Multinomial Logistic Regression for Type of CS

Rainer Lempert

2018-05-02 23:01:15

# Develop The Model

33 dependent variables from the survey were chosen:

* Q6 - How long have you been a member of a CSO
* Reasons for CS use
  + Q7\_1 - To reduce carbon footprint
  + Q7\_2 - To save money
  + Q7\_3 - For convenience
  + Q7\_4 - ‘Just in case’ you need it
  + Q7\_5 - Because there is no Uber/Lyft in town
  + Q7\_6 - Personal safety - safer than public transit
* Q8 - Have you disposed a vehicle and used CS instead
* Q9 - Without CS would you have bought or acquired a vehicle?
* Cost-related reasons for CS use
  + Q10\_1 - CS saves me money because I don’t have to own a vehicle (or another vehicle)
  + Q10\_2 - CS saves me money because it’s cheaper than using a taxi
  + Q10\_3 - I look at the cost of CS vs cost of walking or biking
  + Q10\_4 - I look at the cost of CS vs the cost of public transit
  + Q10\_5 - I rarely evaluate the cost of using CS versus the alternatives
* Q11 - Can I afford another vehicle?
* Material benefits of CS
  + Q13\_1 - Easier to meet up with friends and family
  + Q13\_2 - I go to more places in the city
  + Q13\_3 - I go to more places outside the city
  + Q13\_4 - I get ‘stuff’ done more efficiently; errands, meetings, shopping, etc
* Mental benefits of CS
  + Q14\_1 - I enjoy the freedom
  + Q14\_2 - I like not having to rely on others for a ride
  + Q14\_3 - Peace of mind knowing I have personal mobility when needed
  + Q14\_4 - I like not owning a vehicle (or another vehicle)
  + Q14\_5 - I like having different options for getting around
* Q15 - Has your CS use fallen off?
* Q23 - Age
* Q24 - Gender
* Q26 - Number of drivers per household
* Q27 - Number of vehicles per household
* Q28 - Number of children per household
* Q29 - Are you a homeowner
* Q30 - Affordability index
* Q31 - Income

Variables were chosen to be either numerical, ordinal, or binary.

# Model with ‘Both’ as reference variable

The models predictions. Adding down a column gets actual values, across a row is the predicted values. The misclassification rate is 19.9% (1 minus the sum of the diagonals divided by the sum of all values)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Both | One-Way | Two-Way |
| Both | 122 | 55 | 33 |
| One-Way | 211 | 1339 | 47 |
| Two-Way | 15 | 11 | 33 |

I conducted a two-tailed Z-Test. \*\* means that both variables are under the 0.05 confidence level. \* means that only one is under the 0.05 confidence level.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | One.Way.Coefficient | One.Way.PValue | Two.Way.Coefficient | Two.Way.PValue | Sig |
| (Intercept) | 3.1774979 | 0.0054132 | 0.4876687 | 0.7983172 | \* |
| Q6 | -0.9627702 | 0.0000000 | -0.2816569 | 0.0259186 | \*\* |
| Q7\_1 | -0.1062498 | 0.0863534 | 0.0515245 | 0.6578448 |  |
| Q7\_2 | -0.0650280 | 0.2836991 | 0.3128216 | 0.0129611 | \* |
| Q7\_3 | 0.0045270 | 0.9716610 | -0.4576821 | 0.0079323 | \* |
| Q7\_4 | -0.0300078 | 0.5330914 | -0.1296362 | 0.1185948 |  |
| Q7\_5 | 0.0428778 | 0.3816675 | -0.2656482 | 0.0071603 | \* |
| Q7\_6 | 0.0736375 | 0.1895508 | -0.0632598 | 0.5573339 |  |
| Q8 | 0.1541100 | 0.3371089 | 0.1014458 | 0.7156876 |  |
| Q9 | 0.4596211 | 0.0035435 | 0.6571955 | 0.0179047 | \*\* |
| Q10\_1 | -0.0983048 | 0.1507490 | 0.0631143 | 0.6449575 |  |
| Q10\_2 | 0.0954803 | 0.2300497 | -0.4320813 | 0.0000917 | \* |
| Q10\_3 | 0.0441352 | 0.4080558 | -0.0082756 | 0.9304851 |  |
| Q10\_4 | 0.0138279 | 0.8144187 | -0.1422278 | 0.1532999 |  |
| Q10\_5 | 0.1124147 | 0.0595298 | 0.1869391 | 0.0554716 |  |
| Q11 | 0.2598442 | 0.1267764 | -0.3606576 | 0.2329907 |  |
| Q13\_1 | 0.1377284 | 0.0381115 | -0.1511184 | 0.1775218 | \* |
| Q13\_2 | 0.0385713 | 0.5426054 | -0.0082485 | 0.9417997 |  |
| Q13\_3 | -0.3573171 | 0.0000000 | 0.3185909 | 0.0005073 | \*\* |
| Q13\_4 | -0.1157904 | 0.0814063 | -0.0250179 | 0.8239245 |  |
| Q14\_1 | 0.0936881 | 0.3379510 | -0.0727543 | 0.6624139 |  |
| Q14\_2 | -0.0669127 | 0.3891852 | 0.0373750 | 0.7754612 |  |
| Q14\_3 | -0.0404947 | 0.6909995 | 0.3167424 | 0.1067670 |  |
| Q14\_4 | -0.0492659 | 0.3726175 | 0.0059655 | 0.9549660 |  |
| Q14\_5 | 0.0228250 | 0.8383507 | -0.1625897 | 0.4032343 |  |
| Q15 | -0.0914543 | 0.5620548 | 0.2731014 | 0.3408456 |  |
| Q23 | 0.1287547 | 0.0979064 | 0.5013728 | 0.0000112 | \* |
| Q24 | 0.0524529 | 0.6820213 | 0.1676120 | 0.4341638 |  |
| Q26\_1\_TEXT | 0.0901365 | 0.3999891 | -0.1374438 | 0.5057943 |  |
| Q27\_1\_TEXT | 0.6448321 | 0.0000005 | 0.5094826 | 0.0305368 | \*\* |
| Q28 | -0.1494580 | 0.1799788 | -0.0023641 | 0.9898160 |  |
| Q29 | -0.2310369 | 0.2162795 | -0.0322288 | 0.9175423 |  |
| Q30 | 0.0192102 | 0.8596395 | 0.0214670 | 0.9071109 |  |
| Q31 | -0.0826118 | 0.1672231 | -0.3018808 | 0.0050092 | \* |

