model-building

Sarah McDonald

February 1, 2020

data <- read.csv("C:/Users/Sarah McDonald/Documents/honorsthesis/manipulated\_data/cath\_inst.csv")  
  
#remove 1990 and 1991  
data <- data[data$YEAR!= 1990 & data$YEAR != 1991 & data$YEAR != 1992, ]

# Bishop Accountability

### TOTAL ENROLLMENT

enroll\_total <- lm(EFTOTLT ~ bishop\_pub\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Total", ])   
  
summary(enroll\_total)

##   
## Call:  
## lm(formula = EFTOTLT ~ bishop\_pub\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Total", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10137.7 -1880.7 -718.2 595.5 20954.5   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.620e+03 1.545e+02 10.488 < 2e-16 \*\*\*  
## bishop\_pub\_sum\_3 7.174e+00 3.436e+00 2.088 0.0369 \*   
## Percent.Catholic 4.076e+00 3.727e+00 1.094 0.2742   
## his\_prop 2.175e+03 3.803e+02 5.719 1.13e-08 \*\*\*  
## pop\_dens 6.178e+05 2.831e+04 21.822 < 2e-16 \*\*\*  
## income\_per\_capita 1.663e-02 3.242e-03 5.129 3.01e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3328 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1385, Adjusted R-squared: 0.1377   
## F-statistic: 167 on 5 and 5194 DF, p-value: < 2.2e-16

### GRADUATE ENROLLMENT

enroll\_grad <- lm(EFTOTLT ~ bishop\_pub\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])   
  
summary(enroll\_grad)

##   
## Call:  
## lm(formula = EFTOTLT ~ bishop\_pub\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4472.0 -638.8 -379.7 160.7 9288.6   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.696e+02 6.121e+01 4.404 1.08e-05 \*\*\*  
## bishop\_pub\_sum\_3 2.135e+00 1.362e+00 1.568 0.1169   
## Percent.Catholic -2.798e+00 1.477e+00 -1.895 0.0581 .   
## his\_prop 1.193e+03 1.507e+02 7.912 3.07e-15 \*\*\*  
## pop\_dens 2.593e+05 1.122e+04 23.112 < 2e-16 \*\*\*  
## income\_per\_capita 9.720e-03 1.285e-03 7.566 4.52e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1319 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1594, Adjusted R-squared: 0.1586   
## F-statistic: 197 on 5 and 5194 DF, p-value: < 2.2e-16

### UNDERGRADUATE ENROLLMENT

enroll\_undergrad <- lm(EFTOTLT ~ bishop\_pub\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = EFTOTLT ~ bishop\_pub\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5717.5 -1339.9 -373.2 522.2 13491.6   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.350e+03 1.014e+02 13.322 < 2e-16 \*\*\*  
## bishop\_pub\_sum\_3 5.038e+00 2.255e+00 2.234 0.02550 \*   
## Percent.Catholic 6.874e+00 2.446e+00 2.811 0.00496 \*\*   
## his\_prop 9.825e+02 2.496e+02 3.936 8.38e-05 \*\*\*  
## pop\_dens 3.585e+05 1.858e+04 19.297 < 2e-16 \*\*\*  
## income\_per\_capita 6.909e-03 2.128e-03 3.248 0.00117 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2184 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1101, Adjusted R-squared: 0.1092   
## F-statistic: 128.5 on 5 and 5194 DF, p-value: < 2.2e-16

### TOTAL ENROLLMENT - Square Root Transformation

enroll\_total <- lm(sqrt(EFTOTLT) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_total)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -48.131 -12.389 0.071 9.930 78.354   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.285e+01 9.319e-01 35.248 < 2e-16 \*\*\*  
## sqrt(bishop\_pub\_sum\_3) 7.444e-01 1.758e-01 4.235 2.33e-05 \*\*\*  
## Percent.Catholic 8.579e-02 2.283e-02 3.759 0.000173 \*\*\*  
## his\_prop 9.518e+00 2.302e+00 4.134 3.62e-05 \*\*\*  
## pop\_dens 2.990e+03 1.724e+02 17.349 < 2e-16 \*\*\*  
## income\_per\_capita 6.056e-05 1.982e-05 3.055 0.002264 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.15 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1075, Adjusted R-squared: 0.1066   
## F-statistic: 125.1 on 5 and 5194 DF, p-value: < 2.2e-16

