

Lead__Scoring

Ziyan Zhao

31/08/2020

Attributes:

- 1: Id
- 2: StageName (factor)
- 3: Status_Reason___c 4: RecordType.Name
- 5: RICE_Supported___c
- 6: CreatedDate
- 7: AccountId
- 8: Lead_Faculty___c
- 9: Parent_Opportunity___c (factor)
- 10: RecordType.Name.1
- 11: Industry
- 12: Business_Type___c
- 13: Is_External___c
- 14: ParentId (factor)
- 15: RecordType

```
# Reading Data In
data <- read.csv("/Users/ChantelleChiu/Documents/GitHub/Project-90106-G28/Chantelle/clean_data.csv",
                 header = TRUE, stringsAsFactors = T)
names(data) <- c("X1", "X2", "X3", "X4", "X5", "X6", "X7", "X8", "X9", "X10",
                 "X11", "X12",
                 "X13", "X14", "X15")

summary(data)
```

```
##           X1           X2
## 0062e000002Hc2PAAS:    1   Min.    :0.0000
## 0062e000002HFaaAAG:    1   1st Qu.:0.0000
## 0062e000002HFabAAG:    1   Median :0.0000
## 0062e000002HFaiAAG:    1   Mean   :0.3389
## 0062e000002HFalAAG:    1   3rd Qu.:1.0000
## 0062e000002HFavAAG:    1   Max.    :1.0000
## (Other)                :6046
```

```

##                                     X3
## Other (Lost)                       :2347
## Won                               :2047
## Customer No Longer Interested    : 680
## Proposal/Application Unsuccessful: 337
## Academic Capacity (Lost)         : 202
## Out-Sold                         : 121
## (Other)                          : 318
##                                     X4                                     X5
## Research Contract                 :2976   Not supported   : 484
## Custom Education (MSPACE Included):1171   RIC RE&D and BD&I:3031
## Grants                           : 725   RIC-BD&I         :2241
## Consultancy (Non-research)       : 486   RIC-RE&D         : 296
## Technology Transfer              : 369
## Competitive Bid                  : 89
## (Other)                         : 236
##                                     X6                                     X7                                     X8
## Min.      :2015    0                               : 644   0012e000002ZGfbAAG:1191
## 1st Qu.:2017    0012e000003A6qHAAS: 279   0012e000002ZLg7AAG: 947
## Median :2018    0012e000003A716AAC: 239   0                               : 909
## Mean      :2018    0012e000003A6qZAAS: 186   0012e000002Zmo0AAC: 781
## 3rd Qu.:2019    0012e000003A6bnAAC: 151   0012e000002ZmnxAAC: 671
## Max.      :2020    0012e000003A6ElAAK: 71   0012e000002ZmnzAAC: 310
##                                     (Other)       :4482   (Other)       :1243
##                                     X9                                     X10
## Min.      :0.00000   Administrative       : 1
## 1st Qu.:0.00000   Business Organization:6012
## Median :0.00000   University Department: 39
## Mean      :0.09336
## 3rd Qu.:0.00000
## Max.      :1.00000
##
##                                     X11
## Services                       :2694
## Health                         :1146
## Public Administration          : 693
## Defence                       : 386
## Agriculture & Horticulture     : 212
## Technology: IT & Telephone Services: 205
## (Other)                       : 716
##                                     X12                                     X13                                     X14
## Not for profit                 :1333   External:5368   Min.      :0.0000
## Government: Australia: Local & State :1126   Internal: 40   1st Qu.:0.0000
## Government: Australia: Federal      :1007   missing : 644   Median :0.0000
## SME (small to medium enterprise)    : 756   Mean      :0.1454

```

```

## Multinational / Other Large Corporate: 680          3rd Qu.:0.0000
## University : 446          Max. :1.0000
## (Other) : 704
##
##                                     X15
## Research ContractBusiness Organization :2963
## Custom Education (MSPACE Included)Business Organization:1158
## GrantsBusiness Organization : 721
## Consultancy (Non-research)Business Organization : 482
## Technology TransferBusiness Organization : 364
## Competitive BidBusiness Organization : 89
## (Other) : 275

mod1 <- glm(factor(X2)~X3+X4+X5+X6+X8+factor(X9)+X10+X11+X12+X13+factor(X14)+X15,
            family = "binomial", data = data)

# 10-fold cross_validation
train_control <- trainControl(method = "cv", number = 10)

# train the model on data set
model <- train(factor(X2)~X3+X4+X5+X6+X8+factor(X9)+X10+X11+X12+X13+factor(X14)+X15,
              data = data,
              trControl = train_control,
              method = "glm",
              family="binomial")

# print cv scores
print(model)

## Generalized Linear Model
##
## 6052 samples
## 12 predictor
## 2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 5447, 5447, 5447, 5446, 5446, 5447, ...
## Resampling results:
##
## Accuracy Kappa
## 0.9980168 0.9955805

summary(model)

##
## Call:

