|  |
| --- |
| March 25,2019 - Team 03 (Mackenzie O., Madison R., Madeline S.) |

|  |
| --- |
| Executive Summary |

# Analysis of Potential Land Rover SUV Buyers

This paper analyzes the lifestyle of potential Land Rover SUV buyers through a study of consumer interests. The consumers answered a 30 question survey available in Appendix A that asked their opinions on attributes that may influence their likelihood of buying a Discovery Land Rover. We have analyzed the answers to the survey through correlations, linear regression, distribution of buyers vs non-buyers, and used decision trees to profile a potential buyer.

In our exploratory analysis we found that the mean of Attitude was 5.023 on a Likert scale of 1-9. We found the standard deviation and variance to be 2.587 and 6.694 respectively. As shown in the histogram of attitude in Appendix B, we see spikes at values near 3, 5, and 9, with the lowest value of 7 having a count just above 35. This dataset is not normally distributed. There are just a few outliers in the data, but no extremes. There were no missing values, so no remediation occurred.

Looking at Appendix C, you can see that there are weak correlations between attitude and society fine, no time for charity, no debt, prefer cash, spendthrift, prefer credit, no coupons, low interest buyer, dependable, children important, and introverted. There is a medium correlation between attitude and active. The rest of the correlations are strong according to Pearson’s correlation model. Refer to Appendix C for a clear description on which predictors are strongly correlated with which other predictor variables.

Using SPSS Modeler’s linear regression node to see relationships between attitude and the predictor variables, we chose to use a stepwise regression, as it tests for the possibility of variables becoming non-significant. The estimated regression equation we discovered is Attitude = Fashionable \* 0.3496 + Risk Taker \* 0.6209 + Confident\*0.1194 + Adventurous \* 0.6098 + Skeptics Wrong \*0.35 + -4.015. We found the goodness of fit to be .575, which means that the model accounts for 57.5% variability for the data. Looking at the standard error, we found that the model’s predictions may be off by an average of 1.698228. Lastly, there is a distinct unbalance in the dataset. There are proportionally less buyers than non-buyers based on overall attitude. Refer to Appendix D for the Model Summary and the Buyer Distribution.

The SPSS Modeler c.5 decision tree and its rules can be seen in Appendix E. The tree uses Adventurous as its root to segment the data into buyers vs non-buyers. People who are not adventurous and not risk takers are most likely to be non-buyers. People who are adventurous, determined, prefer American cars, and their skeptics are usually wrong are most likely to be buyers.

To evaluate the performance of the decision tree, a coincidence matrix was produced and can be seen in Appendix E. Since the data set has proportionately less buyers than non-buyers based off of overall attitude, we derived the performance metrics of recall, specificity, and precision for the training set. Recall equals 49.3%, specificity equals 98.5%., and precision equals 91.7%. A gain chart can also be seen in Appendix E, which shows the percentage improvement in the decision tree classification strategy compared to a random chance strategy. At about 15% of the sample, we attain the largest gain with respect to a random sample strategy.

## Conclusions

Based on the correlations and linear regression results, we can conclude that the higher the attitude, the more likely they are to be buyers. Attitude raises mainly for adventurous people and risk-takers. Since there are proportionally less buyers than non-buyers, the marketing team should consider the type of people they are trying to get to become buyers.

Based on the results of the decision tree, marketing for the Discovery Land Rover should be targeted at the type of people who are particularly adventurous risk-takers and love American cars. This would stereotypically mean young to middle-aged men who live in more rural areas of America.

