

1 **A Model Validation and Predicting the Rental Values of**
2 **Residential Properties Using Logistic Regression Model**

3 Tawakalitu Odubiyi¹, Anthony Ugulu¹, Olalekan Oshodi¹, Clinton Aigbavboa¹, and
4 Wellington Thwala¹

5 ¹ SARChI in Sustainable Construction Management and Leadership in the Built Environment
6 Faculty of Engineering and the Built Environment
7 University of Johannesburg, South Africa
8 kalitutawa@gmail.com

9 **Abstract.** The property market is a key contributor to the economic growth of
10 many countries. This makes information from property valuation reports vital for
11 decisions on real estate investments and property tax. Unfortunately, literature
12 reveals that inaccurate property valuation arising from a reliance on traditional
13 methods of valuation remains a major problem facing real estate practice. To
14 improve the prediction accuracy of property valuation estimates, modelling
15 techniques such as neural networks have previously been applied to this problem.
16 This present study uses a logistic regression model to predict the rental values of
17 residential properties in Cape Town, South Africa. Field survey data was divided
18 into two groups: training and test sets. The training set was used for model
19 development while the test set was used for model validation. The results of the
20 study revealed that parking, garden, number of bedrooms and floor area have the
21 most significant impact on the rental values of residential properties.
22 Surprisingly, proximity to a police station has one of the least effects on the rental
23 values of residential properties. With a prediction performance of over 70%
24 accuracy, findings indicate that the logistic regression model is suitable for
25 predicting the rental values of residential properties. This study evaluates the
26 factors that influence the rental values of residential properties located within the
27 study area. The developed model can serve as a decision support tool for
28 estimating the tax payable by property owners.

29 **Keywords:** Residential Property, Rental value, Logistic regression modelling,
30 Prediction, Tax

31 **1 Introduction**

32 Residential property prices are an important reflection of an economy. This is because
33 the housing sector is a major contributor to economic growth [1]. The ranges for rental
34 prices interest property valuers, property owners, the users and the local council that
35 collect property taxes. Property valuers and property owners are interested in rental
36 yields [2], while the potential occupants or renters are concerned about portions of their
37 budget that goes towards rental payments [3]. In all of these, society benefits from the

38 actions of both parties (i.e. valuers and users) through taxes collected by the local
39 authorities. It is, therefore, fundamental that property valuation should follow a certain
40 standard to ensure accuracy in the valuation practices.

41 Accurately predicting residential rental prices has been a major area of concern in
42 mortgage valuations. Consequently, literature have often discussed several
43 methodologies for such property valuations with regression analysis and data mining
44 techniques the most commonly used methodologies [4]. These methods, models, and
45 techniques help researchers to empirically reveal the various factors that influence
46 property rental prices. In general, the influential factors are classified into the
47 neighborhood or locational and structural [5].

48 Globally, locational characteristics like parks and security, have been noted to
49 influence residential rental values [5], [6]. In addition, structural attributes of residential
50 properties like the number of bedrooms available bathrooms, floor size, parking space,
51 and garage also influence rental prices [7]. However, the impact of these identified
52 variables differs across geographic locations. This is because the needs of renters vary,
53 thereby influencing decision-making for property valuation. Property valuation serves
54 as a decision support tool for estimating the tax payable by property owners. Property
55 tax is a revenue-generating source for local governments [8]. An examination of
56 property tax in 21 countries, including South Africa, Botswana, and Brazil shows that
57 rental property tax is significant and annual rental value is used as a basis for tax in
58 many locations [8]. In many of the locations, the payment of rental tax falls on the
59 property owners. This applies in South Africa, Zimbabwe, and Australia. In Botswana,
60 Brazil, and Poland, the payment of rental tax falls on the occupier, or both owner and
61 occupier as in the case of Botswana and The Netherlands [8]. Therefore, the onus of
62 paying property tax by these stakeholders calls for proper property valuation.

