DETERMINANTS OF HOUSE PRICES IN HOUT BAY

STEPHAN VAN DER WALT

Thesis presented in partial fulfilment of the requirements for the degree of Master of Arts at Stellenbosch University.

Supervisor: Dr A van Niekerk
Co-supervisor: Prof JZ Bloom
April 2010

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature: ------------------------------

Date: ------------------------------

SUMMARY

The research problem addressed in this study is how to ascertain the primary determinants of house prices in Hout Bay. This overarching aim encompasses three interwoven aspects. The research attempts first to determine which factors generally affect property prices in Hout Bay; second, to assess the extent to which individual factors affect house prices; and third, to discover the role variables collectively play in determining house prices in Hout Bay. Four objectives emerge from this subdivision of the aim, namely identify potential house price-influencing factors in Hout Bay; quantify the selected locational variables; statistically analyse the variables to distinguish the significant and insignificant ones; and use regression analysis to deduce the collective and individual influences of the significant factors on house prices.

Structured interviews were conducted with representatives of 12 estate agencies in Hout Bay to uncover factors affecting the local property market. Through insights gleaned from the literature, manipulation of municipal valuation and cadastral data and the structured interviews, 39 structural and site-related variables, 18 distance variables and 11 socio-economic variables were constructed.

Several preliminary and descriptive analyses performed on the variables gave a general impression of the distribution of data and assisted in identifying statistically significant variables for determining house prices. These analyses included measures of central tendency (mean, median and mode); measures of dispersion (minimum and maximum values, range, standard deviation, skewness and kurtosis); the compilation of histograms for each variable; analysis of variance (ANOVA) on nominal data variables; and the creation of 2D scatterplots for ordinal data variables. Spearman rank order correlation was performed on the nominal and ordinal data variables. Statistically weak variables and those exhibiting signs of multicollinearity were eliminated. A best-subsets regression analysis was executed on the remaining variables.

The regression model performed adequately, explaining close to 54% of the variation in house prices in Hout Bay. Among the individual factors, the size of the erf was the strongest predictor of the house price dependent variable, house size was the second most important factor, while distance to busy roads and quality of the house shared similar importance. Regression residuals were also mapped to expose spatial patterns. It is recommended that comparable research be conducted on a citywide scale, that variables be quantified differently and that new GIS techniques be incorporated in future studies.

KEYWORDS

GIS, house prices, Hout Bay, property valuation, property values, regression analysis

OPSOMMING

Die navorsingsprobleem wat hierdie studie aanspreek, is hoe om vas te stel wat die primêre faktore is wat huispryse in Houtbaai bepaal. Hierdie oorkoepelende doelwit vervat drie onderling verwante aspekte. Eerstens, poog die navorsing om te bepaal watter faktore in die algemeen huispryse in Houtbaai beïnvloed; tweedens, om te assesseer tot watter mate individuele faktore huispryse affekteer; en derdens, om te ontdek watter kollektiewe rol veranderlikes in die bepaling van huispryse in Houtbaai speel. Vanuit hierdie onderverdeling van die navorsingsdoelwit het vier doelstellings ontstaan, naamlik identifiseer die potensiële faktore wat huispryse in Houtbaai beïnvloed; kwantifiseer die geselekteerde liggingsveranderlikes; voer verskeie analises uit op die veranderlikes om die beduidende en onbeduidende veranderlikes te identifiseer; en benut regressie-analise om die kollektiewe en individuele invloed van beduidende faktore op huispryse in die studiegebied vas te stel.

Gestruktureerde onderhoude is met verkoopslui van 12 eiendomsagentskappe in Houtbaai gevoer om die faktore te bepaal wat die plaaslike eiendomsmark beïnvloed. Deur middel van insigte verkry uit die akademiese literatuur, manipulasie van munisipale waardasie- en kadastrale data en die gestruktureerde onderhoude is 39 strukturele en liggingsverwante veranderlikes, 18 afstandsveranderlikes en 11 sosio-ekonomiese veranderlikes geskep.

Verskeie analises wat op die veranderlikes uitgevoer is, het 'n algemene indruk van die verspreiding van die data verskaf en het die identifisering van statistiesbeduidende veranderlikes bevorder. Hierdie analises het maatstawwe vir sentrale neiging (rekenkundige gemiddelde, mediaan en modus); maatstawwe vir dispersie (minimum en maksimum, variasiewydte, standaardafwyking, skeefheid en kurtose); die samestelling van histogramme vir elke veranderlike; die analise van variansie (ANOVA) op veranderlikes met nominale data; en die skep van 2D-spreidingstippe vir veranderlikes met ordinale data behels. Spearman se rangorde korrelasie is op beide die nominale en ordinale data uitgevoer. Statistiesonbeduidende veranderlikes, of dié wat tekens van multikollineariteit met ander veranderlikes getoon het, is geëlimineer. 'n Beste deelversameling regressie-analise is uitgevoer op die oorblywende veranderlikes.

Die regressiemodel het gepaste resultate behaal deurdat dit byna 54% van die variasie in Houtbaai se huispryse verklaar het. Van die individuele veranderlikes was die grootte van die erf die sterkste voorspeller van die huisprys afhanklike veranderlike, huisgrootte was die tweede belangrikste faktor, terwyl afstand van besige paaie en die kwaliteit van die huis soortgelyke invloed gedeel het. Die regressiemodel se residu's is gekarteer om ruimtelike patrone vas te stel. Dit word aanbeveel dat soortgelyke navorsing op 'n stadswye skaal uitgevoer word, dat die veranderlikes op ander wyses gekwantifiseer word en dat nuwe GIS-tegnieke in toekomstige studies aangewend word.

TREFWOORDE

Eiendomswaardasie, eiendomswaardes, GIS, Houtbaai, huispryse, regressie-analise

ACKNOWLEDGEMENTS

This thesis is the culmination of many years of study at the Department of Geography and Environmental Studies at Stellenbosch University. Due to the long relationship with this department I have come to know many people who have helped me over the years and to whom I owe a great debt of gratitude. The following deserve special mention:

- Dr PH de Necker, retired senior lecturer of the Department of Geography and Environment Studies, for his support throughout my years at the university. His open-door policy, compassion and willingness to help, which extended beyond the call of duty, will always be etched in my memory. It is no exaggeration to say that without his help my studies would never have been completed.
- Retired professors IJ van der Merwe and HL Zietsman, of the Department of Geography and Environment Studies, for their help throughout my studies. Professor Zietsman's assistance and encouragement especially during tough times is deeply appreciated as is his help with some technical aspects of the research.
- My research supervisors, Dr A van Niekerk of the Department of Geology, Geography and Environment Studies, and Professor JZ Bloom of the Department of Business Management, Stellenbosch University, for their unstinting help and advice.

Others who have also been of great help must be singled out:

- Mr P Guise, City of Cape Town municipality, for his invaluable assistance in providing various data for my study area.
- Professor M Kidd, Centre for Statistical Consultation at Stellenbosch University, for always being available and willing to help with the statistical analyses and to answer the many questions I had.
- Mr L van der Westhuizen, who spent many uncomplaining hours helping me in many areas of my life.
- My dear mother who constantly encouraged and exhorted me to finish this thesis, but sadly will not have the opportunity to see it completed. I dedicate this thesis to her memory.
- And last, I thank the triune God who has been at my side throughout my university career.

CONTENTS

TABLES

FIGURES

CHAPTER 1: INTRODUCTION

1.1 THE GEOGRAPHY OF HOUSING

According to Bourne (1981), the study of housing does not fit well within any particular discipline or field of study. However, a geographical perspective on housing research is an imperative as housing can never be divorced from the physical and social environments in which it is located. The geography of housing depends on a unique combination of local and national factors. The local factors are a city's specific location, topography, transport system, income level of the population, social diversity, economic base, and housing stock. To study housing effectively, analysis at a national level is vitally important, for example the analysis of the national economy and the examination of the socio-political system, demographic change, migration, national housing policy, social preferences and constraints on the freedom of choice (Bourne 1981; Smith 2000).

The national socio-political system plays a defining role in the geography of housing because housing may be viewed as the spatial expression of income, wealth, race, and gender disparities in society (Smith 2000). Access to private-sector housing is often dependent on high incomes to sustain mortgage repayments in affluent areas, while poor people are often forced to live in the worst available public sector housing. In Britain, racial stereotypes of reputability are often translated into offers for better tenancies for white applicants and inferior housing for their black counterparts (Smith 2000), while in the USA black and Hispanic homebuyers are regularly subjected to significant levels of adverse treatment compared to whites in both the rental and sales markets of public sector housing (Turner & Ross 2005). Geographic steering also occurs where estate agents systematically show whites houses in predominantly white suburbs, while minority groups are "steered" toward minority or mixed neighbourhoods (Turner & Ross 2005). Access to housing for women in many countries is also very difficult unless they are part of a traditional nuclear family (Johnson 1994). Residential space thus underpins the income, race and gender inequalities inherent to the wider organization of society. Where people live, whether in private or public housing, or in affluent or poor suburbs is thus largely a product of who they are (Smith 2000).

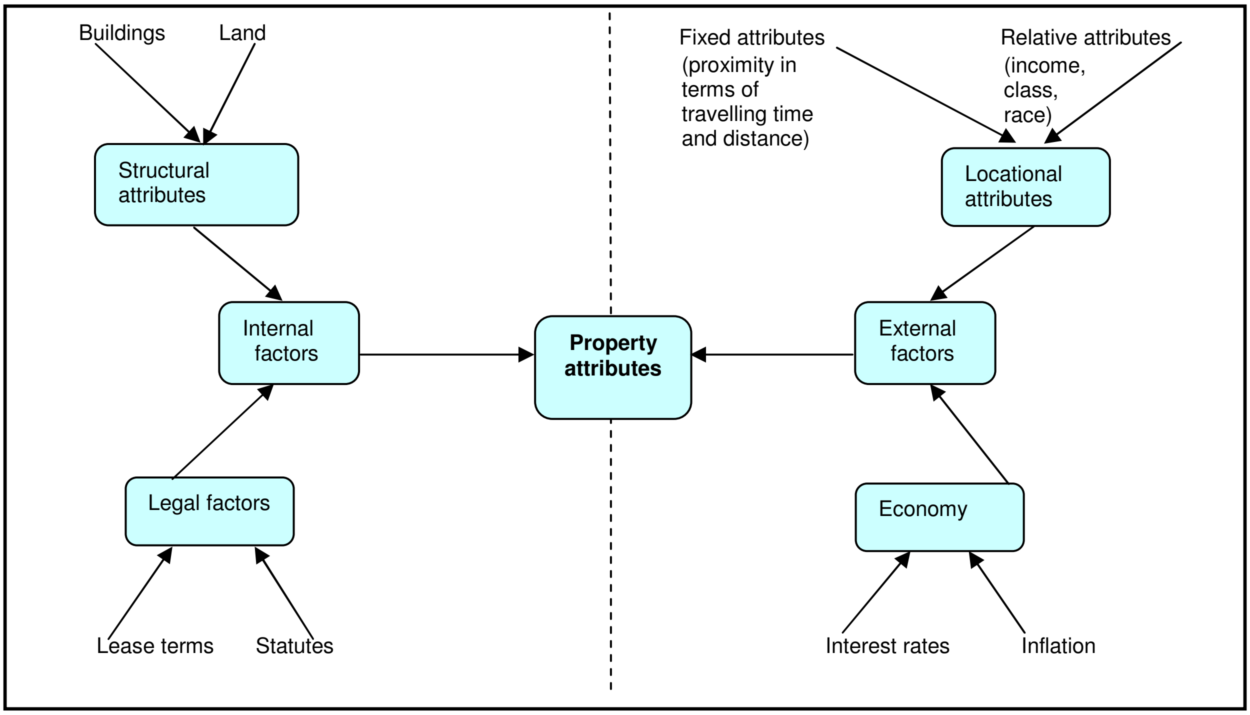
The above factors lead to several important spatial manifestations in the geography of housing of which the particular pattern of urban housing is a foremost expression (Bourne 1981). Spatial regularities in the age of the housing stock, housing density, tenure, structure type and house prices may thus be identified, even on a national scale as Wheeler et al.'s (2007) research shows.

The concentration of specific features of the housing stock, particularly those that are relatively scarce (such as high-value dwellings), and the wide variability and diversity of housing attributes within suburbs previously thought to be homogeneous, are two other spatial manifestations of these factors (Bourne 1981).

Housing is the building-stock equivalent of residential land use and must compete with other types of land use (e.g. manufacturing) within the urban land market for a particular plot of land. Theoretically, this leads to the so-called bid-rent (or location rent) curve which manifests itself spatially with zones of homogeneous land uses radiating from the city centre (Bourne 1981; Carter 1995; Pacione 2005). Three types of residential land use are usually identified in this idealized model, namely high-density housing, low-density housing and exurban residential developments. This model provides a useful spatial framework to analyse the complex working of the urban housing market (Bourne 1981). The factors influencing supply and demand in this market are numerous. These factors lead to certain processes of change within the urban housing market, such as housing occupancy and turnover, vacancy chains, filtering, revitalization and gentrification (Bourne 1981; Smith 2000; Kaplan, Wheeler & Holloway 2008; Hamnett 2009). Finally, the spatial variation of housing prices is a key manifestation of these processes of change (Bourne 1981). It is in this latter focus area within the geography of housing that this study is positioned. In the next section, the attributes of housing and their bearing on house prices are discussed. The reasons why people move home are also briefly investigated.

1.2 ATTRIBUTES OF HOUSING AND RESIDENTIAL RELOCATION

Housing has many features. Primary among these is the provision of shelter – a basic need for human survival. Buying a house is, at its basic level, a means toward achieving shelter from the elements and protection from hostile people and predators. In addition, Beamish, Goss & Immel (2001), drawing on the work of Abraham Maslow in the 1940s, contend that a home can provide a sense of belonging through family interaction, fulfilling self-esteem needs, as well as self-actualization or the fulfilment of one’s potential. A home, whether it is a shack in an informal settlement, a flat or a mansion, can be considered to consist of a complex package of attributes (Dodgson & Topham 1990). According to Orford (2002), property attributes have traditionally been divided into structural and locational attributes, each subcomponent of which may contribute value to (or subtract value from) the property (Sirmans & Macpherson 2003). The structural and locational property attribute dichotomy is schematically depicted in Figure 1.1.



Source: Adapted from Wyatt & Ralphs (2003: 271)

Figure 1.1 A dichotomous taxonomy of factors affecting property value

Structural attributes describe the internal physical structure of the property such as its size, layout and design, construction quality and amenities, and the land parcel within which it is situated, whereas locational attributes address the external geographic characteristics of the property, in fixed and relative terms. Fixed locational attributes are features of location unique to each property, and usually refer to an accessibility measure (Orford 2002), for example proximity in terms of travelling time or distance to the central business district (CBD), employment and transportation opportunities, places of worship, schools and other amenities. A household thus derives utility from accessibility and this benefit is reflected in a house's price (Jackson, Kupke & Rossini 2006). The geographic position of a house also determines its proximity to disamenities such as rubbish dumps, airports, highways, toxic waste sites, factories, and informal settlements, all of which may negatively affect its price.

Relative locational attributes refer to neighbourhood measures that are shared by many properties. These could be measures of neighbourhood characteristics, socio-economic class and racial composition. These measures are typically constructed from aggregated census data and are often very wide-ranging in their measurement (Orford 2002). House prices are not only affected by the quality and quantity of these individual structural and locational attributes, but also by their combinations (Sirmans & Macpherson 2003).

Housing attributes not covered by the above taxonomy are their limited adaptability, durability, the heterogeneity of the housing stock and services it produces, sensitivity to influences exogenous to the housing market, and their subjection to numerous institutional regulations by various levels of government (Bourne 1981). The house itself provides a basic structure plus enhancements added to that structure over time. The structure affords floor area, room volume and an architectural style. Improvements are features such as fireplaces, carports and garages, security measures and swimming pools (Adams 1984), and typically South African amenities like lapas, indoor braai rooms, home theatres and home gyms.

Residential relocation can either be induced or voluntary. Forced relocation may commonly occur in certain parts of a city due to demolition of a building or eviction of its residents. However, most households and individuals move by their own volition. Despite this, the catalyst for voluntary relocation may be brought about externally, for example due to employment opportunities that may arise or changes that occur in the life cycles of households and individuals. Generally, households move with the expectation of improving their living conditions (Pacione 2005). When people make the decision to buy a home, they are in effect purchasing a bundle of structural, locational and neighbourhood housing attributes (Dodgson & Topham 1990) which maximize the potential for living condition improvement. Clark & Onaka's (1983) research found that housing characteristics such as space, quality, housing design and a desire to move from rental to owner-occupation were of primary importance in inducing people to move. Neighbourhood considerations were less important, as were accessibility benefits that arose due to the relocation process.

Pacione (2005) names several factors which affect residential migration behaviour. These are individual and household characteristics (life cycle and household density); spatial differences in economic opportunity; societal and cultural norms; the attitude of particular households and individuals to risk and their adaptability to change; and the information (volume and accuracy) available to them. In combination, the first four factors generate migration goals, while the last factor affects the likelihood of achieving them. The interaction between migration goals and expectancy of achievement results in migration intention. Households thus make compromises based on these factors when they decide where to buy, how much to pay, whether to maintain or renovate, or whether to move or stay where they are (Can 1998; Pacione 2005). Household income constraints and the price at which sellers are willing to sell their property must always be considered in these decisions (Dodgson & Topham 1990).

1.3 PROJECT DESCRIPTION

In the first two parts of this chapter the geographical importance of housing was emphasized, and the attributes of housing and the reasons for residential relocation described. The next section provides an overview of the project, its aim and specific objectives.

1.3.1 Research problem

As noted in the previous section, each house's value is dependent on a unique combination of its structural and locational characteristics. Housing is thus a composite commodity composed of distinct bundles of attributes whose characteristics fluctuate between individual properties. The price of a house is a realization of the combined price of these attributes, although they are never directly observed in property transactions. Although the structural and locational attribute distinction is fundamentally important, there is little theory offering guidance in determining exactly which of these attributes to include in regression models of property valuation (Orford 2002). This is because the results from previous property valuation studies are location-specific and it is difficult to draw generalizations of value-adding or value-reducing factors and applying them to research in other housing submarkets (Sirmans, Macpherson & Ziets 1995). These limitations present difficulties for modelling house prices in the Hout Bay area. The research problem addressed in this study is how to ascertain the primary determinants of house prices in Hout Bay and the relative importance of each factor.

Specifically, the research will endeavour to answer three questions. First, what housing attributes in Hout Bay do prospective homeowners value highly and, as a result, are willing to pay a premium on a house that satisfies them? Second, what housing factors influence homebuyers to purposefully avoid a property in Hout Bay, and which therefore have a depreciating effect on prices? Last, what factors in Hout Bay previously thought to be significant, are in fact negligible?

1.3.2 Research aim

The aim of this research encompasses three interwoven aspects. In essence, the research will attempt to ascertain which factors are generally perceived to affect property prices in Hout Bay; the extent to which individual factors affect house prices; and the role these variables collectively play in determining house prices. Structural as well as fixed and relative locational attributes will be considered in the study, but macro-economic factors such as interest rates and inflation (see Figure 1.1), which can be used to determine general levels of value, will be disregarded as their effects within a geographical area tend to be unvarying (Gallimore, Fletcher & Carter 1996).

1.3.3 Objectives

If one deconstructs the research aim into its constituent parts, the following objectives are identifiable:

- Identify potential house price-influencing factors by conducting a series of structured interviews with estate agents in Hout Bay and by consulting the relevant literature;
- Use a geographical information system (GIS) and other methods to quantify the selected locational variables.
- Determine key explanatory variables by statistically distinguishing between significant and insignificant variables; and
- Apply best-subsets regression analysis to deduce the collective and individual influences of the significant factors affecting house prices in Hout Bay.

Using Mouton's (2003) classification of research design types, this study can be categorised as an empirical study using statistical modelling as its primary mode of analysis. The research design is illustrated in Figure 1.2. The boxes in the left column show the major steps that will be followed to complete the research, while the boxes in the right column provide more details on the specific procedures required during each step to produce the desired results.

1.4 STUDY AREA

Hout Bay is a geographically-isolated suburb of Cape Town, situated on the western seaboard of the Cape Peninsula (see Figure 1.3). Hout Bay's spatial separation is ideal for this study because extraneous influences on the internal property market are likely to be limited, although it's location on the Western Cape coastline may make it a suitable location for second-home development (Visser 2003). Hout Bay is characterized by magnificent mountain and sea vistas and rural charm, making it an area of high aesthetic quality (Oelofse & Dodson 1997). Hout Bay is bisected along the Hout Bay River by a narrow central strip of uncultivated open space and smallholdings. Commercial activities are centrally located in a large area between Princess Street and Main Road. In 2001, roughly 21 000 people were living in Hout Bay of whom nearly 39% were black, approximately 33% white and slightly more than 27% coloured, with Asians comprising a very small minority of less than 0.5% of the population.

