

## Case Study

# Property Market Analysis



**Application of Multi-Factor Regression Analysis to identify the statistically significant factors driving property prices in Geelong Suburbs, 2018.**

## Question / Hypothesis

- A. What variables should I consider when searching for properties to develop?
- B. Which factors are driving the sold price?
- C. How much should I expect to pay?
- D. Which Suburb has the most development potential?

## Data Set

Source: [www.realestate.com.au](http://www.realestate.com.au)

Suburbs: Belmont, Hamlyn Heights, Norlane.

127 Properties sold in Quarter2 of 2018.

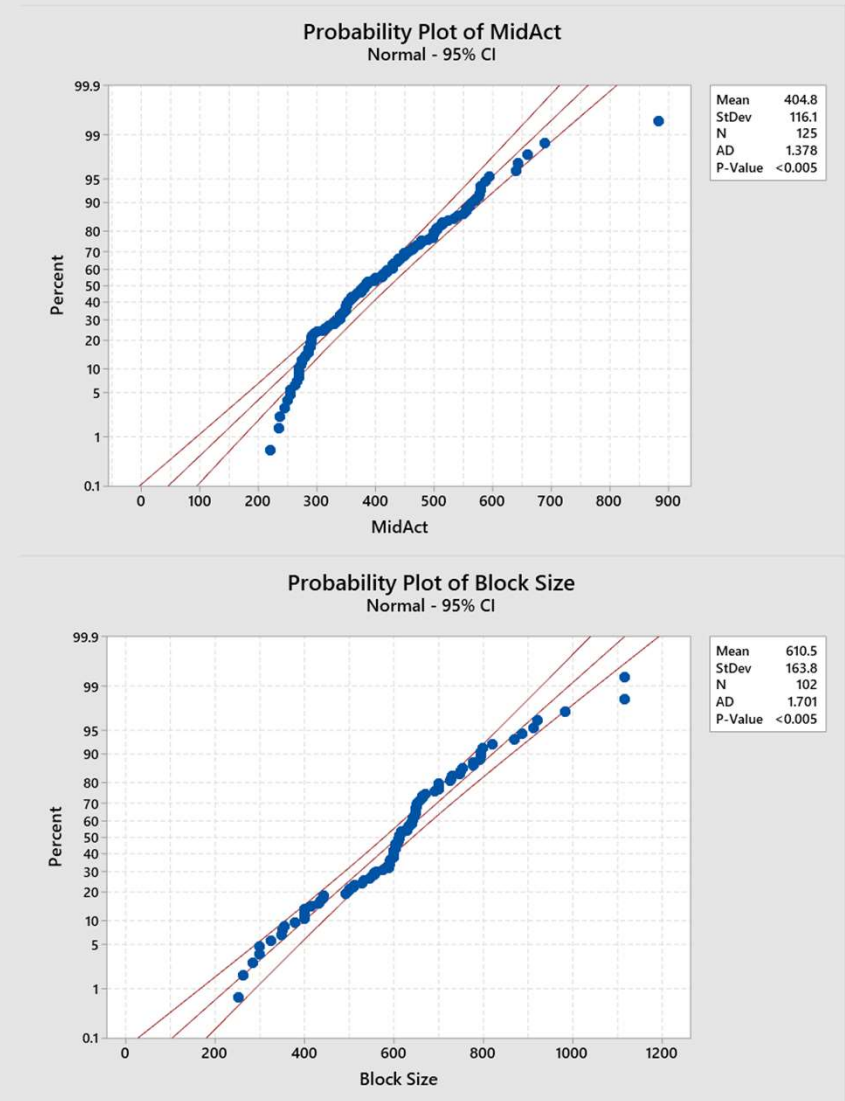
Factors/Variables:

Sold Price (\$000), Number of Bedrooms, Number of Bathrooms, Existing or New Build, Block Size (sq.m), Garage Size.

## Methodology

1. Build simple database with greater than 100 property details.
2. Check data integrity and distribution.
3. Choose the Statistical Tool.
4. Analyse the Data.
5. Record observations.
6. Interpret Statistical Model Equation.
7. Validate Model

1. Database is built.
2. Price and Block Size Data is almost Normally distributed despite a few possible outlying data points.
  - Price is likely not to be zero dollars as block size is also not likely to be zero sq.m. To have a practicably liveable house and meet legal building envelope requirements, land needs to be a certain minimum size and hence price.
3. Multifactor Regression is an appropriate tool to use.