## Appendix A:

## Survey Questions

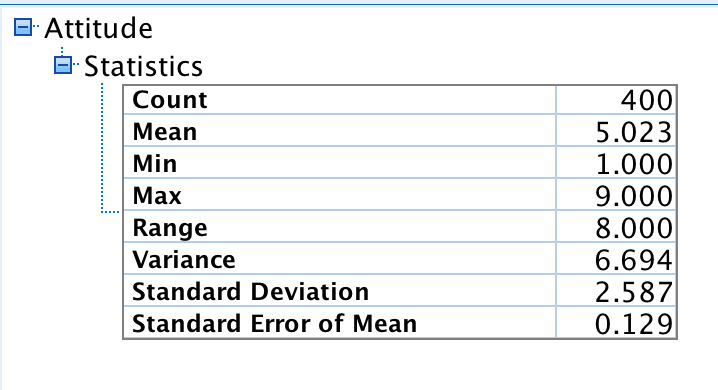
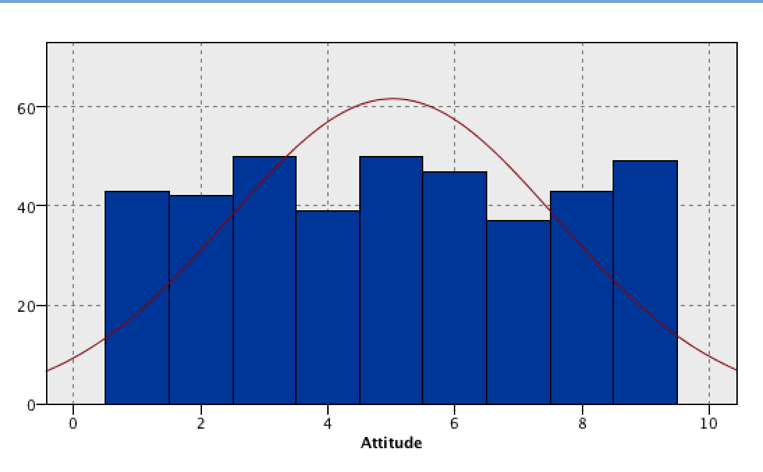
Respondents were asked to use a 9 point Likert scale to give their answer to each question. The value of "1" means the responder disagreed with the statement, and the value of "9" means the responder totally agrees. 400 consumers were surveyed. The profiles were obtained from the mailing lists of Car and Driver, Business Week, and Inc. magazines.

The survey questions (prefixed by their column name in the data file):

1. In Shape - I am in very good physical condition
2. Fashionable - When I must choose between the two, I dress for fashion, not comfort
3. Stylish - I have more stylish clothes than most of my friends
4. Individualistic - I want to look a little different from others
5. Risk Taker - Life is too short not to take some gambles
6. No Ozone Concern - I am not concerned about the ozone layer
7. Right To Pollute- I think the government is doing too much to control pollution
8. Society Fine - Basically, society today is fine
9. No Time For Charity - I don't have time to volunteer for charities
10. No Debt - Our family is not too heavily in debt today
11. Prefer Cash - I like to pay cash for everything I buy
12. Spendthrift - I pretty much spend for today and let tomorrow bring what it will
13. Prefer Credit - I use credit cards because I can pay the bill off slowly
14. No Coupons - I seldom use coupons when I shop
15. Low Interest Buyer - Interest rates are low enough to allow me to buy what I want
16. Confident - I have more self-confidence than most of my friends
17. Leader - I like to be considered a leader
18. Dependable - Others often ask me to help them out of a jam
19. Children Important - Children are the most important thing in a marriage
20. Introverted - I would rather spend a quiet evening at home than go out to a party
21. American Cars Rule - Foreign-made cars can't compare with American-made cars
22. Restrict Japan Imports - The government should restrict imports of products from Japan
23. Buy American - Americans should always try to buy American products
24. Adventurous - I would like to take a trip around the world
25. Midlife Crisis - I wish I could leave my present life and do something entirely different
26. Early Adopter - I am usually among the first to try new products
27. Active - I like to work hard and play hard
28. Skeptics Wrong - Skeptical predictions are usually wrong
29. Determined - I can do anything I set my mind to
30. Optimistic - Five years from now, my income will be a lot higher than it is now
31. Attitude - I would consider buying the Discovery made by Land Rover

## Appendix B:

## Initial Insights of Likelihood of Buying Discovery SUV with 0 extreme outliers and no missing data. No data remediation occurred. Attitude towards buying a car is evenly distributed with no normal curve. At first glance with a mean of 5 on a scale of 1 to 9, there must be further data exploration to determine who should be targeted for marketing based on customer interests and opinions.