### GRADUATE ENROLLMENT - Square Root Transformation

enroll\_grad <- lm(sqrt(EFTOTLT) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])   
  
summary(enroll\_grad)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -57.629 -13.884 -2.206 9.361 66.872   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.219e+01 8.546e-01 14.262 < 2e-16 \*\*\*  
## sqrt(bishop\_pub\_sum\_3) 8.110e-01 1.612e-01 5.032 5.02e-07 \*\*\*  
## Percent.Catholic -7.868e-02 2.093e-02 -3.759 0.000172 \*\*\*  
## his\_prop 2.067e+01 2.111e+00 9.789 < 2e-16 \*\*\*  
## pop\_dens 3.189e+03 1.581e+02 20.173 < 2e-16 \*\*\*  
## income\_per\_capita 1.732e-04 1.818e-05 9.527 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.48 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1612, Adjusted R-squared: 0.1604   
## F-statistic: 199.7 on 5 and 5194 DF, p-value: < 2.2e-16

### UNDERGRADUATE ENROLLMENT - Square Root Transformation

enroll\_undergrad <- lm(sqrt(EFTOTLT) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -48.131 -12.389 0.071 9.930 78.354   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.285e+01 9.319e-01 35.248 < 2e-16 \*\*\*  
## sqrt(bishop\_pub\_sum\_3) 7.444e-01 1.758e-01 4.235 2.33e-05 \*\*\*  
## Percent.Catholic 8.579e-02 2.283e-02 3.759 0.000173 \*\*\*  
## his\_prop 9.518e+00 2.302e+00 4.134 3.62e-05 \*\*\*  
## pop\_dens 2.990e+03 1.724e+02 17.349 < 2e-16 \*\*\*  
## income\_per\_capita 6.056e-05 1.982e-05 3.055 0.002264 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.15 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1075, Adjusted R-squared: 0.1066   
## F-statistic: 125.1 on 5 and 5194 DF, p-value: < 2.2e-16

### TOTAL ENROLLMENT - Log+1 Transformation

enroll\_total <- lm(log(EFTOTLT+1) ~ log(bishop\_pub\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_total)

##   
## Call:  
## lm(formula = log(EFTOTLT + 1) ~ log(bishop\_pub\_sum\_3 + 1) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.7546 -0.3939 0.3001 0.7215 2.5422   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.638e+00 5.559e-02 119.421 < 2e-16 \*\*\*  
## log(bishop\_pub\_sum\_3 + 1) 7.429e-02 1.733e-02 4.285 1.86e-05 \*\*\*  
## Percent.Catholic 6.936e-03 1.363e-03 5.090 3.71e-07 \*\*\*  
## his\_prop 3.809e-01 1.372e-01 2.777 0.00551 \*\*   
## pop\_dens 1.342e+02 1.030e+01 13.029 < 2e-16 \*\*\*  
## income\_per\_capita 2.712e-06 1.185e-06 2.289 0.02213 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.201 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.07512, Adjusted R-squared: 0.07423   
## F-statistic: 84.37 on 5 and 5194 DF, p-value: < 2.2e-16

### GRADUATE ENROLLMENT - Log+1 Transformation

enroll\_grad <- lm(log(EFTOTLT+1) ~ log(bishop\_pub\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])   
  
summary(enroll\_grad)

##   
## Call:  
## lm(formula = log(EFTOTLT + 1) ~ log(bishop\_pub\_sum\_3 + 1) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.6592 -0.8996 0.8979 1.8668 4.1249   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.847e+00 1.234e-01 31.177 < 2e-16 \*\*\*  
## log(bishop\_pub\_sum\_3 + 1) 2.501e-01 3.848e-02 6.498 8.90e-11 \*\*\*  
## Percent.Catholic -1.776e-02 3.025e-03 -5.870 4.62e-09 \*\*\*  
## his\_prop 2.136e+00 3.045e-01 7.014 2.61e-12 \*\*\*  
## pop\_dens 3.390e+02 2.287e+01 14.825 < 2e-16 \*\*\*  
## income\_per\_capita 2.200e-05 2.630e-06 8.364 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.666 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1052, Adjusted R-squared: 0.1044   
## F-statistic: 122.1 on 5 and 5194 DF, p-value: < 2.2e-16