63 Despite the high levels of crime rates as well as the high rate of property ownership
64 of approximately 53.5% in South Africa [9], there has been a paucity of research
65 evidence on the factors influencing rental prices. This study attempts to measure the
66 effect of proximity to the police station on residential rental prices using a logistic
67 regression model. The following structure is adopted in this paper. The next section
68 discussed the methodology applied in the study is discussed. This is followed by the
69 results and discussion of the findings. Lastly, the conclusion of the paper is the last
70 section.

71 **2 Research Methodology**

72 **2.1 Overview**

73 The top high-density cities in South Africa are Cape Town, Pretoria, Port Elizabeth,
74 Durban, and Johannesburg. The increasing population of these cities makes residential
75 rental valuation important. Notably, property ownership is of high national importance
76 [10]. To illustrate, Cape Town is termed the African city of opportunity [11] as it is the
77 second richest city in South Africa [12]. Notwithstanding, the influx of people makes
78 these cities prone to different types of crime. For instance, in Cape Town, residential
79 crime rates drive homeowners to residential fortification measures [13]. Therefore, the

80 presence of crime prevention or control facilities like a Police Station is necessary when
81 addressing rental prices in such areas.

82 For this study, rent and feature data were obtained for 381 residential apartments
83 from the website www.property24.com. Most of the property portal in South Africa
84 have details of real estate firms in the country. The listing price rather than the sales
85 price was taken because compared to transaction data of residential value, price listing
86 provides a better estimate [14]. This is in addition to the fact that price listings are what
87 is provided for properties in Cape Town, South Africa. What informed the data
88 collected was mostly the reviewed literature on the factors that impact rental values.
89 However, in order to examine the role of security, nearness (measured in kilometers)
90 to a police station was included.

91 The dependent variable was the per annum rental prices. It was then grouped into
92 bands A (0- 189,880 South African Rands) and B (189,881-296,540 South African
93 Rand). They were examined against the independent variables which included
94 bedroom, bathroom, the presence of car park, number of parking spaces, dining, lounge,
95 balcony, swimming pool, floor area, furnishing, services, garden, and police presence.
96 Braakman [15] described how security in a neighborhood can influence rental prices.
97 This study assumes that the proximity to a police station in a neighborhood will
98 influence its residential prices.

99 **2.2 Logistics Regression Modelling**

100 Logistic regression is used to predict the class (or category) of individuals based on one
101 or more predictor variables (X). It is used to model a binary outcome which can only
102 have two possible values such as 0 or 1, or yes or no. The simple logistic regression
103 and multiple logistic regression are the approaches to logistic regression modelling.
104 The simple logistic regression forecasts the probability of a dependent factor against
105 one independent factor. The multiple logistic regression models predict the probability
106 of a dependent variable based on multiple independent variables [16]. Past studies have
107 used logistic regression to determine variables that influence rental values, for example,
108 [17]. This present study makes use of multiple logistic regression.

109 The output of a multiple logistic regression model displays four components - the
110 estimate, standard error, z-value, and p-value. The estimate is the intercept, and beta
111 coefficient estimates associated with each predictor variable. It shows the strength of
112 the relationship. Standard error represents the accuracy of the coefficients. Z-value is
113 the result of the division of estimate and standard error. The p-value checks for
114 significance of the estimates. A small p-value shows higher significance for the
115 estimate.

116 3 Results and Discussion

117 3.1 Computing logistic regression model

118 The monthly rental value is based on the 14 independent variables from the collected
 119 data. Hosmer and Lemeshow [18] described that logistic regression modelling tests the
 120 statistical significance of coefficients. In this study, the null hypothesis assumes that
 121 the independent variables/factors do not have statistical significance. The alternate
 122 hypothesis assumes that the independent factors have statistical significance. At 95%
 123 confidence interval, the alpha value is $p = 0.05$. At p-value of less than or equal to 0.05,
 124 accept the alternate hypothesis and vice versa. Table 1 shows that the bathroom,
 125 swimming pool, floor area, and a furnished apartment has values less than 0.05. This
 126 means that the variables are statistically significant. Hence, we reject the null
 127 hypothesis