After doing this analysis, these are the variables that have explanatory power (at an ) over type of CS usage:

* Q6 - How long have you been a member of a CSO
  + *STRONG* correlation between people being members for longer periods of time and using both over one-way
  + Correlation between people being members for longer periods of time and using both over two-way
* Q7\_2 - To save money
  + Correlation between people agreeing and using two-way over both
* Q7\_3 - For convenience
  + Correlation between people agreeing and using both over two-way
* Q7\_5 - Because there is no Uber/Lyft in town
  + Correlation between people agreeing and using both over two-way
* Q9 - Without CS would you have bought or acquired a vehicle?
  + 2=no, 1=yes.
  + Correlation between people disagreeing and using one-way over both
  + Correlation between people disagreeing and using two-way over both
* Q10\_2 - CS saves me money because it’s cheaper than using a taxi
  + Correlation between people agreeing and using both over two-way
* Q13\_1 - Easier to meet up with friends and family
  + Correlation between agreeing and using one-way over both
* Q13\_3 - I go to more places outside the city
  + *STRONG* correlation between people agreeing and using both over one-way
  + Correlation between people agreeing and using two-way over one-way
* Q23 - Age
  + Correlation between people being older and using two-way over both
* Q27 - Number of vehicles per household
  + *STONG* correlation between people having more vehicles and using one-way over both
  + Correlation between people having more vehicles and using two-way over both
* Q31 - Income
  + Correlation between people having more money and using both over two-way

# Model with ‘One-Way’ as reference variable

The models predictions are identical to the previous models. Please refer to the above section for that.