```

```

## NULL
##
## Deviance Residuals:
##      Min        1Q      Median        3Q        Max
## -1.575e-04  -2.100e-08  -2.100e-08   2.100e-08   1.691e-04
##
## Coefficients: (13 not defined because of singularities)
##                                     Estimate
## (Intercept)                        2.283e+08
## `X3Academic No Longer Interested`    1.029e+01
## `X3Awaiting Customer Action`         3.306e+01
## `X3BD Capacity (Lost)`               1.524e+01
## `X3Customer No Longer Interested`    2.411e+00
## `X3Customer Not Responding`         4.749e+00
## `X3Failure to Agree - IP terms`      6.109e+00
## `X3Failure to Agree - Price`        1.985e+01
## `X3Failure to Agree - Scope of Work` 5.384e+00
## `X3Not Ready for Market`            5.595e+00
## `X3On Track`                        4.808e+01
## `X3Other (Lost)`                    3.200e+00
## `X3Other IP Issue`                  -1.089e+00
## `X3Other(Lost)`                     4.721e+00
## `X3Out-Sold`                        -5.484e-01
## `X3Proposal/Application Unsuccessful` 3.420e+00
## `X3Scope of Proposal`               3.074e+01
## `X3Unacceptable Risk`               1.880e+01
## X3Won                              1.143e+02
## `X4Competitive Bid`                 -3.191e+01
## `X4Consultancy (Non-research)`      -2.283e+08
## `X4Custom Education (MSPACE Included)` -2.267e+01
## X4Grants                           -2.205e+01
## X4Internship                       -2.105e+01
## X4Panel                            -2.628e+01
## `X4Parent Grant`                   -1.937e+01
## X4Philanthropic                    -2.516e+01
## `X4Research Contract`               -1.963e+01
## X4Sponsorship                      -1.297e+01
## `X4Technology Transfer`             -2.378e+01
## `X5RIC RE&D and BD&I`               3.247e+00
## `X5RIC-BD&I`                       1.183e+01
## `X5RIC-RE&D`                       1.146e+01
## X6                                 -2.293e+00
## X80012e000002ZFZYAA4               1.148e+01
## X80012e000002ZGfbAAG               2.432e+01
## X80012e000002ZLg7AAG              2.307e+01

```

## X80012e000002ZLgIAAW	1.095e+01
## X80012e000002ZLgJAAW	1.926e+01
## X80012e000002ZmnxAAC	2.112e+01
## X80012e000002ZmnzAAC	1.810e+01
## X80012e000002Zmo0AAC	1.379e+01
## X80012e000002Zm0UAA0	4.875e+01
## X80012e000002Zm0ZAA0	1.347e+01
## X80012e000002ZNY0AA4	2.705e+01
## X80012e000002Zt0MAAS	4.759e+01
## X80012e000002Zt1iAAC	5.145e+01
## X80012e000002Zt1RAAS	2.609e+01
## X80012e000002Zt1zAAC	6.480e+01
## `factor(X9)1`	-5.841e+00
## `X10Business Organization`	-2.283e+08
## `X10University Department`	-2.283e+08
## `X11Agriculture & Horticulture`	-2.486e+01
## `X11Aquaculture & Fisheries`	2.253e+00
## `X11Construction & Construction Materials`	-2.713e+01
## X11Defence	-4.835e+01
## X11Education	NA
## `X11Energy & Utilities`	-2.471e+01
## `X11Forestry & Paper`	-8.194e+01
## X11Health	-2.212e+01
## `X11Manufacturing & R&D`	-3.934e+01
## `X11Materials: Chemicals & Industrial Metals`	-2.509e+01
## `X11Mining & Refining`	-2.489e+01
## `X11Mining Equipment and Technology Services`	-6.055e+01
## `X11Public Administration`	-2.603e+01
## X11Services	-3.178e+01
## `X11Technology: IT & Telephone Services`	-4.320e+01
## X11Transportation	-5.029e+00
## `X11Transportation & Logistics`	-3.526e+01
## X11Veterinary	-5.393e+01
## `X12Government: Australia: Federal`	1.292e+01
## `X12Government: Australia: Local & State`	-8.666e+00
## `X12Government: International`	4.360e+00
## `X12Large Australian Corporate`	2.562e+01
## `X12Multinational / Other Large Corporate`	2.583e+00
## `X12Not for profit`	2.860e+01
## `X12PFR0 (Publicly-Funded Research Organisation)`	2.609e+01
## `X12RDC (Rural Research and Development Corporations)`	2.763e+01
## `X12SME (small to medium enterprise)`	2.059e+01
## X12University	-1.079e+01
## X13Internal	-5.285e+01
## X13missing	-2.352e+01

