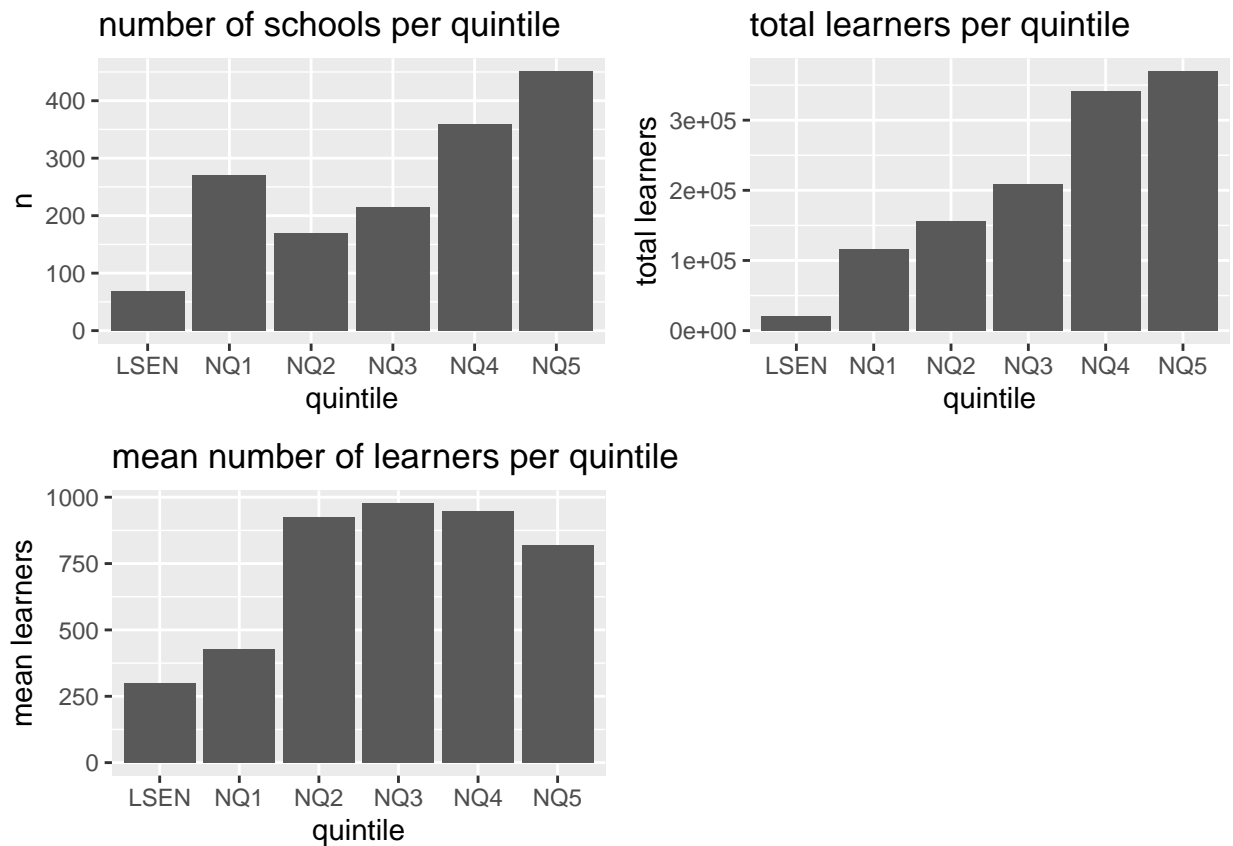# WCED Data Analysis

Gareth Edwards
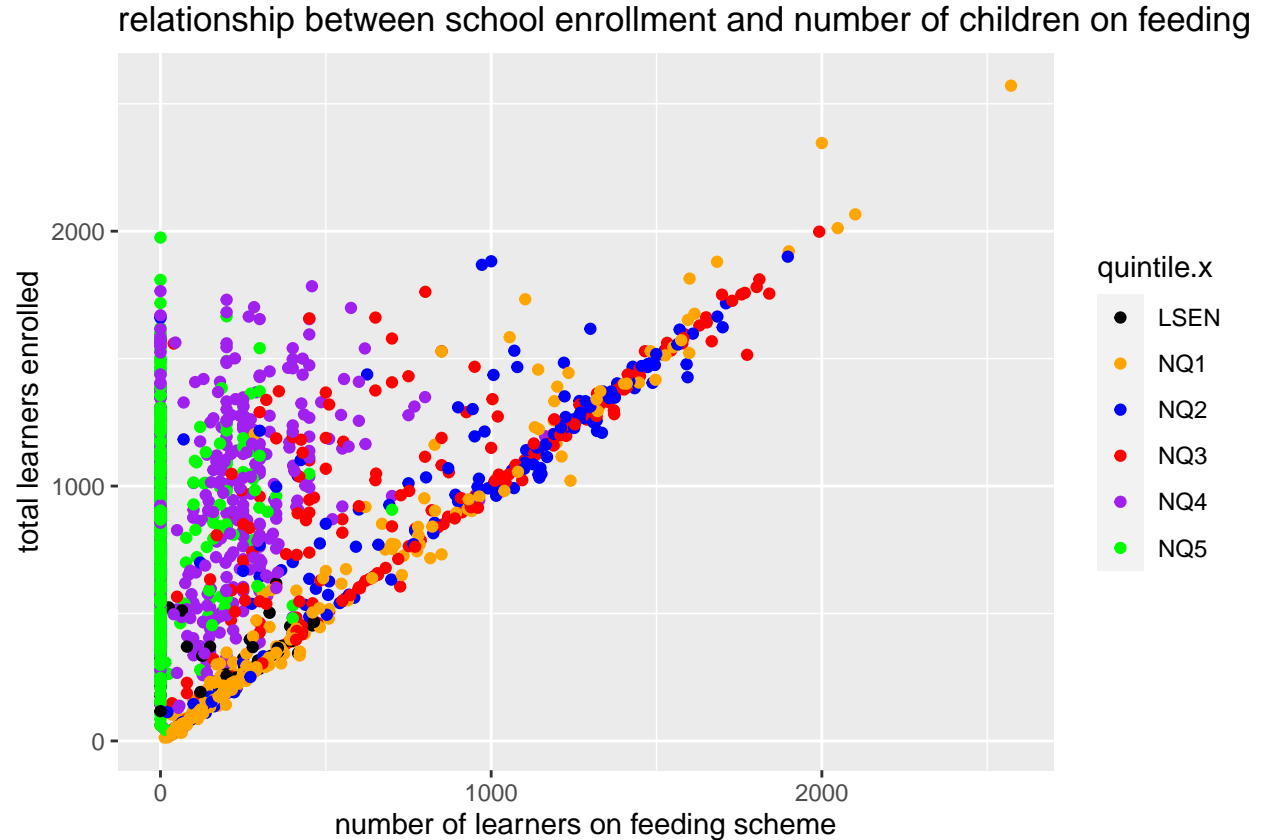
WCED School Data Analysis

Bar Plots

The barplots show there are more quintile 5 schools compared to the rest. There are also more students, in total, going to NQ5 schools. However, we see average school sizes for NQ2, NQ3 and NQ4 are higher than average school sizes for NQ5.