### UNDERGRADUATE ENROLLMENT - Log+1 Transformation

enroll\_undergrad <- lm(log(EFTOTLT+1) ~ log(bishop\_pub\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = log(EFTOTLT + 1) ~ log(bishop\_pub\_sum\_3 + 1) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.7546 -0.3939 0.3001 0.7215 2.5422   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.638e+00 5.559e-02 119.421 < 2e-16 \*\*\*  
## log(bishop\_pub\_sum\_3 + 1) 7.429e-02 1.733e-02 4.285 1.86e-05 \*\*\*  
## Percent.Catholic 6.936e-03 1.363e-03 5.090 3.71e-07 \*\*\*  
## his\_prop 3.809e-01 1.372e-01 2.777 0.00551 \*\*   
## pop\_dens 1.342e+02 1.030e+01 13.029 < 2e-16 \*\*\*  
## income\_per\_capita 2.712e-06 1.185e-06 2.289 0.02213 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.201 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.07512, Adjusted R-squared: 0.07423   
## F-statistic: 84.37 on 5 and 5194 DF, p-value: < 2.2e-16

# Nexis-Lexus

### TOTAL ENROLLMENT

enroll\_total <- lm(EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Total", ])   
  
summary(enroll\_total)

##   
## Call:  
## lm(formula = EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop +   
## pop\_dens + income\_per\_capita, data = data[data$LINE == "Total",   
## ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10161 -1863 -713 588 20940   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.603e+03 1.551e+02 10.335 < 2e-16 \*\*\*  
## nexis\_sum\_3 8.607e-01 1.479e+00 0.582 0.561   
## Percent.Catholic 5.096e+00 3.696e+00 1.379 0.168   
## his\_prop 2.201e+03 3.816e+02 5.768 8.51e-09 \*\*\*  
## pop\_dens 6.192e+05 2.836e+04 21.839 < 2e-16 \*\*\*  
## income\_per\_capita 1.714e-02 3.258e-03 5.261 1.49e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3329 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1378, Adjusted R-squared: 0.137   
## F-statistic: 166 on 5 and 5194 DF, p-value: < 2.2e-16

### GRADUATE ENROLLMENT

enroll\_grad <- lm(EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])   
  
summary(enroll\_grad)

##   
## Call:  
## lm(formula = EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop +   
## pop\_dens + income\_per\_capita, data = data[data$LINE == "Graduate",   
## ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4477.1 -637.5 -382.3 163.2 9269.7   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.688e+02 6.145e+01 4.375 1.24e-05 \*\*\*  
## nexis\_sum\_3 5.940e-01 5.859e-01 1.014 0.3107   
## Percent.Catholic -2.562e+00 1.464e+00 -1.750 0.0803 .   
## his\_prop 1.192e+03 1.512e+02 7.883 3.88e-15 \*\*\*  
## pop\_dens 2.593e+05 1.123e+04 23.082 < 2e-16 \*\*\*  
## income\_per\_capita 9.761e-03 1.291e-03 7.562 4.65e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1319 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1592, Adjusted R-squared: 0.1584   
## F-statistic: 196.7 on 5 and 5194 DF, p-value: < 2.2e-16

### UNDERGRADUATE ENROLLMENT

enroll\_undergrad <- lm(EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop +   
## pop\_dens + income\_per\_capita, data = data[data$LINE == "Undergraduates",   
## ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5726.4 -1347.4 -373.4 523.0 13485.7   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.334e+03 1.018e+02 13.107 < 2e-16 \*\*\*  
## nexis\_sum\_3 2.667e-01 9.705e-01 0.275 0.783512   
## Percent.Catholic 7.658e+00 2.426e+00 3.157 0.001602 \*\*   
## his\_prop 1.009e+03 2.505e+02 4.030 5.66e-05 \*\*\*  
## pop\_dens 3.599e+05 1.861e+04 19.342 < 2e-16 \*\*\*  
## income\_per\_capita 7.381e-03 2.138e-03 3.452 0.000561 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2185 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1093, Adjusted R-squared: 0.1084   
## F-statistic: 127.4 on 5 and 5194 DF, p-value: < 2.2e-16