## `factor(X14)1`	1.093e+00	
## `X15Competitive BidBusiness Organization`	NA	
## `X15Consultancy (Non-research)Administrative`	NA	
## `X15Consultancy (Non-research)Business Organization`	2.283e+08	
## `X15Consultancy (Non-research)University Department`	2.283e+08	
## `X15Custom Education (MSPACE Included)Business Organization`	3.449e-01	
## `X15Custom Education (MSPACE Included)University Department`	NA	
## `X15GrantsBusiness Organization`	-2.892e+00	
## `X15GrantsUniversity Department`	NA	
## `X15InternshipBusiness Organization`	6.243e-02	
## `X15InternshipUniversity Department`	NA	
## `X15PanelBusiness Organization`	NA	
## `X15Parent GrantBusiness Organization`	NA	
## `X15PhilanthropicBusiness Organization`	NA	
## `X15Research ContractBusiness Organization`	-2.280e+00	
## `X15Research ContractUniversity Department`	NA	
## `X15SponsorshipBusiness Organization`	NA	
## `X15Technology TransferBusiness Organization`	NA	
## `X15Technology TransferUniversity Department`	NA	
##	Std. Error	z value
## (Intercept)	4.019e+16	0.000
## `X3Academic No Longer Interested`	5.629e+04	0.000
## `X3Awaiting Customer Action`	2.802e+05	0.000
## `X3BD Capacity (Lost)`	6.636e+04	0.000
## `X3Customer No Longer Interested`	2.667e+04	0.000
## `X3Customer Not Responding`	4.162e+04	0.000
## `X3Failure to Agree - IP terms`	1.513e+05	0.000
## `X3Failure to Agree - Price`	1.025e+05	0.000
## `X3Failure to Agree - Scope of Work`	8.558e+04	0.000
## `X3Not Ready for Market`	5.991e+04	0.000
## `X3On Track`	2.346e+04	0.002
## `X3Other (Lost)`	2.358e+04	0.000
## `X3Other IP Issue`	1.265e+05	0.000
## `X3Other(Lost)`	1.064e+05	0.000
## `X3Out-Sold`	3.119e+04	0.000
## `X3Proposal/Application Unsuccessful`	2.554e+04	0.000
## `X3Scope of Proposal`	3.568e+05	0.000
## `X3Unacceptable Risk`	1.186e+05	0.000
## X3Won	2.435e+04	0.005
## `X4Competitive Bid`	1.441e+05	0.000
## `X4Consultancy (Non-research)`	4.019e+16	0.000
## `X4Custom Education (MSPACE Included)`	4.679e+05	0.000
## X4Grants	5.301e+05	0.000
## X4Internship	5.807e+05	0.000
## X4Panel	1.642e+05	0.000

## `X4Parent Grant`	1.284e+05	0.000
## X4Philanthropic	1.792e+05	0.000
## `X4Research Contract`	4.750e+05	0.000
## X4Sponsorship	1.640e+05	0.000
## `X4Technology Transfer`	1.248e+05	0.000
## `X5RIC RE&D and BD&I`	2.842e+04	0.000
## `X5RIC-BD&I`	2.563e+04	0.000
## `X5RIC-RE&D`	3.222e+04	0.000
## X6	7.385e+03	0.000
## X80012e000002ZFZYAA4	4.856e+05	0.000
## X80012e000002ZGfbAAG	2.971e+04	0.001
## X80012e000002ZLg7AAG	5.153e+04	0.000
## X80012e000002ZLgIAAW	4.055e+04	0.000
## X80012e000002ZLgJAAW	5.120e+04	0.000
## X80012e000002ZmnxAAC	3.147e+04	0.001
## X80012e000002ZmnzAAC	1.193e+05	0.000
## X80012e000002Zmo0AAC	5.488e+04	0.000
## X80012e000002Zm0UAA0	2.601e+05	0.000
## X80012e000002Zm0ZAA0	2.690e+04	0.001
## X80012e000002ZNY0AA4	1.948e+04	0.001
## X80012e000002Zt0MAAS	2.112e+05	0.000
## X80012e000002Zt1iAAC	3.632e+05	0.000
## X80012e000002Zt1RAAS	6.752e+04	0.000
## X80012e000002Zt1zAAC	3.812e+05	0.000
## `factor(X9)1`	3.584e+04	0.000
## `X10Business Organization`	4.019e+16	0.000
## `X10University Department`	4.019e+16	0.000
## `X11Agriculture & Horticulture`	6.578e+04	0.000
## `X11Aquaculture & Fisheries`	2.466e+05	0.000
## `X11Construction & Construction Materials`	8.612e+04	0.000
## X11Defence	9.296e+04	-0.001
## X11Education	NA	NA
## `X11Energy & Utilities`	6.724e+04	0.000
## `X11Forestry & Paper`	2.619e+05	0.000
## X11Health	3.434e+04	-0.001
## `X11Manufacturing & R&D`	6.030e+04	-0.001
## `X11Materials: Chemicals & Industrial Metals`	9.737e+04	0.000
## `X11Mining & Refining`	9.709e+04	0.000
## `X11Mining Equipment and Technology Services`	1.597e+05	0.000
## `X11Public Administration`	5.292e+04	0.000
## X11Services	2.147e+04	-0.001
## `X11Technology: IT & Telephone Services`	4.126e+04	-0.001
## X11Transportation	3.684e+05	0.000
## `X11Transportation & Logistics`	5.750e+04	-0.001
## X11Veterinary	1.255e+05	0.000