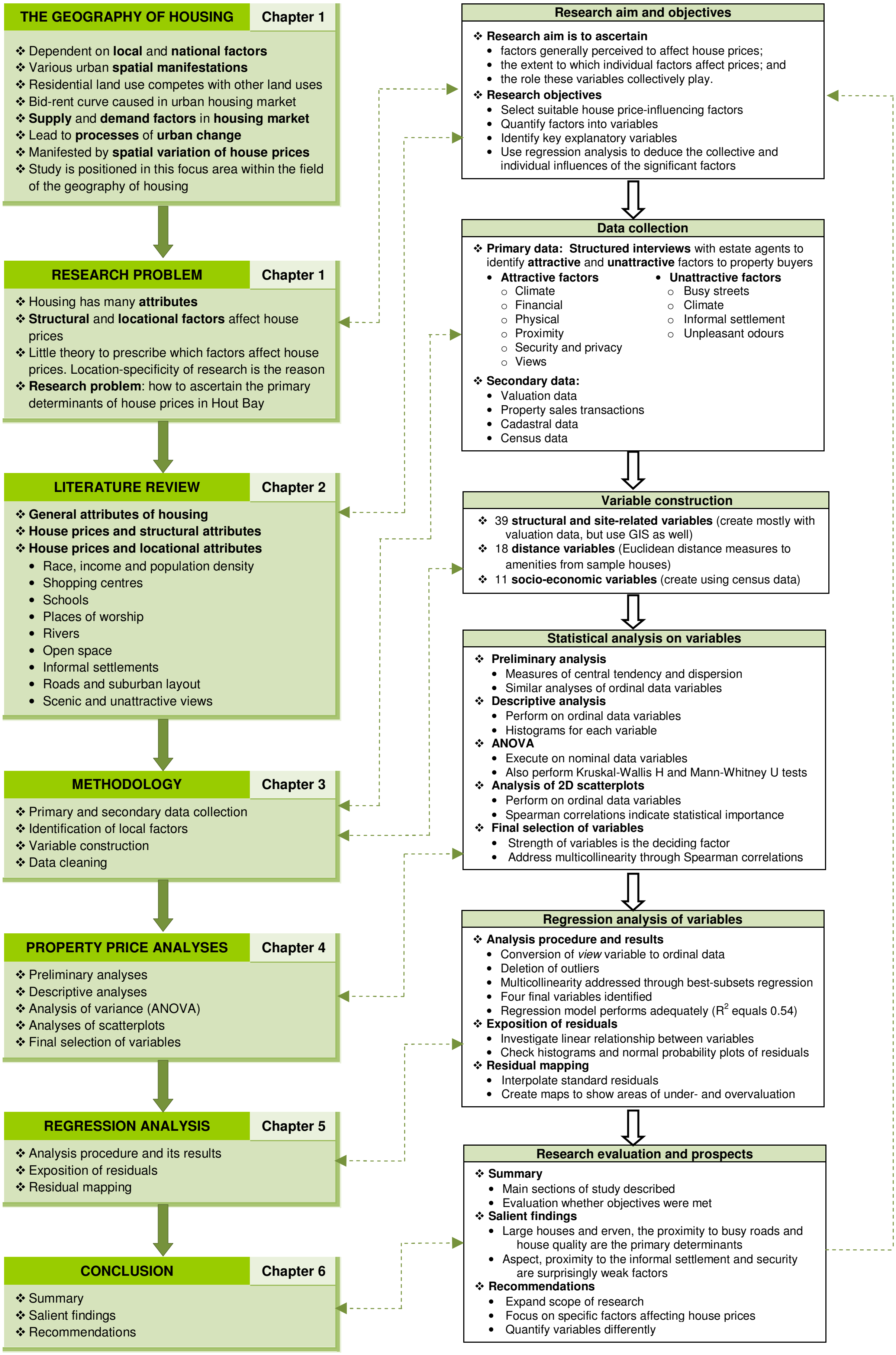


Figure 1.2 Research design for assessing house price determinants

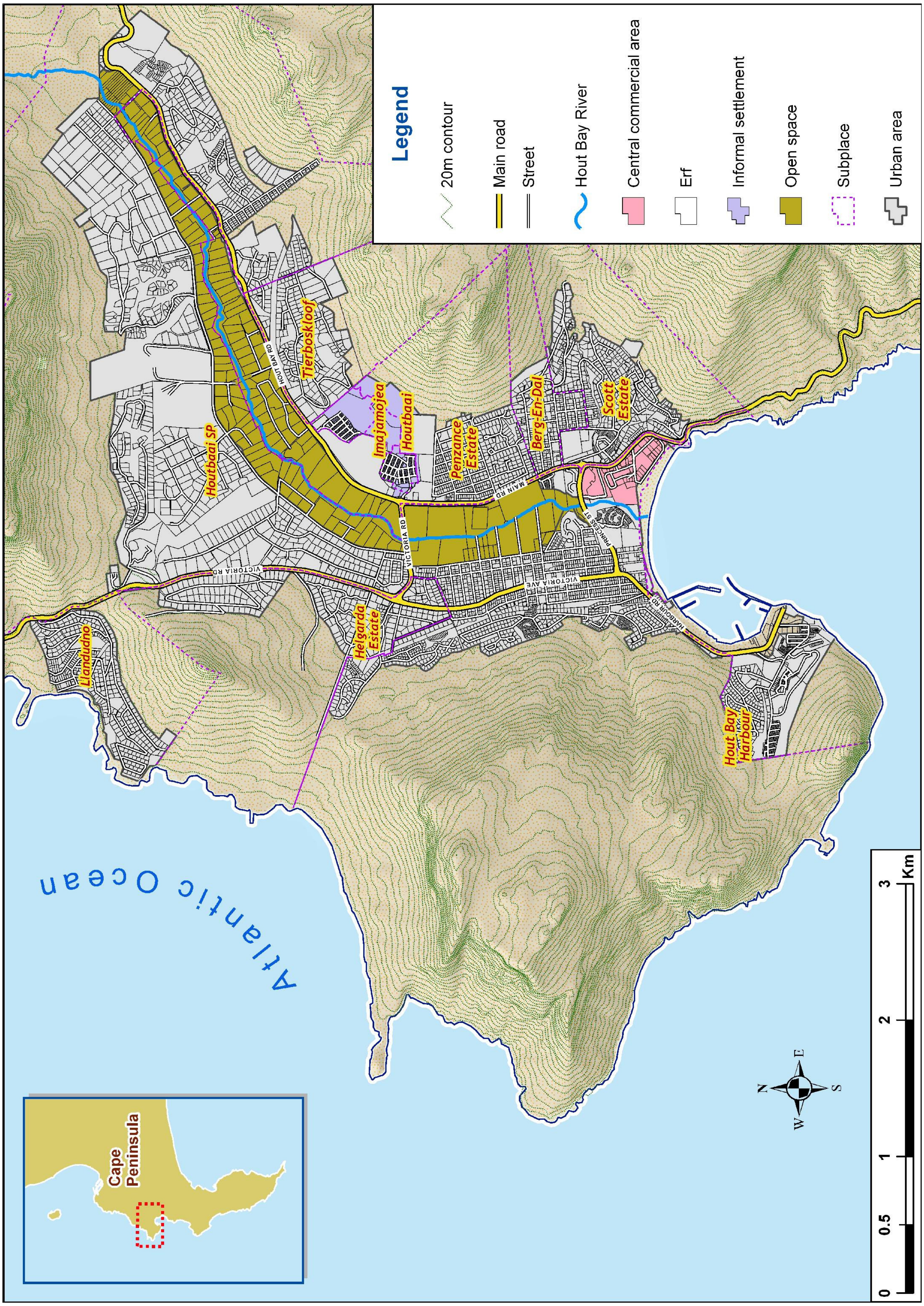


Figure 1.3 The Hout Bay study area

Hout Bay accommodates three demographically distinct communities. One community occupies the largest part of Hout Bay's physical area, which includes all the census 2001 subplaces in Figure 1.3, except Imajamojea, Houtbaai, Hout Bay Harbour and Llandudno. This community lives in medium-density, middle- to upper-class housing, with the residents having a high socio-economic status (Scheepers & Bloom 2005). The majority of the nearly 8000 people living here are white (nearly 92%) and English-speaking (almost 84%). Indicators of this area's socio-economic status are its residents' low unemployment rate (less than 3%) and high level of education (nearly 40% had post-matric qualifications and less than 3% had no schooling in 2001). Commercial activities in this area are located in Victoria Avenue.

Another demographically distinct community lives in the Hout Bay Harbour area, which largely corresponds to the Hout Bay Harbour subplace (see Figure 1.3). This area includes the Hout Bay industrial area and its fish factories, the beach, as well as some restaurants, shops and other tourist-related attractions. Housing mainly consists of high-density public housing. Of the about 5500 people in this area, nearly 94% of the residents are coloured and approximately 85% speak Afrikaans as their home language. Residents have a low socio-economic status (Scheepers & Bloom 2005) characterized by the poor education levels of the residents with almost 7% having no education and only 1.5% possessing a post-matric qualification in 2001. The unemployment rate was relatively high at approximately 18% in 2001.

The third community lives in the Imizamo Yethu township which consists of the Imajamojea and Houtbaai subplaces (see Figure 1.3) and it is wedged between the Penzance Estate and Tierboskloof subplaces. In 2001 this area had almost 8100 residents who live in housing conditions typical of an informal settlement. Its residents are overwhelmingly African (approximately 95%) and Xhosa-speaking (nearly 83%) and have a low socio-economic status (Scheepers & Bloom 2005). Large-scale unemployment existed in 2001 (approximately 40%) with close to 14% of residents having had no schooling and only 1.5% having obtained a post-matric qualification.

1.5 CONCLUSION

The first four sections of this chapter introduced the research by establishing the importance of housing, its attributes and their bearing on house prices. The research problem, research aim and objectives were stated and the study area demarcated. The structure of the remainder of the thesis is as follows: Chapter 2 reviews the literature on the structural and locational attributes that affect house prices; Chapter 3 is devoted to determining the specific factors that affect house prices in

Hout Bay and explains the transformation of these factors into variables. The most significant variables are distinguished through various statistical analyses described in Chapter 4, while Chapter 5 discusses the methods used to determine the collective and individual impacts of the variables on house prices. The thesis is concluded with a summary of the study's salient findings followed by recommendations for further study (Chapter 6).

CHAPTER 2: LITERATURE REVIEW

2.1 INTRODUCTION

This chapter partially addresses the first objective of the study, namely to identify general house price-influencing factors by evaluating the literature on structural and locational determinants of house prices. The purpose of this chapter is not to review the entire body of research pertaining to factors influencing property values or techniques used to analyse them, as it is too voluminous. Instead, it appraises the literature that focuses on residential property valuation and on the factors expected to affect house prices in Hout Bay.

2.2 DETERMINANTS OF HOUSE PRICES

A lack of a theoretical foundation for prescribing which variables to include in regression models of property valuation is evident in the literature so that disagreement on this issue has resulted in a wide range of variables being included in such models (Orford 2002). An indication of the many possible variables that can be used is given by Tables 2.1 and 2.2. Swanepoel's (1996) list of variables that may affect property values includes both structural and locational attributes (see Table 2.1). Mackmin's (1994) inventory of negative factors contains items that discerning buyers may view as advantageous (see Table 2.2).

Table 2.1 Factors which affect the value of a property

Structural attributes	**Locational attributes**			
	Accessibility	**Social**	**Infrastructure**	**Physical**
Plot area	Distance to central business district	Socio-economic status	Neighbourhood condition	Pollution levels
House area	Distance to other nodes	Ethnic composition	Noise level	Slope
Age of property	Distance to businesses (small or large)	Crime rate	Traffic	Soil type
General condition of property	Distance to place of work	Population density	Condition of roads	Bad smells/smoke
Foundation type	Distance to nearest hospital	Quality of adjoining properties	Availability of services	Land use for: • Single families
Wall type	Distance to other medical institutions			• Multi-families
Roof type	Distance to highways			• Offices
Number of living rooms	Distance to railways			• Trade and industry
Number of bedrooms	Distance to airports			• Transport
Number of bathrooms	Distance to schools			• Parks
Garages	Distance to universities			• Sport facilities
Cellars	Distance to sport and recreation centres			• Other land uses
Heating type and/or air conditioning				
Swimming pools				Views
Gardens				
Fireplaces				
Fencing of property				

Source: Swanepoel (1996: no page)

Table 2.2 Factors which negatively influence property values

Proximity to	Poor planning	Other
Abattoir	Site layout	Poor condition
Airport – civil or military	Position of bathroom	Lack of adequate damp course, ventilation, daylight
Builder's yard, builder's merchants	Position of toilet	Lack of water supply
Bus depot	Awkward shape of bedrooms	Lack of street lighting
Bus stop	Access through other rooms	Land liable to flooding
Car park	Poorly-designed kitchens	Land liable to subsidence/sea erosion etc.
Caravan site	Behind adjoining building lines	Isolated position or very poor or restricted access
Cemetery	Back-to-front design	Environmental – radon gas etc.
Chapel of rest	Poor frontage	
Church halls	Lack of storage	
Cinemas/bingo halls	Differences in floor levels	
Clubhouses – general/jazz	High heating costs due to poor/inadequate insulation	
Council estate	Lack of drainage	
Derelict land and properties	Lack of electricity	
Discotheque		
Electricity pylons		
Factory – noise, smell or other nuisance		
Footpaths		
Fried-fish shop/takeaway		
Funfair		
Garage and petrol station		
Gasholders		
Henhouses		
Hospital emergency halls		
Loading depots		
Nursing home		
Piggeries		
Refuse tip – household disposal		
Rifle range – clay-pigeon shoots		
Silage		
Telegraph poles		
Traffic/traffic fumes		
Undertakers		
Warehouse – retail warehousing		

Source: Mackmin (1994: 32)

The following sections provide an overview of the literature relating to structural and locational attributes affecting housing values. Studies about structural attributes are briefly investigated first, followed by a more comprehensive discussion of locational attributes. These studies are location-specific, making it difficult to draw inferences from them across geographic boundaries. By comparing studies in different areas one can single out those characteristics that are consistently valued (either positively or negatively) by homebuyers (Sirmans, Macpherson & Ziets 1995). The discussion in the next section demonstrates this point.

2.3 HOUSE PRICES AND STRUCTURAL ATTRIBUTES

Gallimore, Fletcher & Carter (1996) only considered data about the physical characteristics (e.g. property type, age, size, number of storeys and bedrooms) of residential properties in their study of the valuation of residential properties in Stafford, United Kingdom. They determined the variance between selling prices and prices predicted by multiple regression analysis and identified areas of over- and under-valuation. The ratios of over- and under-valuation were plotted to identify positive- and negative-valued factors, the impact of which could be determined by measuring the distance from the influence to the property. Figueroa (1999) disregarded locational factors in a study of property prices in Regina, Canada by combining actual property sales data with structural variables (e.g. property size, age and number of fireplaces) of properties to create two hedonic regression models. He found that each square foot of total living area of a property added more than $7 to each property's value in the first model and more than $60 in the second model. Attached garages, fireplaces and decks also had a value-adding effect. Sarip's (2005) study in Kuala Lumpur, Malaysia used an artificial neural network (ANN) approach to model several structural attributes including land area, building age, number of storeys, number of bedrooms and quality of finishes. The results of the analysis compared favourably with actual property prices.

Sirmans et al. (2006) performed a meta analysis of research in the USA on the estimated regression coefficients of the nine housing characteristics (building size, erf extent, age, bedrooms, bathrooms, garage, swimming pool, fireplace, and air conditioning) that have appeared most often in hedonic pricing models for single-family homes. Of these the building size, erf extent, house age, bathrooms, swimming pool and air conditioning coefficients were found to be sensitive to geographical location, while bedrooms, garage, and fireplace were not affected by this feature. Wolverton's (1997) study which was not included in the meta analysis found that erf extent was a significant determinant of house price, but with the noteworthy feature that residential erf price per square foot weakens as erf extent increases.

McMillen (2008) found that house prices in Chicago increase with erf extent, building size, and the number of bathrooms; and prices decrease with increasing age of the housing structure. House prices are also higher for brick homes with air conditioning, a fireplace, and a garage. After controlling for building size, he discovered that with respect to the number of rooms and bedrooms in a house, subdividing a specified area into smaller rooms did not automatically increase a home's sales price. Zietz, Zietz & Sirmans' (2008) research in Utah showed that buyers of higher-priced houses regard features such as building size, erf extent, bathrooms, and floor type differently than purchasers of cheaper houses. For example, additional bathrooms have a much stronger value-adding effect in more expensive houses than in lower-priced ones.

2.4. HOUSE PRICES AND LOCATIONAL ATTRIBUTES

The literature is replete with examples of research explaining the amenity or disamenity effect of various local factors on property values. Some recent studies focus on oil and gas facilities (Boxall, Chan & McMillan 2005), derelict industrial areas (Kaufman & Cloutier 2006), golf courses (Nicholls & Crompton 2007), pollution sites (Brasington & Hite 2005) and rubbish dumps (Kiel & Williams 2007). This section focuses on factors presumptively and intuitively seen to influence Hout Bay house prices.

2.4.1 Race, income and population density

Race, income and population density are relative locational attributes which have a neighbourhood-level influence on house prices. Lynch & Rasmussen (2004) investigated the influence of income and racial composition on property prices in Jacksonville, Florida, and found that property values declined in the presence of relatively low-income and black households in a wide area around a property. Cervero & Duncan's (2004) research analysed the effects of a neighbourhood's racial diversity on property prices and found that houses in racially-diverse suburbs sold for substantially less than houses in more racially-homogeneous areas. Harris (1999) found that house prices decreased by up to 16% in areas with more than 10% blacks, confirming previous findings that whites generally prefer not to have blacks as neighbours. This aversion apparently stems from socio-economic class considerations rather than an inherent dislike of blacks. People generally prefer to have prosperous and well-educated neighbours, traits perceived by some whites to be less common among blacks. Regarding population density as a neighbourhood influence, Bae, Jun & Park (2003) found that property prices were negatively associated with this influence but positively with employment opportunities.

2.4.2 Shopping centres

According to Des Rosiers et al. (1996), a shopping centre holds both attraction and aversion to homeowners in its vicinity. The desirability owes to the convenience of the relative trouble-free access to the numerous shopping and entertainment facilities located in the centre, while the dislike stems from the noise, overcrowding and pollution which proximity may engender. Depending on the shopping centre's relative closeness and size, house prices should reveal the collective effect of positive and negative externalities. Des Rosiers et al.'s (1996) case study in the Quebec region of Canada found that larger shopping centres were conducive to higher residential property values. Song & Sohn (2006) report that easy access to retail centres had a similar positive effect on house prices with the effect rapidly decreasing with increased distance from the centre.

2.4.3 Schools

Given that education is the cornerstone of future social and economic progress, the quality of schools in an area plays a vital role in the decisions of households to locate, buy or move to or from a particular property (Can 1998). Many parents prefer to send their children to a school close to home because of safety and travel-time considerations. Choosing a nearby school also allows parents to supervise and influence a child's school activities so that prospective homeowners will be willing to pay more for a home in order to gain the advantage of good accessibility to a school (Chin & Foong 2006). However, accessibility to a school is not the only consideration for home-choice decisions as the quality of the school is also of prime importance. Clark & Herrin's (2000) study in Fresno County in California concluded that particular school attributes significantly increase the selling price of homes. Some of the school attributes they tested were teacher-student ratio, number of schools per district, school size and percentage of ethnic enrolments. Chin & Foong (2006) found that good accessibility to prestigious schools in Singapore significantly increased property prices.

2.4.4 Places of worship

Places of worship, as with shopping centres, have characteristics which may cause potentially negative effects on nearby house prices. Do, Wilbur & Short (1994) mention negative issues associated with churches. These include traffic and parking congestion, bell clangs and other loud sounds emanating from a church, loss of scenic views, disturbances caused by the provision of meals and temporary housing for the poor, and church architecture perceived to be incompatible with the surrounding area. Conversely, places of worship have features which may be perceived in a positive light by some homeowners: acting as moral beacons, gathering places for social interaction and sometimes operating as schools are such examples (Do, Wilbur & Short 1994). Their study in Chula Vista, California, revealed that churches have a negative effect on surrounding house values with the effect increasing as the distance from the church decreases. Carroll, Clauretie & Jensen's (1996) research in Henderson, Nevada, however found that house prices decrease as distance from a suburban church increases. Their results were supported by the findings that large churches have a greater positive effect on house prices and churches can therefore be regarded as an amenity rather than a disamenity for the surrounding suburb. As Hout Bay has a number of churches and a mosque, cognizance needs to be taken of these findings.