This shows that there are an abundance of NQ5 schools so learners can be more evenly dispersed, but there aren't as much NQ2, 3 and 4 schools so learners who have to attend these schools don't have as many options which lead to a potentially crowded school in these areas. This could imply that if schools are to be built then the WCED must focus on NQ2, 3 and 4 schools (especially NQ2 and NQ3 schools).

On the other hand we see there are more NQ1 schools than NQ2 and NQ3 schools and average class sizes for NQ1 schools are very low compared to the other quintile levels. This suggests an over abundance of NQ1 schools. There could be less NQ1 schools and that funding can go into developing more NQ2 and 3 schools.

Scatter Plot

relationship between school enrollment and number of children on feeding

The scatter plot shows that quintile 1 schools have to cater to almost all students. As the number of students in a NQ1 school increases so does the number of students that need a feeding scheme increase (direct linear relationship). The same can be said for NQ2 and NQ3 schools, but they don't follow this trend as strictly.

If a Q1 and Q2 school is being built there must be significant focus on infrastructure to support a feeding scheme. Perhaps one could suggest focusing on feeding schemes at NQ3 schools if we are suggesting more NQ3 schools need to be built.

NQ4 and NQ5 show no real trend when it comes to total learners enrolled and number of learners on a feeding scheme. The plot shows that you can have large numbers of learners in a NQ4 and NQ5 school but have only a few learners on feeding schemes, if any at all.

Multiple Linear Regression Model

```
##
## Call:
## lm(formula = total.learner.enrolled ~ ., data = merged_subset)
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -1.62980 -0.32264 -0.01654  0.28598  2.69837
##
## Coefficients: (2 not defined because of singularities)
##                                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)                            -0.789169   0.674998  -1.169   0.2425
## education.districtEDEN AND CENTRAL KAROO 0.578290   0.702375   0.823   0.4105
## education.districtMETRO CENTRAL         0.316973   0.397990   0.796   0.4259
## education.districtMETRO EAST            0.339782   0.392131   0.867   0.3864
## education.districtMETRO NORTH           0.514160   0.393162   1.308   0.1912
## education.districtMETRO SOUTH           0.589219   0.397980   1.481   0.1389
## education.districtOVERBERG              0.110511   0.238931   0.463   0.6438
## education.districtWEST COAST            0.308741   0.404649   0.763   0.4456
## correspondence.language.xENGLISH        0.176951   0.045320   3.904 9.87e-05
## institution.type.xIntermediate School   0.001729   0.119944   0.014   0.9885
## institution.type.xPrimary School        0.015466   0.101160   0.153   0.8785
## institution.type.xSchool of Skills     -1.042483   0.152867  -6.820 1.33e-11
## institution.type.xSecondary School      0.427684   0.103075   4.149 3.53e-05
## institution.type.xSpecial School       -1.211365   0.127769  -9.481  < 2e-16
## fee.status.xNo Fee                      0.015399   0.052053   0.296   0.7674
## quintile.xNQ1                          -0.965859   0.075258 -12.834  < 2e-16
## quintile.xNQ2                          -0.776490   0.078738  -9.862  < 2e-16
## quintile.xNQ3                          -0.746272   0.073927 -10.095  < 2e-16
## quintile.xNQ4                           0.086910   0.048653   1.786   0.0743
## quintile.xNQ5                                 NA         NA      NA       NA
## magisterial.district.xBEAUFORT WEST    -0.208734   0.260707  -0.801   0.4235
## magisterial.district.xBELLVILLE         0.341243   0.543585   0.628   0.5303
## magisterial.district.xBREDASDORP        0.334741   0.725274   0.462   0.6445
## magisterial.district.xCALEDON           0.313966   0.715462   0.439   0.6608
## magisterial.district.xCALITZDORP       -0.311218   0.344803  -0.903   0.3669
## magisterial.district.xCAPE             -0.113414   0.546903  -0.207   0.8357
## magisterial.district.xCERES             0.361386   0.671710   0.538   0.5907
## magisterial.district.xCLANWILLIAM       0.091046   0.573473   0.159   0.8739
## magisterial.district.xGEORGE            0.280269   0.235366   1.191   0.2339
## magisterial.district.xGOODWOOD         -0.188313   0.544585  -0.346   0.7295
## magisterial.district.xHEIDELBERG (WC)  -0.543398   0.324386  -1.675   0.0941
## magisterial.district.xHERMANUS          0.564835   0.723154   0.781   0.4349
## magisterial.district.xHOPEFIELD        -0.104496   0.603517  -0.173   0.8626
## magisterial.district.xKNYSNA            0.006312   0.243255   0.026   0.9793
## magisterial.district.xKUILS RIVER       0.244226   0.765737   0.319   0.7498
## magisterial.district.xKUILSRIVIER       0.683039   0.546340   1.250   0.2114
## magisterial.district.xLADISMITH        -0.200914   0.276666  -0.726   0.4678
## magisterial.district.xLAINGSBURG       -0.254881   0.346490  -0.736   0.4621
## magisterial.district.xMALMESBURY        0.259145   0.552440   0.469   0.6391
## magisterial.district.xMITCHELL'S PLAIN  0.484931   0.622255   0.779   0.4359
## magisterial.district.xMITCHELLS PLAIN   0.210049   0.547965   0.383   0.7015
## magisterial.district.xMONTAGU           0.411772   0.677365   0.608   0.5433
```

```
## magisterial.district.xMOORREESBURG          0.023796   0.597208    0.040   0.9682
## magisterial.district.xMOSSEL BAY            0.120817   0.246392    0.490   0.6240
## magisterial.district.xMURRAYSBURG           0.020067   0.438769    0.046   0.9635
## magisterial.district.xOUDTSHOORN           -0.022869   0.239302   -0.096   0.9239
## magisterial.district.xPAARL                 0.573911   0.671194    0.855   0.3927
## magisterial.district.xPIKETBERG             0.067403   0.576123    0.117   0.9069
## magisterial.district.xPRINCE ALBERT        -0.177335   0.346666   -0.512   0.6090
## magisterial.district.xRIVERSDALE           -0.212717   0.266447   -0.798   0.4248
## magisterial.district.xROBERTSON             0.374169   0.677083    0.553   0.5806
## magisterial.district.xSIMONS TOWN          -0.140053   0.558744   -0.251   0.8021
## magisterial.district.xSOMERSET WEST         0.776016   0.560367    1.385   0.1663
## magisterial.district.xSTELLENBOSCH          0.527059   0.675472    0.780   0.4354
## magisterial.district.xSTRAND                0.449198   0.556264    0.808   0.4195
## magisterial.district.xSWELLENDAM            0.219331   0.697948    0.314   0.7534
## magisterial.district.xTULBAGH               0.261030   0.680350    0.384   0.7013
## magisterial.district.xUNIONDALE                   NA         NA       NA       NA
## magisterial.district.xVANRHYNSDORP          0.016812   0.590229    0.028   0.9773
## magisterial.district.xVREDENBURG            0.495185   0.572571    0.865   0.3873
## magisterial.district.xVREDENDAL             0.129340   0.577378    0.224   0.8228
## magisterial.district.xWELLINGTON            0.728507   0.679587    1.072   0.2839
## magisterial.district.xWORCESTER             0.460188   0.670469    0.686   0.4926
## magisterial.district.xWYNBERG              -0.166237   0.547426   -0.304   0.7614
## bus.route.learners                          0.114475   0.015519    7.376 2.71e-13
## feeding.scheme.learners                     0.625898   0.023360   26.794  < 2e-16
## mobile.schoolYES                           -0.144731   0.093737   -1.544   0.1228
## connectivityYES                             0.530591   0.052261   10.153  < 2e-16
##
## (Intercept)
## education.districtEDEN AND CENTRAL KAROO
## education.districtMETRO CENTRAL
## education.districtMETRO EAST
## education.districtMETRO NORTH
## education.districtMETRO SOUTH
## education.districtOVERBERG
## education.districtWEST COAST
## correspondence.language.xENGLISH       ***
## institution.type.xIntermediate School
## institution.type.xPrimary School
## institution.type.xSchool of Skills     ***
## institution.type.xSecondary School     ***
## institution.type.xSpecial School       ***
## fee.status.xNo Fee
## quintile.xNQ1                          ***
## quintile.xNQ2                          ***
## quintile.xNQ3                          ***
## quintile.xNQ4                          .
## quintile.xNQ5
## magisterial.district.xBEAUFORT WEST
## magisterial.district.xBELLVILLE
## magisterial.district.xBREDASDORP
## magisterial.district.xCALEDON
## magisterial.district.xCALITZDORP
## magisterial.district.xCAPE
## magisterial.district.xCERES
```

```
## magisterial.district.xCLANWILLIAM
## magisterial.district.xGEORGE
## magisterial.district.xGOODWOOD
## magisterial.district.xHEIDELBERG (WC)        .
## magisterial.district.xHERMANUS
## magisterial.district.xHOPEFIELD
## magisterial.district.xKNYSNA
## magisterial.district.xKUILS RIVER
## magisterial.district.xKUILSRIVIER
## magisterial.district.xLADISMITH
## magisterial.district.xLAINGSBURG
## magisterial.district.xMALMESBURY
## magisterial.district.xMITCHELL'S PLAIN
## magisterial.district.xMITCHELLS PLAIN
## magisterial.district.xMONTAGU
## magisterial.district.xMOORREESBURG
## magisterial.district.xMOSSEL BAY
## magisterial.district.xMURRAYSBURG
## magisterial.district.xOUDTSHOORN
## magisterial.district.xPAARL
## magisterial.district.xPIKETBERG
## magisterial.district.xPRINCE ALBERT
## magisterial.district.xRIVERSDALE
## magisterial.district.xROBERTSON
## magisterial.district.xSIMONS TOWN
## magisterial.district.xSOMERSET WEST
## magisterial.district.xSTELLENBOSCH
## magisterial.district.xSTRAND
## magisterial.district.xSWELLENDAM
## magisterial.district.xTULBAGH
## magisterial.district.xUNIONDALE
## magisterial.district.xVANRHYNSDORP
## magisterial.district.xVREDENBURG
## magisterial.district.xVREDENDAL
## magisterial.district.xWELLINGTON
## magisterial.district.xWORCESTER
## magisterial.district.xWYNBERG
## bus.route.learners                      ***
## feeding.scheme.learners                 ***
## mobile.schoolYES
## connectivityYES                         ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5333 on 1467 degrees of freedom
## Multiple R-squared:  0.7277, Adjusted R-squared:  0.7156
## F-statistic: 60.31 on 65 and 1467 DF,  p-value: < 2.2e-16
```