## `X12Government: Australia: Federal`	2.384e+05	0.000
## `X12Government: Australia: Local & State`	2.392e+05	0.000
## `X12Government: International`	2.786e+05	0.000
## `X12Large Australian Corporate`	2.401e+05	0.000
## `X12Multinational / Other Large Corporate`	2.427e+05	0.000
## `X12Not for profit`	2.390e+05	0.000
## `X12PFR0 (Publicly-Funded Research Organisation)`	2.506e+05	0.000
## `X12RDC (Rural Research and Development Corporations)`	2.606e+05	0.000
## `X12SME (small to medium enterprise)`	2.392e+05	0.000
## X12University	2.400e+05	0.000
## X13Internal	3.575e+05	0.000
## X13missing	3.378e+04	-0.001
## `factor(X14)1`	2.881e+04	0.000
## `X15Competitive BidBusiness Organization`	NA	NA
## `X15Consultancy (Non-research)Administrative`	NA	NA
## `X15Consultancy (Non-research)Business Organization`	4.019e+16	0.000
## `X15Consultancy (Non-research)University Department`	4.019e+16	0.000
## `X15Custom Education (MSPACE Included)Business Organization`	4.521e+05	0.000
## `X15Custom Education (MSPACE Included)University Department`	NA	NA
## `X15GrantsBusiness Organization`	5.163e+05	0.000
## `X15GrantsUniversity Department`	NA	NA
## `X15InternshipBusiness Organization`	5.779e+05	0.000
## `X15InternshipUniversity Department`	NA	NA
## `X15PanelBusiness Organization`	NA	NA
## `X15Parent GrantBusiness Organization`	NA	NA
## `X15PhilanthropicBusiness Organization`	NA	NA
## `X15Research ContractBusiness Organization`	4.595e+05	0.000
## `X15Research ContractUniversity Department`	NA	NA
## `X15SponsorshipBusiness Organization`	NA	NA
## `X15Technology TransferBusiness Organization`	NA	NA
## `X15Technology TransferUniversity Department`	NA	NA
##	Pr(> z)	
## (Intercept)	1.000	
## `X3Academic No Longer Interested`	1.000	
## `X3Awaiting Customer Action`	1.000	
## `X3BD Capacity (Lost)`	1.000	
## `X3Customer No Longer Interested`	1.000	
## `X3Customer Not Responding`	1.000	
## `X3Failure to Agree - IP terms`	1.000	
## `X3Failure to Agree - Price`	1.000	
## `X3Failure to Agree - Scope of Work`	1.000	
## `X3Not Ready for Market`	1.000	
## `X3On Track`	0.998	
## `X3Other (Lost)`	1.000	
## `X3Other IP Issue`	1.000	

## `X3Other(Lost)`	1.000
## `X3Out-Sold`	1.000
## `X3Proposal/Application Unsuccessful`	1.000
## `X3Scope of Proposal`	1.000
## `X3Unacceptable Risk`	1.000
## X3Won	0.996
## `X4Competitive Bid`	1.000
## `X4Consultancy (Non-research)`	1.000
## `X4Custom Education (MSPACE Included)`	1.000
## X4Grants	1.000
## X4Internship	1.000
## X4Panel	1.000
## `X4Parent Grant`	1.000
## X4Philanthropic	1.000
## `X4Research Contract`	1.000
## X4Sponsorship	1.000
## `X4Technology Transfer`	1.000
## `X5RIC RE&D and BD&I`	1.000
## `X5RIC-BD&I`	1.000
## `X5RIC-RE&D`	1.000
## X6	1.000
## X80012e000002ZFZYAA4	1.000
## X80012e000002ZGfbAAG	0.999
## X80012e000002ZLg7AAG	1.000
## X80012e000002ZLgIAAW	1.000
## X80012e000002ZLgJAAW	1.000
## X80012e000002ZmnxAAC	0.999
## X80012e000002ZmnzAAC	1.000
## X80012e000002Zmo0AAC	1.000
## X80012e000002ZmOUAA0	1.000
## X80012e000002ZmOZAA0	1.000
## X80012e000002ZNY0AA4	0.999
## X80012e000002Zt0MAAS	1.000
## X80012e000002Zt1iAAC	1.000
## X80012e000002Zt1RAAS	1.000
## X80012e000002Zt1zAAC	1.000
## `factor(X9)1`	1.000
## `X10Business Organization`	1.000
## `X10University Department`	1.000
## `X11Agriculture & Horticulture`	1.000
## `X11Aquaculture & Fisheries`	1.000
## `X11Construction & Construction Materials`	1.000
## X11Defence	1.000
## X11Education	NA
## `X11Energy & Utilities`	1.000