2.4.5 Rivers

The Department of Water Affairs and Forestry (2003) has reported faecal contamination from sewerage in the Hout Bay River. Leggett & Bockstael's (2000) study in Arundel County, Maryland, examined the effects of water quality on waterfront house prices and observed that faecal coliform levels had a negative effect on property prices. Urban homeowners may also derive benefits from rivers adjacent to or near their properties such as scenic views, open space, bird and wildlife viewing, and buffering from urban noise (Colby and Wishart 2002). Despite the water pollution of the Hout Bay River, some of these benefits may be present, especially in the upper reaches of the river. Colby and Wishart's (2002) research in the Tucson metropolitan area found that riparian corridors contribute significantly to nearby house prices, particularly those homes located within the first 0.8 km of the riparian area.

2.4.6 Open space

Open space as a land use may have important benefits for urban dwellers. These include recreation opportunities, scenic views, the absence of negative side effects associated with development, conservation of the natural landscape, reduction of water run-off and erosion, serving as a buffer from incompatible land uses and provision of pedestrian and cycle routes (Geoghegan 2002; Van Wyk 2005). Van Wyk (2005) also mentions problems associated with open space, such as illegal dumping of refuse and open spaces serving as havens for vagrants and criminals. Studies draw a distinction between preserved open space, examples being public parks and land under conservation, and developable open space such as privately-owned agricultural land. This is important because the effect of open space on property values mirrors both current benefits and potential developments (Geoghegan 2002). Geoghegan's (2002) study in Howard County, Maryland, demonstrated that preserved open space increased nearby residential land values over three times as much as an equivalent amount of developable open space. Irwin (2002) also used the preserved and developable open space distinction in her research which revealed that surrounding open space significantly influenced the sales price of houses, so corroborating Geoghegan's (2002) findings that the spillover effects of preserved open space were noticeably higher than developable open space.

Locally, Van Zyl & Leiman (2002) discovered that the Zandvlei wetland, a unique example of preserved open space, did not have an identifiable influence on surrounding house prices in the southern suburbs of Cape Town. Hout Bay has an attractive beach, which is another form of preserved open space. According to Pompe & Rinehart (1995), beaches offer aesthetic pleasure as well as recreational opportunities and they found in South Carolina that people were willing to pay a premium for properties close to wide beaches. Although Anderson & West's (2006) study in Minneapolis–St. Paul did not explicitly use the preserved–developable open space distinction, their research showed that the influence of proximity to open space is greater in densely built-up areas, those near the CBD and areas suffering from high crime rates. Similarly, open spaces are more important in high-income areas and suburbs which are home to many children.

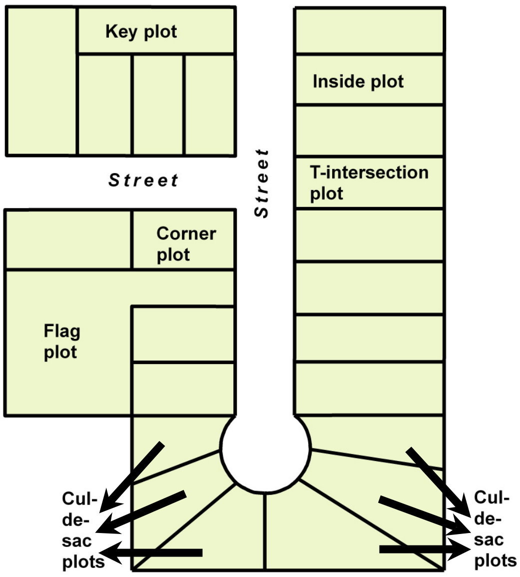
2.4.7 Informal settlements

An informal settlement is a land use type many people deem to have disamenity value, especially if the neighbouring suburbs are middle-class areas. These settlements are often perceived to be a form of visual pollution, places which lead to increased crime in adjacent suburbs, and areas which decrease property values and cause environmental degradation (Dixon, Reicher & Foster 1997; Saff 1998). Given the presence of Hout Bay's large Imizamo Yethu informal settlement, Saff's (1998) study to determine whether the Marconi Beam informal settlement in Cape Town affected house prices in the neighbouring suburb of Milnerton is pertinent. Interestingly, he found that the informal settlement had a negligible effect on property prices, but that there was some evidence that properties close to the settlement were difficult to sell.

2.4.8 Roads and suburban layout

The benefit or disamenity values of urban road systems are dependent on their relative location to a house and the perceived advantage they generate for homeowners. For example, Lake et al.'s (1998) study in Glasgow showed that each decibel increase in road noise depressed property prices by slightly more than 1%. In contrast, Cervero & Duncan (2004) found that in Santa Clara, California, properties accessible to jobs via highway and transit networks commanded higher prices. The suburban plot layout shown in Figure 2.1 has important consequences for existing and prospective homeowners.

Of particular significance are corner plots and plots in cul-de-sacs. Corner plots front on two intersecting streets and usually have access to both streets (Betts & Ely 2001). Theoretically, a corner plot should command a premium as it provides easy access to the rear yard, a sought-after feature in suburbs where the parking of trailers and caravans in front of a house is prohibited. Such plots are popular with people operating home-based businesses, for example doctors and lawyers who need a side entrance for access by clients or patients (Betts & Ely 2001). However, Asabere (1990) noted that such plots are targets for burglars as they offer convenient escape routes.



Source: Adapted from Betts & Ely (2001: 113)

Figure 2.1 Common types of urban plots

A cul-de-sac plot is situated at the end of a dead-end street. These plots generally have little street frontage, which may make parking difficult, but this potential disadvantage is often outweighed by large backyards and additional privacy (Betts & Ely 2001). Another disadvantage is the lack of exits for emergency and delivery vehicles (Asabere 1990). These plots are popular for family housing because of the large play areas at the back and greater safety owing to reduced street traffic (Betts & Ely 2001). Asabere's (1990) research in Halifax, Canada, found that a location in a cul-de-sac commanded a price premium of between 22 and 28%, while Matthews & Turnbull's (2007) study in King County, Washington, revealed that in automobile-orientated neighbourhoods based on curvilinear and cul-de-sac street patterns there was no discernible effect on house prices.

2.4.9 Scenic and unattractive views

According to Bond, Seiler & Seiler (2002), research on the impact of views on property values dates from the early 1970s with Darling's (1973) pioneer study. Despite this relatively long history of research effort, determining the influences that visibility and scenic views have on house prices is not a straightforward process. This is because views of the same landscape object tend to vary in quality for different properties and therefore do not impart a constant monetary value to a property. Related to this is the subjective nature of what constitutes a good, average, or bad view (Bond, Seiler & Seiler 2002). These difficulties are further compounded by the reality that property owners, especially in coastal areas, have many view possibilities, particularly of water, such as views of the sea, estuaries, and marshes (Rinehart & Pompe 1999).

Attempts to circumvent these problems have led to the use of GIS techniques such as viewshed analysis to impartially determine the extent of a view and its influence. However, Bourassa, Hoesli & Sun (2004) believe that these techniques should be approached with caution as they are prone to measurement error. Lake et al.'s (1998) study is regarded as the first to incorporate these methods (Yu, Han & Chai 2007). They examined the importance of negative impacts associated with road development, specifically noise and visual intrusion, on house prices in Glasgow, United Kingdom. Using various techniques, including viewshed analysis, they found that the visual impact of roads reduced the average property price by nearly 3%. Lake et al. (2000) later followed a similar methodology as Lake et al. (1998) to investigate the impact of visibility of land use composition on house prices in the same city. One of the visual impact variables in their research, namely railroads visible from the back of properties, was found to have a significantly negative impact on property values. Both Paterson & Boyle's (2002) and Yu, Han & Chai's (2007) research applied viewshed and hedonic regression analysis. Paterson & Boyle (2002) found that visibility of land use and land cover were important determinants of residential property prices in Simsbury and Avon, Connecticut. Yu, Han & Chai's (2007) work on the effect of sea views in Singapore showed that an unobstructed sea view added an average of 15% to the house price.

The view variables used in earlier research were derived from data obtained by site inspection either personally by the researcher or from a property appraiser or assessor's database. Using assessor data, Rodriguez & Sirmans (1994) found in Fairfax County, Virginia, that views were positively correlated to sale price. No specific types of view were studied, rather any view a potential buyer would find appealing was considered. Seiler, Bond & Seiler (2001) and Bond, Seiler & Seiler (2002) used basic measures of view in their research. In both studies the view variables were derived by simply determining if properties had a view of Lake Erie or not. Seiler, Bond & Seiler (2001) found that a lake view added approximately 56% to a home's value, while Bond, Seiler & Seiler (2002) concluded that a lake view added close to a 90% premium to a house's value.

The generic view variables used in the above studies may not be suitable where views in an area vary by type or quality (Benson et al. 1998). To counteract this shortcoming Benson et al. (1998) personally inspected properties in Bellingham, Oregon, and distinguished three types of view (lake, sea and mountain) as well as three quality categories (partial, superior and full) to indicate the extent of each view from a particular property. By means of hedonic regression analysis it was found that the highest-quality ocean views increased a house's value by almost 60%, while the lowest-quality ocean views added around 8%. Wolverton (1997) raised concerns about valuation models that allow only for the existence of view, rather than the quality of view. He considered such models as crude and inaccurate. Consequently, his study in Tucson, Arizona, measured the quality of city view, but did not distinguish between different types of view. He determined the width of each property's city view angle and made adjustments for obstructions caused by nearby properties. The results of the regression analysis showed that property values are positively influenced by quality of view. Fraser & Spencer's (1998) regression analysis of property values in Western Australia incorporated three dimensions of ocean view quality, namely degree of view panorama, potential loss of view through obstructions and elevation of the property above sea level, the first two of which were incorporated in Wolverton's (1997) research.

Cellphone towers are visually-intrusive urban features widely perceived to reduce house prices. As Hout Bay has a high telecommunications tower located in its suburban area, Bond & Wang's (2005) research on the influence of cellphone towers on property values in Christchurch, New Zealand, is pertinent. No view measures were used in their analysis, but a questionnaire survey indicated that around 20% of the respondents perceived cellphone towers to affect neighbourhood aesthetics and more than 10% said they negatively affect views from their houses. Bond & Wang's (2005) analysis confirmed that property values were reduced by about 21% after the erection of a cellphone base station in a neighbourhood.

All the studies reviewed in this subsection confirm the positive link between attractive views and increased house prices. Lake et al.'s (1998; 2000) and Bond & Wang's (2005) studies also verify the correlation between visually-intrusive landscape features and decreased property values.

2.5 CONCLUSION

This chapter has reviewed the literature on residential property valuation and on factors used as variables in this study. It has succeeded in providing the necessary theoretical guidance on whether or not a locational property attribute can be considered to be an amenity or disamenity or if unique local conditions cause uncertainty about an attribute's status. This has partially supported the attainment of the first objective of the study, namely to identify house price-influencing factors. The next chapter is devoted to investigating which factors affect house prices in Hout Bay and to the transformation of these factors into variables useful for statistical analysis.

CHAPTER 3: DEVELOPING QUANTITATIVE MEASURES OF HOUSE PRICE DETERMINANTS

3.1 INTRODUCTION

This chapter explains the methodology used to accomplish the first and second objectives of the study, namely to identify factors influencing Hout Bay's residential property values and to quantify these factors as appropriate explanatory variables. The collection of primary and secondary data is first discussed, followed by a description of the techniques used to transform identified factors into measures suitable for statistical analysis. Finally, the data-cleaning procedures that were followed are described.

3.2 DATA COLLECTION

The main source of data in this study is a database obtained from the City of Cape Town municipality containing the registered property transactions in Hout Bay and Llandudno from 1981 to 2006. The 2000 general valuation data, also obtained from the municipality, was linked to the property transactions file, together forming an extremely rich database. The database contains the Surveyor-General's 21-digit code for each house sold which enabled it to be linked to digital cadastral data also containing the 21-digit code. Aerial photographs taken in 2005, as well as roads, rivers and zoning data in ESRI shapefile format and a 20-metre digital elevation model (DEM) were obtained. The researcher visited the study area and street guides, tourism maps and local telephone directories were consulted to extract information on the existence and location of schools, churches, factories, the post office, cemeteries, the informal settlement, public libraries, open space areas and shopping centres. Although these data are available digitally from various sources, it was considered important to verify their completeness and accuracy in the field. Compact discs containing census 2001 data were obtained from Statistics South Africa. In addition to all the above secondary data and information, primary data were obtained from structured interviews with local estate agents. This process and the selection of variables from the interview data are discussed next.

3.3 IDENTIFICATION OF LOCAL FACTORS

The selection of appropriate local factors which affect house prices in Hout Bay is essential to this study. In addition to the factors identified in the literature review, structured interviews were conducted with one estate agent at each of 12 estate agencies in Hout Bay to get a general impression of the local property market. This was done on the assumption that they have firsthand knowledge of factors that prospective property owners either look for or avoid when buying property in Hout Bay. This information was used to compile a list of important determinants of house prices in the area. The structured interview topics were formulated in questionnaire format (see Appendix) which was emailed or faxed to estate agents prior to the interview. Twenty-six estate agencies were identified by consulting Hout Bay's local telephone directory and they were contacted during September 2005, but due to the busy schedules of agents only 12 estate agencies participated. The interviewees often consulted their colleagues during the interviews, therefore one interview per estate agency was regarded as sufficient for data-collection purposes. The interviews did not require interviewees to rank the factors in any order of importance, however they have been grouped in related categories in Table 3.1, which summarizes the estate agents' responses concerning attractive factors.

Table 3.1 Factors that are attractive to property buyers in Hout Bay according to estate agents

Factors	**Estate agent**												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total**
Views													
Mountain views	X	X								X		X	4
Other views	X			X									2
Sea views	X	X	X	X	X	X		X	X	X	X		10
Total													**16**
Security and privacy													
Gated communities	X		X				X						3
Neighbourhood/area					X				X				2
Privacy									X				1
Security	X	X		X	X	X			X				6
Spacious grounds					X								1
Total													**13**
Proximity factors													
Far from busy Cape Town CBD										X			1
Not near busy roads/road noise					X				X				2
Proximity to commercial and other amenities			X		X	X	X				X		5
Sports fishing										X			1
Total													**9**
Physical factors													
Beach	X										X		2
Mountain	X												1
River	X												1
Sea							X					X	2
Total													**6**
Financial factors													
Affordability/price								X		X	X		3
Value for money						X							1
Total													**4**
Climate factors													
North-facing areas			X										1
Sunny areas												X	1
Wind-protected areas			X										1
Total													**3**

By adding the total numbers of responses per category, one can conclude that views, security and privacy, and proximity factors are the most attractive features to potential buyers. Because the questions in the structured interviews specifically asked estate agents to identify factors linked to a house's location that prospective homeowners either want or avoid, most of the factors they named are spatial in nature. Factors that are not explicitly spatial are the financial factors and privacy, although it is reasonable to argue that affordable houses are generally located in the same area and that the larger the properties' grounds the more likely they provide a sense of privacy to owners.

Table 3.2 summarizes the estate agents' responses regarding unattractive factors which have been grouped in related categories. There were relatively few avoidance factors compared to factors that property buyers find attractive.

Table 3.2 Factors that property buyers avoid in Hout Bay according to estate agents

Factors	**Estate agent**												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total**
Informal settlement factors													
Proximity to informal settlement		X		X	X	X	X	X	X	X	X	X	10
View of informal settlement			X		X								2
Total													12
Street factors													
Busy streets	X				X	X	X	X	X				6
New road developments					X								1
Noise	X												1
Total													8
Climate factors													
Areas not facing north										X			1
Dark areas	X												1
Gloomy forest areas												X	1
Windy areas			X							X		X	3
Total													6
Security factors													
Areas with little security									X				1
Undeveloped areas	X												1
Total													2
Unpleasant odours													
Factories (fish smells)					X								1
Harbour (fish smells)							X						1
Total													2

From Table 3.2 it is clear that factors related to the informal settlement and Hout Bay's busy streets are those which prospective homeowners wish to avoid most. Climate-related factors are also noteworthy.

3.4 STRUCTURAL AND SITE-RELATED VARIABLES

The variables discussed in this subsection relate to aspects of the house structure itself or to the land on which the house is built. House sales were extracted for 2005 and 2006 to obtain a snapshot view of the Hout Bay property market. A two-year sales period was used as it was regarded to be a sufficiently long sample period and it obviated the need to adjust sales prices to accommodate the increase in house values. All other sales data were excluded from the database obtained from the City of Cape Town. Only those houses with a database code of “single-dwelling residential” were selected. Sales in Llandudno were excluded after identifying them by means of a field in the file which specified the neighbourhood in which a sale occurred. A total of 396 house sales were selected for analysis. The footprints of the selected houses were digitized from aerial photographs and their centroids were determined in ESRI shapefile format using GIS. The general valuation data were excellent indicators of the structural attributes of a house and their descriptions are given in Table 3.3.

Table 3.3 Structural and site-related variables for determining house prices

Variable name	Description
Aspect	Slope direction of cell in degrees
Attached carport	Attached carport (dummy variable of 1=yes; 0=no)
Attached garage	Attached garage (dummy variable of 1=yes; 0=no)
Building size	Size of the main building on the erf in square metres
Building style	1=sub-economic, 2=conventional, 3=unconventional, 4=Cape Dutch, 5=Victorian, 6=maisonette, 7=group housing
Common walls	1=one, 2= two, 3=none
Condition	Overall condition of the building (e=excellent, g=good, a=average, f=fair, p=poor)
Corner plot	Location of house on corner plot (dummy variable of 1=yes; 0=no)
Covered area	Covered area present (dummy variable of 1=yes; 0=no)
Cul-de-sac	Location of house in cul-de-sac (dummy variable of 1=yes; 0=no)
Detached carport	Free-standing carport (dummy variable of 1=yes; 0=no)
Detached covered area	Detached covered area present (dummy variable of 1=yes; 0=no)
Detached garage	Free-standing garage present (dummy variable of 1=yes; 0=no)
Detached servant's quarters	Detached servant's quarters present (dummy variable of 1=yes; 0=no)
Elevation	Height of house centroid above sea level (metres)
Erf extent	Registered extent of the land in square metres
Month of sale	Month of the year in which the property was sold
Number bedrooms	Number of bedrooms in house
Pergola	Pergola present (dummy variable of 1=yes; 0=no)
Pool	Pool present (dummy variable of 1=yes; 0=no)
Porch	Porch present (dummy variable of 1=yes; 0=no)
Previous purchase price	Purchase price when house was previously sold (Rand)
Quality	Overall quality of the building (e=excellent, g=good, a=average, f=fair, p=poor)
Roofing	1=tile, 2=sheeting, 3=mazista, 4=thatch, 5=other
Security	1=high, 2=medium, 3=low, 4=none
Slope	Maximum rate of change in elevation over cell in degrees
Solar radiation (1 day)	Incoming solar radiation (watt hours per square metre (WH/m^2)) at winter solstice
Solar radiation (5 days)	Incoming solar radiation (watt hours per square metre (WH/m^2)) for five days including winter solstice
Store	Store present (dummy variable of 1=yes; 0=no)
Storeys	Number of dwelling storeys
Terrace/balcony	Terrace/balcony present (dummy variable of 1=yes; 0=no)
Topography	1=flat, 2=undulating/uneven, 3=steep, 4=low & wet
Total bathroom fixtures	Total bathroom fixtures (bath, toilet, shower, basin = 1 fixture)
Total rooms	Total number of rooms in house
Traffic	1=expressway, 2=busy, 3=medium, 4=light, 5=none
View	e=excellent, pa=panoramic, po=partially obstructed, aa=above average, a=average, ba=below average, p=poor

In spite of bedrooms and garages being identified by Sirmans et al. (2006) as factors not affected by geographical location, the *number bedrooms*, *attached garage* and *detached garage* variables were nevertheless included to test if this was the case locally. Despite the caveats given in Section 2.4.9 about using the all-inclusive view measures shown in Table 3.3, the fact that property assessors had already impartially determined the general view quality of houses in Hout Bay made the use of complicated viewshed analysis to derive the *view* variable unnecessary. The same section also discussed the inherent problems of viewshed analysis.

Several of the structural variables in Table 3.3 were constructed by the researcher. By using the DEM as input, the *slope* and *aspect* variables were derived by means of the “Slope” and “Aspect” tools in the ArcGIS 9.2 Spatial Analyst extension. The *solar radiation (1 day)* and *solar radiation (5 days)* variables were calculated by means of the “points solar radiation” tool in the same software extension. The rationale behind the creation of these two variables is that the estate agents had remarked that sunny areas were popular and dark areas were unpopular among prospective homeowners. Tovar-Pescador et al. (2006) note that in areas of complex topography (such as found in Hout Bay) strong local gradients of insolation may occur and a DEM thus affords the ability to examine the role of topographic features in the spatial and temporal distribution of solar radiation.