I applied a linear model to the data with total learners enrolled to a school as the target variable. The independent variables and their effects on the target variable are displayed in the summary above. The model shows that quintile level is one of the more significant variables that explain total enrollment numbers. From the bar graph seen previously we can see that the effect it exerts is that for NQ2, NQ3 and NQ4 you have more learners than in NQ1 and NQ5 schools.

Institution type also has an effect on total learners enrolled. Secondary schools tend to have more learners

than other school types, with school of skills and special schools having much fewer learners than others.

Connectivity also has a positive impact on school enrollment numbers. Schools that have internet connection attract more students.

We also see that schools with an English correspondence language have a positive effect on number of student enrollments.

Location Insights

```
## # A tibble: 8 x 2
##   education.district     mean_learners
##   <chr>                          <dbl>
## 1 CAPE WINELANDS                   606.
## 2 EDEN AND CENTRAL KAROO           670.
## 3 METRO CENTRAL                    694.
## 4 METRO EAST                      1147.
## 5 METRO NORTH                      961.
## 6 METRO SOUTH                      948.
## 7 OVERBERG                         608.
## 8 WEST COAST                       587.


## # A tibble: 6 x 2
##   institution.type.x  mean_learners
##   <chr>                       <dbl>
## 1 Combined School              706.
## 2 Intermediate School          754.
## 3 Primary School               728.
## 4 School of Skills             410.
## 5 Secondary School            1069.
## 6 Special School               250.
```

The output tables show that there are higher school sizes, on average, for the Metro East region. We also see that the average secondary school sizes are larger than the rest. Perhaps focus on Metro East for a school building project can be the start as it seems they have large numbers of students per school, on average. And a focus can be on Secondary schools as they also seem to be overburdened.

Predicting School Closure

Support Vector Machine (SVM)

```
## Confusion Matrix and Statistics
##
##                   Reference
## Prediction       Closed Open Pending Closure
##   Closed            253   37               0
##   Open               32  464               0
##   Pending Closure    40    0               0
##
## Overall Statistics
##
##                Accuracy : 0.868
##                  95% CI : (0.843, 0.8904)
##     No Information Rate : 0.6065
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.7348
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Closed Class: Open Class: Pending Closure
## Sensitivity                 0.7785      0.9261                     NA
## Specificity                 0.9261      0.9015                0.95157
## Pos Pred Value              0.8724      0.9355                     NA
## Neg Pred Value              0.8657      0.8879                     NA
## Prevalence                  0.3935      0.6065                0.00000
## Detection Rate              0.3063      0.5617                0.00000
## Detection Prevalence        0.3511      0.6005                0.04843
## Balanced Accuracy           0.8523      0.9138                     NA
```

Using an SVM to classify rows according to school status (opened or closed), it shows that if we use the variables correspondence language, sector, institution type, quintile, fee status and magisterial district, we can predict a school closure with approximately 87% accuracy.