### TOTAL ENROLLMENT - Square Root Transformation

enroll\_total <- lm(sqrt(EFTOTLT) ~ sqrt(nexis\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_total)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT) ~ sqrt(nexis\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.489 -12.433 0.078 9.911 78.336   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.336e+01 9.490e-01 35.154 < 2e-16 \*\*\*  
## sqrt(nexis\_sum\_3) 3.311e-01 1.205e-01 2.747 0.00604 \*\*   
## Percent.Catholic 9.947e-02 2.247e-02 4.428 9.71e-06 \*\*\*  
## his\_prop 9.479e+00 2.310e+00 4.103 4.14e-05 \*\*\*  
## pop\_dens 2.994e+03 1.738e+02 17.221 < 2e-16 \*\*\*  
## income\_per\_capita 5.732e-05 2.064e-05 2.777 0.00550 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.17 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1057, Adjusted R-squared: 0.1048   
## F-statistic: 122.8 on 5 and 5194 DF, p-value: < 2.2e-16

### GRADUATE ENROLLMENT - Square Root Transformation

enroll\_grad <- lm(sqrt(EFTOTLT) ~ sqrt(nexis\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])   
  
summary(enroll\_grad)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT) ~ sqrt(nexis\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -58.064 -13.977 -2.173 9.443 66.309   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.259e+01 8.710e-01 14.456 < 2e-16 \*\*\*  
## sqrt(nexis\_sum\_3) 2.541e-01 1.106e-01 2.296 0.02170 \*   
## Percent.Catholic -6.155e-02 2.062e-02 -2.985 0.00285 \*\*   
## his\_prop 2.080e+01 2.120e+00 9.811 < 2e-16 \*\*\*  
## pop\_dens 3.217e+03 1.595e+02 20.166 < 2e-16 \*\*\*  
## income\_per\_capita 1.756e-04 1.894e-05 9.270 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.52 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.158, Adjusted R-squared: 0.1572   
## F-statistic: 194.9 on 5 and 5194 DF, p-value: < 2.2e-16

### UNDERGRADUATE ENROLLMENT - Square Root Transformation

enroll\_undergrad <- lm(sqrt(EFTOTLT) ~ sqrt(nexis\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT) ~ sqrt(nexis\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.489 -12.433 0.078 9.911 78.336   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.336e+01 9.490e-01 35.154 < 2e-16 \*\*\*  
## sqrt(nexis\_sum\_3) 3.311e-01 1.205e-01 2.747 0.00604 \*\*   
## Percent.Catholic 9.947e-02 2.247e-02 4.428 9.71e-06 \*\*\*  
## his\_prop 9.479e+00 2.310e+00 4.103 4.14e-05 \*\*\*  
## pop\_dens 2.994e+03 1.738e+02 17.221 < 2e-16 \*\*\*  
## income\_per\_capita 5.732e-05 2.064e-05 2.777 0.00550 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.17 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.1057, Adjusted R-squared: 0.1048   
## F-statistic: 122.8 on 5 and 5194 DF, p-value: < 2.2e-16

### TOTAL ENROLLMENT - Log + 1 Transformation

enroll\_total <- lm(log(EFTOTLT+1) ~ log(nexis\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_total)

##   
## Call:  
## lm(formula = log(EFTOTLT + 1) ~ log(nexis\_sum\_3 + 1) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.7496 -0.4004 0.2978 0.7127 2.5193   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.693e+00 5.655e-02 118.364 < 2e-16 \*\*\*  
## log(nexis\_sum\_3 + 1) 5.748e-02 1.402e-02 4.101 4.18e-05 \*\*\*  
## Percent.Catholic 7.648e-03 1.337e-03 5.723 1.11e-08 \*\*\*  
## his\_prop 3.697e-01 1.374e-01 2.691 0.00715 \*\*   
## pop\_dens 1.318e+02 1.041e+01 12.664 < 2e-16 \*\*\*  
## income\_per\_capita 1.706e-06 1.261e-06 1.353 0.17602   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.201 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.07485, Adjusted R-squared: 0.07396   
## F-statistic: 84.04 on 5 and 5194 DF, p-value: < 2.2e-16