## Statistical Tool

### Multi-factor Regression Analysis.

#### Key Statistical Indicators

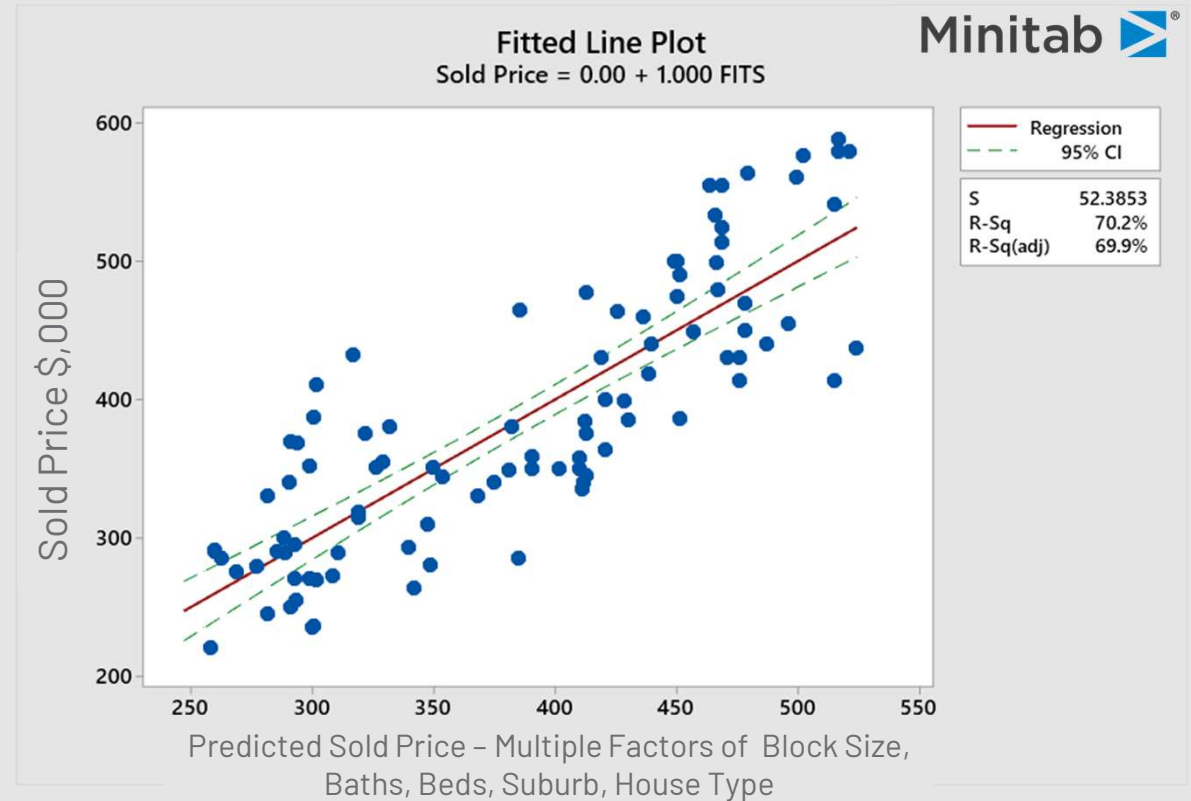
- Confidence: 95%.
- Rsq – strength of correlation. 0-100%, > 60% is usable.
- P-Value – statistical significance (less than 0.05 for each factor is acceptable).
- Variance Inflation Factor – degree of confounding or interaction between factors.

Following 7 iterations of analysis, we have an acceptable model equation with 94 data points, a correlation strength of 70% for Sold Price, containing 5 key factors:

- a) Block Size,
- b) Bathrooms,
- c) Bedrooms,
- d) Suburb
- e) House Type.

Insignificant Factors:

- Garage Size



## Predictive Model Equation

Coefficients multiplied by the variables.

$$\begin{aligned}
 \text{Sold Price } \$000 = & 312.6 \\
 & + 0.21 \times \text{Block Size} \\
 & + 59.9 \times \text{Baths} \\
 & + 0.0 \times \text{Existing} + 61.1 \times \text{New Build} \\
 & - 28.8 \times \text{Beds} \\
 & + 0.0 \times \text{Belmont} \\
 & - 10.4 \times \text{Hamlyn Heights} \\
 & - 119.7 \times \text{Norlane}
 \end{aligned}$$

A new build commands \$61k more than an existing property within this scope.

Coefficients  
-ve reduces the price  
+ve increases the price

Suburb



## Contribution of Factors

By inputting some typical average numbers for the variables, the contribution of each factor to Sold Price can be approximated.