Multinomial Logistic Regression

```
## # weights:  237 (156 variable)
## initial  value 2700.389006
## iter  10 value 897.409662
## iter  20 value 761.518914
## iter  30 value 752.981017
## iter  40 value 752.393065
## iter  50 value 752.040118
## iter  60 value 751.988712
## iter  70 value 751.977429
## iter  80 value 751.962903
## iter  90 value 751.958371
## iter 100 value 751.932878
```

```
## final  value 751.932878
## stopped after 100 iterations


## Warning in sqrt(diag(vc)): NaNs produced


## Call:
## multinom(formula = status.x ~ ., data = training_set)
##
## Coefficients:
##                  (Intercept) correspondence.language.xAFR/ENG
## Closed              -3.289674                         18.07673
## Pending Closure   -23.508449                          1.91721
##                  correspondence.language.xAFRIKAANS sector.xLSEN
## Closed                                    0.5635603            0
## Pending Closure                          -0.6293415            0
##                  sector.xORDINARY institution.type.xCombined School
## Closed                  -2.569771                        -0.1848964
## Pending Closure        -12.249842                        -2.6413367
##                  institution.type.xHospital School
## Closed                                           0
## Pending Closure                                  0
##                  institution.type.xIndustrial School
## Closed                                             0
## Pending Closure                                    0
##                  institution.type.xIntermediate School
## Closed                                        -1.113525
## Pending Closure                               -4.495272
##                  institution.type.xPlaces of Safety
## Closed                                            0
## Pending Closure                                   0
##                  institution.type.xPrimary School
## Closed                                  -0.7680149
## Pending Closure                          6.3039725
##                  institution.type.xReform School
## Closed                                         0
## Pending Closure                                0
##                  institution.type.xSchool of Skills
## Closed                                            0
## Pending Closure                                   0
##                  institution.type.xSecondary School
## Closed                                   -0.5033352
## Pending Closure                         -11.4172054
##                  institution.type.xSpecial School
## Closed                                          0
## Pending Closure                                 0
##                  institution.type.xSpecial School Resource Cen
## Closed                                                       0
## Pending Closure                                              0
##                  institution.type.xSpecial Youth Centre
## Closed                                                0
## Pending Closure                                       0
##                  institution.type.xSpecialized School
## Closed                                             0
## Pending Closure                                    0
```

```
##                institution.type.xTraining Centre
## Closed                                         0
## Pending Closure                                0
##                institution.type.xYouth Centre quintile.xNQ2 quintile.xNQ3
## Closed                                       0      4.486635    -0.1299928
## Pending Closure                              0     19.356727    -3.9818067
##                quintile.xNQ4 quintile.xNQ5 fee.status.xFee charging
## Closed           -0.04980484     0.4211653                -1.708283
## Pending Closure   9.95417810    10.8209742                -6.369279
##                magisterial.district.xBEAUFORT-WES/WEST
## Closed                                               0
## Pending Closure                                      0
##                magisterial.district.xBEAUFORT WEST
## Closed                                    3.792325
## Pending Closure                          -2.331807
##                magisterial.district.xBELLVILLE
## Closed                                2.584720
## Pending Closure                       4.691782
##                magisterial.district.xBREDASDORP magisterial.district.xCALEDON
## Closed                                 3.270277                      1.776502
## Pending Closure                       -2.518546                     -2.697172
##                magisterial.district.xCALITZDORP magisterial.district.xCAPE
## Closed                                 3.560195                   2.585888
## Pending Closure                       10.775162                   7.560723
##                magisterial.district.xCERES magisterial.district.xCLANWILLIAM
## Closed                            1.358236                          2.635800
## Pending Closure                  -4.125780                         -1.740384
##                magisterial.district.xGEORGE magisterial.district.xGOODWOOD
## Closed                             2.096182                       2.784960
## Pending Closure                   -2.305694                       6.877713
##                magisterial.district.xHEIDELBERG (WC)
## Closed                                     1.810054
## Pending Closure                           -4.508108
##                magisterial.district.xHEIDELBERG(C)
## Closed                                14.927653778
## Pending Closure                        0.004863897
##                magisterial.district.xHERMANUS magisterial.district.xHOPEFIELD
## Closed                               0.6104845                       0.2570585
## Pending Closure                     -6.0967818                      -6.9576440
##                magisterial.district.xKNYSNA magisterial.district.xKUILS RIVER
## Closed                            0.8722635                          3.029620
## Pending Closure                  -5.5472450                         -4.766544
##                magisterial.district.xKUILSRIVIER
## Closed                                  1.543762
## Pending Closure                         5.988104
##                magisterial.district.xLADISMITH
## Closed                                3.8446134
## Pending Closure                       0.1892695
##                magisterial.district.xLAINGSBURG
## Closed                                 1.998743
## Pending Closure                       -5.045125
##                magisterial.district.xMALMESBURY
## Closed                                 1.758004
## Pending Closure                       -1.076154
```

```
