

CHOICE MODELLING

EXECUTIVE SUMMARY

Conjoint analysis is a popular market research approach for measuring the value that consumers place on individual and packages of features of a product. In this exercise, I have analysed the 'HousingMarket.csv' dataset, which contains information on people who have previously purchased houses who are relocating to a high value neighbourhood in Sunnyvale, California. Consumers provided input on their preferences for houses in different models (such as colonials, craftsman style, condominiums) and indicated their preferred choices. Information on comparable house sales from the school district or zip code is provided on a square foot basis. Several attributes that are important in the purchase of a house. Also part of the dataset such as number of bedrooms, price if purchased without a mortgage, etc.

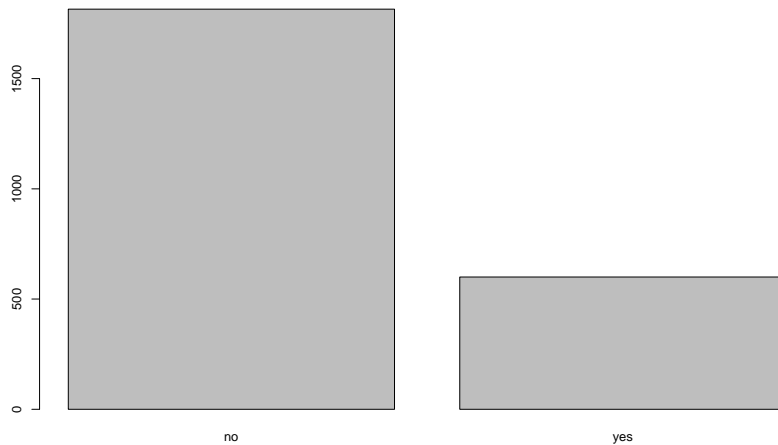
First glance at the dataset revealed the following:

```
• str(housingmarket, give.attr = FALSE)
'data.frame': 7245 obs. of 9 variables:
 $ caseid      : int  1 2 3 4 5 6 7 8 9 10 ...
 $ respondentid: int  1 1 1 1 1 1 1 1 1 1 ...
 $ houseviewed : int  1 1 1 2 2 2 3 3 3 4 ...
 $ option      : chr  "A" "B" "C" "A" ...
 $ fireplaces  : chr  "yes" "yes" "yes" "yes" ...
 $ bedrooms    : chr  "3 bedroom" "2 bedroom" "3 bedroom" "3 bedroom" ...
 $ preference   : int  0 0 1 0 1 0 1 0 0 1 ...
 $ averagecostpersquarefootcomparisons: int  90 110 90 90 100 90 110 100 110 100 ...
 $ cashprice   : int  420000 370000 370000 370000 420000 420000 420000 370000 470000 470000 ...
• head(housingmarket)
  caseid respondentid houseviewed option fireplaces bedrooms preference averagecostpersquarefootcomparisons cashprice
1      1             1           1     A    yes 3 bedroom      0              90      420000
2      2             1           1     B    yes 2 bedroom      0             110      370000
3      3             1           1     C    yes 3 bedroom      1              90      370000
4      4             1           2     A    yes 3 bedroom      0              90      370000
5      5             1           2     B    yes 3 bedroom      1             100      420000
6      6             1           2     C    yes 4 bedroom      0              90      420000
```

I proceeded to crosstab the features based on the customer's preferences:

1.FIREPLACES

fireplaces	
no	yes
1815	600

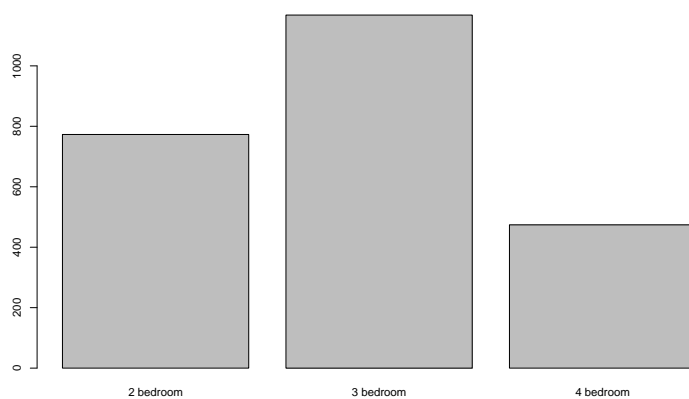


Clearly, fireplaces are not a popular feature going by the numbers with more than two-thirds prefer houses without them.

2. BEDROOMS

bedrooms

2 bedroom 3 bedroom 4 bedroom
773 1168 474



When it comes to number of bedrooms, 3-beds are the most popular, followed by 2-beds.