## Appendix C:

## Relationship of Attitude with Survey Answers

## Relationship among Predictors

|  |  |
| --- | --- |
| Predictor | Strongly Correlated to |
| 1. In Shape | fashionable, stylish, individualistic, risk taker, adventurous, midlife crisis, skeptics wrong, optimistic, and attitude |
| 1. Fashionable | In shape, stylish, individualistic, risk taker, adventurous, midlife crisis, skeptics wrong, optimistic, and attitude |
| 1. Stylish | In shape, fashionable, individualistic, risk taker, skeptics wrong, determined, optimistic, and attitude |
| 1. Individualistic | In shape, fashionable, stylish, risk taker, adventurous, skeptics wrong, determined, optimistic, and attitude |
| 1. Risk Taker | In shape, fashionable, stylish, risk taker, individualistic, no ozone concern, right to pollute, American cars rule, restrict japan imports, buy American, adventurous, midlife crisis, early adopter, skeptics wrong, determined, optimistic, and attitude |
| 1. No Ozone Concern | Risk taker, right to pollute, society fine, no time for charity, confident, leader, adventurous, and attitude |
| 1. Right To Pollute | Risk taker, no ozone concern, society fine, no time for charity, and attitude |
| 1. Society Fine | No ozone concern, right to pollute, no time for charity |
| 1. No Time For Charity | no ozone concern, right to pollute, society fine |
| 1. No Debt | prefer cash, spendthrift, prefer credit, no coupons, low interest buyer |
| 1. Prefer Cash | No debt, spendthrift, prefer credit, no coupons, low interest buyer |
| 1. Spendthrift | No debt, prefer cash, prefer credit, no coupons, low interest buyer |
| 1. Prefer Credit | No debt, spendthrift, prefer cash, no coupons, low interest buyer |
| 1. No Coupons | No debt, spendthrift, prefer credit, prefer cash, low interest buyer |
| 1. Low Interest Buyer | No debt, spendthrift, prefer credit, no coupons, prefer cash |
| 1. Confident | No ozone concern, leader, dependable, American cars rule, restrict japan imports, and attitude |
| 1. Leader | No ozone concern, confident, dependable, restrict japan imports, skeptics wrong, determined, optimistic, and attitude |
| 1. Dependable | confident, leader |
| 1. Children Important | Introverted |
| 1. Introverted | Children important |
| 1. American Cars Rule | Risk taker, confident, restrict japan imports, buy American, adventurous, midlife crisis, and attitude |
| 1. Restrict Japan Imports | Risk taker, confident, leader, American cars rule, buy American, adventurous, midlife crisis, skeptics wrong, optimistic, and attitude |
| 1. Buy American | Risk taker, American cars rule, restrict japan imports, adventurous, midlife crisis, optimistic, and attitude |
| 1. Adventurous | In shape, fashionable, individualistic, risk taker, no ozone concern, American cars rule, restrict japan imports, buy American, midlife crisis, early adopter, active, skeptics wrong, determined, and attitude |
| 1. Midlife Crisis | In shape, fashionable, risk taker, American cars rule, restrict japan imports, buy American, adventurous, early adopter, active, skeptics wrong, determined, optimistic, and attitude |
| 1. Early Adopter | Risk taker, adventurous, midlife crisis, active, skeptics wrong, determined, optimistic, and attitude |
| 1. Active | Adventurous, midlife crisis, early adopter, skeptics wrong, determined, optimistic |
| 1. Skeptics Wrong | In shape, fashionable, stylish, individualistic, risk taker, leader, restrict japan imports, adventurous, midlife crisis, early adopter, skeptics wrong, determined, optimistic. |
| 1. Determined | Stylish, individualistic, risk taker, leader, adventurous, midlife crisis, early adopter, skeptics wrong, determined, optimistic. |
| 1. Optimistic | In shape, fashionable, stylish, individualistic, risk taker, leader, restrict japan imports, buy American, midlife crisis, early adopter, skeptics wrong, determined, optimistic. |