## `X11Forestry & Paper`	1.000
## X11Health	0.999
## `X11Manufacturing & R&D`	0.999
## `X11Materials: Chemicals & Industrial Metals`	1.000
## `X11Mining & Refining`	1.000
## `X11Mining Equipment and Technology Services`	1.000
## `X11Public Administration`	1.000
## X11Services	0.999
## `X11Technology: IT & Telephone Services`	0.999
## X11Transportation	1.000
## `X11Transportation & Logistics`	1.000
## X11Veterinary	1.000
## `X12Government: Australia: Federal`	1.000
## `X12Government: Australia: Local & State`	1.000
## `X12Government: International`	1.000
## `X12Large Australian Corporate`	1.000
## `X12Multinational / Other Large Corporate`	1.000
## `X12Not for profit`	1.000
## `X12PFR0 (Publicly-Funded Research Organisation)`	1.000
## `X12RDC (Rural Research and Development Corporations)`	1.000
## `X12SME (small to medium enterprise)`	1.000
## X12University	1.000
## X13Internal	1.000
## X13missing	0.999
## `factor(X14)1`	1.000
## `X15Competitive BidBusiness Organization`	NA
## `X15Consultancy (Non-research)Administrative`	NA
## `X15Consultancy (Non-research)Business Organization`	1.000
## `X15Consultancy (Non-research)University Department`	1.000
## `X15Custom Education (MSPACE Included)Business Organization`	1.000
## `X15Custom Education (MSPACE Included)University Department`	NA
## `X15GrantsBusiness Organization`	1.000
## `X15GrantsUniversity Department`	NA
## `X15InternshipBusiness Organization`	1.000
## `X15InternshipUniversity Department`	NA
## `X15PanelBusiness Organization`	NA
## `X15Parent GrantBusiness Organization`	NA
## `X15PhilanthropicBusiness Organization`	NA
## `X15Research ContractBusiness Organization`	1.000
## `X15Research ContractUniversity Department`	NA
## `X15SponsorshipBusiness Organization`	NA
## `X15Technology TransferBusiness Organization`	NA
## `X15Technology TransferUniversity Department`	NA
##	
## (Dispersion parameter for binomial family taken to be 1)	

```
##
##      Null deviance: 7.7502e+03   on 6051   degrees of freedom
## Residual deviance: 1.8979e-07   on 5964   degrees of freedom
## AIC: 176
##
## Number of Fisher Scoring iterations: 25
```

Conclusion:

No attributes are significant (p-value < 0.05)

Next step:

Remove attribute X3: "Status_Reason__c"

```
mod2 <- glm(factor(X2)~X4+X5+X6+X8+factor(X9)+X10+X11+X12+X13+factor(X14)+X15,
             family = "binomial", data = data)

# 10-fold cross_validation
train_control <- trainControl(method = "cv", number = 10)

# train the model on data set
model <- train(factor(X2)~X4+X5+X6+X8+factor(X9)+X10+X11+X12+X13+factor(X14)+X15,
               data = data,
               trControl = train_control,
               method = "glm",
               family="binomial")

summary(model)
```

```
##
## Call:
## NULL
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6514  -0.8524  -0.2888   0.9322   3.5448
##
## Coefficients: (11 not defined because of singularities)
##
##              Estimate
## (Intercept)    -3.083e+13
## `X4Competitive Bid`    -3.936e+00
## `X4Consultancy (Non-research)`    -1.906e+00
## `X4Custom Education (MSPACE Included)`    2.177e+11
## X4Grants    -1.803e+00
## X4Internship    -5.796e+13
```

## X4Panel	-1.839e+00
## `X4Parent Grant`	5.249e+00
## X4Philanthropic	-1.730e+00
## `X4Research Contract`	-1.637e+00
## X4Sponsorship	-1.891e+00
## `X4Technology Transfer`	-1.950e+00
## `X5RIC RE&D and BD&I`	2.066e+00
## `X5RIC-BD&I`	1.222e+00
## `X5RIC-RE&D`	7.120e-01
## X6	4.168e-01
## X80012e000002ZFZYAA4	3.817e+00
## X80012e000002ZGfbAAG	3.755e+00
## X80012e000002ZLg7AAG	3.501e+00
## X80012e000002ZLgIAAW	2.348e+00
## X80012e000002ZLgJAAW	3.254e+00
## X80012e000002ZmnxAAC	2.426e+00
## X80012e000002ZmnzAAC	3.027e+00
## X80012e000002Zmo0AAC	4.021e+00
## X80012e000002ZmOUAA0	3.624e+00
## X80012e000002ZmOZAA0	3.498e+00
## X80012e000002ZNYOAA4	3.132e+00
## X80012e000002Zt0MAAS	-3.341e+01
## X80012e000002Zt1iAAC	-1.451e+01
## X80012e000002Zt1RAAS	3.359e+00
## X80012e000002Zt1zAAC	-5.056e-01
## `factor(X9)1`	-4.230e-01
## `X10Business Organization`	3.083e+13
## `X10University Department`	3.083e+13
## `X11Agriculture & Horticulture`	-1.570e+00
## `X11Aquaculture & Fisheries`	-1.094e+00
## `X11Construction & Construction Materials`	-1.197e+00
## X11Defence	-6.141e-01
## X11Education	NA
## `X11Energy & Utilities`	-1.161e+00
## `X11Forestry & Paper`	-7.524e-02
## X11Health	-1.365e+00
## `X11Manufacturing & R&D`	-1.796e+00
## `X11Materials: Chemicals & Industrial Metals`	-1.930e+00
## `X11Mining & Refining`	-1.677e+00
## `X11Mining Equipment and Technology Services`	-2.370e+00
## `X11Public Administration`	-1.537e+00
## X11Services	-1.491e+00
## `X11Technology: IT & Telephone Services`	-1.623e+00
## X11Transportation	-2.767e+01
## `X11Transportation & Logistics`	-1.775e+00