### GRADUATE ENROLLMENT - Log + 1 Transformation

enroll\_grad <- lm(log(EFTOTLT+1) ~ log(nexis\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])   
  
summary(enroll\_grad)

##   
## Call:  
## lm(formula = log(EFTOTLT + 1) ~ log(nexis\_sum\_3 + 1) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.5419 -0.8739 0.8975 1.8297 4.1861   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.941e+00 1.260e-01 31.287 < 2e-16 \*\*\*  
## log(nexis\_sum\_3 + 1) 7.286e-02 3.122e-02 2.334 0.0197 \*   
## Percent.Catholic -1.414e-02 2.977e-03 -4.748 2.11e-06 \*\*\*  
## his\_prop 2.187e+00 3.060e-01 7.146 1.02e-12 \*\*\*  
## pop\_dens 3.486e+02 2.319e+01 15.033 < 2e-16 \*\*\*  
## income\_per\_capita 2.282e-05 2.808e-06 8.129 5.38e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.675 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.09888, Adjusted R-squared: 0.09802   
## F-statistic: 114 on 5 and 5194 DF, p-value: < 2.2e-16

### UNDERGRADUATE ENROLLMENT - Log + 1 Transformation

enroll\_undergrad <- lm(log(EFTOTLT+1) ~ log(nexis\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = log(EFTOTLT + 1) ~ log(nexis\_sum\_3 + 1) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.7496 -0.4004 0.2978 0.7127 2.5193   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.693e+00 5.655e-02 118.364 < 2e-16 \*\*\*  
## log(nexis\_sum\_3 + 1) 5.748e-02 1.402e-02 4.101 4.18e-05 \*\*\*  
## Percent.Catholic 7.648e-03 1.337e-03 5.723 1.11e-08 \*\*\*  
## his\_prop 3.697e-01 1.374e-01 2.691 0.00715 \*\*   
## pop\_dens 1.318e+02 1.041e+01 12.664 < 2e-16 \*\*\*  
## income\_per\_capita 1.706e-06 1.261e-06 1.353 0.17602   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.201 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.07485, Adjusted R-squared: 0.07396   
## F-statistic: 84.04 on 5 and 5194 DF, p-value: < 2.2e-16

# Bishop Accountability Percent change

### TOTAL ENROLLMENT

enroll\_total <- lm(EFTOTLT\_3\_change ~ bishop\_pub\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Total", ])   
  
summary(enroll\_total)

##   
## Call:  
## lm(formula = EFTOTLT\_3\_change ~ bishop\_pub\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Total", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -99.15 -10.07 -2.99 5.54 968.27   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.093e+01 1.392e+00 7.854 4.84e-15 \*\*\*  
## bishop\_pub\_sum\_3 1.358e-01 3.096e-02 4.387 1.17e-05 \*\*\*  
## Percent.Catholic -5.818e-02 3.357e-02 -1.733 0.0832 .   
## his\_prop -4.251e-01 3.427e+00 -0.124 0.9013   
## pop\_dens -3.249e+02 2.551e+02 -1.274 0.2028   
## income\_per\_capita -1.202e-04 2.921e-05 -4.114 3.95e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 29.98 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.00717, Adjusted R-squared: 0.006214   
## F-statistic: 7.502 on 5 and 5194 DF, p-value: 5.011e-07

### GRADUATE ENROLLMENT

enroll\_grad <- lm(EFTOTLT\_3\_change ~ bishop\_pub\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])   
  
summary(enroll\_grad)

##   
## Call:  
## lm(formula = EFTOTLT\_3\_change ~ bishop\_pub\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -138.40 -26.08 -15.43 3.40 2869.94   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.758e+01 4.161e+00 9.032 < 2e-16 \*\*\*  
## bishop\_pub\_sum\_3 -5.043e-02 9.256e-02 -0.545 0.5859   
## Percent.Catholic 1.750e-01 1.004e-01 1.743 0.0814 .   
## his\_prop -1.191e+01 1.024e+01 -1.162 0.2451   
## pop\_dens -1.057e+03 7.625e+02 -1.386 0.1658   
## income\_per\_capita -5.519e-04 8.732e-05 -6.321 2.82e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 89.64 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.009239, Adjusted R-squared: 0.008285   
## F-statistic: 9.687 on 5 and 5194 DF, p-value: 3.197e-09