## Appendix D:

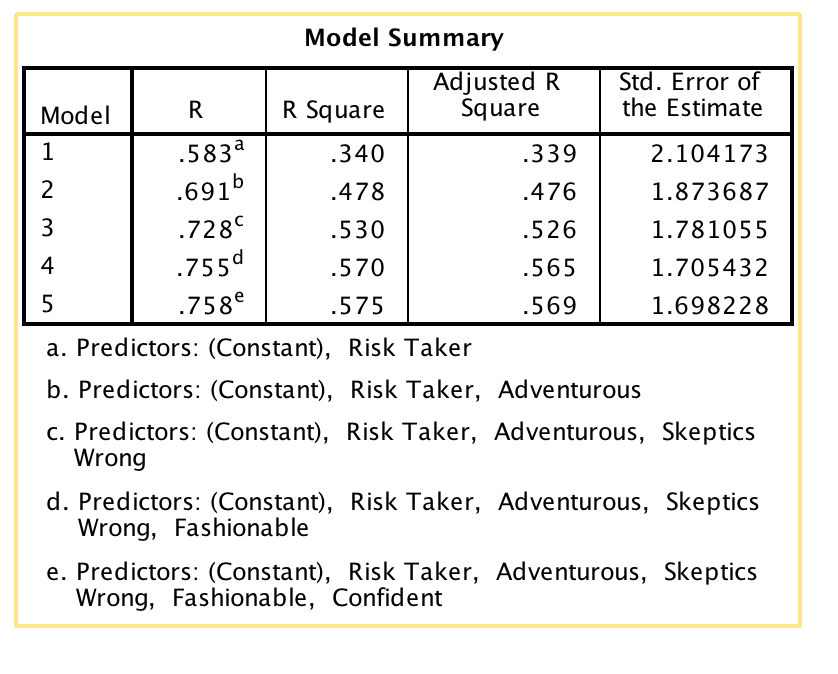
## Linear Regression

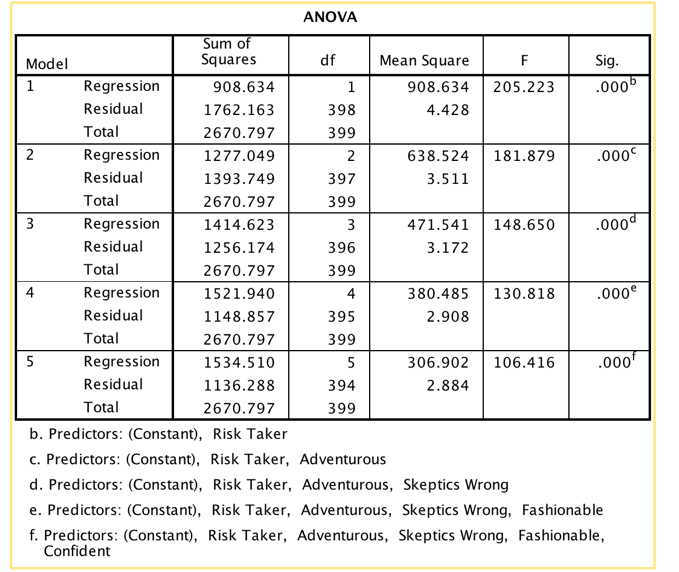
The Linear regression equation using a stepwise regression method results in:

Attitude = Fashionable \* 0.3496 + Risk Taker \* 0.6209 + Confident\*0.1194 + Adventurous \* 0.6098 + Skeptics Wrong \* 0.35 + -4.015.

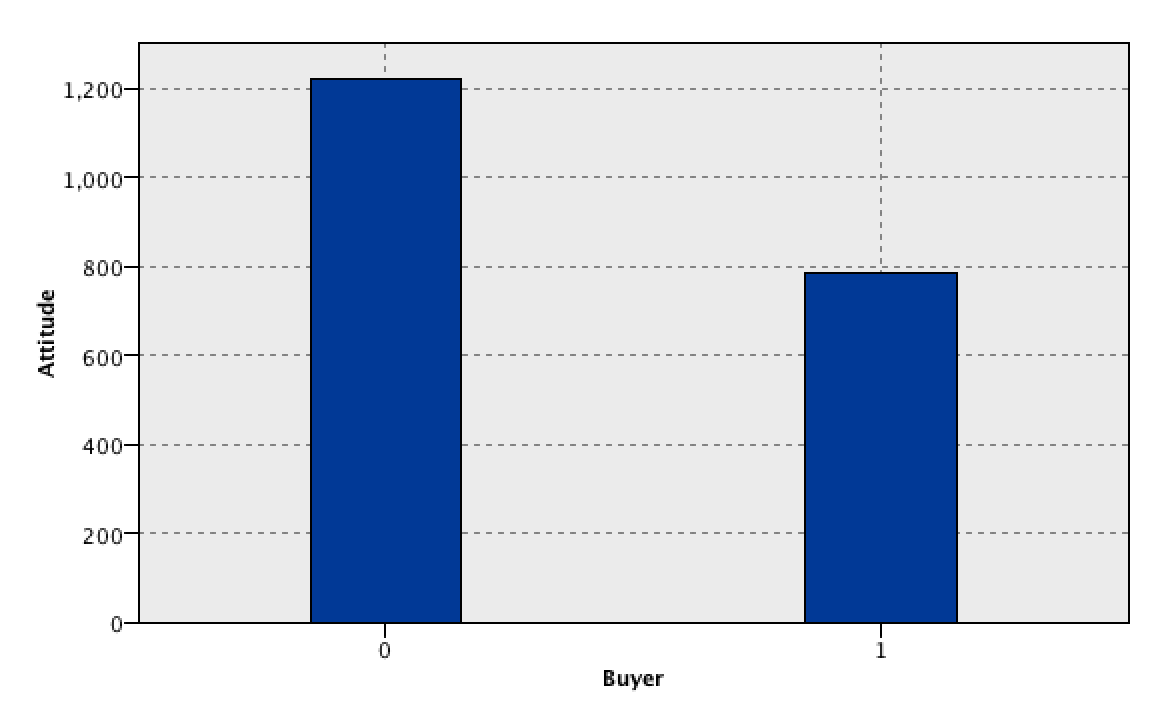
This means that if all the variables are 0, the attitude is -4.015 as a base value, however this is impossible as attitude must be at least 1. As Fashionable increases on the Likert scale, it increases the likelihood of increasing attitude by 0.3496. This is true for the other variables and their slopes.