The *solar radiation (1 day)* variable calculates the insolation (in watt hours per square metre) for the house centroids on the winter solstice, while the *solar radiation (5 days)* variable calculates the insolation (also in watt hours per square metre) during a five-day period, two days before, two days after, and inclusive of the winter solstice (21 June). A skysize of 4000 was selected in both variables. The software package recommends a large skysize where day intervals of less than 14 days are used, so the maximum allowable option was selected. As typical transmitivity values for very clear sky conditions are either 0.6 or 0.7, a value of 0.3 was used as the diffuse radiation model of standard overcast sky was selected in both variables. A diffuse proportion of 0.5 was also selected. The diffuse proportion indicates the proportion of global normal radiation flux that is diffuse (Environmental Systems Research Institute 2008). The cell values of the slope, aspect and solar radiation rasters were transferred to the house centroids shapefile. To create the *elevation* variable, the elevation values of the DEM’s cells were similarly transferred.

The *cul-de-sac* and *corner plot* variables were created by virtue of their importance reported in the literature (see Section 2.4.8). They were derived by overlaying Hout Bay’s road network on aerial photographs of the area using GIS and then visually determining if a house is located on a corner plot or on a cul-de-sac plot. Although the structured interviews explicitly asked estate agents to identify factors specific to a house's location that may affect house prices, several estate agents mentioned factors that were related more to a house's site-related characteristics. The privacy factor is one of these and, as previously stated, it is presumably related to the size of a house's plot. The *erf extent* variable accommodates this influence by using data on the plot's size. Another two site-related factors identified by the estate agents are whether a house is located in a gated community or in a wind-protected area. The problem with creating a variable for the former is that many gated communities are of the sectional-title variety with no individual erf numbers available per home, while the lack of available data precluded the use of the latter as a variable.

3.5 DISTANCE VARIABLES

Fixed locational attributes of houses can best be determined by means of distance measures from the selected houses to specific amenities or disamenities in Hout Bay. The "point distance" function in the commercially-available ArcGIS software tool ET Geowizards was used for this purpose. Euclidean (straight line) distance measures were used because according to Humavindu & Stage (2003) they provide a reasonable approximation of actual travel distance in dense road networks, as found in Hout Bay. These distance variables are summarized in Table 3.4.

Table 3.4 Distance variables for determining house prices

Variable name	Description
250m contour distance	Euclidean distance to 250-metre contour
All commercial distance	Euclidean distance to all commercial areas
All open space distance	Euclidean distance to all open space areas
Beach distance	Euclidean distance to beach
Cemetery distance	Euclidean distance to cemetery
Central open space distance	Euclidean distance to central open space area
Churches distance	Euclidean distance to closest church or mosque
Fish factories distance	Euclidean distance to fish factories
Harbour distance	Euclidean distance to harbour
Informal settlement distance	Euclidean distance to informal settlement
Post office distance	Euclidean distance to post office
Public libraries distance	Euclidean distance to public libraries
River distance	Euclidean distance to Hout Bay River
Roads distance	Euclidean distance to closest busy road
Schools distance	Euclidean distance to closest school
Shopping centres distance	Euclidean distance to shopping centres
Tower distance	Euclidean distance to telecommunications tower
Victoria Avenue distance	Euclidean distance to Victoria Avenue commercial area

More than half of the estate agents stated that proximity to street-related factors is potentially unattractive to property buyers. The busiest roads in Hout Bay are Harbour Road, Princess Street, Victoria Avenue, Victoria Road and Main Road (see Figure 3.1 which shows the large polygon and line features, such as roads, used to create variables).

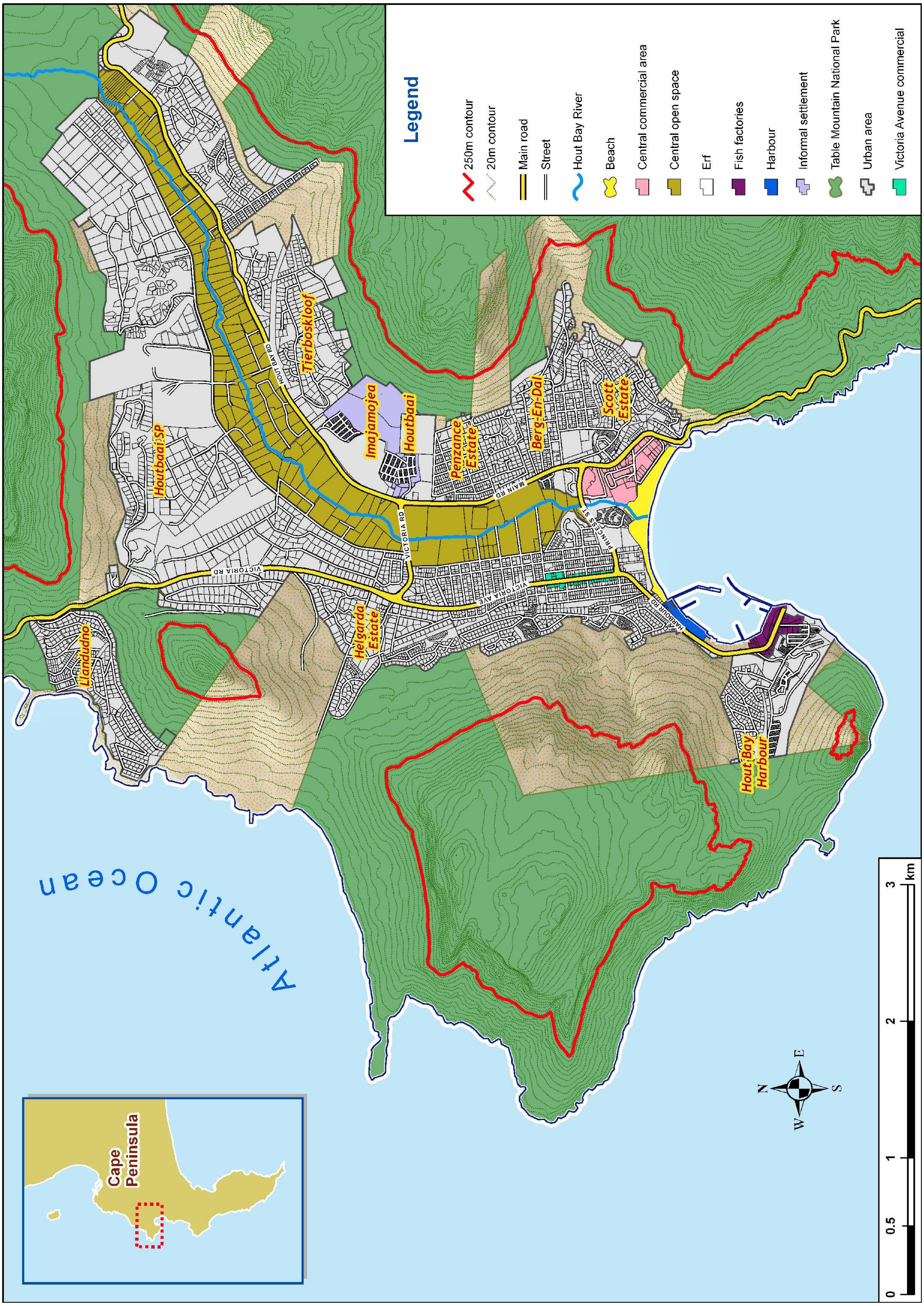


Figure 3.1 Location of large features in Hout Bay used to create variables of house-price determinants

The distances were calculated from houses to these roads and the *roads distance* variable was derived from these distances. One estate agent also stated that proximity to the Hout Bay River was attractive to potential house buyers, despite the faecal pollution present in some parts of the river (Department of Water Affairs and Forestry 2003). The *river distance* variable was thus created to determine the extent of Hout Bay River's influence.

An estate agent mentioned that Hout Bay's mountains are appealing to potential homebuyers. Much of the mountainous areas surrounding Hout Bay forms part of the greater Table Mountain National Park (TMNP) which, according to South African National Parks (2004), has a significant influence on the property market in Cape Town. Their study notes that areas in Cape Town's suburbs with no or distant sea views, but having a view of the TMNP had higher property values than zones in these areas without such views. To quantify the combined influence of Hout Bay's mountains and the TMNP on house prices, the distance of houses to an arbitrarily chosen contour is used. The 250-metre contour was selected to create the *250m contour distance* variable as it includes all the mountains towering around Hout Bay and is close to the last line of houses between the 205- and 210-metre contours. The distance to the TMNP itself is not used as a variable because its boundaries are not uniformly distributed across all of Hout Bay's mountainous areas (see Figure 3.1).

One estate agent identified proximity to Hout Bay's fish factories with their unpleasant odours as an important avoidance factor for prospective house buyers, while another stated that the smells emanating from the adjoining harbour are similarly distasteful. The resultant *fish factories distance* variable was generated by creating a polygon with GIS around all the erven zoned as "general industrial" and which corresponds to the area of Hout Bay where fish processing takes place (see Figure 3.1). The *beach distance, informal settlement distance* and *harbour distance* variables were similarly derived. For the latter variable, a polygon was created that includes only the part of the harbour area that has amenity value. This area includes the well-known Mariner's Wharf and adjoining shops, the Sea Fisheries Museum and the National Sea Rescue Institute (NSRI). The section of the harbour that possibly has disamenity value was excluded as it is best approximated by the *fish factories distance* variable.

The structured interviews suggested that the distance of houses to commercial amenities is important. Three variables were created to accommodate the variegated character of this factor. The first, the *Victoria Avenue distance* variable was created for erven in Victoria Avenue zoned as commercial or those having a municipal database code of "commercial/business". Second, the *shopping centres distance* variable was constructed by measuring distances to the three shopping centres in Hout Bay, namely Victoria Mall, Mainstream Centre and Melkhout Centre (see Figure 3.2). Third, the *all commercial distance* is an all-encompassing variable that calculates distances to all the above commercial areas in Hout Bay. The distance of properties in Hout Bay to the eight primary and secondary schools (see Figure 3.2) is encapsulated in the *schools distance* variable (which excludes crèches and day schools in the area). The *central open space distance* and *all open space distance* variables were created as one estate agent noted that undeveloped areas were a locational feature prospective homeowners avoided when looking a house. The *central open space distance* variable encapsulates only the large open space in central Hout Bay, while the *all open space distance* covers all open space areas. These open spaces were determined by aerial photograph analysis and from digital zoning data.

Some other locational features in Hout Bay's built environment which very likely have an influence on house prices are incorporated as distance variables despite the estate agents not having explicitly identified them during the structured interviews. Their inclusion is endorsed by their relevance having been recorded in the academic literature (recall Chapter 2) or by virtue of their intuitive importance. These variables are: *cemetery distance*, *churches distance*, *public libraries distance* and *tower distance* (distance to the telecommunications tower). Three locational factors identified by estate agents as important house-price drivers and which are quantifiable by accessibility and proximity measures, namely distance to the Cape Town CBD, sports fishing, and distance to the sea, were excluded as variables. The *beach distance* variable, however, is regarded to be an adequate proxy for the influence of the distance to sea. The sports-fishing factor is too vague to quantify as a variable.

3.6 SOCIO-ECONOMIC VARIABLES

Relative locational attributes refer to neighbourhood measures common to many properties. These are important because people tend to view neighbourhoods having particular income, racial and ethnic mixes to be more desirable to live in than other suburbs (Can 1998). This study thus incorporates several socio-economic variables derived from census data as shown in Table 3.5. The census data were extracted, manipulated and linked by means of a table join through subplace codes to a polygon shapefile of the Hout Bay area. These census data were transferred through a spatial join to the house centroids shapefile.

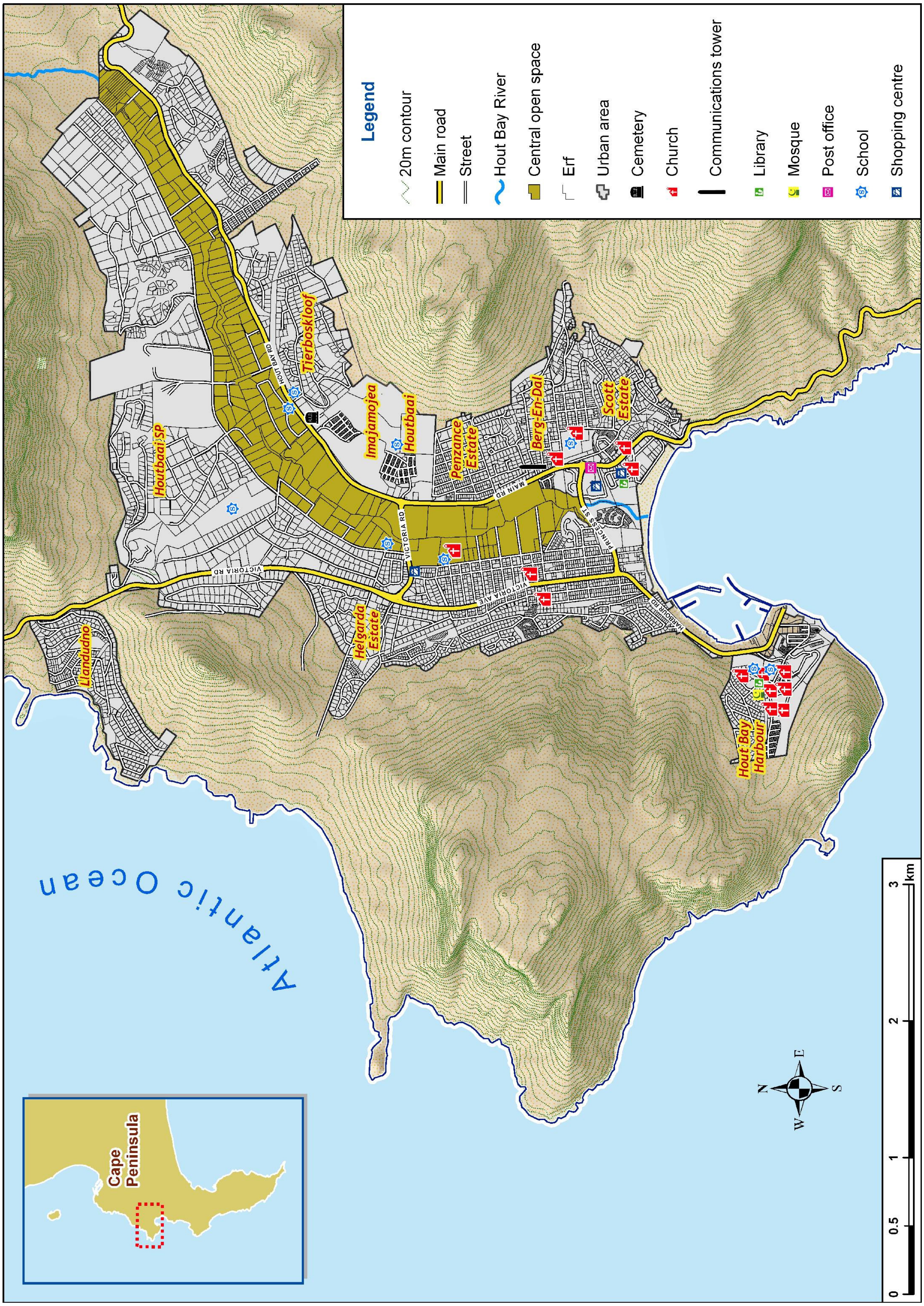


Figure 3.2 Location of small features in Hout Bay used to create variables of house-price determinants

Table 3.5 Socio-economic variables for determining house prices

Variable name	Description
Age 1 average	Sum of the central point of age intervals (0-18, 19-30, 31-40, 41-50, 51-65, and over 66) multiplied by proportion of age group present in subplace
Age 2 average	Sum of the central point of age intervals (0-14,15-24, 25-44, 45-64 and over 65) multiplied by proportion of age group present in subplace
Age 3 average	Sum of the central point of age intervals (0-54 and over 55) multiplied by proportion of age group present in subplace
Annual household income	Average annual household income in subplace (Rand)
Employed	Percentage employed people in subplace
Household size	Average household size in subplace (number of persons)
Individual monthly income	Average individual monthly income in subplace (Rand)
Population density	Population density in subplace (persons per km^2)
Qualification	Qualification levels ranked according to level and allocated to each record according to most prevalent in subplace
Race	Racial group with highest proportion in subplace
Unemployed	Percentage unemployed people in subplace

The *age 1 average* variable's intervals indicated in Table 3.5 were created according to the researcher's own judgment of important age categories, while Thériault et al. (2003) employed the age intervals used in the *age 2 average* variable and Lynch & Rasmussen (2004) those in the *age 3 average* variable.

3.7 FINAL DATA CLEANING

As a last step, the attribute table of the house centroids shapefile was inspected to test the integrity of the data. Nineteen records were found to have missing or zero values for a number of fields and when plotted in a GIS it was noted that they were randomly located throughout Hout Bay. These records were subsequently deleted due to their incompleteness which made them unsuitable for inclusion in the data set. This final step resulted in a data set of 377 sales records to be subjected to property price analyses.

3.8 CONCLUSION

This chapter has succeeded in fulfilling the first and second objectives of the study, namely to identify the factors affecting Hout Bay house prices and to quantify these factors as variables. The primary and secondary data sources which assist in describing the structural and locational characteristics of houses in Hout Bay were discussed, followed by an explanation of the important techniques used to convert these features into appropriate house-price variables. The next chapter is devoted to describing the techniques used to identify the essential descriptive variables from all the variables generated above.

CHAPTER 4: PROPERTY PRICE ANALYSES

4.1 INTRODUCTION

This chapter describes the steps followed to reach the third objective of the study, namely to identify key house-price explanatory variables for regression analysis. The final data-cleaning step discussed in the previous chapter prepared the final records for mapping to show the spatial distribution of property sales values. This is illustrated in Figure 4.1 where the value categories have been classified using the natural-breaks method. The cleaned data were imported into the Statistica statistical analysis software package, and several analyses were performed on the variables. Descriptions of these analyses and the results are presented in this chapter, starting with a preliminary analysis of central tendency and dispersion, followed by descriptive analyses using histograms, then an analysis of variance and 2D scatterplots. The chapter concludes with the selection of variables for regression analysis.

4.2 PRELIMINARY ANALYSIS

As a first step, a cursory analysis was performed on the property sales values. This involved the calculation of measures of central tendency (mean, median and mode) and measures of dispersion (minimum and maximum values, range, standard deviation, skewness and kurtosis). The purpose of these statistics is to describe the characteristics of the *house price* dependent variable. The results are shown in Table 4.1.

Table 4.1 Summary statistics of residential property sales prices in Hout Bay in 2005-06

Valid N	Mean	Median	Minimum	Maximum	Range	Standard deviation	Skewness	Kurtosis
377	R1 976 932	R1 750 000	R80 000	R10 000 000	R9 920 000	R1 058 668	2.8	15.8

The average and median house prices in Hout Bay for the 2005-06 period were indicated respectively by values close to R2 million and R1.8 million. The minimum (R80 000) and maximum (R10 million) sales prices represented a wide price range of about R9.9 million. The sales values are also positively skewed and have a marked peaked distribution. A similar preliminary statistical investigation of the ordinal data variables produced the results shown in Table 4.2. Due to their inherent nature, it was not possible to perform this type of analysis on nominal data variables.