##                 magisterial.district.xMITCHELL'S PLAIN
## Closed                                  1.166133
## Pending Closure                         -6.739270
##                 magisterial.district.xMITCHELLS PLAIN
## Closed                                  0.4106659
## Pending Closure                         4.8472469
##                 magisterial.district.xMONTAGU
## Closed                                  1.419061
## Pending Closure                         -4.848932
##                 magisterial.district.xMONTANA EXTENSION
## Closed                                          0
## Pending Closure                                 0
##                 magisterial.district.xMOOREESBURG
## Closed                                  13.8407875
## Pending Closure                         -0.9251347
##                 magisterial.district.xMOORREESBURG
## Closed                                  2.799071
## Pending Closure                         -1.482689
##                 magisterial.district.xMOSSEL BAY
## Closed                                  1.596308
## Pending Closure                         -5.084176
##                 magisterial.district.xMOSSELBAAI/BAY
## Closed                                  14.5486938
## Pending Closure                         -0.1762253
##                 magisterial.district.xMURRAYSBURG
## Closed                                  2.177566
## Pending Closure                         -3.082665
##                 magisterial.district.xOUDTSHOORN magisterial.district.xPAARL
## Closed                                  1.937048                    0.5549962
## Pending Closure                         5.535117                    3.7628362
##                 magisterial.district.xPIKETBERG
## Closed                                  4.4772258
## Pending Closure                         0.4049478
##                 magisterial.district.xPRINCE ALBERT
## Closed                                  4.6514396
## Pending Closure                         0.1077826
##                 magisterial.district.xRIVERSDAL
## Closed                                  16.0038572
## Pending Closure                         0.2769506
##                 magisterial.district.xRIVERSDALE
## Closed                                  3.8711405
## Pending Closure                         0.6678375
##                 magisterial.district.xROBERTSON
## Closed                                  1.964384
## Pending Closure                         -3.887003
##                 magisterial.district.xSIMON'S TOWN
## Closed                                          0
## Pending Closure                                 0
##                 magisterial.district.xSIMONS TOWN
## Closed                                  3.456983
## Pending Closure                         7.779133
##                 magisterial.district.xSIMONSTAD
## Closed                                          0
## Pending Closure                                 0
```

```
##                magisterial.district.xSOMERSET-WES/WEST
## Closed                                             0
## Pending Closure                                    0
##                magisterial.district.xSOMERSET WEST
## Closed                               2.896279
## Pending Closure                      6.382417
##                magisterial.district.xSTELLENBOSCH magisterial.district.xSTRAND
## Closed                               1.453438                       -0.1897797
## Pending Closure                     -5.067350                        4.0662533
##                magisterial.district.xSWELLENDAM magisterial.district.xTULBAGH
## Closed                              2.744421                       1.843600
## Pending Closure                    -1.219113                      -3.896264
##                magisterial.district.xUNIONDALE
## Closed                              2.597792
## Pending Closure                    -3.833496
##                magisterial.district.xVANRHYNSDORP
## Closed                              3.3262670
## Pending Closure                    -0.5285818
##                magisterial.district.xVREDENBURG
## Closed                              2.802977
## Pending Closure                    -3.579215
##                magisterial.district.xVREDENDAL
## Closed                              3.347098
## Pending Closure                    -1.329374
##                magisterial.district.xWELLINGTON
## Closed                              3.350899
## Pending Closure                    -2.562559
##                magisterial.district.xWORCESTER magisterial.district.xWYNBERG
## Closed                              1.948374                       1.833871
## Pending Closure                    -2.319359                       6.764071
##
## Std. Errors:
##                (Intercept) correspondence.language.xAFR/ENG
## Closed           28.13371                     3.956356e-05
## Pending Closure  19.59727                     1.222348e-05
##                correspondence.language.xAFRIKAANS sector.xLSEN
## Closed                                  0.2798805          NaN
## Pending Closure                         0.5230948 3.219147e-14
##                sector.xORDINARY institution.type.xCombined School
## Closed                0.1906702                        0.3048677
## Pending Closure      77.2766356                      145.9996533
##                institution.type.xHospital School
## Closed                              2.059929e-08
## Pending Closure                     1.147292e-14
##                institution.type.xIndustrial School
## Closed                              4.005692e-08
## Pending Closure                              NaN
##                institution.type.xIntermediate School
## Closed                              0.3661746
## Pending Closure                     8.5310704
##                institution.type.xPlaces of Safety
## Closed                               3.93347e-08
## Pending Closure                              NaN
##                institution.type.xPrimary School
```

```
## Closed                                            0.179682
## Pending Closure                                   77.279486
##              institution.type.xReform School
## Closed                                            2.244727e-08
## Pending Closure                                   7.775858e-15
##              institution.type.xSchool of Skills
## Closed                                                 NaN
## Pending Closure                                        NaN
##              institution.type.xSecondary School
## Closed                                            0.25584057
## Pending Closure                                   0.02113296
##              institution.type.xSpecial School
## Closed                                                 NaN
## Pending Closure                                   4.705282e-14
##              institution.type.xSpecial School Resource Cen
## Closed                                                 1.278583e-08
## Pending Closure                                        5.902700e-15
##              institution.type.xSpecial Youth Centre
## Closed                                                 NaN
## Pending Closure                                        NaN
##              institution.type.xSpecialized School
## Closed                                                 NaN
## Pending Closure                                        NaN
##              institution.type.xTraining Centre
## Closed                                            3.364827e-09
## Pending Closure                                        NaN
##              institution.type.xYouth Centre quintile.xNQ2 quintile.xNQ3
## Closed                             5.803504e-09    0.3742196    0.44161227
## Pending Closure                    7.599205e-14    9.7359684    0.00938727
##              quintile.xNQ4 quintile.xNQ5 fee.status.xFee charging
## Closed            0.5359947    0.5446674                0.3739827
## Pending Closure  20.0653545    15.5224383               20.2764934
##              magisterial.district.xBEAUFORT-WES/WEST
## Closed                                            3.028374e-09
## Pending Closure                                        NaN
##              magisterial.district.xBEAUFORT WEST
## Closed                                            28.14288
## Pending Closure                                   177.91744
##              magisterial.district.xBELLVILLE
## Closed                                            28.13368
## Pending Closure                                   24.20347
##              magisterial.district.xBREDASDORP magisterial.district.xCALEDON
## Closed                             28.14459                         28.13619
## Pending Closure                    172.26032                        66.43524
##              magisterial.district.xCALITZDORP magisterial.district.xCAPE
## Closed                             28.16247                         28.13426
## Pending Closure                    24.25482                         24.18825
##              magisterial.district.xCERES magisterial.district.xCLANWILLIAM
## Closed                             28.13626                         28.14122
## Pending Closure                    102.91536                        130.53325
##              magisterial.district.xGEORGE magisterial.district.xGOODWOOD
## Closed                             28.13495                         28.13349
## Pending Closure                    46.93485                         24.18701
##              magisterial.district.xHEIDELBERG (WC)
```