128 Estimate coefficient (b₀) shows the strength of relationships between independent
 129 and dependent variables. The B-coefficients describes the extent to which a dependent
 130 variable (rental price) is influenced by a unit change in each of the independent
 131 variables [19]. Positive coefficient estimate increases the probability and vice versa.
 132 The coefficient for the variable bathroom is $b = -1.058707$ and negative. This means
 133 that an increase in the number of bathrooms will be associated with a decreased
 134 probability of residential rental price rise. Also, the variable furnished (which implies
 135 whether a property is furnished or not) has a negative $b = -1.884781$. This shows that
 136 an inverse relationship exists between furnished apartment and decreased probability
 137 of a rental rise. However, the coefficient of the variable Swimming Pool is positive b
 138 $= 0.940621$. This means that the availability of a swimming pool is linked with a
 139 probability of a residential rental value increase. These three variables have p-values
 140 less than or equal to 0.05, which make them statistically significant compared to other
 141 independent variables that have a p-value greater than 0.05.

142 **Table 1.** Coefficients of the logistic regression model

Variable	Estimate	Std. Error	z value	p-value
(Intercept)	6.84371617	1.093829473	6.2566573	0.00000
Bedroom	-0.677291	0.385797	-1.756	0.07916
Bathroom	-1.058707	0.427533	-2.476	0.01327
Parking	-0.547072	0.302566	-1.808	0.07059
Park	-0.653368	0.568963	-1.148	0.25082
Dining	0.202445	0.366713	0.552	0.58091
Lounge	-0.067425	0.3176	-0.212	0.83188
Balcony	-0.341771	0.344761	-0.991	0.32153
Swimming Pool	0.940621	0.369129	2.548	0.01083
Floor area	-0.026086	0.009759	-2.673	0.00752
Furnished	-1.884781	0.369715	-5.098	0.00000

Services	-0.602996	0.576351	-1.046	0.29545
Garden	0.271286	0.477845	0.568	0.57022
Police	0.096143	0.063208	1.521	0.12824

143 **3.2 Model validation and predictive performance of the developed logistic**
 144 **regression model**

145 This study predicts the rental values of residential properties in Cape Town using
 146 logistic regression. Predictive performance value ranges between a percentage of 0 and
 147 100 [20]. Generally, a value close to 100% indicates that the model can correctly
 148 classify all the test data set. In this study, the prediction of the data sets based on the
 149 groupings A and B as earlier described to reveal an overall accuracy of 75.21% as
 150 shown in Table 2. This implies that the model can adequately classify the data set. The
 151 result of model validation is summarized and presented in Table 2.

152 **Table 2.** Summary of model validation logistic regression

Observed	Predicted		Accuracy
	A	B	
A	59	8	88.06%
B	21	29	58.00%
	Overall		75.21%

153 **4 Discussion**

154 As earlier discussed in the literature review, inherent attributes of properties that are
 155 valued by consumers' influence rental values. The result from this study is evidence
 156 that each independent variable has a different impact on the rental price of residential
 157 properties (see Table 1). Findings from the study reveal that the number of bathrooms,
 158 the presence of a swimming pool, floor area, and furnished property have the most
 159 significant effect on residential property prices in Cape Town, South Africa while
 160 police station is insignificant.

161 Bathroom and conveniences are vital parts of a building. This study also indicated
 162 that the number of bathrooms (toilet inclusive) can influence rental value. Bathrooms
 163 serve as a fundamental human function [21]. Number of bathroom can have significant
 164 effects on residential house prices as was reported by [22]. Likewise, the number of
 165 bedrooms can influence property prices as found in this study. Abidoeye and Chan [7]
 166 stated that the number of bedrooms increases the residential rental price. Other
 167 documented evidence reveals that the number of bedrooms can significantly influence
 168 residential rental prices [23]. Surprisingly, this contradicts the findings of [19] who
 169 noted that the number of bedrooms do not have a significant effect on rental prices.