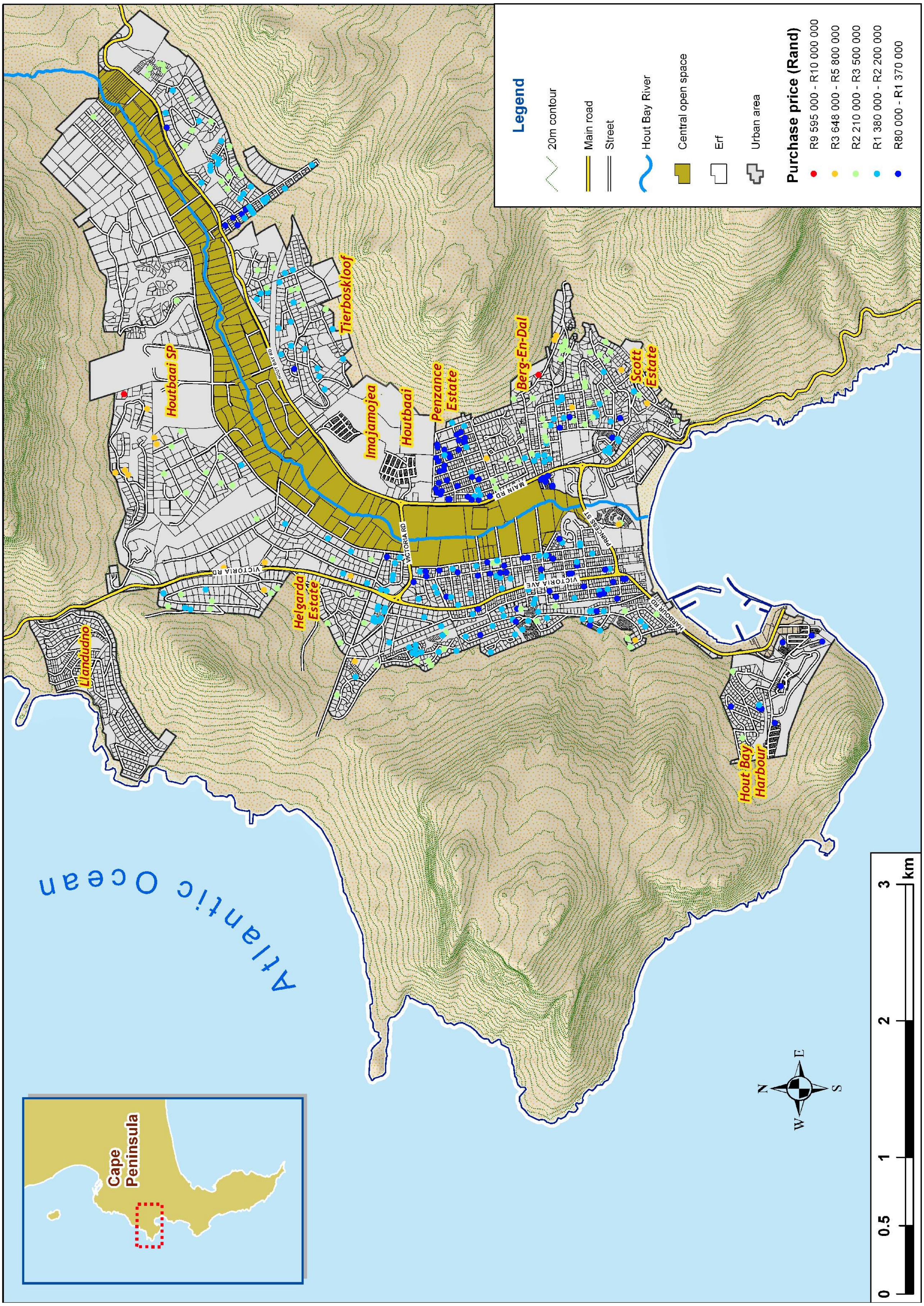


Figure 4.1 Distribution of houses in Hout Bay according to sales price

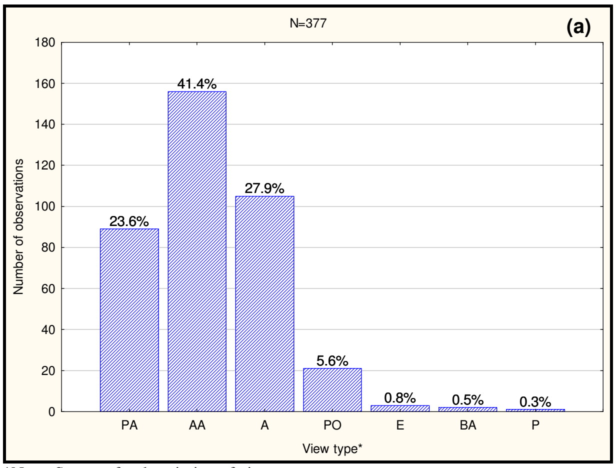
Table 4.2 Summary statistics for ordinal data variables

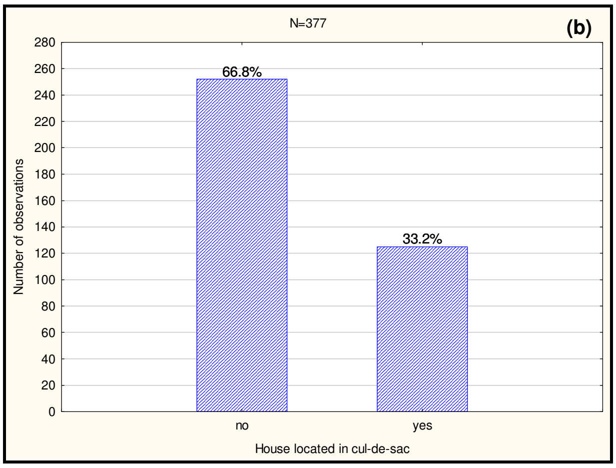
Variable	Valid N	Mean	Median	Minimum	Maximum	Range	Standard deviation	Skewness	Kurtosis
250m contour distance (m)	377	711.7	690.1	126.4	1339.2	1212.8	251.1	0.2	-0.5
Age 1 average	377	34.4	34.0	27.7	37.0	9.4	1.9	-1.5	3.6
Age 2 average	377	34.6	34.2	27.4	37.9	10.5	2.0	-1.4	4.0
Age 3 average	377	31.2	31.0	28.0	33.0	5.0	1.2	-0.8	1.3
All commercial distance (m)	377	490.8	385.8	8.0	1821.9	1813.9	395.7	1.4	1.7
All open space distance (m)	377	92.3	71.2	4.8	356.5	351.8	76.3	1.2	1.1
Annual household income(R)	377	316992	287147	54900	444094	389194	72961	-0.5	3.5
Beach distance (m)	377	1667.7	1281.8	59.5	4805.2	4745.7	1175.4	0.8	-0.4
Building size (m^2)	377	187.1	173.0	31.0	628.0	597.0	81.8	1.2	2.6
Cemetery distance (m)	377	1713.0	1703.4	162.2	4036.0	3873.8	652.5	0.7	1.9
Central open space distance (m)	377	919.9	560.7	19.4	3698.9	3679.5	899.6	1.3	0.8
Churches distance (m)	377	922.8	483.0	7.4	4013.5	4006.1	1027.3	1.4	0.8
Elevation (metres above sea level)	377	55.7	47.9	4.4	207.7	203.3	40.8	1.0	1.1
Employed	377	0.6	0.7	0.5	0.7	0.1	0.1	-0.9	-0.4
Erf extent (m^2)	377	1005.6	694.0	67.0	8565.0	8498.0	937.4	3.4	17.8
Fish factories distance (m)	377	2678.8	2329.7	44.7	6106.6	6061.9	1320.7	0.7	-0.3
Harbour distance (m)	377	2094.8	1746.3	68.4	5481.1	5412.7	1276.1	0.9	-0.2
Household size	377	2.9	2.7	2.7	4.6	1.9	0.3	4.2	18.5
Individual monthly income (R)	377	11706	12672	1615	15146	13531	2520	-2.0	5.4
Informal settlement distance (m)	377	1158.1	1120.9	121.3	3160.3	3039.0	605.9	0.9	1.3
Number bedrooms	377	3.3	3.0	1.0	6.0	5.0	0.8	0.5	1.4
Population density (persons per km^2)	377	432.4	215.1	215.1	4552.5	4337.5	730.6	5.3	27.1
Post office distance (m)	377	1606.1	1205.5	147.3	4436.3	4289.0	1038.6	0.8	-0.4
Previous purchase price (R)	377	682181	540000	0	4500000	4500000	607536	2.0	6.5
Public libraries distance (m)	377	1667.0	1238.9	138.8	4654.6	4515.8	1111.2	0.8	-0.4
Qualification	377	2.9	3.0	1.0	3.0	2.0	0.4	-3.7	13.1
River distance (m)	377	579.6	529.8	106.7	1793.8	1687.1	312.0	1.2	1.9
Roads distance (m)	377	271.0	218.5	21.6	1107.0	1085.4	213.5	1.4	2.0
Schools distance (m)	377	623.0	512.7	30.7	2560.8	2530.1	503.5	1.8	3.3
Security	376	2.4	2.0	1.0	4.0	3.0	0.8	0.1	-0.5
Shopping centres distance (m)	377	1125.7	715.9	19.3	4119.9	4100.6	971.1	1.4	1.0
Slope (degrees)	377	8.2	7.7	0.1	25.1	25.0	4.2	1.0	1.6
Solar radiation (1 day) (WH/m^2)	377	805.9	808.6	535.4	954.4	419.0	64.4	-0.8	1.9
Solar radiation (5 days) (WH/m^2)	377	2576.3	2584.6	1408.1	3083.7	1675.6	221.4	-1.0	3.1
Storeys	377	1.4	1.0	1.0	3.0	2.0	0.5	0.7	-1.0
Total bathroom fixtures	377	8.8	8.0	2.0	22.0	20.0	3.2	0.9	1.4
Total rooms	377	2.1	2.0	1.0	8.0	7.0	1.0	1.4	4.4
Tower distance (m)	377	1412.2	1072.6	43.9	4134.3	4090.5	961.4	0.9	-0.2
Traffic	377	4.1	4.0	2.0	5.0	3.0	0.7	-0.8	2.0
Unemployed	377	0.03	0.02	0.00	0.18	0.18	0.03	3.12	10.85
Victoria Avenue distance (m)	377	1465.6	1125.9	8.0	4698.7	4690.7	1171.5	1.0	0.1

There are 41 ordinal data variables in Table 4.2. The descriptive statistics for two of them, namely the *building size* and *roads distance* variables, are discussed in the next section.

4.3 DESCRIPTIVE ANALYSIS

Following the preliminary analyses of the sales values and ordinal data variables, a descriptive analysis was performed on all the variables. This involved the compilation of histograms for each variable to visually determine how the values in the data are distributed. Two selected histograms are illustrated in Figure 4.2 and described below. The x-axis (horizontal) in each histogram indicates the data category and the y-axis (vertical) the number of observations. As views were identified by estate agents as attractive to homebuyers, the *view* variable was selected for discussion. Figure 4.2(a), representing the *view* variable's distribution, shows that a large cross section (more than 41%) of houses had above average (AA) views, with a substantial proportion (almost 28%) having average (A) views, and around 24% having panoramic (PA) views. Partially obstructed (PO) views were fourth in importance (nearly 6%), while the excellent (E), below average (BA) and poor (P) view classes were negligible. Because location on cul-de-sac plots is important according to the literature, the *cul-de-sac* variable is selected for discussion. Figure 4.2(b) shows that a substantial number (one third) of Hout Bay houses are located in a cul-de-sac.





*Note: See text for description of view types

Figure 4.2 Histograms of two nominal data variables, *view* (a) and *cul-de-sac* (b)

The histograms described above are both based on nominal data, whereas the *building size* variable contains ordinal data. This structural variable warrants further description owing to its intuitive importance. Figure 4.3 is a histogram of the variable and it provides some descriptive statistics. The size of houses was spread around a median of 173 square metres and averaged 187 square metres. About 64% of the houses measured up to 200 square metres and overall the size of houses varied considerably between a diminutive 31 and a very large 628 square metres.

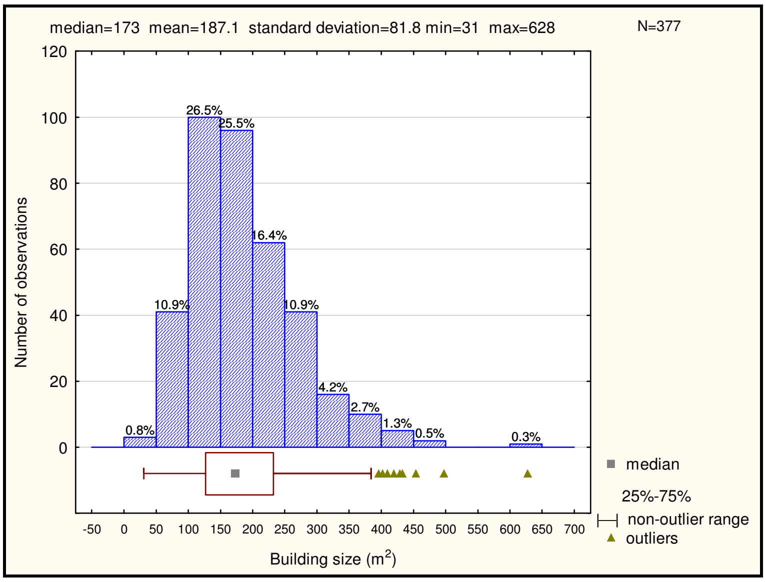


Figure 4.3 Histogram of the *building size* variable

The *roads distance* ordinal data variable is selected for discussion because most of the interviewed estate agents in Hout Bay identified street-related factors as important. A histogram and descriptive statistics of the variable are shown in Figure 4.4. Houses are located, on average, about 270 metres away from busy roads, with a median distance of nearly 220 metres. Eightyseven per cent of the houses lie within 500 metres of busy roads, while the house farthest away was slightly more than 1.1 km distant.

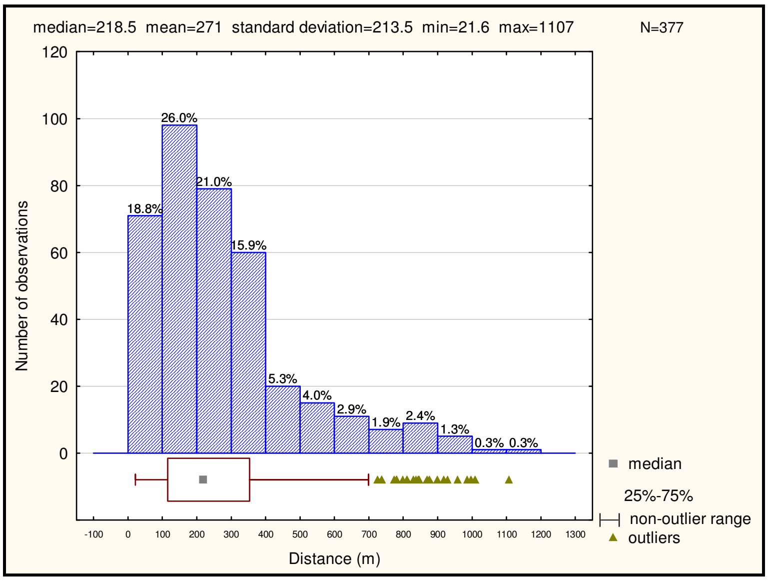


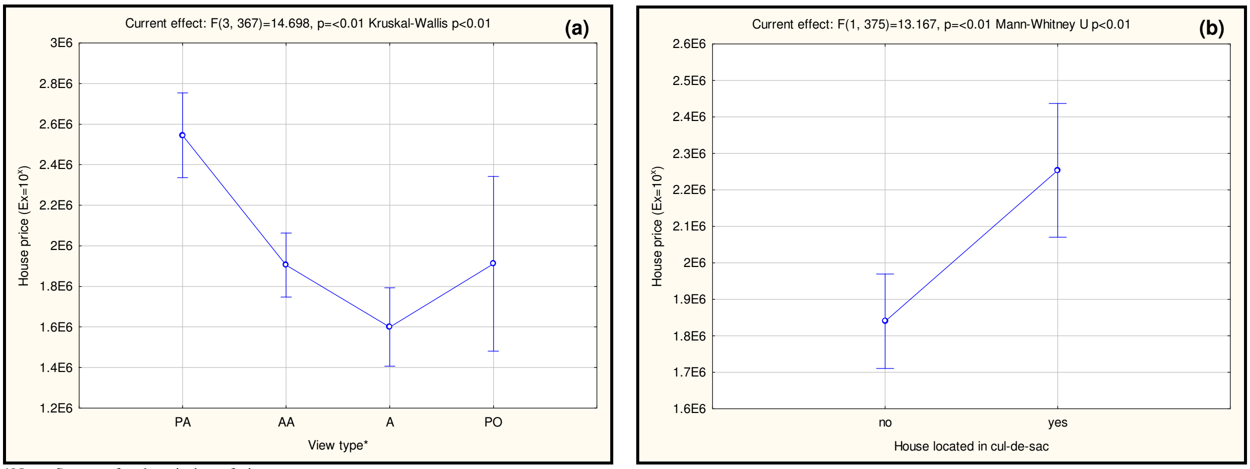
Figure 4.4 Histogram of the *roads distance* variable

All of the variables were examined similarly but they are not discussed here. The methods used to ascertain the statistical significance of variables are discussed in Sections 4.4 to 4.6.

4.4 ANALYSIS OF VARIANCE (ANOVA)

Analysis of variance (ANOVA) on nominal data variables was performed to determine if there is a statistical relationship between the *house price* dependent variable and the predictor or independent variables. To determine this, the p-values' statistical importance were assessed using a 5% significance level. Two examples of ANOVA examinations are provided in Figure 4.5 where the *view* and *cul-de-sac* variable cases are illustrated. The vertical bar of the *view* variable's "panoramic" (PA) category depicted in Figure 4.5(a) indicates that 95% of the houses sold which have such views were in the highest price range (R2.3 to R2.7 million). The same percentage of houses with above average (AA) views fetched prices ranging from R1.7 to R2.1 million, and those houses with average (A) views were in the R1.4 to R1.8 million range. Houses with partially obstructed (PO) views fell in the widest price range (R1.5 to R2.4 million). The excellent, below average and poor view categories were omitted as they are negligible. The significance level (p=<0.01) points to views being significant in determining house prices.

The vertical bar of the "yes" category in Figure 4.5(b) shows that 95% of houses located on cul-de-sac plots sold for between R2.1 and R2.4 million and the same percentage of houses not located in cul-de-sacs were in the lower R1.7 to R2 million range. The significance level of 1% shows that location in a cul-de-sac is a dependable house-price predictor. Each of the other nominal data variables were similarly assessed using ANOVA. The Kruskal-Wallis H test was also applied. This nonparametric test is useful for determining if there is a significant difference between three or more groups (Ebdon 1985). It tests the null hypothesis that the samples were drawn from populations with identical distributions. In Figure 4.5(a) the H test's p-value for the view variable is less than 1% and thus there are insufficient grounds to accept the null hypothesis at the 99% confidence level. The Mann-Whitney U test was used instead of the Kruskal-Wallis test where there were only two groups present for a variable as shown in Figure 4.5(b).



*Note: See text for description of view types

Figure 4.5 One-way ANOVA graphs of two variables, *view* (a) and *cul-de-sac* (b)

Table 4.3 summarizes the F- and P-values for all the nominal data variables. *Pool*, *topography*, *view*, *cul-de-sac* and *race* are the five most statistically significant variables.

Table 4.3 Summary statistics of ANOVA tests for nominal data variables

Variable	F-value	P-value
Pool	25.47	p=<0.01
Topography	21.22	p=<0.01
View	14.70	p=<0.01
Cul-de-sac	13.17	p=<0.01
Race	12.06	p=<0.01
Attached garage	4.82	p=0.03
Aspect	4.77	p=<0.01
Building style	4.50	p=0.03
Common walls	4.31	p=0.01
Storeys	3.34	p=0.07
Corner plot	3.28	p=0.07
Terrace/balcony	3.16	p=0.08
Pergola	3.14	p=0.08
Detached covered area	2.00	p=0.16
Detached servants quarters	1.90	p=0.17
Roofing	1.63	p=0.20
Month of sale	0.95	p=0.49
Covered area	0.83	p=0.36
Detached garage	0.78	p=0.38
Attached carport	0.57	p=0.45
Store	0.53	p=0.47
Porch	0.31	p=0.58
Detached carport	0.00	p=0.99

A further part of the analysis involved the creation and analysis of 2D scatterplots of variables containing ordinal data. These analyses are described next.

4.5 ANALYSIS OF 2D SCATTERPLOTS

Scatterplots enable one to visualize the correlation between two variables (Statsoft 2009). Figure 4.6 shows the relationship between the *house price* dependent variable and four selected independent variables. The strength of the relationship between *house price* and the independent variables are indicated on the scatterplots by Spearman rank correlation statistics. Figures 4.6(a) and 4.6(b) show the observations positioned in close proximity to each other around the trend lines (regression lines fitting the data), visually suggesting relatively strong relationships.