```
## Closed                                             28.16434
## Pending Closure                                     8.28741
##                 magisterial.district.xHEIDELBERG(C)
## Closed                                            1.756510564
## Pending Closure                                   0.000748393
##                 magisterial.district.xHERMANUS magisterial.district.xHOPEFIELD
## Closed                                   28.14185                    28.176415
## Pending Closure                         240.15032                     2.153995
##                 magisterial.district.xKNYSNA magisterial.district.xKUILS RIVER
## Closed                                   28.13752                    28.306173
## Pending Closure                         145.63045                     9.584934
##                 magisterial.district.xKUILSRIVIER
## Closed                                             28.13560
## Pending Closure                                    24.19048
##                 magisterial.district.xLADISMITH
## Closed                                             28.14729
## Pending Closure                                   134.10289
##                 magisterial.district.xLAINGSBURG
## Closed                                            28.262161
## Pending Closure                                    5.000904
##                 magisterial.district.xMALMESBURY
## Closed                                             28.13501
## Pending Closure                                    30.37101
##                 magisterial.district.xMITCHELL'S PLAIN
## Closed                                            28.245545
## Pending Closure                                    1.803571
##                 magisterial.district.xMITCHELLS PLAIN
## Closed                                             28.13273
## Pending Closure                                    24.18436
##                 magisterial.district.xMONTAGU
## Closed                                             28.15183
## Pending Closure                                   299.83336
##                 magisterial.district.xMONTANA EXTENSION
## Closed                                           2.310989e-10
## Pending Closure                                  0.000000e+00
##                 magisterial.district.xMOOREESBURG
## Closed                                           0.341369521
## Pending Closure                                  0.001590443
##                 magisterial.district.xMOORREESBURG
## Closed                                             28.15143
## Pending Closure                                   203.90874
##                 magisterial.district.xMOSSEL BAY
## Closed                                             28.13993
## Pending Closure                                   225.53797
##                 magisterial.district.xMOSSELBAAI/BAY
## Closed                                           3.70531851
## Pending Closure                                  0.01404883
##                 magisterial.district.xMURRAYSBURG
## Closed                                             28.59200
## Pending Closure                                    16.93857
##                 magisterial.district.xOUDTSHOORN magisterial.district.xPAARL
## Closed                                   28.13553                    28.13606
## Pending Closure                         24.20420                     24.20385
##                 magisterial.district.xPIKETBERG
```
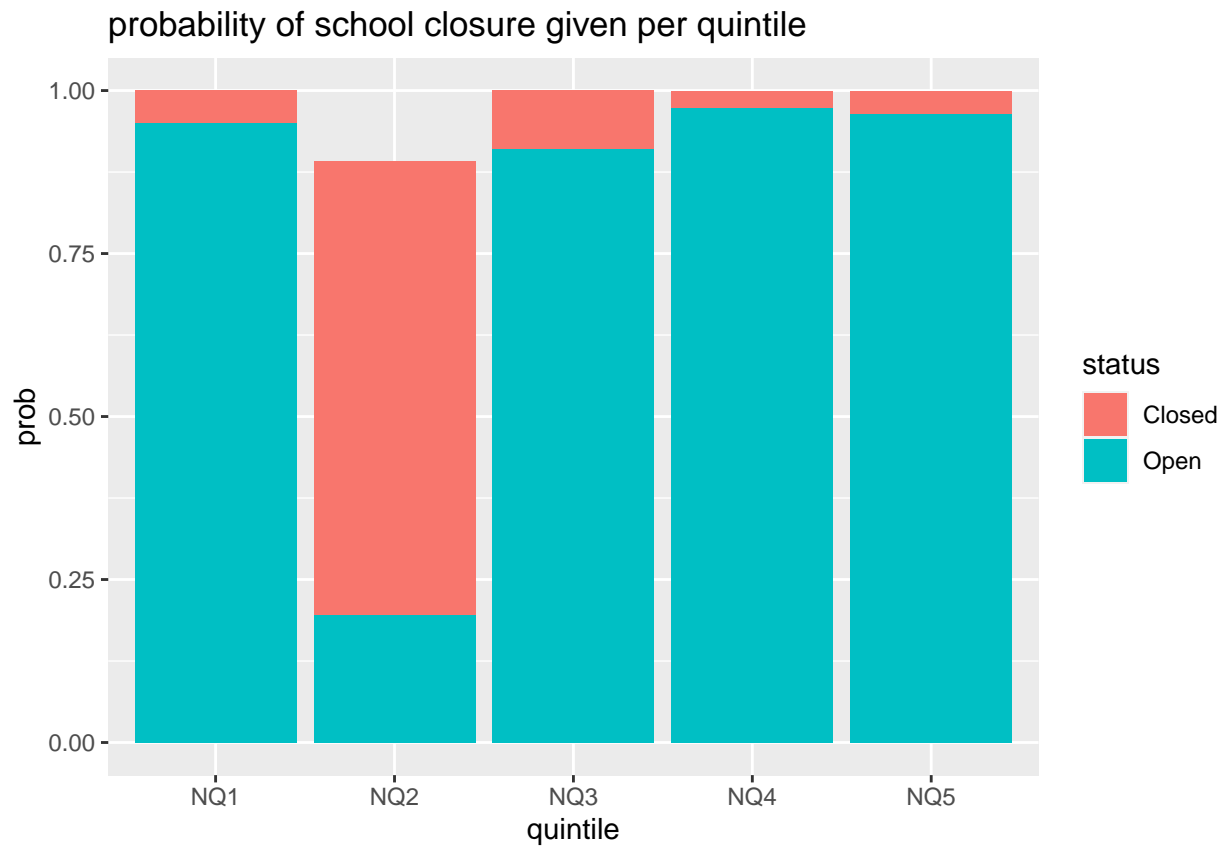
```
## Closed                                        28.13999
## Pending Closure                               87.40691
##                    magisterial.district.xPRINCE ALBERT
## Closed                                            28.16389
## Pending Closure                                  241.47729
##                    magisterial.district.xRIVERSDAL
## Closed                                      1.4400102398
## Pending Closure                             0.0003101596
##                    magisterial.district.xRIVERSDALE
## Closed                                        28.14160
## Pending Closure                               96.28338
##                    magisterial.district.xROBERTSON
## Closed                                         28.14433
## Pending Closure                               238.95909
##                    magisterial.district.xSIMON'S TOWN
## Closed                                             NaN
## Pending Closure                                      0
##                    magisterial.district.xSIMONS TOWN
## Closed                                        28.13690
## Pending Closure                               24.19801
##                    magisterial.district.xSIMONSTAD
## Closed                                             NaN
## Pending Closure                                      0
##                    magisterial.district.xSOMERSET-WES/WEST
## Closed                                         3.952231e-11
## Pending Closure                                0.000000e+00
##                    magisterial.district.xSOMERSET WEST
## Closed                                        28.13515
## Pending Closure                               24.20698
##                    magisterial.district.xSTELLENBOSCH magisterial.district.xSTRAND
## Closed                                        28.14366                      28.13776
## Pending Closure                              148.54384                      24.19394
##                    magisterial.district.xSWELLENDAM magisterial.district.xTULBAGH
## Closed                                        28.14081                     28.14314
## Pending Closure                              124.72084                    258.06454
##                    magisterial.district.xUNIONDALE
## Closed                                        28.14948
## Pending Closure                              283.06387
##                    magisterial.district.xVANRHYNSDORP
## Closed                                        28.15417
## Pending Closure                              163.91971
##                    magisterial.district.xVREDENBURG
## Closed                                        28.14588
## Pending Closure                              189.54344
##                    magisterial.district.xVREDENDAL
## Closed                                        28.14425
## Pending Closure                              149.80604
##                    magisterial.district.xWELLINGTON
## Closed                                        28.14838
## Pending Closure                              184.15515
##                    magisterial.district.xWORCESTER magisterial.district.xWYNBERG
## Closed                                         28.1390                     28.13189
## Pending Closure                                48.0806                     24.18355
##
```
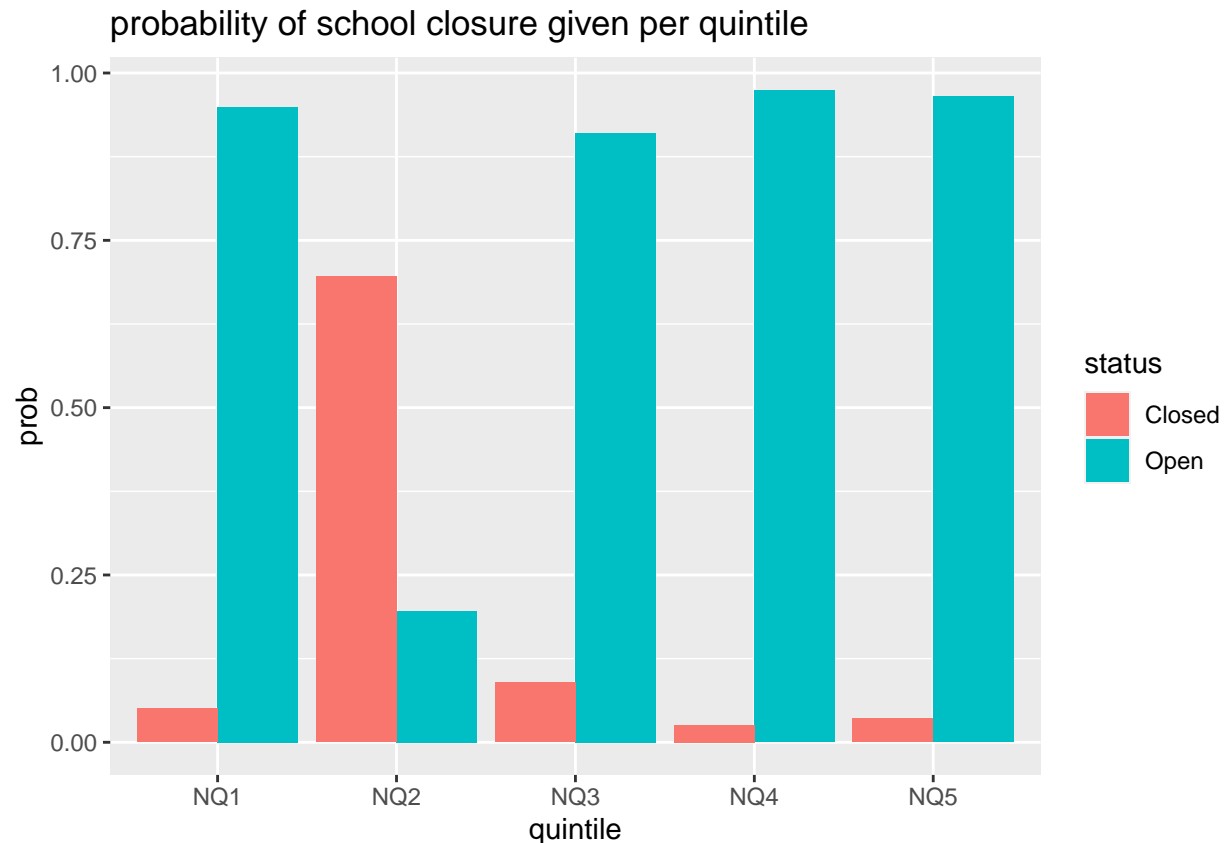
```
## Residual Deviance: 1503.866
## AIC: 1743.866
```