Two tailed z-test: \*\* means that both variables are under the 0.05 confidence level. \* means that only one is under the 0.05 confidence level.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Both.Coefficient | Both.PValue | Two.Way.Coefficient | Two.Way.PValue | Sig |
| (Intercept) | -3.1774476 | 0.0054139 | -2.6895479 | 0.1587605 | \* |
| Q6 | 0.9627769 | 0.0000000 | 0.6811106 | 0.0000001 | \*\* |
| Q7\_1 | 0.1062515 | 0.0863485 | 0.1577932 | 0.1749887 |  |
| Q7\_2 | 0.0650239 | 0.2837293 | 0.3778534 | 0.0026874 | \* |
| Q7\_3 | -0.0045243 | 0.9716777 | -0.4622159 | 0.0073351 | \* |
| Q7\_4 | 0.0300045 | 0.5331360 | -0.0996119 | 0.2304368 |  |
| Q7\_5 | -0.0428823 | 0.3816176 | -0.3084912 | 0.0017901 | \* |
| Q7\_6 | -0.0736349 | 0.1895670 | -0.1368798 | 0.2041864 |  |
| Q8 | -0.1541283 | 0.3370516 | -0.0526975 | 0.8499333 |  |
| Q9 | -0.4596279 | 0.0035430 | 0.1975232 | 0.4767198 | \* |
| Q10\_1 | 0.0983098 | 0.1507282 | 0.1614213 | 0.2386020 |  |
| Q10\_2 | -0.0954776 | 0.2300629 | -0.5275686 | 0.0000018 | \* |
| Q10\_3 | -0.0441366 | 0.4080408 | -0.0523927 | 0.5807559 |  |
| Q10\_4 | -0.0138286 | 0.8144098 | -0.1560821 | 0.1170985 |  |
| Q10\_5 | -0.1124154 | 0.0595280 | 0.0745298 | 0.4451373 |  |
| Q11 | -0.2598601 | 0.1267533 | -0.6205310 | 0.0401610 | \* |
| Q13\_1 | -0.1377308 | 0.0381080 | -0.2888622 | 0.0099515 | \*\* |
| Q13\_2 | -0.0385783 | 0.5425326 | -0.0468218 | 0.6785658 |  |
| Q13\_3 | 0.3573249 | 0.0000000 | 0.6759448 | 0.0000000 | \*\* |
| Q13\_4 | 0.1157901 | 0.0814072 | 0.0907691 | 0.4195129 |  |
| Q14\_1 | -0.0936780 | 0.3380034 | -0.1664597 | 0.3178483 |  |
| Q14\_2 | 0.0669166 | 0.3891573 | 0.1042728 | 0.4261547 |  |
| Q14\_3 | 0.0404849 | 0.6910705 | 0.3572210 | 0.0689085 |  |
| Q14\_4 | 0.0492682 | 0.3725953 | 0.0552214 | 0.6011473 |  |
| Q14\_5 | -0.0228296 | 0.8383189 | -0.1854351 | 0.3404355 |  |
| Q15 | 0.0914609 | 0.5620268 | 0.3645988 | 0.2035112 |  |
| Q23 | -0.1287575 | 0.0978993 | 0.3726237 | 0.0010966 | \* |
| Q24 | -0.0524477 | 0.6820508 | 0.1151786 | 0.5909711 |  |
| Q26\_1\_TEXT | -0.0901536 | 0.3998996 | -0.2276004 | 0.2705174 |  |
| Q27\_1\_TEXT | -0.6448265 | 0.0000005 | -0.1353295 | 0.5655927 | \* |
| Q28 | 0.1494617 | 0.1799678 | 0.1470585 | 0.4272022 |  |
| Q29 | 0.2310399 | 0.2162736 | 0.1987978 | 0.5230782 |  |
| Q30 | -0.0192126 | 0.8596227 | 0.0022531 | 0.9902290 |  |
| Q31 | 0.0826125 | 0.1672195 | -0.2192637 | 0.0415109 | \* |

After doing this analysis, these are the variables that have explanatory power (at an ) over type of CS usage:

* Q6 - How long have you been a member of a CSO
  + *STRONG* correlation between using carshare longer and using both over one-way
  + *STRONG* correlation between using carshare longer and using two-way over one-way
* Q7\_2 - To save money
  + Correlation between agreeing and using two-way over one-way
* Q7\_3 - For convenience
  + Correlation between agreeing and using one-way over two-way
* Q7\_5 - Because there is no Uber/Lyft in town
  + Correlatoin between agreeing and using one-way over two-way
* Q9 - Without CS would you have bought or acquired a vehicle?
  + 2=no, 1=yes.
  + Correlation between disagreeing and using one-way over both
* Q10\_2 - CS saves me money because it’s cheaper than using a taxi
  + Correlation between people agreeing with this and using one-way, not two-way.
* Q11 - Can I afford another vehicle?
  + 2=no, 1=yes.
  + Correlation between people disagreeing with this and using one-way over two-way
* Q13\_1 - Easier to meet up with friends and family
  + Correlation between agreeing and using one-way over two-way.
  + Correlation between agreeing and using one-way over both
* Q13\_3 - I go to more places outside the city
  + *STRONG* correlation between agreeing and using both over one-way
  + *STRONG* correlation between agreeing and using two-way over one-way
* Q23 - Age
  + Correlation between being older and using two-way over one-way
* Q27 - Number of vehicles per household
  + Correlation between having more vehicles and using one-way over two-way
* Q31 - Income
  + Correlation between being wealthier and using one-way over two-way