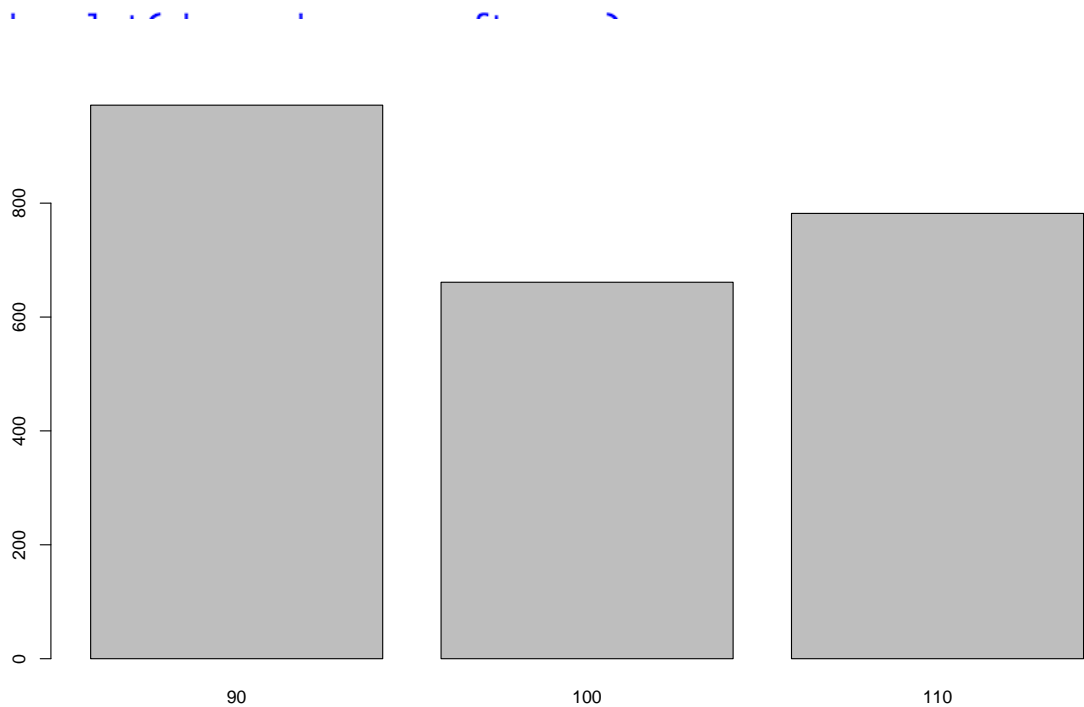
3. AVE_SQFT_COMPARISON

Average comparison of cost based on square foot in similar neighbourhood falls under 3 categories: 90,100,110

averagecostpersquarefootcomparisons

90 100 110

972 661 782



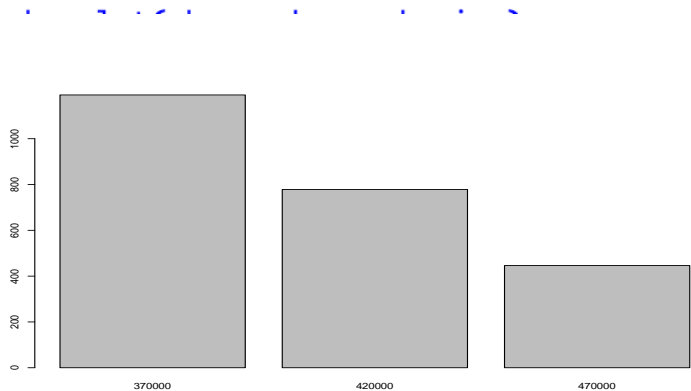
\$90 is the most preferred choice, followed by \$110.

4. CASHPRICE

cashprice

370000 420000 470000

1191 778 446



Predictably, higher the price, lower the demand.

mLogit Model:

Although, conjoint analysis is an excellent way to determine the customer preferences based on their choices, having too many choices can bring a questionnaire fatigue which will result less than ideal quality data. For that reason, and the fact there is overwhelming majority for houses without fireplaces, I have decided to drop the 'fireplaces' variable from the mLogit model.

```
Call:
mlogit(formula = preference ~ 0 + bedrooms + cashprice + averagecostpersquarefootcomparisons,
      data = housingmarket.ml, method = "nr")

Frequencies of alternatives:choice
      A      B      C
0.32588 0.33251 0.34161

nr method
5 iterations, 0h:0m:0s
g'(-H)^-1g = 2.08E-05
successive function values within tolerance limits

Coefficients :
              Estimate      Std. Error z-value      Pr(>|z|)
bedrooms3 bedroom      0.78586464779    0.06627634195  11.8574 < 0.00000000000000022 ***
bedrooms4 bedroom     -0.76197901012    0.07266532547 -10.4861 < 0.00000000000000022 ***
cashprice              -0.00001717392    0.00000076668 -22.4005 < 0.00000000000000022 ***
averagecostpersquarefootcomparisons -0.02064356878    0.00345941989  -5.9673    0.000000002411 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -2111.6
> |
```

INTERPRETATION

Using 2-bedroom houses as base line, customers statistically prefer more 3-bedroom and less of 4-bedroom houses, compared with 2-bedroom houses. Price has marginally negative effect on the choice, meaning higher the price lesser demand, but less impactful. Similarly average cost comparison is not that important a feature.