I applied a multinomial logistic regression model to the data to find out which explanatory variables carried the most significance in predicting a school closure or if a school would remain open. The results show that quintile level and institution type play the biggest role in predicting school closures.

```
## # weights:  18 (10 variable)
## initial  value 2700.389006
## iter  10 value 1203.887547
## iter  20 value 1161.966845
## iter  30 value 1161.524670
## iter  40 value 1161.461716
## final  value 1161.461663
## converged
```


probability of school closure given per quintile

probability of school closure given per quintile


```
## # A tibble: 10 x 3
##    quintile.x status   prob
##    <fct>      <chr>   <dbl>
##  1 NQ2        Open    0.196
##  2 NQ2        Closed  0.696
##  3 NQ3        Open    0.911
##  4 NQ3        Closed  0.0894
##  5 NQ5        Open    0.965
##  6 NQ5        Closed  0.0353
##  7 NQ4        Open    0.974
##  8 NQ4        Closed  0.0259
##  9 NQ1        Open    0.950
## 10 NQ1        Closed  0.0504
```

The multinomial logistic regression model shows that NQ2 schools are almost 70% likely to close down. That is only based on this dataset though. We don't know what history looks like. But, based on the previous graphs showing NQ2 schools are experiencing higher average school sizes compared to NQ1 and NQ5 it could emphasize a need to build more NQ2 schools.

```
## # weights:  51 (32 variable)
## initial  value 2700.389006
## iter  10 value 1523.798086
## iter  20 value 1516.497976
## iter  30 value 1516.475802
## iter  40 value 1516.464223
```

```
## final  value 1516.464171
## converged

## Warning in sqrt(diag(vc)): NaNs produced

## Call:
## multinom(formula = status.x ~ institution.type.x, data = training_set)
##
## Coefficients:
##                  (Intercept) institution.type.xCombined School
## Closed            0.9307004                         -2.484092
## Pending Closure  -0.7094142                        -13.336599
##                  institution.type.xHospital School
## Closed                                           0
## Pending Closure                                  0
##                  institution.type.xIndustrial School
## Closed                                             0
## Pending Closure                                    0
##                  institution.type.xIntermediate School
## Closed                                       -2.503087
## Pending Closure                             -30.972118
##                  institution.type.xPlaces of Safety
## Closed                                            0
## Pending Closure                                   0
##                  institution.type.xPrimary School
## Closed                                    -2.535252
## Pending Closure                           -5.988178
##                  institution.type.xReform School
## Closed                                          0
## Pending Closure                                 0
##                  institution.type.xSchool of Skills
## Closed                                            0
## Pending Closure                                   0
##                  institution.type.xSecondary School
## Closed                                     -3.263649
## Pending Closure                           -14.090964
##                  institution.type.xSpecial School
## Closed                                           0
## Pending Closure                                  0
##                  institution.type.xSpecial School Resource Cen
## Closed                                                       0
## Pending Closure                                              0
##                  institution.type.xSpecial Youth Centre
## Closed                                                 0
## Pending Closure                                        0
##                  institution.type.xSpecialized School
## Closed                                               0
## Pending Closure                                      0
##                  institution.type.xTraining Centre
## Closed                                            0
## Pending Closure                                   0
##                  institution.type.xYouth Centre
## Closed                                         0
## Pending Closure                                0
```

```