## X11Veterinary	-7.881e-01	
## `X12Government: Australia: Federal`	-3.580e-01	
## `X12Government: Australia: Local & State`	-3.059e-01	
## `X12Government: International`	-6.537e-01	
## `X12Large Australian Corporate`	-6.422e-01	
## `X12Multinational / Other Large Corporate`	-1.118e+00	
## `X12Not for profit`	-3.328e-01	
## `X12PFR0 (Publicly-Funded Research Organisation)`	-1.709e+00	
## `X12RDC (Rural Research and Development Corporations)`	-6.720e-01	
## `X12SME (small to medium enterprise)`	-6.588e-01	
## X12University	5.674e-01	
## X13Internal	2.600e+01	
## X13missing	-2.498e+00	
## `factor(X14)1`	-1.831e-01	
## `X15Competitive BidBusiness Organization`	NA	
## `X15Consultancy (Non-research)Administrative`	3.083e+13	
## `X15Consultancy (Non-research)Business Organization`	-1.024e-01	
## `X15Consultancy (Non-research)University Department`	NA	
## `X15Custom Education (MSPACE Included)Business Organization`	-2.177e+11	
## `X15Custom Education (MSPACE Included)University Department`	-2.177e+11	
## `X15GrantsBusiness Organization`	-9.126e-01	
## `X15GrantsUniversity Department`	NA	
## `X15InternshipBusiness Organization`	5.796e+13	
## `X15InternshipUniversity Department`	5.796e+13	
## `X15PanelBusiness Organization`	NA	
## `X15Parent GrantBusiness Organization`	NA	
## `X15PhilanthropicBusiness Organization`	NA	
## `X15Research ContractBusiness Organization`	-6.506e-01	
## `X15Research ContractUniversity Department`	NA	
## `X15SponsorshipBusiness Organization`	NA	
## `X15Technology TransferBusiness Organization`	NA	
## `X15Technology TransferUniversity Department`	NA	
##	Std. Error	z value
## (Intercept)	5.356e+13	-0.576
## `X4Competitive Bid`	8.180e-01	-4.812
## `X4Consultancy (Non-research)`	1.698e+00	-1.122
## `X4Custom Education (MSPACE Included)`	1.902e+13	0.011
## X4Grants	1.681e+00	-1.073
## X4Internship	1.663e+14	-0.348
## X4Panel	7.453e-01	-2.467
## `X4Parent Grant`	1.229e+00	4.273
## X4Philanthropic	8.064e-01	-2.145
## `X4Research Contract`	1.383e+00	-1.184
## X4Sponsorship	8.096e-01	-2.335
## `X4Technology Transfer`	7.032e-01	-2.774

## `X5RIC RE&D and BD&I`	2.071e-01	9.974
## `X5RIC-BD&I`	1.954e-01	6.256
## `X5RIC-RE&D`	2.633e-01	2.704
## X6	3.954e-02	10.541
## X80012e000002ZFZYAA4	7.567e-01	5.045
## X80012e000002ZGfbAAG	2.575e-01	14.582
## X80012e000002ZLg7AAG	2.568e-01	13.633
## X80012e000002ZLgIAAW	3.107e-01	7.555
## X80012e000002ZLgJAAW	2.786e-01	11.678
## X80012e000002ZmnxAAC	2.645e-01	9.171
## X80012e000002ZmnzAAC	2.915e-01	10.385
## X80012e000002Zmo0AAC	2.580e-01	15.588
## X80012e000002ZmOUAA0	9.831e-01	3.686
## X80012e000002ZmOZAA0	2.880e-01	12.145
## X80012e000002ZNYOAA4	2.783e-01	11.253
## X80012e000002Zt0MAAS	4.680e+05	0.000
## X80012e000002Zt1iAAC	1.558e+04	-0.001
## X80012e000002Zt1RAAS	3.117e-01	10.776
## X80012e000002Zt1zAAC	8.976e+00	-0.056
## `factor(X9)1`	1.515e-01	-2.792
## `X10Business Organization`	5.356e+13	0.576
## `X10University Department`	5.356e+13	0.576
## `X11Agriculture & Horticulture`	1.675e+00	-0.938
## `X11Aquaculture & Fisheries`	2.063e+00	-0.530
## `X11Construction & Construction Materials`	1.678e+00	-0.713
## X11Defence	1.670e+00	-0.368
## X11Education	NA	NA
## `X11Energy & Utilities`	1.673e+00	-0.694
## `X11Forestry & Paper`	2.043e+00	-0.037
## X11Health	1.664e+00	-0.820
## `X11Manufacturing & R&D`	1.676e+00	-1.071
## `X11Materials: Chemicals & Industrial Metals`	1.713e+00	-1.126
## `X11Mining & Refining`	1.695e+00	-0.990
## `X11Mining Equipment and Technology Services`	1.842e+00	-1.287
## `X11Public Administration`	1.667e+00	-0.922
## X11Services	1.664e+00	-0.896
## `X11Technology: IT & Telephone Services`	1.674e+00	-0.969
## X11Transportation	3.191e+05	0.000
## `X11Transportation & Logistics`	1.689e+00	-1.051
## X11Veterinary	1.711e+00	-0.461
## `X12Government: Australia: Federal`	4.212e-01	-0.850
## `X12Government: Australia: Local & State`	4.186e-01	-0.731
## `X12Government: International`	4.489e-01	-1.456
## `X12Large Australian Corporate`	4.288e-01	-1.498
## `X12Multinational / Other Large Corporate`	4.227e-01	-2.646