### UNDERGRADUATE ENROLLMENT

enroll\_undergrad <- lm(EFTOTLT\_3\_change ~ bishop\_pub\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = EFTOTLT\_3\_change ~ bishop\_pub\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -97.70 -10.44 -3.01 5.20 969.62   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.911e+00 1.495e+00 5.962 2.66e-09 \*\*\*  
## bishop\_pub\_sum\_3 1.313e-01 3.325e-02 3.950 7.92e-05 \*\*\*  
## Percent.Catholic -7.553e-02 3.606e-02 -2.094 0.03627 \*   
## his\_prop 1.322e+00 3.680e+00 0.359 0.71942   
## pop\_dens -2.510e+02 2.740e+02 -0.916 0.35950   
## income\_per\_capita -8.274e-05 3.137e-05 -2.637 0.00838 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 32.2 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.004785, Adjusted R-squared: 0.003827   
## F-statistic: 4.994 on 5 and 5194 DF, p-value: 0.0001444

### TOTAL ENROLLMENT - Square Root Transformation

enroll\_total <- lm(sqrt(EFTOTLT\_3\_change) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])

## Warning in sqrt(EFTOTLT\_3\_change): NaNs produced

summary(enroll\_total)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT\_3\_change) ~ sqrt(bishop\_pub\_sum\_3) +   
## Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita,   
## data = data[data$LINE == "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.6153 -1.4093 -0.3906 0.8068 27.7192   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.728e+00 1.411e-01 26.427 < 2e-16 \*\*\*  
## sqrt(bishop\_pub\_sum\_3) 5.701e-02 2.610e-02 2.184 0.029054 \*   
## Percent.Catholic -8.579e-04 3.636e-03 -0.236 0.813507   
## his\_prop 3.693e-02 3.593e-01 0.103 0.918142   
## pop\_dens -1.059e+02 2.397e+01 -4.418 1.03e-05 \*\*\*  
## income\_per\_capita -1.102e-05 3.077e-06 -3.582 0.000347 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.304 on 2857 degrees of freedom  
## (2362 observations deleted due to missingness)  
## Multiple R-squared: 0.01301, Adjusted R-squared: 0.01128   
## F-statistic: 7.531 on 5 and 2857 DF, p-value: 4.917e-07

### GRADUATE ENROLLMENT - Square Root Transformation

enroll\_grad <- lm(sqrt(EFTOTLT\_3\_change) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])

## Warning in sqrt(EFTOTLT\_3\_change): NaNs produced

summary(enroll\_grad)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT\_3\_change) ~ sqrt(bishop\_pub\_sum\_3) +   
## Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita,   
## data = data[data$LINE == "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.323 -3.482 -0.747 1.906 49.647   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.827e+00 2.377e-01 20.309 < 2e-16 \*\*\*  
## sqrt(bishop\_pub\_sum\_3) 2.078e-02 4.434e-02 0.469 0.6393   
## Percent.Catholic -3.754e-03 5.835e-03 -0.643 0.5200   
## his\_prop -3.539e-01 5.862e-01 -0.604 0.5461   
## pop\_dens 1.111e+02 4.914e+01 2.262 0.0238 \*   
## income\_per\_capita -2.637e-05 5.330e-06 -4.949 7.82e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.331 on 3597 degrees of freedom  
## (1622 observations deleted due to missingness)  
## Multiple R-squared: 0.008844, Adjusted R-squared: 0.007466   
## F-statistic: 6.419 on 5 and 3597 DF, p-value: 6.03e-06

### UNDERGRADUATE ENROLLMENT - Square Root Transformation

enroll\_undergrad <- lm(sqrt(EFTOTLT\_3\_change) ~ sqrt(bishop\_pub\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])