170 This could be attributed to differences in the density of geographical locations [24]. In
171 this study, while previous studies posited that the number of bedrooms was a factor; it
172 was insignificant to residential rental values in Cape Town, South Africa.

173 The presence of a swimming pool raises property value and benefits the users in
174 terms of health and recreation. It gives a higher level of relaxation compared to
175 traditional bathroom facilities. This study reveals that a swimming pool is a significant
176 variable that influences rental prices in the study area. Evidence shows that swimming
177 pool presence is valued in properties across the country [25] especially if it is a lower
178 floor building where the swimming pool can be easily accessed [26]. However, a
179 swimming pool is a luxury that renters or homeowners could choose whether to have
180 or not.

181 Floor area or floor size represents the space in a property. Floor size is a significant
182 variable that affects rental prices as we have demonstrated in this study [5] [27]. Floor
183 number or level can also affect rental values. Some consumers prefer higher floor levels
184 since it gives a wider view [7]. This is down to consumer choice or preference. Chin
185 and Chan [5] however stated that floor area is the most important attribute of a building
186 influencing rental values. In addition, the results of this study found that a furnished
187 apartment is one of the most important/significant factors impacting rental values. In a
188 furnished apartment, there is furniture in place before the property is rented out or
189 mortgaged out. In general, furnished apartments are more expensive than an
190 unfurnished apartment [28]. Findings from this study agree with that of Buiga and
191 Toth, [29] who examined properties in Romania.

192 In terms of security, a police station is described as a social amenity that is relevant
193 to the proper function of society. Police presence, provides a sense of security for house
194 owners and intending buyers. The fear of crime or sensation of safety influences
195 property value [30]. Police presence, therefore is expected to reduce property crime
196 [31]. Other studies have reported that police presence influenced residential property
197 values [32]. This shows that there is a correlation between crime, property price, and
198 police presence. In this study, while nearness to a police station was a factor, it was
199 insignificant to residential rental values in Cape Town, South Africa. This study has
200 shown that logistic regression can produce a reliable prediction of the factors that
201 influence the rental values of residential properties. The results support the findings of
202 [7], who reported that logistic regression modelling technique has a reliable predictive
203 ability that can help address the property attributes and valuation challenges.

204 **5 Conclusion**

205 This study predicts the rental values of residential properties in Cape Town, South
206 Africa using logistic regression. The predictive accuracy of the developed model
207 suggests that logistic regression modelling can produce reliable estimates of the rental
208 values of residential properties. In addition, it was discovered that aside from a
209 furnished apartment, number of bathrooms and the presence of a swimming pool,
210 parking space, garden, number of bedrooms and floor area all impact rental values of
211 residential properties. The findings imply that in Cape Town, South Africa, residential

212 property rental prices are determined by these significant factors. However, the
213 nearness to a police station does not significantly influence rental values. This means
214 that consumers' choice of residential property is not affected by the presence or absence
215 of a police station. An explanation could be that this is due to the high prevalence of
216 private security firms in South Africa. This study predicts the rental values of residential
217 properties in Cape Town, South Africa using logistic regression. The predictive
218 accuracy of the developed model suggests that logistic regression modelling can
219 produce reliable estimates of the rental values of residential properties. In addition, it
220 was discovered that aside a furnished apartment, number of bathrooms and the presence
221 of a swimming pool, parking space, garden, number of bedrooms and floor area all
222 impact rental values of residential properties. The findings imply that in Cape Town,
223 South Africa, residential property rental prices are determined by these significant
224 factors. However, the nearness to a police station does not significantly influence rental
225 values. This means that consumers' choice of residential property is not affected by the
226 presence or absence of a police station. An explanation could be that this is due to the
227 high prevalence of private security firms in South Africa.