## `X12Not for profit`	4.220e-01	-0.789
## `X12PFR0 (Publicly-Funded Research Organisation)`	5.413e-01	-3.157
## `X12RDC (Rural Research and Development Corporations)`	5.946e-01	-1.130
## `X12SME (small to medium enterprise)`	4.194e-01	-1.571
## X12University	4.364e-01	1.300
## X13Internal	3.945e+05	0.000
## X13missing	2.186e-01	-11.425
## `factor(X14)1`	1.000e-01	-1.831
## `X15Competitive BidBusiness Organization`	NA	NA
## `X15Consultancy (Non-research)Administrative`	5.356e+13	0.576
## `X15Consultancy (Non-research)Business Organization`	1.554e+00	-0.066
## `X15Consultancy (Non-research)University Department`	NA	NA
## `X15Custom Education (MSPACE Included)Business Organization`	1.902e+13	-0.011
## `X15Custom Education (MSPACE Included)University Department`	1.902e+13	-0.011
## `X15GrantsBusiness Organization`	1.544e+00	-0.591
## `X15GrantsUniversity Department`	NA	NA
## `X15InternshipBusiness Organization`	1.663e+14	0.348
## `X15InternshipUniversity Department`	1.663e+14	0.348
## `X15PanelBusiness Organization`	NA	NA
## `X15Parent GrantBusiness Organization`	NA	NA
## `X15PhilanthropicBusiness Organization`	NA	NA
## `X15Research ContractBusiness Organization`	1.200e+00	-0.542
## `X15Research ContractUniversity Department`	NA	NA
## `X15SponsorshipBusiness Organization`	NA	NA
## `X15Technology TransferBusiness Organization`	NA	NA
## `X15Technology TransferUniversity Department`	NA	NA
##	Pr(> z)	
## (Intercept)	0.564908	
## `X4Competitive Bid`	1.50e-06	***
## `X4Consultancy (Non-research)`	0.261824	
## `X4Custom Education (MSPACE Included)`	0.990867	
## X4Grants	0.283271	
## X4Internship	0.727478	
## X4Panel	0.013620	*
## `X4Parent Grant`	1.93e-05	***
## X4Philanthropic	0.031913	*
## `X4Research Contract`	0.236426	
## X4Sponsorship	0.019531	*
## `X4Technology Transfer`	0.005545	**
## `X5RIC RE&D and BD&I`	< 2e-16	***
## `X5RIC-BD&I`	3.94e-10	***
## `X5RIC-RE&D`	0.006850	**
## X6	< 2e-16	***
## X80012e000002ZFZYAA4	4.55e-07	***
## X80012e000002ZGfbAAG	< 2e-16	***