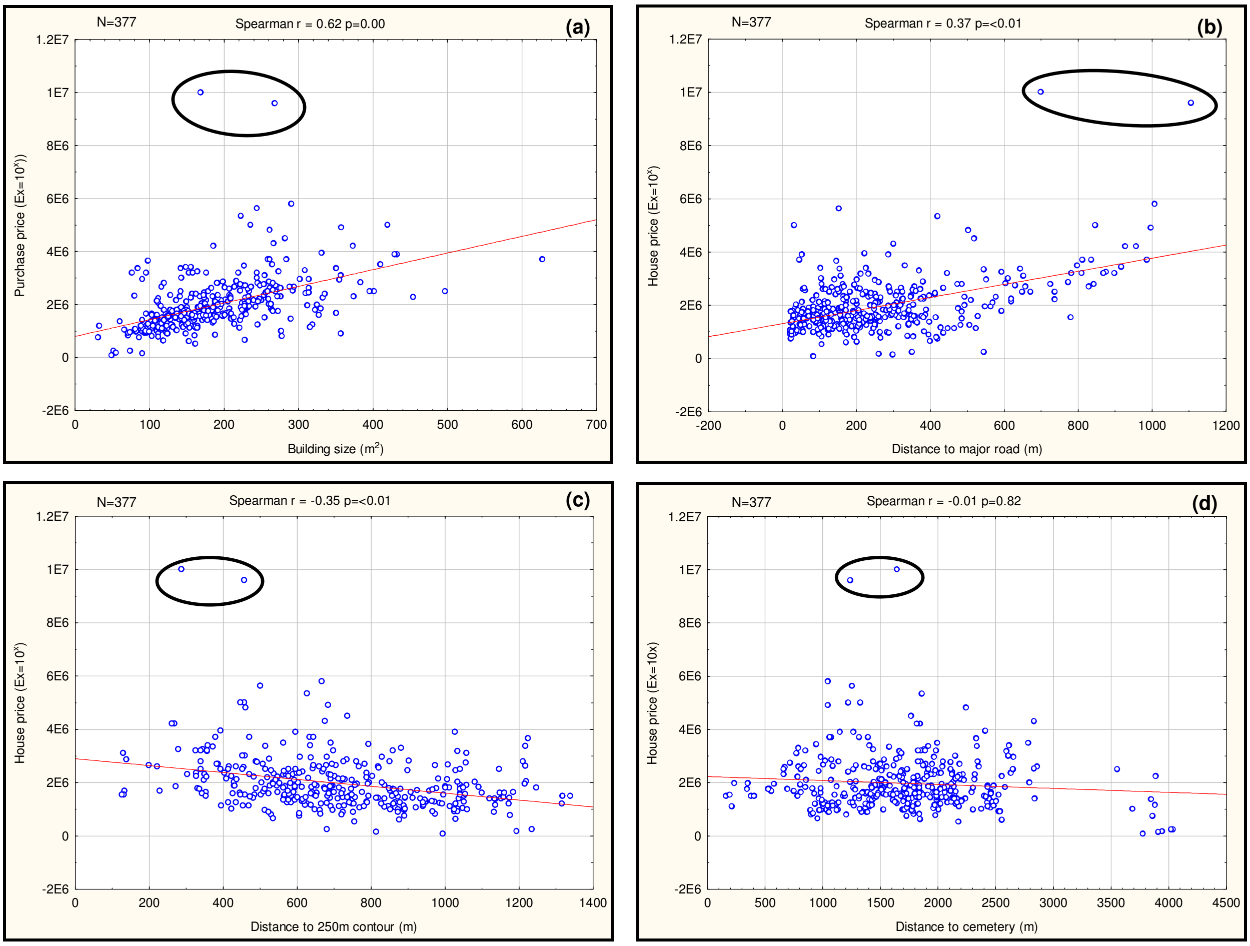


Figure 4.6 Scatterplots of *building size* (a), *roads distance* (b), *250m contour distance* (c), and *cemetery distance* (d)

In Figure 4.6(a) the trend line indicates that house prices predictably increase as the building sizes increase and in Figure 4.5(b) the trend line points to a similar direct relationship between house prices and increased distance from major roads. Also observable in both graphs are two encircled outliers. The Spearman coefficient or r-value of 0.62 for the *building size* variable (Figure 4.6(a)) indicates a substantial positive correlation, while the r-value (0.37) for the *roads distance* variable

(Figure 4.6(b)) is low, indicating a definite but small positive correlation. In both variables p is less than 0.01. The coefficients and p-values point to a statistically important relationship between *house price* and the two independent variables. In Figures 4.6(c) and 4.6(d) the opposite trend is found for the *250m contour distance* and *cemetery distance* variables. The two outliers are again present. It is clear from both graphs' negative-sloping trend line that house prices decrease as distance to the 250m contour and the cemetery increases. The Spearman r-value of -0.35 and p of <0.01 for the *250m contour distance* and the Spearman r-value of -0.01 and p of 0.82 for the *cemetery distance* variable imply an inverse relationship between these two explanatory variables and house prices. This trend is expected with the *250m contour distance* variable, but unanticipated in the *cemetery distance* variable, as one would expect houses to be more expensive further away from the cemetery. However, the low Spearman r-value and the gentleness of the slope of the trend line for the *cemetery distance* variable denote a negligible correlation. Table 4.4 summarizes the Spearman rank correlation coefficients (r-values) and the associated p-values for the ordinal data variables.

Table 4.4 Spearman statistics for ordinal data variables

Variable	Spearman r-value	Spearman p-value
Building size	r = 0.62	p=<0.01
Erf extent	r = 0.52	p=<0.01
Total bathroom fixtures	r = 0.51	p=<0.01
All commercial distance	r = 0.49	p=<0.01
Elevation	r = 0.47	p=<0.01
Total rooms	r = 0.42	p=<0.01
Central open space distance	r = 0.40	p=<0.01
Victoria Avenue distance	r = 0.38	p=<0.01
Roads distance	r = 0.37	p=<0.01
250m contour distance	r = -0.35	p=<0.01
Number bedrooms	r = 0.35	p=<0.01
Quality	r = 0.35	p=<0.01
Churches distance	r = 0.33	p=<0.01
Harbour distance	r = 0.33	p=<0.01
Condition	r = 0.31	p=<0.01
River distance	r = 0.31	p=<0.01
Slope	r = 0.30	p=<0.01
Fish factories distance	r = 0.29	p=<0.01
Tower distance	r = 0.28	p=<0.01
Solar radiation (1 day)	r = -0.27	p=<0.01
Solar radiation (5 days)	r = -0.26	p=<0.01
Household size	r = 0.26	p=<0.01
Public libraries distance	r = 0.25	p=<0.01
Security	r = -0.24	p=<0.01
Beach distance	r = 0.22	p=<0.01
Previous purchase price	r = 0.22	p=<0.01
Informal settlement distance	r = 0.20	p=<0.01
Post office distance	r = 0.20	p=<0.01
Shopping centres distance	r = 0.19	p=<0.01
Individual monthly income	r = 0.18	p=<0.01
Annual household income	r = 0.17	p=<0.01
Schools distance	r = 0.16	p=<0.01
Age 1 average	r = 0.15	p=<0.01
Age 2 average	r = 0.15	p=<0.01
Qualification	r = 0.10	p=0.05
Employed	r = -0.09	p=0.09
Age 3 average	r = 0.08	p=0.10
Traffic	r = 0.07	p=0.18
Population density	r = 0.06	p=0.24
Unemployed	r = -0.05	p=0.36
All open space distance	r = -0.01	p=0.88
Cemetery distance	r = -0.01	p=0.82

From the coefficients of the ordinal data variables and the p-values listed in Table 4.4, it is clear that most of these variables are statistically significant. The steps followed to select the most suitable variables for the regression analysis are discussed next.

4.6 FINAL SELECTION OF VARIABLES

As it was necessary to use the most appropriate nominal and ordinal data variables for the regression analysis, several issues needed to be addressed. Because too many variables tend to erode a regression model's predictive ability (Statsoft 2008), only 15 variables were used in the regression analysis. Care had to be taken that the number of observations were at least ten times more than the number of variables to ensure that the model's estimates are reliable (Statsoft 2008). As there were 377 observations in the data set, this reliability was not in danger of being compromised.

High correlations between variables also had to be addressed. When explanatory variables are highly correlated, the issue of multicollinearity is present (Dielman 2005) so that dependable estimates of variables' individual regression coefficients are compromised (ResearchConsultation.com 2007). Variables that are highly correlated essentially measure the same phenomenon and convey the same information (ResearchConsultation.com 2007), so that it is undesirable to have marked relationships existing between explanatory variables. To overcome this problem, Spearman rank correlations were calculated between the nominal data variables because these correlations are effective guides for detecting multicollinearity. These correlations do, however, have the limitation of only detecting correlations between two explanatory variables, whereas the relationships causing multicollinearity may be more complex than basic pairwise correlations. For example, if there are three explanatory variables in a model, one variable may not be correlated with the two other variables individually, but instead with a linear combination of these two variables (Dielman 2005). Following an inspection of the correlation coefficients, an arbitrary value of 0.6 was chosen to identify variables showing signs of multicollinearity. Dielman (2005) notes that some researchers recommend a value of 0.5, but this is not a hard and fast rule. None of the nominal data variables qualified for exclusion using 0.6 as a cut-off value.

Next, the most statistically significant variables had to be identified. Based on their F- and P-values derived from the ANOVA tests described in Section 4.4, the five most statistically significant nominal data variables were selected for the regression analysis and the rest were excluded. The Spearman correlation coefficients for these final nominal data variables are listed in Table 4.5. Among the excluded nominal data variables, the low F-value of 4.77 for the *aspect* variable was unexpected given that one estate agent stated that north-facing slopes are attractive to prospective homeowners, while another estate agent noted that any areas not facing north are unappealing. The low F-value of 3.28 for the excluded *corner plot* variable is quite likely due to the negative qualities of such plots outweighing their positive characteristics (see Section 2.4.8) in the minds of Hout Bay property buyers.

Table 4.5 Spearman rank correlation coefficients of the selected nominal data variables

Variable	Cul-de-sac	Pool	Race	Topography	View
Cul-de-sac	1.00	0.04	-0.11	0.07	-0.07
Pool	0.04	1.00	0.06	0.05	0.06
Race	-0.11	0.06	1.00	0.04	0.17
Topography	0.07	0.05	0.04	1.00	-0.30
View	-0.07	0.06	0.17	-0.30	1.00

Because the data set included twice as many ordinal data variables as nominal data variables only ten final ordinal data variables were selected. The selection process for ordinal data variables was similar to that used for the selection of nominal data variables. First, a Spearman rank correlation examination was performed on all the ordinal data variables. An arbitrary value of 0.6 was chosen to identify variables indicating possible multicollinearity. These Spearman correlations were similar to those described in Section 4.5, except that they were calculated between the independent variables, whereas previously the correlations were computed between the *house price* dependent variable and the other ordinal data variables. The analysis showed that there were several variables that shared coefficients of 0.6 or greater and therefore qualified for elimination. Second, the statistically significant ordinal data variables were identified by using the Spearman rank correlation coefficients and associated p-values (see Table 4.4) derived from the analysis performed in Section 4.5.

The twelve identified ordinal data variables in decreasing order of statistical importance, are: *building size*, *erf extent*, *total bathroom fixtures*, *all commercial distance*, *elevation, total rooms*, *central open space distance*, *Victoria Avenue distance*, *roads distance* and *250m contour distance, quality* and *number bedrooms*. These variables were mainly retained as Spearman correlations between them were in most cases less than 0.6, but two variables were judiciously excluded by examining the strength of their correlations and statistical significance in combination. After thorough consideration, the *Victoria Avenue distance* variable was omitted because its Spearman r-value was lower than that of the *all commercial distance* variable (0.38 versus 0.49) and these two variables shared a Spearman correlation of 0.71 indicating possible multicollinearity. Although the *quality*, *number bedrooms* and *250m contour distance* variables share equal statistical importance (see Table 4.4), the *250m contour distance* variable was excluded as the very high Spearman correlation of -0.85 indicated a strong likelihood of multicollinearity with the *elevation* variable. The *central open space distance* variable was retained despite high correlations with the *all commercial distance* and *elevation* variables (see Table 4.6) as the significance of *central open space distance* was noted in the literature (see Chapter 2) and one estate agent had noted its importance as an avoidance factor. A correlation of just above the cut-off value of 0.6 was observed between the *building size* and *total bathroom fixtures*. As these two variables were the first and third most significant ordinal data variables based on the Spearman r-values (see Table 4.4) both were retained. The ten ordinal data variables selected for the regression analysis and their Spearman coefficients are shown in Table 4.6.

Table 4.6 Spearman rank correlation coefficients of the selected ordinal data variables

Variable	All commer-cial distance	Building size	Central open space distance	Eleva-tion	Erf extent	Number bedrooms	Quality	Roads distance	Total bathroom fixtures	Total rooms
All commercial distance	1.00	0.37	0.66	0.78	0.43	0.19	0.13	0.58	0.30	0.30
Building size	0.37	1.00	0.30	0.40	0.52	0.51	0.33	0.28	0.63	0.53
Central open space distance	0.66	0.30	1.00	0.81	0.37	0.11	0.11	0.46	0.23	0.20
Elevation	0.78	0.40	0.81	1.00	0.44	0.20	0.14	0.56	0.30	0.28
Erf extent	0.43	0.52	0.37	0.44	1.00	0.38	0.09	0.21	0.37	0.44
Number bedrooms	0.19	0.51	0.11	0.20	0.38	1.00	0.19	0.08	0.50	0.38
Quality	0.13	0.33	0.11	0.14	0.09	0.19	1.00	0.07	0.39	0.20
Roads distance	0.58	0.28	0.46	0.56	0.21	0.08	0.07	1.00	0.21	0.25
Total bathroom fixtures	0.30	0.63	0.23	0.30	0.37	0.50	0.39	0.21	1.00	0.49
Total rooms	0.30	0.53	0.20	0.28	0.44	0.38	0.20	0.25	0.49	1.00

The omission of the relatively weak (Spearman r-value of 0.20) *informal settlement distance* variable was unexpected given that 11 of the 12 estate agents named the presence of the informal settlement as an important factor that depreciated property values. However, Saff (1998) found that proximity to informal settlements, widely perceived to cheapen house prices in their close vicinity, did not always have such an effect. When the house sales in Hout Bay were mapped, it was noticeable that there were many house sales immediately south of the informal settlement in the Penzance Estate subplace (see Figure 4.1), which suggests that the informal settlement has an effect on housing turnover but not price. The estate agents also confirmed that Penzance Estate is an area where a high turnover of property sales occurs.

Because six estate agents identified security as being important for new homeowners, the relative weakness (Spearman r-value of -0.24) of the omitted *security* variable is notable. Seeing that the data used to quantify this structural data variable dates from the 2000 general property valuation, security conditions might have changed by the time of the structured interviews. The weakness (Spearman r-values of -0.27 and -0.26) of the two *solar radiation* variables and the *aspect* variable, confirms that although sunny, north-facing areas may have been more popular among prospective homeowners, they did not have a marked effect on house prices.

4.7 CONCLUSION

Ten ordinal data variables along with five selected nominal data variables qualify for use in the regression analysis. The selection of these explanatory variables fulfils the third objective of this study, namely to conduct a statistical analysis to identify key variables for regression analysis. The next chapter describes the implementation of regression analysis and presents its results.

CHAPTER 5: REGRESSION ANALYSIS

5.1 INTRODUCTION

Regression analysis is a technique for modelling the relationships between two or more variables (Miles & Shevlin 2001). An ordinary least squares (OLS) multiple regression analysis approach is used here as it offers a relatively straightforward method to fulfil the fourth objective of the study, namely to assess the collective effect of multiple variables on property sales prices, as well as their relative contributions. This approach also makes it possible to evaluate whether other variables could have played a role in influencing house prices in Hout Bay. The discussion of the multiple regression analysis in this chapter is presented in three sections, namely a description of the analysis procedure and its results; an exposition of the residuals from the regression; and the mapping of the residuals.

5.2 REGRESSION ANALYSIS AND RESULTS

The five nominal data variables had to be transformed to ordinal data variables before they could be used in the regression model. The *view* variable was modified to contain three categories, namely average–poor, good and panoramic. It was belatedly discovered that coloureds were the predominant category for a negligibly small number of records (11), while whites were the principal category for the remainder (366 records). The *race* variable was therefore excluded from the analysis. Two outliers were deleted as they were expected to have an unacceptable distorting effect on the regression analysis. In the first stage of the analysis all 14 final variables were used to determine which are important predictors of house prices. However, the presence of multicollinearity remained problematic so that best-subsets regression was chosen to model the effects of the variables. This type of regression analysis eliminates multicollinearity and selects the subset of predictor variables which provides the best statistical results. The selected variables were *building size*, *erf extent*, *quality* and *roads distance*. The results of this analysis are summarized in Table 5.1 and some outcomes are discussed next.

Table 5.1 Results of the best-subsets regression analysis of house prices in Hout Bay

Variable	Beta (β) coefficient	Standard error of Beta	B coefficient	Standard error of B	t(370)	p-level
Intercept			-198 310.00	157 803.38	-1.26	0.21
Building size (m^2)	0.29	0.04	3 139.07	451.74	6.95	p=<0.01
Erf extent (m^2)	0.33	0.04	310.39	38.01	8.17	p=<0.01
Quality	0.22	0.04	416 330.64	71 956.91	5.79	p=<0.01
Roads distance (metres)	0.27	0.04	1 160.67	160.75	7.22	p=<0.01
R=0 .73; R^2= 0.54; F(4,370)=107.64; p<0.00; Standard error of estimate: 610478						

Based on the values in Table 5.1 the regression equation was derived and used to model house prices in Hout Bay. The equation is:

House price = α+B1X1+B2X2+B3X3+B4X4+*e*,
where α = Regression intercept
X1 = Building size
X2 = Erf extent
X3 = Quality
X4 = Roads distance
e = error term

The coefficient of determination (R^2) is a measure of how well the regression line fits the data, or stated differently, what proportion of the variance of the dependent variable is explained or accounted for by the predictor variables (Dielman 2005). The R^2 value of the regression analysis was 0.54, suggesting that the model performed adequately as 54% of the house price variation in Hout Bay is explained by the regression. The results of the model are satisfactory as only local factors in Hout Bay are incorporated in the model while macro-economic factors such as inflation, interest rates and economic growth that could influence property prices were excluded. There is a possibility that important price-influencing factors exist in Hout Bay that were not considered in the model. This is unlikely as the knowledgeable estate agents would have emphasized these during the structured interviews. Another possibility is that the variables may have been misspecified when they were constructed.

The beta coefficients (β) shown in Table 5.1 were obtained by standardizing all of the variables to a mean of zero and a standard deviation of one. The advantage of beta coefficients, in contrast to B coefficients (described below) which are not standardized, is that the magnitude of beta coefficients allows the comparison of each independent variable's relative contribution to the prediction of the dependent variable (Statsoft 2009). The beta coefficient of 0.33 for the *erf extent* variable indicates that this variable has the strongest relative effect compared to the other variables in determining house prices. The *building size* variable accounted for the next most important contribution based on its beta coefficient of 0.29, while the *roads distance* and the *quality* variables were respectively the third and fourth most important determinants of Hout Bay house prices, with beta coefficients of 0.27 and 0.22.

The B coefficients in Table 5.1 signify the contribution of each independent variable to the prediction of the dependent variable. The B values of the variables are not comparable because the units of measurement of the respective variables differ (Statsoft 2009). Therefore, for the *building size* variable, each square-metre increase in a house's size added close to R3140 to a house's selling price, while for the *erf extent* variable, each square-metre increase in erf size added roughly R310 to the value of a house in Hout Bay. The four categories for the *quality* variable predict large incremental increases of just over R400 000 for each improvement in the quality of a house. Similarly, for each increase by one metre in distance from busy roads, a house's value increases by approximately R1160. The findings support those of Lake et al. (1998) who noted that road noise depreciates house prices. The p-values shown in Table 5.1 test the hypothesis that the beta coefficients are equal to zero. If the p-value is smaller than 0.05, the hypothesis can be rejected that the particular beta coefficient is equal to zero. All the p-values shown in Table 5.1 are important at the 1% significance level. An analysis of the residuals derived from the regression analysis is essential and this is discussed next.

5.3 RESIDUAL ANALYSIS

Residuals are numerical measures of the correspondence between observed and predicted values (Pelosi & Sandifer 2003). Multiple regression analysis is based on a number of assumptions and residual analysis is a useful tool to determine if these assumptions have been violated (Dielman 2005). One assumption is that the relationship between the dependent variable and the independent variables is linear (Dielman 2005). According to the assumptions of linear regression, for all values of the independent variable the residuals of the model should be randomly distributed with a mean value of zero (Pelosi & Sandifer 2003). Figure 5.1 plots the residuals and the predicted values of the independent variable and shows that the residuals are randomly distributed around the mean value of zero.

