change on the previous year of residential property prices in Ireland from 2020 to 2022. [online] https://www.statista.com/statistics/1174799/residential-real-estate-price-forecast-change-in-ir eland/ (Accessed: 21 November 2021).

[7] Lyu, G., 2021. Dublin renting dataset scraped from Daft.ie (2020.9). [online] Available at: https://www.kaggle.com/d17129765/predicting-dublin-rental-daftie [Accessed 20 November 2021)].

[8] Hehir, C., 2021. CA675 Cloud Technologies Assignment 2 Report.

[9] Statistics, L., 2013. Hypothesis testing. The Null and Alternative Hypothesis.

[10] Bakshi, C., 2020. Random Forest Regression. [online and image] Available at: https://levelup.gitconnected.com/random-forest-regression-209c0f354c84 [Accessed 22 November 2021)].

[11] Marascuilo, L. A., & McSweeney, M., 1977. Nonparametric and distribution-free methods for the social sciences. Belmont, CA: Wadsworth Publishing Company.

[12] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New York: Psychology Press.

[13] Agresti, A. (2013). Categorical data analysis (3rd ed.). Hoboken, NJ: Wiley.

[14] Agresti, A., & Franklin, C. (2014). Statistics: The art of learning from data (3rd ed.). Harlow, England: Pearson.

[15] Dunn, O. J. (1964). Multiple comparisons using rank sums. Technometrics, 6, 241-252.

[16] Frey, F., 2017. SPSS (software). The International Encyclopedia of Communication Research Methods, pp.1-2.

[17] Sanner, M.F., 1999. Python: a programming language for software integration and development. J Mol Graph Model, 17(1), pp.57-61.

RANDOM FOREST REGRESSION – PRICE PER BED PER MONTH

K-FOLD CROSS VALIDATION RESULTS

K-Fold Cross Validation #1	Mean Absolute Error: 193.05
	Accuracy: 83.39 %.
K-Fold Cross Validation #2	Mean Absolute Error: 136.54
	Accuracy: 86.82 %.
K-Fold Cross Validation #3	Mean Absolute Error: 147.6
	Accuracy: 77.75 %.
K-Fold Cross Validation #4	Mean Absolute Error: 140.03
	Accuracy: 88.72 %.
K-Fold Cross Validation #5	Mean Absolute Error: 191.63
	Accuracy: 84.91 %.
K-Fold Cross Validation #6	Mean Absolute Error: 177.52
	Accuracy: 86.4 %.
K-Fold Cross Validation #7	Mean Absolute Error: 193.1
	Accuracy: 85.99 %.
K-Fold Cross Validation #8	Mean Absolute Error: 194.5
	Accuracy: 81.07 %.
K-Fold Cross Validation #9	Mean Absolute Error: 158.64
	Accuracy: 86.34 %.
K-Fold Cross Validation #10	Mean Absolute Error: 175.5
	Accuracy: 86.74 %.

RANDOM FOREST REGRESSION – PRICE PER BED PER MONTH

DECISION TREE

All the random forest's algorithm decisions are represented in the diagram, encompassing divisions, branches, leaves, mean standard errors, sample sizes, and anticipated values for each divide.