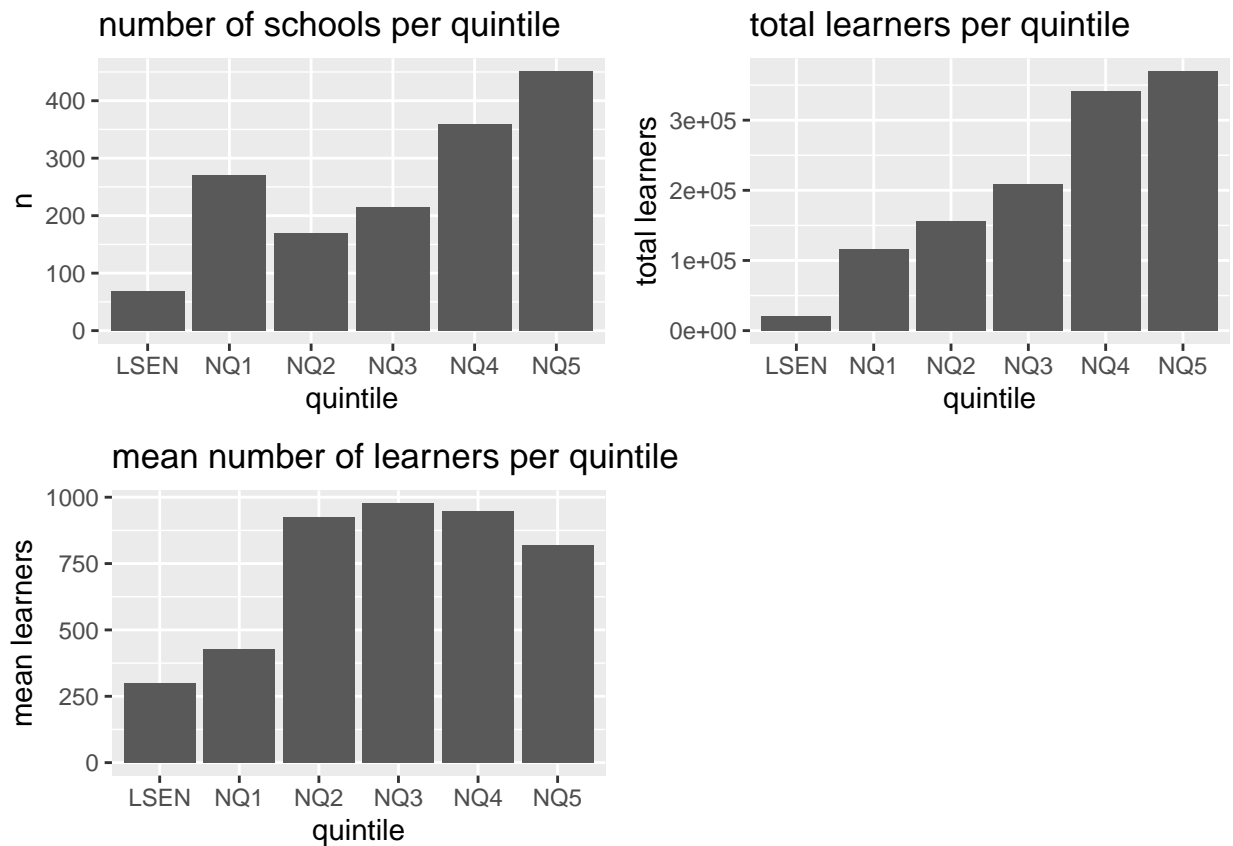


WCED Data Analysis

Gareth Edwards

WCED School Data Analysis

Bar Plots

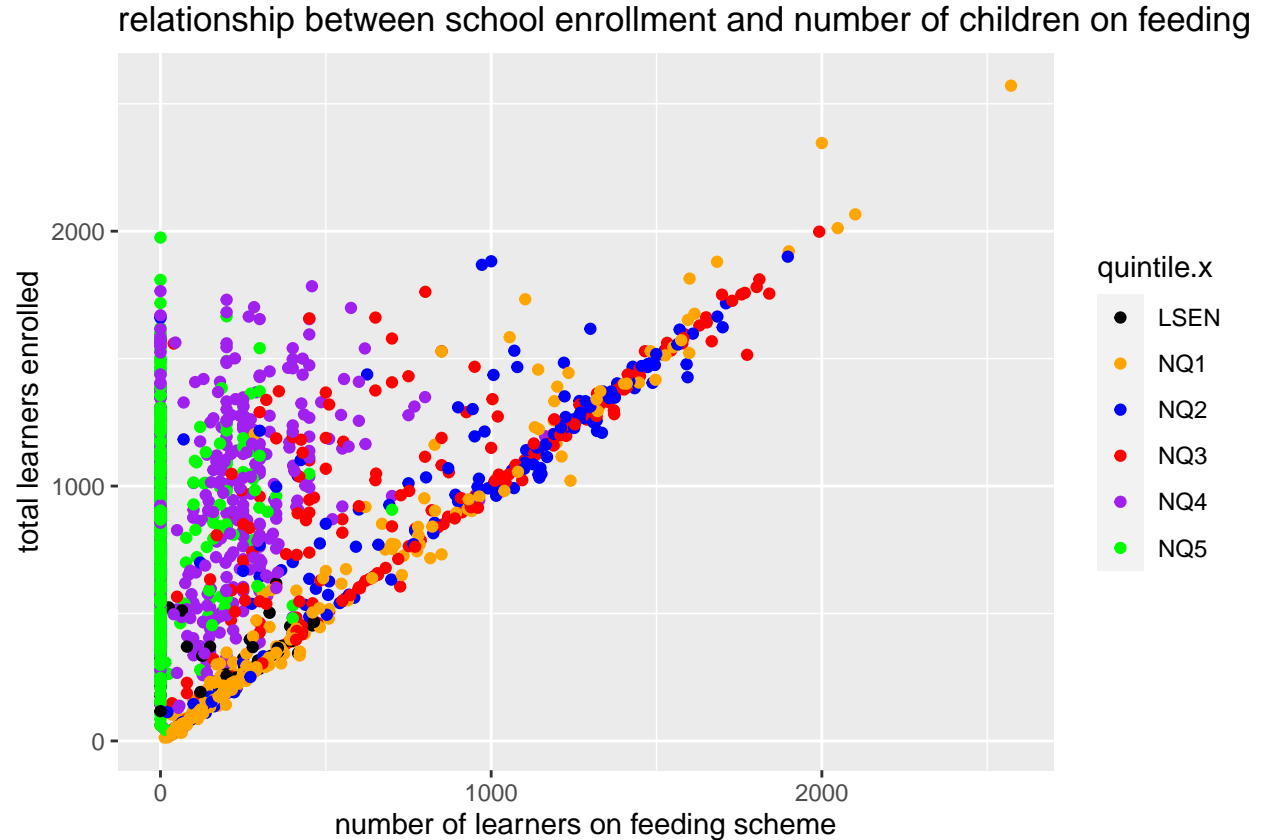


The barplots show there are more quintile 5 schools compared to the rest. There are also more students, in total, going to NQ5 schools. However, we see average school sizes for NQ2, NQ3 and NQ4 are higher than average school sizes for NQ5.

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



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.