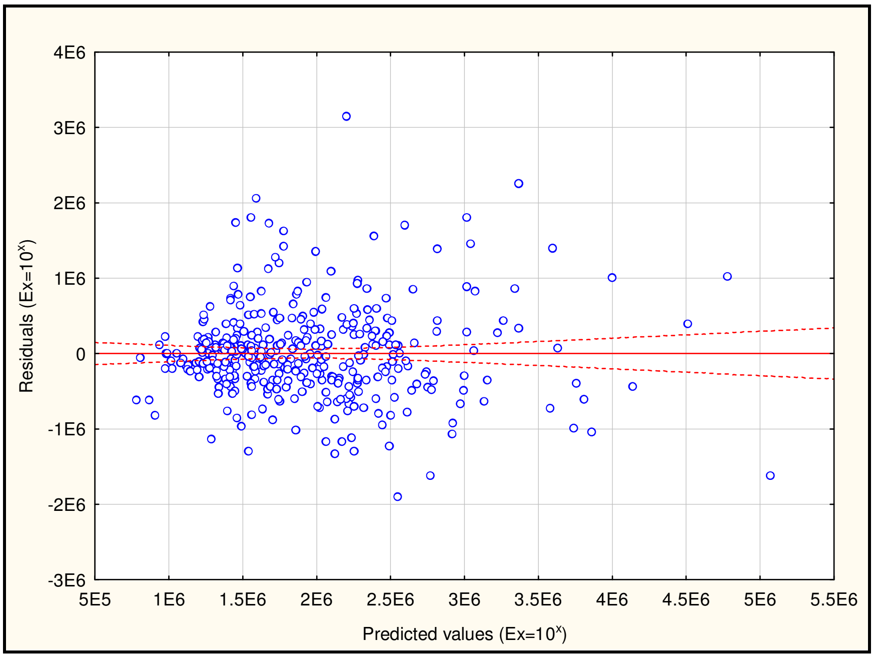


Figure 5.1 Predicted versus residual scores of house prices

The plot in Figure 5.1 has no discernible pattern, indicating that a linear relationship exists in the model. In addition to verifying that the residuals have a mean of zero and are randomly distributed around this value, the shape of the distribution must be checked to see if it is normal. This is because the normal distribution of residuals is another assumption of multiple regression analysis (Dielman 2005).

The histogram of the residuals (Figure 5.2(a)) is an informal method of assessing whether the residuals are derived from a normal distribution (Pelosi & Sandifer 2003). Here the histogram is approximately bell-shaped with skewness to the right. The normal probability plot of the residuals shown in Figure 5.2(b) also indicates a possible deviation from normality due to the long tail which indicates a number of high-priced houses. The Shapiro-Wilkes test's p value is less than 0.01, therefore the assumption of normally distributed residuals can be rejected in this case. A normal probability plot of residuals is a more formal method to assess whether the data are normally distributed (Pelosi & Sandifer 2003).

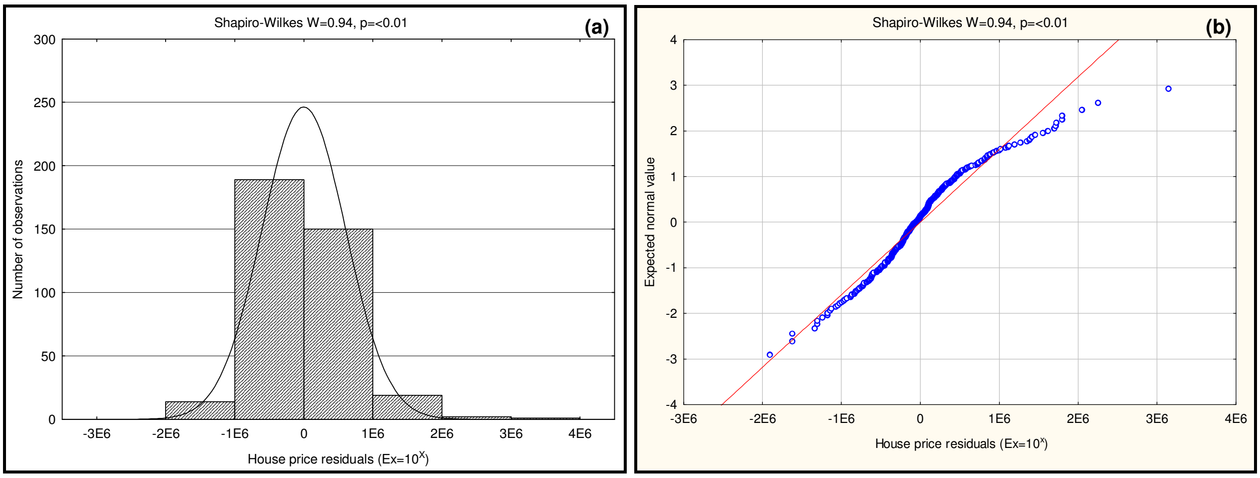


Figure 5.2 Histogram (a) and normal probability plot (b) of house price residuals

When mapped, residuals can be an excellent visualization tool for interpreting regression analysis results spatially. The next subsection describes this mapping procedure.

5.4 RESIDUAL MAPPING

The residuals for each data record were derived by means of the Statistica statistical program. Table 5.2 summarizes the results. Each erf's 21-digit code was retained in this Statistica file, which enabled the standard residuals to be mapped as shown in Figure 5.3.

Table 5.2 Summary of predicted and residual values of house prices

Summary statistic	Observed value	Predicted value	Residual	Standard predicted value	Standard residual	Standard error predicted value	Deleted residual
Minimum	80000	782452	-1901689	-1.8	-3.1	35531	-2041057
Maximum	5800000	5069758	3146642	4.8	5.2	277356	3164144
Mean	1935222	1935222	0.0	0.0	-0.0	64225	-384
Median	1750000	1805686	-44991	-0.2	-0.1	55514	-45609

The negative residuals are evidence that the regression model has overestimated the sales price of a house, whereas the positive residuals indicate an underestimation by the model. A map depicting the residuals represents the spatial variation of deviations from the regression line. The residuals are that portion of the variation of a house's sales price not explained by the regression model. The spatial patterns discernible in a residual map give an indication of additional variables that could possibly contribute to a better understanding of the factors responsible for the spatial variation of house prices in an urban area (Zietsman 1975).

It was assumed that the regression model predicts house values quite accurately if the standard residuals lie between arbitrarily chosen values of -0.5 and 0.5, which represents a narrow band of over- and undervalued houses. Other categories used were: highly overestimated (-3.12 to -1.01), moderately overestimated (-1 to -0.5), moderately underestimated (0.51 to 1) and highly underestimated (1.01 to 5.15). In Figure 5.3 the light green dots represent the accurately modelled band which accounts for more than half of the valued houses (see Table 5.3), while the bright red and darks blue dots indicate highly overestimated and highly underestimated house prices respectively. Table 5.3 also shows that the model tends toward overvaluation as the total overestimated and underestimated observations comprised about 27% and 22% respectively.

Table 5.3 Observations of over- and underestimations per standard residual category

Residual categories	Band	Observations	%
Highly overestimated	-3.12 – -1.01	43	11.5
Moderately overestimated	-1 – -0.5	59	15.7
Accurately modelled	-0.5 – 0.5	190	50.7
Moderately underestimated	0.51 – 1	40	10.7
Highly underestimated	1.01 – 5.15	43	11.5

The standard residuals were interpolated using the inverse distance weighted (IDW) tool in ArcGIS 9.2's Spatial Analyst extension to create a standard residual raster surface (see Figure 5.4). As no house sales occurred in the large central open space area, the informal settlement and the central commercial area, these areas were excluded from the interpolation process. An output cell size of 20m, a power of two and variable-search radius with the default settings of 12 observation points and no maximum distance were used as parameters in the IDW process.

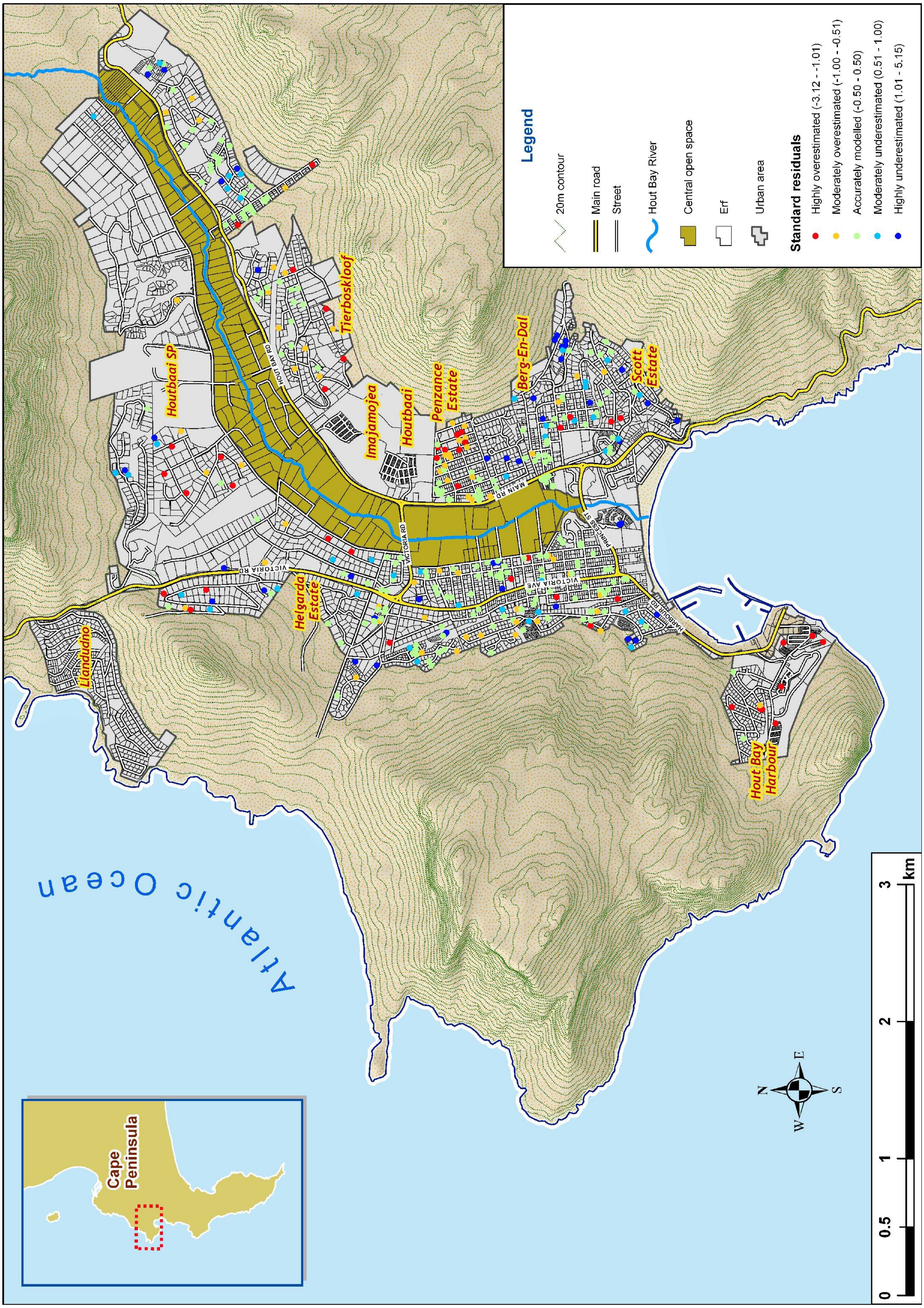


Figure 5.3 Spatial distribution of standard regression residuals of house prices in Hout Bay

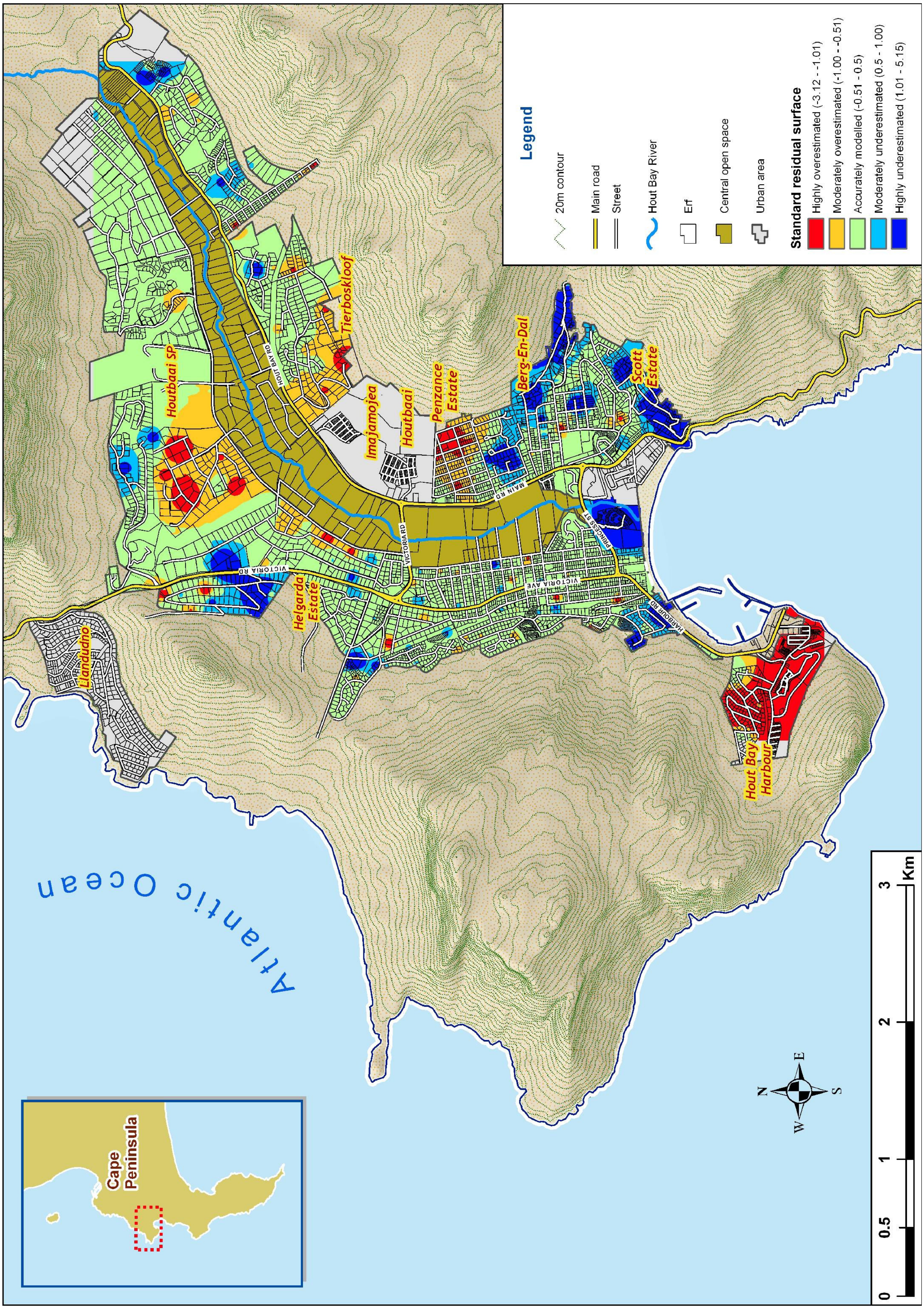


Figure 5.4 Interpolated standard regression residuals of house prices in Hout Bay

In Figure 5.4 there are noticeable areas of high overvaluation (indicated by bright red shading) in the Hout Bay Harbour, Houtbaai SP and Penzance Estate subplaces, where actual house prices are lower than those predicted by the regression model. In these areas some missing variables inhibit the explanatory power of the regression model. It also appears that specific variables may be unique to certain areas. Should they be included in the regression analysis, the predictability of the house prices may be more in line with the actual house values. The task of identifying potential price-reducing factors is difficult. In the zone of high overvaluation in Penzance Estate the proximity to the informal settlement could be a value-reducing factor. Although *informal settlement distance* was not a significant variable in Hout Bay in general, it appears to have a localised effect on this area, with an effective threshold of influence of about 1 km. Likewise, the omitted *fish factories distance* and *harbour distance* variables may have a localised effect on the Hout Bay Harbour area, but not on Hout Bay house prices in general. Higher predicted than actual prices in the Hout Bay Harbour area may also be explained by the fact that this area's historically low socio-economic status (see Section 1.4) has not been adequately captured in the model.

There are many small areas shown in Figure 5.4 where the regression model has underestimated house prices, i.e. where actual house prices are higher than those predicted by the model. These areas are indicated by dark blue shading and are mostly located in the Scott Estate and Berg-en-Dal subplaces, while there is a large area in Houtbaai SP close to the beach and another in the Helgarda subplace. In these underestimated zones the regression model has not captured the effects of key variables that might explain the larger than normal deviations between predicted and actual house prices. As these zones of underestimation are mostly close to the sea, the value-adding effects of proximity to the beach and sea views were most probably inadequately explained by the regression model. Had the *beach distance* and *view* variables been included in the regression model they could have added explanatory power to the differences in the predicted and actual house prices in these underestimated areas. However, both the *beach distance* and *view* variables are weak predictors in Hout Bay's property market in general (see Tables 4.3 and 4.4).

5.5 CONCLUSION

This chapter concludes the statistical analyses done in the research. It described the regression analysis and its results led to the fulfilment of the fourth objective of the study, namely to assess the collective effect of multiple variables on property sales prices, as well as their relative contributions. In the next chapter the thesis is concluded by underscoring its salient features and findings, and by recommending some avenues of future research.

CHAPTER 6: SYNTHESIS AND CONCLUSIONS

6.1 INTRODUCTION

The purpose of this chapter is to conclude the thesis by highlighting and reviewing its salient features. In the first subsection the main sections of the thesis are summarized and the study's objectives evaluated to see if they have been achieved. This is followed by a reaffirmation of the major findings. Finally, recommendations for further research are proposed.

6.2 SUMMARY

In the first chapter, the geographical foundation of the research was laid by showing that residential properties can never be disconnected from the physical environment in which they are embedded. These environmental influences can be at national or local level and may lead to various manifestations in the spatial arrangement of the housing stock. The competition of housing with other land uses in the urban housing market is dependent on factors of supply and demand, which in turn cause certain processes of change in this market. One manifestation of these processes, the spatial variation of house prices, was the focus of this study. With the various attributes of housing and the human needs they satisfy as foundation, the research problem addressed in this study was how to empirically ascertain the primary determinants of house prices in Hout Bay and the relative importance of each factor. The research attempted to answer several related questions concerning this research problem: What are the housing attributes prospective homeowners value highly, or deliberately avoid when looking to buy a house Hout Bay and, what factors, previously regarded as being significant, are actually negligible?

Several objectives were pursued to unravel the research problem. The first objective was to select suitable house price-influencing factors in Hout Bay. The literature review (presented in Chapter 2), and the collection of primary and secondary data (described in Chapter 3) assisted this selection process. The literature review provided the necessary theoretical grounding to the study. The academic literature generally differentiated between structural and locational attributes of properties as determinants of house prices. Studies which focus on structural attributes were scarce so that little theoretical guidance was forthcoming on this topic. The many studies which analyse locational attributes provided useful insights into the choice and use of variables used for this study. Three studies in the USA found race to be an important determinant of house prices, with the presence of black minorities being a specific depreciating factor. Two other studies in North America revealed that proximity to shopping centres has a positive influence, despite the fact that some homeowners in the vicinity of such centres often disapproved of their presence. Proximity to good schools is seen to affect house prices positively, while evidence of the influence of churches was inconclusive as the two available studies produced diametrically opposite results. With respect to the influence of water aspects on house prices, one study predictably found that relatively higher water pollution levels had a depreciating effect on house prices, while riparian corridors with their associated benefits were found to increase house prices in the surrounding area.

Open space, especially that which is designated for future preservation, was reported to have a positive effect on house prices. A South African case study did not support this view. The literature on the influence of roads and suburban layout on house prices yielded conflicting viewpoints. Road noise predictably decreases house prices, while the accessibility benefit of highways increases property values. Cul-de-sacs were said to have a significant effect on house prices, while automobile-orientated suburbs with curvilinear and cul-de-sac street patterns had no noticeable effect. Studies on the effect of views were unanimous: all the research reported a correlation between either scenic views and increased property values or between unattractive views and decreased property values. These studies were important because they provided empirical evidence confirming whether or not the locational variables used in this thesis were amenities or disamenities or if there was uncertainty about their status.

Because the determinants of property value are potentially numerous, it was crucial to reduce their number by identifying only specific house price-influencing factors applicable to Hout Bay. Consequently, primary data on these factors were collected by conducting structured interviews with estate agents in Hout Bay to ascertain their expert opinions about the locational attributes potential property owners found either desirable or which they avoided. This interview procedure was described in Chapter 3. Views, security and privacy, and proximity factors were singled out by estate agents as those which prospective homeowners found most attractive, while informal settlement-related factors and Hout Bay's busy roads were features potential buyers most avoided. These pointers, combined with insights gleaned from the literature, as well as secondary data of the registered property transactions and 2000 general valuation data enabled the identification of specific structural as well as fixed and relative locational variables.