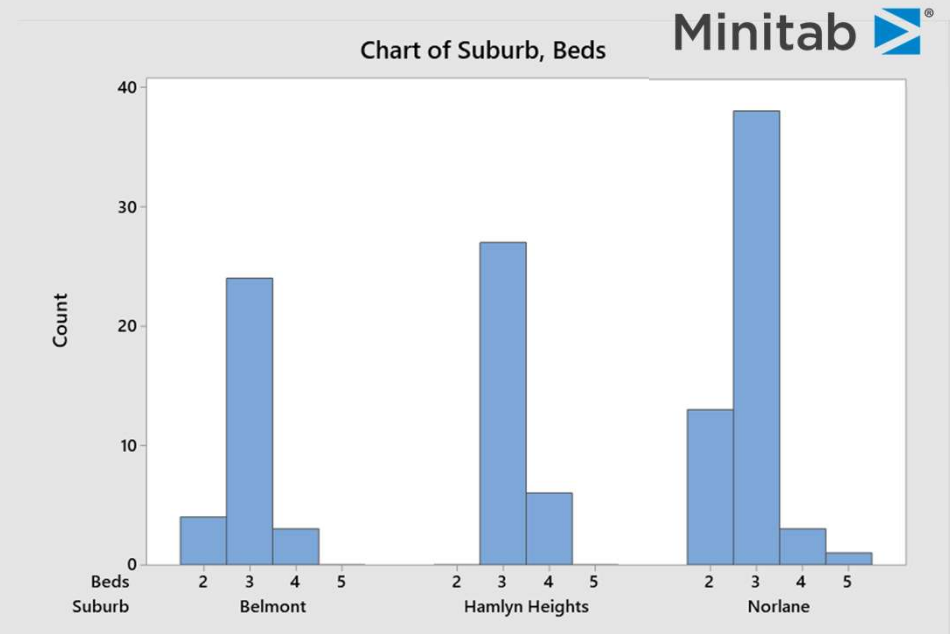
Block Size	+ 26%	
Suburb	25%	
Bathrooms	+ 19%	More Baths has Positive impact
Bedrooms	- 18%	More Beds has Negative impact.
Build (New)	+ 13%	

Once a Suburb is selected, Block Size and Number of Bathrooms have the highest contribution to the Sold Price.

## Interpretation of the Model Equation

Comparing coefficients, for all three suburbs, on average:

- a New Build commands \$61k more than an existing property.
- Norlane is \$119k cheaper than Belmont.
- More bedrooms has a negative impact on Sold Price.



## Application of the Predictive Model

Norlane & Belmont have many large Blocks with 2-3 Bedroom houses suitable for subdivision.

There is development potential by purchasing a block with

- an existing 2 or 3 bedroom house
- on the largest block available,
- subdividing into 2 or 3 blocks and
- building 2 or 3 new 2 or 3 bedroom duplex/townhouses or units, each with an ensuite bathroom in addition to the family bathroom.

## Property Development Feasibility – 1. Norlane

Factor	Buy	Sell	Notes
Block Size	600	2 x 300	Subdivided.
Baths	1	1 + 0.5	Family + ensuite
Beds	2	2	New Small starter homes.
Build	Existing	New	New build is positive.
Price	\$295k	\$348k x 2 = \$746k	
Opportunity	\$746 less \$295 less \$290k Development Costs		
Profit	\$161k = <b>FEASIBLE with Acceptable Margin</b>		

## Property Development Feasibility – 2. Norlane

Factor	Buy	Sell	Notes
Block Size	750	3 x 250	Subdivided.
Baths	1	1 + 0.5	Family + ensuite
Beds	2	3 each	New Small family homes.
Build	Existing	New	
Price	\$330k	\$309k x 3 = \$927k	
Opportunity	\$927 less \$350 less \$460k Development Costs		
Profit	\$117k = <b>FEASIBLE but narrow margin.</b>		

## Property Development Feasibility – 3. Belmont

Factor	Buy	Sell	Notes
Block Size	680	2 x 340	Subdivided.
Baths	1	2	Family + full ensuite
Beds	3	4	Large family home
Build	Existing	New	
Price	\$410k	\$450k x 2 = \$900k	
Opportunity	\$900 less \$410 less \$500k Development Costs		
Profit	- \$10k <b>LOSS – NOT FEASIBLE</b>		

## Conclusion

One would think that larger existing blocks have greater subdivision & profit potential. However, this is not what the model demonstrates.

For properties in Norlane and Belmont, by subdividing larger blocks and tailoring the new house specification to the market actually yields lower profits, has greater investment of time and resource, poorer cashflow and hence, a feasibility risk.

The Predictive Model shows that Norlane has the most profit potential by selecting an existing 2 bedroom + 1 bathroom house on approx. 600 sqm land and then subdividing into 2 smaller blocks each with new 2 bedroom units, a family bathroom and an en-suite shower-room.

We hope this Case Study has helped you appreciate the power of statistical analysis in delivering business opportunities and an overall strategy from a small data set.  
Imagine the opportunities with more data!

Data Analysis using Minitab 

**Please return to 1iDEA's website for more tutorials and case studies.**



*All you need is...*



**1iIDEA**

innovation / design / analytics

**Thank You**