## Warning in sqrt(EFTOTLT\_3\_change): NaNs produced

summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT\_3\_change) ~ sqrt(bishop\_pub\_sum\_3) +   
## Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita,   
## data = data[data$LINE == "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.6153 -1.4093 -0.3906 0.8068 27.7192   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.728e+00 1.411e-01 26.427 < 2e-16 \*\*\*  
## sqrt(bishop\_pub\_sum\_3) 5.701e-02 2.610e-02 2.184 0.029054 \*   
## Percent.Catholic -8.579e-04 3.636e-03 -0.236 0.813507   
## his\_prop 3.693e-02 3.593e-01 0.103 0.918142   
## pop\_dens -1.059e+02 2.397e+01 -4.418 1.03e-05 \*\*\*  
## income\_per\_capita -1.102e-05 3.077e-06 -3.582 0.000347 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.304 on 2857 degrees of freedom  
## (2362 observations deleted due to missingness)  
## Multiple R-squared: 0.01301, Adjusted R-squared: 0.01128   
## F-statistic: 7.531 on 5 and 2857 DF, p-value: 4.917e-07

### TOTAL ENROLLMENT - Log+1 Transformation

enroll\_total <- lm(log(EFTOTLT\_3\_change+1) ~ log(bishop\_pub\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])

## Warning in log(EFTOTLT\_3\_change + 1): NaNs produced

summary(enroll\_total)

##   
## Call:  
## lm(formula = log(EFTOTLT\_3\_change + 1) ~ log(bishop\_pub\_sum\_3 +   
## 1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita,   
## data = data[data$LINE == "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.0226 -0.6795 0.1328 0.8100 4.7676   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.254e+00 7.801e-02 28.894 < 2e-16 \*\*\*  
## log(bishop\_pub\_sum\_3 + 1) 4.943e-02 2.457e-02 2.011 0.044362 \*   
## Percent.Catholic -1.562e-03 2.006e-03 -0.779 0.436314   
## his\_prop 1.492e-01 1.989e-01 0.750 0.453264   
## pop\_dens -4.277e+01 1.335e+01 -3.204 0.001371 \*\*   
## income\_per\_capita -6.436e-06 1.697e-06 -3.793 0.000151 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.315 on 3056 degrees of freedom  
## (2163 observations deleted due to missingness)  
## Multiple R-squared: 0.00881, Adjusted R-squared: 0.007188   
## F-statistic: 5.433 on 5 and 3056 DF, p-value: 5.562e-05

### GRADUATE ENROLLMENT - Log+1 Transformation

enroll\_grad <- lm(log(EFTOTLT\_3\_change+1) ~ log(bishop\_pub\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])

## Warning in log(EFTOTLT\_3\_change + 1): NaNs produced

summary(enroll\_grad)

##   
## Call:  
## lm(formula = log(EFTOTLT\_3\_change + 1) ~ log(bishop\_pub\_sum\_3 +   
## 1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita,   
## data = data[data$LINE == "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.6388 -1.9134 0.1939 1.4204 5.9198   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.331e+00 9.705e-02 24.018 < 2e-16 \*\*\*  
## log(bishop\_pub\_sum\_3 + 1) 7.919e-02 3.059e-02 2.589 0.00967 \*\*   
## Percent.Catholic -4.802e-03 2.394e-03 -2.006 0.04489 \*   
## his\_prop 2.707e-01 2.387e-01 1.134 0.25685   
## pop\_dens 7.376e+01 2.007e+01 3.676 0.00024 \*\*\*  
## income\_per\_capita -8.562e-06 2.180e-06 -3.928 8.73e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.787 on 3676 degrees of freedom  
## (1543 observations deleted due to missingness)  
## Multiple R-squared: 0.01159, Adjusted R-squared: 0.01024   
## F-statistic: 8.618 on 5 and 3676 DF, p-value: 3.95e-08

### UNDERGRADUATE ENROLLMENT - Log+1 Transformation

enroll\_undergrad <- lm(log(EFTOTLT\_3\_change+1) ~ log(bishop\_pub\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])