The second objective, namely to quantify the variables, was met by using GIS and other techniques to generate spatial variables, i.e. the distance, slope, aspect, and solar radiation variables. The third objective, to statistically analyse the *house price* and independent variables, was accomplished using various methods including ANOVA tests, 2D scatterplots and Spearman rank order correlations. These property price analyses are described in Chapter 4. This facilitated the derivation of key explanatory variables which constituted the third objective. The fourth objective, namely to deduce the collective and individual influences of the key variables for use in the regression analysis was successfully done as described in Chapter 5. The results of this analysis are briefly revisited in the next section.

6.3 SALIENT FINDINGS

Answering the questions aimed at resolving the research problem produced some noteworthy results. The factors prospective homeowners in Hout Bay valued highly or avoided were determined by regression analysis of the statistically significant variables. To avoid the problem of multicollinearity between variables, best-subsets regression was used. The regression model performed adequately and the selected variables collectively accounted for 54% of the variation in Hout Bay's house prices. The *erf extent* independent variable is the best predictor of the *house price* dependent variable. The *building size* variable was the second-best predictor, although there was only a slight difference between its explanatory power and that of *erf extent*. The *roads distance* and the *quality* variables were respectively the third and fourth most important determinants of Hout Bay house prices.

These findings support some of the views expressed by the estate agents, namely that spacious grounds and privacy (presumably best expressed by having a big erf) are desirable to property buyers, while noisy and busy roads are undesirable. It is no surprise that the *roads distance* variable is statistically significant given that extant research confirms this and more than half of the estate agents mentioned that street-related factors are unattractive to buyers. The *building size* and *quality* variables were not identified by the estate agents as the structured interviews focussed only on locational factors. The structured interviews emphasized the importance of interrogating estate agents' knowledge of the local property market to ascertain the determinants of house prices in an area, while the statistical analyses proved useful in identifying both the important and insignificant variables. The regression analysis was successful in determining the individual and collective effects of variables on house prices.

Among the variables initially thought to be reliable indicators of house prices, several were shown by the statistical analyses to be poor predictors. The weakness of the *aspect* and the two *solar radiation* variables was unexpected as sunny, north-facing areas were identified by the estate agents as being popular among prospective homeowners. It was similarly unanticipated that the *informal settlement distance* variable is such a weak predictor, taking into account the estate agents' agreement that proximity to the local informal settlement has a negative influence on house prices. However, this study's findings correspond to Saff's (1998) that the presence of an informal settlement in a South African context does not necessarily have a deleterious effect on house prices in its vicinity. The weakness of the *security* variable was surprising in the light of South Africa's high crime rate, although the somewhat outdated nature of the valuation data could have accounted for this.

6.4 RECOMMENDATIONS

Several recommendations for further improving the research are in order. First, similar research could be undertaken on a larger, perhaps citywide scale, to obtain results applicable to large urban areas. Research on a larger scale may produce results having greater value for urban and real estate planning purposes. Such research could employ census data to fuller potential due to the greater variation of socio-economic data on an urban scale. The urban study area could be fruitfully subdivided into various submarkets based on, for example, voting wards or other spatial subdivisions to examine the variable influences of price-determining factors within a city. An attractive variation on this theme is explicit attention to the influence on urban house prices of specific features such as informal settlements, power lines, gated communities or golf estates, and landfill sites. Despite the caveats related to location specificity of property-valuation studies mentioned in Chapter 1, a comparative study between this one and another to be done in a coastal town or suburb of similar size can be initiated. The specific variables that should be compared are *erf extent*, *building size, quality* and *roads distance* as it will be meaningful to verify whether these variables play such defining roles in determining house prices in other places.

A further recommendation is that the variables be quantified more accurately. The use of outdated municipal valuation data could be eliminated by site-inspection of properties. For example, the view measures could be created with more sophistication by employing the methods described by Wolverton (1997) and Benson et al. (1998) or, alternatively, by using advanced viewshed analysis of data obtained from highly accurate digital terrain models (DTMs) created by means of photogrammetric techniques or from LIDAR data. The limitations of viewshed analysis (alluded to in Chapter 2) can be averted by using these methods. The accuracy of existing data on property security measures could also be enhanced by personally visiting sampled properties to establish what security measures are actually in place (e.g. high walls, palisade fences, electrified fences, security service notices, and spikes on walls).

It is now possible with the latest release (9.3.1) of the GIS program ArcGIS to perform ordinary least squares (OLS) regression analysis directly without first exporting the data to a statistical package such as Statistica and then performing the analyses in this external package. Another useful new tool available in ArcGIS 9.3.1 is geographically weighted regression (GWR), one of a number of spatial regression techniques increasingly used in geography and other fields (Environmental Systems Research Institute 2009). Future studies should take cognizance of these new tools and techniques and incorporate them wherever possible in analyses of house prices in an urban context.

At the beginning of this study it was declared that the unavailability of a convenient method to know which structural and locational factors affect house prices, and what these factors' relative importance is, pose difficulties in determining house prices. By empirically establishing which factors determine house prices in Hout Bay and what each one's relative importance is, this study has enlarged the body of knowledge pertaining to property-valuation research.

REFERENCES

Adams JS 1984. The meaning of housing in America. *Annals of the Association of American Geographers* 74: 515-526.

Anderson ST & West SE 2006. Open space, residential property values and spatial context. *Regional Science and Urban Economics* 36: 773-789.

Asabere PK 1990. The value of a neighborhood street with reference to the cul-de-sac. *Journal of Real Estate Finance and Economics* 3: 185-193.

Bae CC, Jun M & Park H 2003. The impact of Seoul's subway Line 5 on residential property values. *Transport Policy* 10: 85-89.

Beamish JO, Goss RC & Immel J 2001. Lifestyle influences on housing preferences. *Housing and Society* 28, 1-2: 1-28.

Benson ED, Hansen LJ, Schwartz AL & Smersh GT 1998. Pricing residential amenities: The value of a view. *Journal of Real Estate Finance and Economics* 16: 55-73.

Betts RM & Ely SJ 2001. *Basic real estate appraisal*. 5th ed. Cincinnati: South-Western.

Bond MT, Seiler VL & Seiler MJ 2002. Residential real estate prices: A room with a view. *Journal of Real Estate Research* 23: 129-137.

Bond S and Wang K 2005. The impact of cell phone towers on house prices in residential neighborhoods. *The Appraisal Journal* Summer: 256-277.

Bourassa SC, Hoesli M & Sun J 2004. What's in a view? *Environment and Planning A* 36: 1427-1450.

Bourne LS 1981. *The geography of housing*. London: Edward Arnold.

Boxall PC, Chan WC & McMillan ML 2005. The impact of oil and natural gas facilities on rural residential property values: A spatial hedonic analysis. *Resource and Energy Economics* 27: 248-269.

Brasington DM & Hite D 2005. Demand for environmental quality: A spatial hedonic analysis. *Regional Science and Urban Economics* 35: 57-82.

Can A 1998. GIS and spatial analysis of housing and mortgage markets. *Journal of Housing Research* 9: 61-86.

Carroll TM, Clauretie TM & Jensen J 1996. Living next to godliness: Residential property values and churches. *Journal of Real Estate Finance and Economics* 12: 319-330.

Carter 1995. *The study of urban geography*. 4th ed. London: Arnold.

Cervero R & Duncan M 2004. Neighbourhood composition and residential land values: Does exclusion raise or lower values? *Urban Studies* 41: 299-315.

Chin HC & Foong KW 2006. Influence of school accessibility on housing values. *Journal of Urban Planning and Development* 132: 120-129.

Clark DE & Herrin WE 2000. The impact of public school attributes on home sale prices in California. *Growth and Change* 31: 385-407.

Clark W & Onaka J 1983. Life cycle and housing adjustment as explanations of residential mobility. *Urban Studies* 20: 47-57.

Colby BG & Wishart S 2002. Quantifying the influence of desert riparian areas on residential property values. *The Appraisal Journal* Summer: 304-308.

Darling A 1973. Measuring benefits generated by urban water parks. *Land Economics* 49: 22-34.

Department of Water Affairs and Forestry 2003. *State-of-rivers report: Diep, Hout Bay, Lourens and Palmiet River systems, 2003*. Pretoria: Department of Water Affairs and Forestry.

Des Rosiers F, Lagana A, Thériault M & Beaudoin M 1996. Shopping centres and house values: An empirical investigation. *Journal of Property Valuation and Investment* 14: 41-62.

Dielman TE 2005. *Applied regression analysis: A second course in business and economic statistics*. 4th ed. Belmont: Thomson Learning.

Dixon JA, Reicher S & Foster DH 1997. Ideology, geography, racial exclusion: The squatter camp as "blot on the landscape". *Text* 17: 317-348.

Do AQ, Wilbur RW & Short JL 1994. An empirical examination of the externalities of neighborhood churches on housing values. *Journal of Real Estate Finance and Economics* 9: 127-136.

Dodgson JS & Topham N 1990. Valuing residential properties with the hedonic method: A comparison with the results of professional valuations. *Housing Studies* 5: 209-213.

Ebdon D 1985. *Statistics in geography*. 2nd ed. Oxford: Blackwell.

Environmental Systems Research Institute 2008. *ArcGIS 9.2 Desktop help: Points solar radiation* [online]. Redlands: ESRI. Available from http: // webhelp. esri.com/arcgisdesktop/ 9.2/index.cfm? TopicName=Points_ Solar_Radiation [Accessed 9 July 2009].

Environmental Systems Research Institute 2009. *ArcGIS 9.3 Desktop help: How GWR regression works* [online]. Redlands: ESRI. Available from http://webhelp.esri.com/ arcgisdesktop/9.3/index.cfm?TopicName=How%20GWR%20Regression%20works [Accessed 26 May 2009].

Figueroa RA 1999. Modelling the value of location in Regina using GIS and spatial autocorrelation statistics. *Assessment Journal* 6: 29-37.

Fraser R & Spencer G 1998. The value of an ocean view: An example of a hedonic property amenity valuation. *Australian Geographical Studies* 36:94-98.

Gallimore P, Fletcher M & Carter M 1996. Modelling the influence of location on value. *Journal of Property Valuation and Investment* 14: 6-19.

Geoghegan J 2002. The value of open spaces in residential land use. *Land Use Policy* 19: 91-98.

Hamnett C 2009. Spatially displaced demand and the changing geography of house prices in London, 1996-2006. *Housing Studies* 24: 301-320.

Harris DR 1999. "Property values drop when blacks move in, because…": Racial and socio-economic determinants of neighborhood desirability. *American Sociological Review* 64: 461-479.

Humavindu MN & Stage J 2003. Hedonic pricing in Windhoek townships. *Environment and Development Economics* 8: 391-404.

Irwin EG 2002. The effects of open space on residential property values. *Land Economics* 78: 465-480.

Jackson E, Kupke V & Rossini P 2006. The relationship between socio-economic indicators and residential property values in Darwin. Paper delivered at the 13th annual PRRES conference, Fremantle, Australia.

Johnson LC 1994. What future for feminist geography? *Gender, Place and Culture* 1: 103-113.

Kaplan D, Wheeler JO & Holloway S 2008. *Urban geography*. 2nd ed. New York: John Wiley & Sons.

Kaufman DA & Cloutier NR 2006. The impact of small brownfields and greenspaces on residential property values. *Journal of Real Estate Finance and Economics* 33: 19-30.

Kiel KA & Williams M 2007. The impact of Superfund sites on local property values: Are all sites the same? *Journal of Urban Economics* 61: 170-192.

Lake IR, Lovett AA, Bateman IJ & Day B 2000. Using GIS and large-scale digital data to implement hedonic pricing studies. *International Journal of Geographical Information Science* 14: 521-541.

Lake IR, Lovett AA, Bateman IJ & Langford IH 1998. Modelling environmental influences on property prices in an urban environment. *Computers, Environment and Urban Systems* 22: 121-136.

Leggett CG & Bockstael NE 2000. Evidence of the effect of water quality on residential land values. *Journal of Environmental Economics and Management* 39: 121-144.

Lynch AK & Rasmussen DW 2004. Proximity, neighbourhood and the efficacy of exclusion. *Urban Studies* 41: 285-298.

Mackmin D 1994. *The valuation and sale of residential property*. 2nd ed. London: Routledge.

Matthews JW & Turnbull GK 2007. Neighborhood street layout and property value: The interaction of accessibility and land use mix. *Journal of Real Estate Finance and Economics* 35: 111-141.

McMillen DP 2008. Changes in the distribution of house prices over time: Structural characteristics, neighborhood, or coefficients? *Journal of Urban Economics* 64: 573-589.

Miles JM & Shevlin M 2001. *Applying regression and correlation: A guide for students and researchers*. London: Sage Publications.

Mouton J 2003. *How to succeed in your Master's and Doctoral studies: A South African guide and resource book*. Pretoria: Van Schaik.

Nicholls S & Crompton JL 2007. The impact of a golf course on residential property values. *Journal of Sport Management* 2007: 555-570.

Oelofse C & Dodson B 1997. Community, place and transformation: A perceptual analysis of residents' responses to an informal settlement in Hout Bay, South Africa. *Geoforum* 98: 91-101.

Orford S 2002. Valuing locational externalities: A GIS and multilevel modelling approach. *Environment and Planning B: Planning and Design* 29: 105-127.

Pacione M 2005. *Urban geography: A global perspective*. 2nd ed. London: Routledge.

Paterson RW & Boyle KJ 2002. Out of sight, out of mind? Using GIS to incorporate visibility in hedonic property value models. *Land Economics* 78: 417-425.

Pelosi MK & Sandifer TM 2003. *Elementary statistics: From discovery to decision*. Hoboken: John Wiley & Sons.

Pompe JJ & Rinehart JR 1995. Beach quality and the enhancement of recreational property values. *Journal of Leisure Research* 27: 143-154.

ResearchConsultation.com 2007. *What is multicollinearity in multiple regression? Statistics help for dissertation students and researchers* [online]. S.l.: ResearchConsultation.com. Available from http://www.researchconsultation.com/what-is-multicollinearity-multiple-regression.asp [Accessed 8 July 2009].

Rinehart JR & Pompe JJ 1999. Estimating the effect of a view on undeveloped property values. *The Appraisal Journal* January: 57-61.

Rodriguez M & Sirmans CF 1994. Quantifying the value of a view in single-family housing markets. *The Appraisal Journal* 62: 600-603.

Saff G 1998. The effects of informal settlement on suburban property values in Cape Town, South Africa. *Professional Geographer* 50: 449-464.

Sarip AG 2005. Integrating artificial neural networks and GIS for single-property valuation. Paper delivered at the 11th annual PRRES conference, Melbourne, Australia.

Scheepers MJ & Bloom JZ 2005. Resident perceptions of mixed-use development in Hout Bay, Cape Town. *South African Journal of Economic and Management Sciences* 8: 1-17.

Seiler MJ, Bond MT & Seiler VL 2001. The impact of world class Great Lakes views on residential property values. *The Appraisal Journal* 69: 287-296.

Sirmans GS, MacDonald L, Macpherson DA & Zietz EN 2006. The value of housing characteristics: A meta analysis. *Journal of Real Estate Finance and Economics* 3: 215-240.

Sirmans GS & Macpherson DA 2003. The value of housing characteristics. Research project. s.l.: National Association of Realtors.

Sirmans GS, Macpherson DA & Ziets EN 1995. The composition of hedonic pricing models. *Journal of Real Estate Literature* 13: 3-43.

Smith S 2000. Housing studies. In Johnston RJ, Gregory D, Pratt G, Watts M (eds) *The dictionary of human geography*. 4th ed., 346-349. London: Credo Reference.

Song Y & Sohn J 2006. Valuing spatial accessibility to retailing: A case study of the single family housing market in Hillsboro, Oregon. *Journal of Retailing and Consumer Services* 14: 279-288.

South African National Parks 2004. The economic contribution of Table Mountain National Park. Cape Town: University of Cape Town (Graduate School of Business).

Statsoft 2008. *Multiple regression* [online]. Tulsa: Statsoft Inc. Available from http://www.statsoft.com/textbook/stmulreg.html [Accessed 27 July 2009].

Statsoft 2009. *Statistics glossary* [online]. Tulsa: Statsoft Inc. Available from http://www.statsoft.com/textbook/glosfra.HTML [Accessed 9 June 2009].

Swanepoel PA 1996. Ruimtelik-vergelykende ontleding van eiendomswaardasie en -verkoopswaardes in Stellenbosch. Honours research report. Stellenbosch: University of Stellenbosch, Department of Geography and Environmental Studies.

Thériault M, Des Rosiers F, Villeneuve P & Kestens Y 2003. Modelling interactions of location with specific value of housing attributes. *Property Management* 21: 25-62.

Tovar-Pescador J, Pozo-Vazquez D, Ruiz-Ariaz JA, Batlles J, Lopez G & Bosch JL 2006. On the use of the digital elevation model to estimate the solar radiation in areas of complex topography. *Meteorological Applications* 13: 279-287.

Turner MA & Ross SL 2005. How racial discrimination affects the search for housing. In De Souza Briggs X (ed) *The geography of opportunity: Race and housing choice in metropolitan America*, 81-10. Washington: Brookings Institution Press.

Van Wyk J 2005. Open-space systems in urban land-use planning: Invaluable assets in conserving the environment and enhancing the quality of life. *Tydskrif vir die Suid-Afrikaanse Reg* 2: 256-273.

Van Zyl H & Leiman A 2002. Hedonic approaches to estimating the impacts of open spaces: A case study in the Cape. *South African Journal of Economic and Management Sciences* 5: 379-394.

Visser G 2003. Visible, yet unknown: Reflections on second-home development in South Africa. *Urban Forum* 14: 379-406.

Wheeler B, Mitchell R, Shaw M & Dorling D 2007. The geography of housing in Britain. *Geography Review* 20, 3: 2-6.

Wolverton M 1997. Empirical study of the relationship between residential lot price, size and view. *Journal of Property Valuation and Investment* 15: 48-57.

Wyatt P & Ralphs M 2003. *GIS in land and property management*. London: Spon Press.

Yu SM, Han SS & Chai CH 2007. Modeling the value of view in high-rise apartments: A 3D GIS approach. *Environment and Planning B: Planning and Design* 34: 139-153.

Zietsman HL 1975. ‘n Ontleding van die grondwaardepatroon van Oudtshoorn. *South African Geographer* 5,1: 47-55.

Zietz J, Zietz EN & Sirmans GS 2008. Determinants of house prices: A quantile regression approach. *Journal of Real Estate Finance and Economics* 37: 317-333.

APPENDIX: ESTATE AGENT STRUCTURED INTERVIEW FORM

UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot · your knowledge partner

Department of Geography and Environmental Studies
Private Bag XI
Matieland 7602
Tel: +27 21 808 3218 • Student's tel: 072 382 3958
Student's email address: 11126833@sun.ac.za

1. WHAT ARE THE PRIMARY FACTORS RELATING TO A RESIDENTIAL PROPERTY'S LOCATION THAT PROSPECTIVE BUYERS IN GENERAL WANT WHEN LOOKING FOR RESIDENTIAL PROPERTY IN HOUT BAY? (for example, mountain views or proximity to shops)

1. ______ 2. ______
3. ______ 4. ______
5. ______ 6. ______
7. ______ 8. ______

2. WHAT ARE THE PRIMARY FACTORS RELATING TO A RESIDENTIAL PROPERTY'S LOCATION THAT PROSPECTIVE BUYERS IN GENERAL AVOID WHEN LOOKING FOR RESIDENTIAL PROPERTY IN HOUT BAY? (for example, noise or smells)

1. ______ 2. ______
3. ______ 4. ______
4. ______ 6. ______
7. ______ 8. ______

3. IN WHICH AREAS (OR STREETS) OF HOUT BAY ARE THERE A HIGH TURNOVER OF RESIDENTIAL PROPERTY SALES?

○ Don't know ○ No areas 1. ______
2. ______ 3. ______
4. ______ 5. ______

4. IF ANY AREAS (OR STREETS) WERE IDENTIFIED IN QUESTION 3, WHAT ARE POSSIBLE REASONS FOR THIS HIGH TURNOVER?

○ Don't know 1. ______
2. ______ 3. ______
4. ______ 5. ______

5. IN WHICH AREAS (OR STREETS) OF HOUT BAY ARE RESIDENTIAL PROPERTIES DIFFICULT TO SELL?

○ Don't know ○ No areas 1. ______
2. ______ 3. ______
4. ______ 5. ______

6. IF ANY AREAS (OR STREETS) WERE IDENTIFIED IN QUESTION 5, WHAT ARE POSSIBLE REASONS FOR THIS DIFFICULTY ?

○ Don't know 1. ______
2. ______ 3. ______
4. ______ 5. ______