228 This study contributes to the body of existing knowledge in property economics. It
229 adds to current knowledge on the impact of the attributes of a residential property on
230 its rental value. However, it must be reiterated that the availability of data remains a
231 challenge to researchers in the fields of construction economics and property
232 economics. It is also important to point out that the geographical scope of this study
233 was limited to Cape Town, South Africa. However, results arrived at would easily apply
234 in similar locales. It would be necessary for further research to be carried out on
235 property valuation using logistic regression model so as to better understand the
236 identified trends.

237 References

238 This study contributes to the body of existing knowledge in property economics. It
239 adds to current knowledge on the impact of the attributes of a residential property on
240 its rental value. However, it must be reiterated that the availability of data remains a
241 challenge to researchers in the fields of construction economics and property
242 economics. It is also important to point out that the geographical scope of this study
243 was limited to Cape Town, South Africa. However, results arrived at would easily apply
244 in similar locales. It would be necessary for further research to be carried out on
245 property valuation using logistic regression model so as to better understand the
246 identified trends.

247 References

- 248 1. Hadavandi, E., Ghanbari, A., Mirjani, S.M., Abbasian, S. An econometric panel data-based
249 approach for housing price forecasting in Iran. *International Journal of Housing Markets and*
250 *Analysis*, 4(1), 70-83 (2011)
- 251 2. Bracke, P. House Prices and Rents: Micro Evidence from a Matched Dataset in Central
252 London. 53rd Congress of the European Regional Science Association: "Regional
253 Integration: Europe, the Mediterranean and the World Econom (pp. 1-52). Palermo, Italy:
254 European Regional Science Association (ERSA) (2013)

- 255 3. Berger, D., Guerrieri, V., Lorenzoni, G., Vavra, J. House Prices and Consumer Spending.
256 NBER Working Paper No. 21667 (2017)
- 257 4. Guan, J., Shi, D., Zuranda, J.M., Levitan, A.S. Analyzing Massive Data Sets: An Adaptive
258 Fuzzy Neural Approach For Prediction, with a Real Estate Illustration. *Journal of*
259 *Organizational Computing and Electronic Commerce*, 24, 94-112 (2014)
- 260 5. Chin, T. L., Chau, K. W., 2002, A critical review of literature on the hedonic price model.
261 *International Journal for Housing Science and Its Applications*, 27(2), 145–165. (2002)
- 262 6. Won, J., Lee, J. Investigating How the Rents of Small Urban Houses are Determined: Using
263 Spatial Hedonic Modeling for Urban Residential Housing in Seoul. *Sustainability*, 10(31),
264 1-15 (2018)
- 265 7. Abidoye, R. B., Chan, A. P. C., Modeling property values in Nigeria using artificial neural
266 network. *Journal of Property Research*, 34(1), 36-53. (2017)
- 267 8. McCluskey, W. *Property Tax: An International comparative Review*. New York: Routledge
268 (2018)
- 269 9. Statistics South Africa. General Household Survey 2017. Statistical Release P0318. Pretoria,
270 South Africa. (2018, June 21)
- 271 10. Luüs, C. The Absa residential property market database for South Africa- key data trends
272 and implications. IMF/BIS Conference on Real Estate Indicators and Financial Stability (pp.
273 149-170). Washignton, DC, USA: BIS Papers No 21. (2003)
- 274 11. PwC. Cape Town: African City of Opportunity. South Africa: PwC
275 (PricewaterhouseCoopers). (2018)
- 276 12. New World Wealth. The South Africa Wealth Report. Johannesburg: AFRASIA. (2018)
- 277 13. Cronje, C.J., Spocter, M. Open-plan Suburb to fortified suburb: Home fortification in
278 Soneike, Cape Town, South Africa. *Journal of Housing and the Built Environment* 32, 713-
279 732 (2017)
- 280 14. Abidoye, R.B., Chan, A.P. C. Hedonic Valuation of Real Estate Properties in Nigeria.
281 *Jouranal of African Real Estate Research*, 1(1), 122-140 (2018)
- 282 15. Braakmann, N. The link between crime risk and property prices in England and Wales:
283 Evidence from street-level data. *Urban studies*, 54(8), 1990-2007 (2017)
- 284 16. Tranmer, Elliot. *Binary Logistic Regression*. Cathie Marsh center for Census and Survey
285 Reseach (2008)
- 286 17. Yu, H., Wu, J. Real Estate Price Prediction with Regression and Classification. CS 229
287 Autumn 2016 Project Final Report. Standford university. (2016)
- 288 18. Hosmer D., Lemeshow S. *Applied logistic regression*. 2nd ed. New (2nd ed.). New York:
289 Wiley (2000)
- 290 19. Liman, H.S., Sipan, I., Olatunji, I.A., Afrane, E. . Hedonic Modelling Of Determinants Of
291 House Price In Minna, Nigeria. ASIA International Conferences on Emerging Issues in
292 Economics and Finance (pp. Hassan Shuaibu Liman1, Dr. Ibrahim Sipan2, Dr. Isaac
293 Ayodele Olatunji1 and Ebenezer Afrane3). Kuala Lumpur, Malaysia: International
294 Conferences on Emerging Issues in Economics and Finance. (2015)
- 295 20. Cortez, P., A tutorial on the rminer R package for data mining tasks. Teaching Report,
296 Department of Information Systems, Algoritmi Research Centre, Engineering School,
297 University of Minho, Guimaraes, Portugal. (2015)
- 298 21. Margolin, S., Poggiali, J. "Where are the Bathrooms?": Academic Library Restrooms and
299 Student Needs. *Journal of Library Administration*, 57, 481-489. (2017)
- 300 22. Musa, U., Yusoff, W.Z.N. The influencing components on prices of residential houses: A
301 review of literature. *the social sciences*, 12(4), 625-632. (2017)