##
## Std. Errors:
##                (Intercept) institution.type.xCombined School
## Closed          0.0749805                        0.2463607
## Pending Closure 0.1105848                      110.0361022
##                institution.type.xHospital School
## Closed                                       NaN
## Pending Closure                                0
##                institution.type.xIndustrial School
## Closed                                  4.644427e-15
## Pending Closure                         0.000000e+00
##                institution.type.xIntermediate School
## Closed                                  3.397034e-01
## Pending Closure                         1.677836e-12
##                institution.type.xPlaces of Safety
## Closed                                  1.527453e-14
## Pending Closure                         0.000000e+00
##                institution.type.xPrimary School
## Closed                                    0.1139748
## Pending Closure                           1.0063704
##                institution.type.xReform School
## Closed                                            0
## Pending Closure                                   0
##                institution.type.xSchool of Skills
## Closed                                            0
## Pending Closure                                   0
##                institution.type.xSecondary School
## Closed                                     0.218671
## Pending Closure                           99.952403
##                institution.type.xSpecial School
## Closed                                            0
## Pending Closure                                   0
##                institution.type.xSpecial School Resource Cen
## Closed                                                     0
## Pending Closure                                            0
##                institution.type.xSpecial Youth Centre
## Closed                                                0
## Pending Closure                                       0
##                institution.type.xSpecialized School
## Closed                                             0
## Pending Closure                                    0
##                institution.type.xTraining Centre
## Closed                                          0
## Pending Closure                                 0
##                institution.type.xYouth Centre
## Closed                                       0
## Pending Closure                              0
##
## Residual Deviance: 3032.928
## AIC: 3052.928


## Warning in confusionMatrix.default(test_set[, 3], inst_preds): Levels are not
## in the same order for reference and data. Refactoring data to match.
```
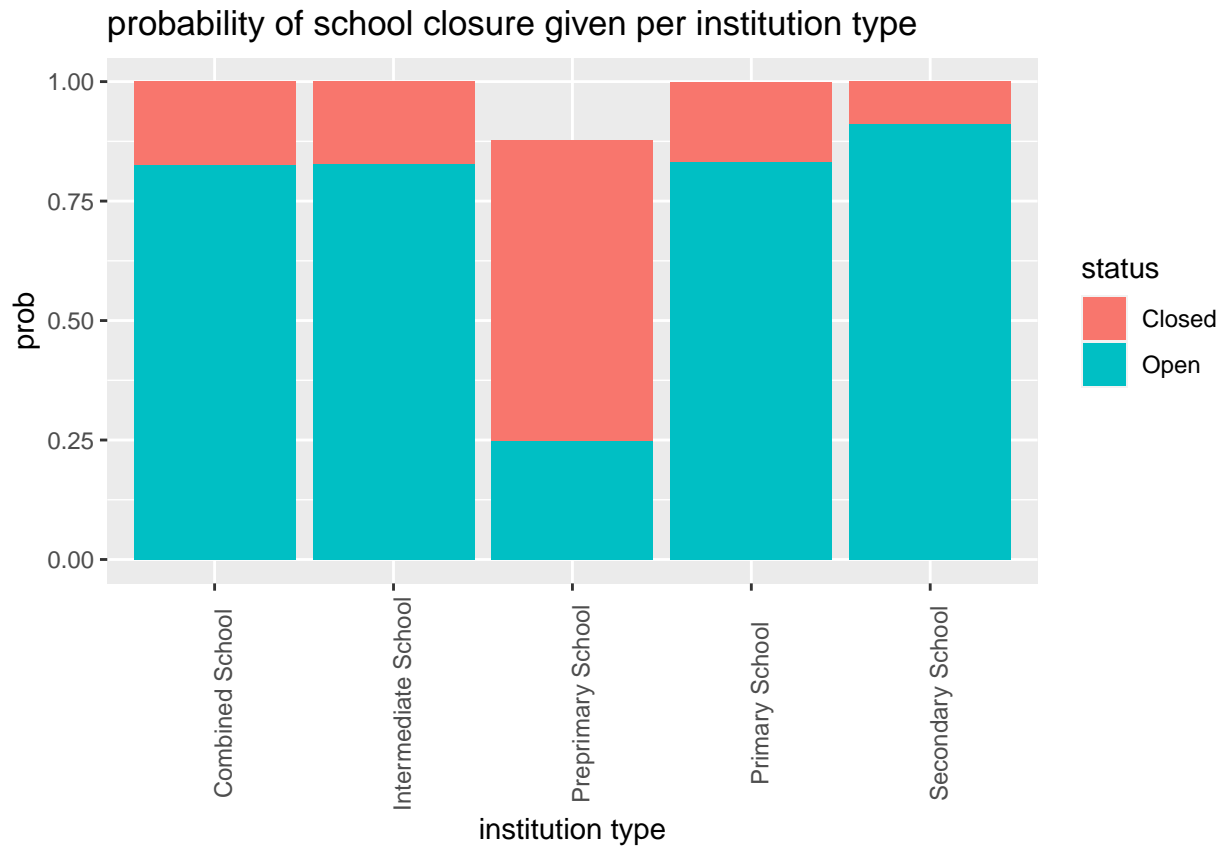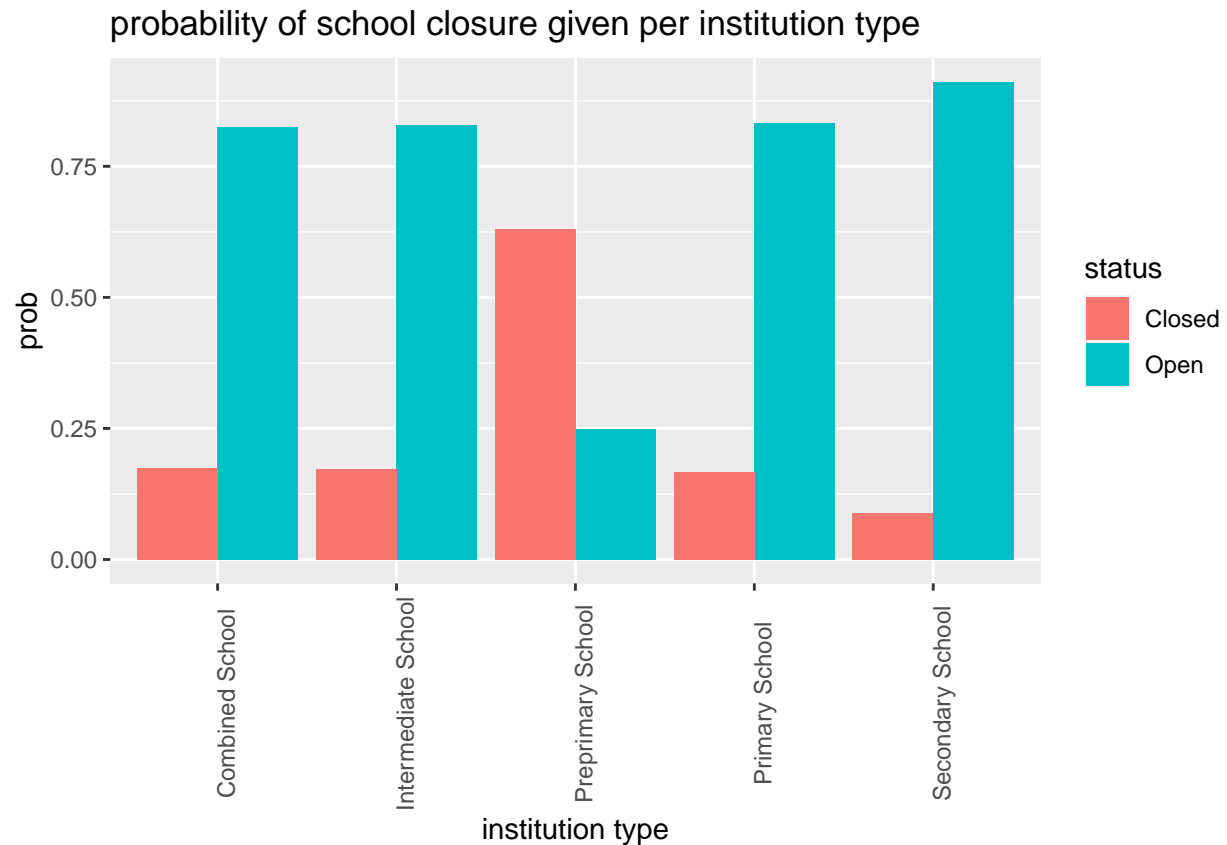
```
## Confusion Matrix and Statistics
##
##                   Reference
## Prediction         Open Closed Pending Closure
##   Open              423    73                0
##   Closed             81   209                0
##   Pending Closure     0    40                0
##
## Overall Statistics
##
##                Accuracy : 0.7651
##                  95% CI : (0.7347, 0.7937)
##     No Information Rate : 0.6102
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.5272
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Open Class: Closed Class: Pending Closure
## Sensitivity               0.8393        0.6491                    NA
## Specificity               0.7733        0.8393               0.95157
## Pos Pred Value            0.8528        0.7207                    NA
## Neg Pred Value            0.7545        0.7892                    NA
## Prevalence                0.6102        0.3898               0.00000
## Detection Rate            0.5121        0.2530               0.00000
## Detection Prevalence      0.6005        0.3511               0.04843
## Balanced Accuracy         0.8063        0.7442                    NA
```

probability of school closure given per institution type

probability of school closure given per institution type


```
## # A tibble: 10 x 3
##    institution.type.x  status    prob
##    <fct>               <chr>    <dbl>
##  1 Preprimary School   Open    0.248
##  2 Preprimary School   Closed  0.630
##  3 Combined School     Open    0.825
##  4 Combined School     Closed  0.175
##  5 Secondary School    Open    0.912
##  6 Secondary School    Closed  0.0884
##  7 Primary School      Open    0.832
##  8 Primary School      Closed  0.167
##  9 Intermediate School Open    0.828
## 10 Intermediate School Closed  0.172
```

Here it shows that Preprimary school is more likely to be closed down with a probability of 63%. Another factor that could be worth mentioning that perhaps focus needs to be placed on Preprimary schools. Again, this is only based on this dataset. Historical data could alter the story.

The graphs below show number of closures per quintile and number of closures per institution type compared to number of schools that have remained open.

number of school closures per quintile

Number of school closures per institution type