## X80012e000002ZLg7AAG	< 2e-16 ***
## X80012e000002ZLgIAAW	4.18e-14 ***
## X80012e000002ZLgJAAW	< 2e-16 ***
## X80012e000002ZmnxAAC	< 2e-16 ***
## X80012e000002ZmnzAAC	< 2e-16 ***
## X80012e000002Zmo0AAC	< 2e-16 ***
## X80012e000002ZmOUAAO	0.000228 ***
## X80012e000002ZmOZAAO	< 2e-16 ***
## X80012e000002ZNY0AA4	< 2e-16 ***
## X80012e000002Zt0MAAS	0.999943
## X80012e000002Zt1iAAC	0.999257
## X80012e000002Zt1RAAS	< 2e-16 ***
## X80012e000002Zt1zAAC	0.955081
## `factor(X9)1`	0.005236 **
## `X10Business Organization`	0.564908
## `X10University Department`	0.564908
## `X11Agriculture & Horticulture`	0.348467
## `X11Aquaculture & Fisheries`	0.595775
## `X11Construction & Construction Materials`	0.475842
## X11Defence	0.712977
## X11Education	NA
## `X11Energy & Utilities`	0.487695
## `X11Forestry & Paper`	0.970620
## X11Health	0.411985
## `X11Manufacturing & R&D`	0.284006
## `X11Materials: Chemicals & Industrial Metals`	0.259988
## `X11Mining & Refining`	0.322313
## `X11Mining Equipment and Technology Services`	0.198056
## `X11Public Administration`	0.356643
## X11Services	0.370060
## `X11Technology: IT & Telephone Services`	0.332423
## X11Transportation	0.999931
## `X11Transportation & Logistics`	0.293175
## X11Veterinary	0.645125
## `X12Government: Australia: Federal`	0.395395
## `X12Government: Australia: Local & State`	0.464845
## `X12Government: International`	0.145367
## `X12Large Australian Corporate`	0.134203
## `X12Multinational / Other Large Corporate`	0.008156 **
## `X12Not for profit`	0.430334
## `X12PFR0 (Publicly-Funded Research Organisation)`	0.001596 **
## `X12RDC (Rural Research and Development Corporations)`	0.258450
## `X12SME (small to medium enterprise)`	0.116185
## X12University	0.193590
## X13Internal	0.999947


```

## X13missing < 2e-16 ***
## `factor(X14)1` 0.067108 .
## `X15Competitive BidBusiness Organization` NA
## `X15Consultancy (Non-research)Administrative` 0.564908
## `X15Consultancy (Non-research)Business Organization` 0.947471
## `X15Consultancy (Non-research)University Department` NA
## `X15Custom Education (MSPACE Included)Business Organization` 0.990867
## `X15Custom Education (MSPACE Included)University Department` 0.990867
## `X15GrantsBusiness Organization` 0.554601
## `X15GrantsUniversity Department` NA
## `X15InternshipBusiness Organization` 0.727478
## `X15InternshipUniversity Department` 0.727478
## `X15PanelBusiness Organization` NA
## `X15Parent GrantBusiness Organization` NA
## `X15PhilanthropicBusiness Organization` NA
## `X15Research ContractBusiness Organization` 0.587853
## `X15Research ContractUniversity Department` NA
## `X15SponsorshipBusiness Organization` NA
## `X15Technology TransferBusiness Organization` NA
## `X15Technology TransferUniversity Department` NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 7750.2 on 6051 degrees of freedom
## Residual deviance: 5904.9 on 5980 degrees of freedom
## AIC: 6048.9
##
## Number of Fisher Scoring iterations: 25

```

Conclusion:

Significant Attributes:

Very significant (p-value < 0.001):

X4Competitive Bid
 X4Parent Grant
 X5RIC RE&D and BD&I
 X5RIC-BD&I
 X6
 X80012e000002ZFZYAA4
 X80012e000002ZGfbAAG
 X80012e000002ZLg7AAG
 X80012e000002ZLgIAAW

X80012e000002ZLgJAAW
X80012e000002ZmnxAAC
X80012e000002ZmnzAAC
X80012e000002Zmo0AAC
X80012e000002ZmOUAA0
X80012e000002ZmOZAA0
X80012e000002ZNYOAA4 X80012e000002Zt1RAAS X13missing

Significant ($0,001 < \text{p-value} < 0.01$)

X4Technology Transfer
X5RIC-RE&D
factor(X9)1
X12Multinational / Other Large Corporate
X12PFRO (Publicly-Funded Research Organisation)

Somehow significant ($0.01 < \text{p-value} < 0.05$)

X4Panel
X4Philanthropic
X4Sponsorship

$0.05 < \text{p-value} < 0.1$ (keep or not?)

factor(X14)1

```
print(model)
```

```
## Generalized Linear Model
##
## 6052 samples
##    11 predictor
##    2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 5447, 5447, 5447, 5447, 5446, 5447, ...
## Resampling results:
##
##   Accuracy   Kappa
##  0.7219205  0.3716663
```

from website <https://machinelearningmastery.com/machine-learning-evaluation-metrics-in-r/>

Accuracy is the percentage of correctly classifies instances out of all instances. It is more useful on a binary classification than multi-class classification problems because it can be less

clear exactly how the accuracy breaks down across those classes.

Kappa or Cohen's Kappa is like classification accuracy, except that it is normalized at the baseline of random chance on your dataset. It is a more useful measure to use on problems that have an imbalance in the classes (e.g. 70-30 split for classes 0 and 1 and you can achieve 70% accuracy by predicting all instances are for class 0).

Suggested guidelines for Kappa:

values ≤ 0 as indicating no agreement

0.01–0.20 as none to slight

0.21–0.40 as fair, 0.41– 0.60 as moderate,

0.61–0.80 as substantial

0.81–1.00 as almost perfect agreement.