- 302 23. Sanga, S.A. The Impact of Traditional House-type on Rental Values in Kinondoni
303 Municipality Dar es Salaam Tanzania. *Nordic Journal of Surveying and Real Estate*
304 *Research*, 2(1), 7-37 (2017)
- 305 24. Sirmans, G. S., MacDonald, L., Macpherson, D. A., Zietz, E. N. The value of housing
306 characteristics: a meta analysis. *The Journal of Real Estate Finance and Economics*, 33(3),
307 215-240 (2006)
- 308 25. Gnagey, M., Tans, R. Property-Price Determinants in Indonesia. *Bulletin of Indonesian*
309 *Economic Studies*, 54(1), 61-84 (2018)
- 310 26. Keskin, B. Hedonic analysis of price in the Istanbul housing market. *International journal of*
311 *Strategic property Management*, 12, 125-138.(2008)
- 312 27. Zietz, J., Zietz, E. N., Sirmans, G. S. Determinants of house prices: A quantile regression
313 approach. *The Journal of Real Estate Finance and Economics*, 38, 317-333.(2008)
- 314 28. Hasanah, A.N., Yudhistira, M.H. Landscape view, height preferences and apartment prices:
315 evidence from major urban area in Indonesia. *International Journal of Housing market and*
316 *Analysis*, 11(4), 701-715.(2017)
- 317 29. Buiga, A., Toth, G. Hedonic Analysis of Apartments' Price in Romania. *Review of Economic*
318 *Studies and Research*, 8(2), 5-23.(2015)
- 319 30. Klimova, A., Lee, A.D. Does a Nearby Murder Affect Housing Prices and Rents? The Case
320 of Sydney. *Economic Record*, 19(June), 6-40.(2014)
- 321 31. Andresen, M.A., Lau, K.C. An evaluation of police foot patrol in Lower Lonsdale, British
322 Columbia. *Police Practice and Research*, 15(6), 476-489 (2014)
- 323 32. Vetter, D.M., Beltrao, K.I., Massena, R.M.R. Working paper: The impact of the sense of
324 security from crime on Residential Property Value in Brazilian Metropolitan Areas.
325 Washington, DC: Inter-American Development Bank (2013)