Willingness To Pay

It is great to know what features customers prefer, but it would be better to put a price on that. For example, we can deduce that the 3-bedroom houses goes at a premium of \$1000 compared to 2- bedroom houses. But there is a discount of \$969 with the 4-bedroom houses. As mentioned before, the premium or discount for cash price is almost negligible, while there is a discount of \$26 for increase in comparative average cost.

```

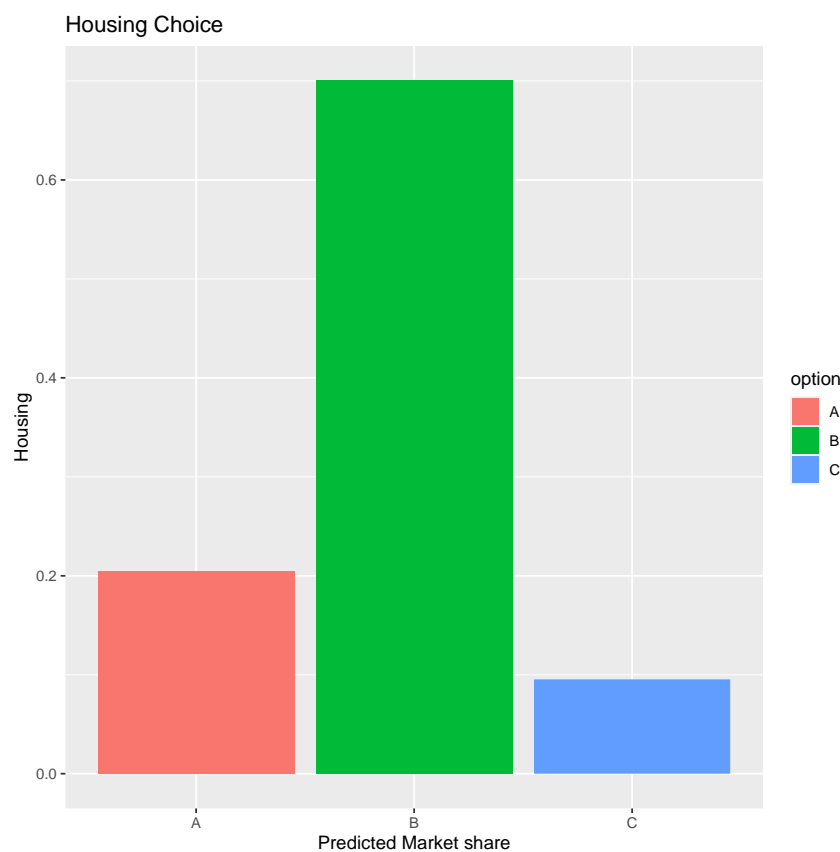
> WTP<-coef(housingmarket_model)/coef(housingmarket_model)[1]
> WTP
                bedrooms3 bedroom          bedrooms4 bedroom          cashprice
averagecostpersquarefootcomparisons          -0.96960591402          -0.00002185353
                -0.02626860597
>

```

MARKET SHARE

Based on the test data, if a product is introduced with a suite of above features, we can predict the market share of the 3 possible options of A,B and C.

	share	option	averagecostpersquarefootcomparisons	fireplaces	bedrooms	cashprice
1	0.20434206	A	90	yes	2 bedroom	420000
2	0.70028295	B	110	no	3 bedroom	370000
3	0.09537499	C	90	no	4 bedroom	420000



Option B has all the alluring features to garner a market share of 70%, compared with 20% for Option A, and less than 10% for C.

CONCLUSIONS

With choice modelling we can glean into the minds of consumers as to what appeals the most to them just by looking at the choices they have made. From this dataset of HousingMarket, we understood, that consumers prefer houses without fireplaces rather than with fireplaces. And 3-bedroom houses command at least \$1000 premium compared to 2-bedroom houses. More expensive the house, lesser the demand, but the trade-off is not very sensitive.

Beyond these insights we can also say with confidence, how much premium a feature can command. For example, a 3-bedroom house can command a premium of \$1000. Based on the preferences we have also predicted the Market Share of a new entrant with a suite of features.