The goodness of fit (R2) for the final model in the step-wise regression (e/f) is .575, meaning that the model accounts for 57.5% variability of the data. We get this number from SSR/SST. The standard error of the estimate is 1.698228, meaning that the model prediction may be off by an average of 1.698228.





Buyer Distribution (o for non-buyers, 1 for buyers)

There is a distinct unbalance is the dataset in which there are proportionately less buyers that non- buyers based off of overall attitude.

## Appendix E:

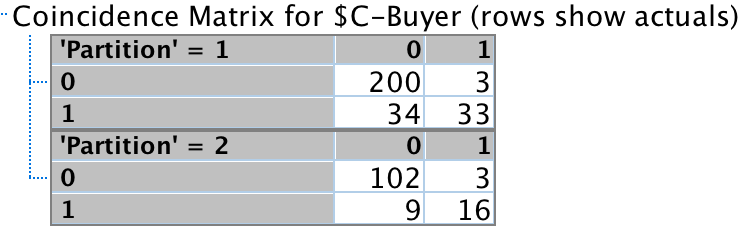
Decision Tree Classification



Decision Tree Rules

|  |  |  |  |
| --- | --- | --- | --- |
| **Antecedent** | **Consequence** | **Support** | **Confidence** |
| If (Adventurous<=5.5) ^ (Risk Taker<=5.5) | Then Buyer=0 | 229/400 | 93.5% |
| If (Adventurous<=5.5) ^ (Risk Taker>5.5) ^ (Risk Taker<=7.5) | Then Buyer=0 | 52/400 | 69.3% |
| If (Adventurous<=5.5) ^ (Risk Taker>5.5) ^ (Risk Taker>7.5) | Then Buyer=1 | 8/400 | 88.9% |
| If (Adventurous>5.5) ^ (Determined<=2.5) | Then Buyer=0 | 8/400 | 88.9% |
| If (Adventurous>5.5) ^ (Determined>2.5) ^ (American Cars Rule<=3.5) | Then Buyer=0 | 8/400 | 72.7% |
| If (Adventurous>5.5) ^ (Determined>2.5) ^ (American Cars Rule>3.5) ^ (Skeptics Wrong>4.5) | Then Buyer=1 | 31/400 | 91.2% |
| If (Adventurous>5.5) ^ (Determined>2.5) ^ (American Cars Rule>3.5) ^ (Skeptics Wrong<=4.5) ^ (Risk Taker<=5.5) | Then Buyer=0 | 5/400 | 100% |
| If (Adventurous>5.5) ^ (Determined>2.5) ^ (American Cars Rule>3.5) ^ (Skeptics Wrong<=4.5) ^ (Risk Taker>5.5) | Then Buyer=1 | 10/400 | 83.3% |

Performance Evaluation -- Coincidence Matrix



Since the data set is unbalanced, with proportionately less buyers than non-buyers based off of overall attitude, we derived the performance metrics of recall, specificity, and precision for the training set. Recall is equal to 33/(33+34)=0.493, or 49.3%. Specificity is equal to 200/(200+3)=0.985, or 98.5%. Precision is equal to 33/(33+3)=0.917, or 91.7%.

Performance Evaluation -- Gain Chart



This gain chart shows the percentage improvement in the decision tree classification strategy compared to a random chance strategy. At about 15% of the sample, we attain the largest gain with respect to a random sample strategy.