## Warning in log(EFTOTLT\_3\_change + 1): NaNs produced

summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = log(EFTOTLT\_3\_change + 1) ~ log(bishop\_pub\_sum\_3 +   
## 1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita,   
## data = data[data$LINE == "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.0226 -0.6795 0.1328 0.8100 4.7676   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.254e+00 7.801e-02 28.894 < 2e-16 \*\*\*  
## log(bishop\_pub\_sum\_3 + 1) 4.943e-02 2.457e-02 2.011 0.044362 \*   
## Percent.Catholic -1.562e-03 2.006e-03 -0.779 0.436314   
## his\_prop 1.492e-01 1.989e-01 0.750 0.453264   
## pop\_dens -4.277e+01 1.335e+01 -3.204 0.001371 \*\*   
## income\_per\_capita -6.436e-06 1.697e-06 -3.793 0.000151 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.315 on 3056 degrees of freedom  
## (2163 observations deleted due to missingness)  
## Multiple R-squared: 0.00881, Adjusted R-squared: 0.007188   
## F-statistic: 5.433 on 5 and 3056 DF, p-value: 5.562e-05

# Percent Change, Nexis

### TOTAL ENROLLMENT

enroll\_total <- lm(EFTOTLT\_3\_change ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Total", ])   
  
summary(enroll\_total)

##   
## Call:  
## lm(formula = EFTOTLT\_3\_change ~ nexis\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Total", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -99.41 -10.13 -3.15 5.44 968.42   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.096e+01 1.398e+00 7.839 5.49e-15 \*\*\*  
## nexis\_sum\_3 4.379e-02 1.333e-02 3.285 0.00103 \*\*   
## Percent.Catholic -4.430e-02 3.331e-02 -1.330 0.18362   
## his\_prop -6.246e-01 3.440e+00 -0.182 0.85591   
## pop\_dens -3.314e+02 2.556e+02 -1.297 0.19475   
## income\_per\_capita -1.196e-04 2.936e-05 -4.073 4.71e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 30.01 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.005558, Adjusted R-squared: 0.0046   
## F-statistic: 5.805 on 5 and 5194 DF, p-value: 2.367e-05

### GRADUATE ENROLLMENT

enroll\_grad <- lm(EFTOTLT\_3\_change ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])   
  
summary(enroll\_grad)

##   
## Call:  
## lm(formula = EFTOTLT\_3\_change ~ nexis\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -137.91 -26.08 -15.45 3.35 2870.08   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.763e+01 4.176e+00 9.011 < 2e-16 \*\*\*  
## nexis\_sum\_3 -1.141e-02 3.982e-02 -0.286 0.7745   
## Percent.Catholic 1.688e-01 9.951e-02 1.697 0.0898 .   
## his\_prop -1.196e+01 1.028e+01 -1.164 0.2446   
## pop\_dens -1.060e+03 7.635e+02 -1.389 0.1649   
## income\_per\_capita -5.537e-04 8.772e-05 -6.313 2.97e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 89.64 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.009198, Adjusted R-squared: 0.008244   
## F-statistic: 9.644 on 5 and 5194 DF, p-value: 3.537e-09

### UNDERGRADUATE ENROLLMENT

enroll\_undergrad <- lm(EFTOTLT\_3\_change ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])   
  
summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = EFTOTLT\_3\_change ~ nexis\_sum\_3 + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -97.95 -10.39 -3.00 5.02 969.76   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.958e+00 1.501e+00 5.967 2.57e-09 \*\*\*  
## nexis\_sum\_3 4.399e-02 1.431e-02 3.074 0.00212 \*\*   
## Percent.Catholic -6.244e-02 3.577e-02 -1.745 0.08096 .   
## his\_prop 1.088e+00 3.694e+00 0.295 0.76838   
## pop\_dens -2.594e+02 2.744e+02 -0.945 0.34454   
## income\_per\_capita -8.274e-05 3.153e-05 -2.624 0.00871 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 32.22 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.003608, Adjusted R-squared: 0.002648   
## F-statistic: 3.761 on 5 and 5194 DF, p-value: 0.002115

### TOTAL ENROLLMENT - Square Root Transformation - this does not work for negative percent changes.

enroll\_total <- lm(sqrt(EFTOTLT\_3\_change) ~ sqrt(nexis\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])

## Warning in sqrt(EFTOTLT\_3\_change): NaNs produced

summary(enroll\_total)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT\_3\_change) ~ sqrt(nexis\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.6287 -1.3964 -0.3912 0.8114 27.7469   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.784e+00 1.433e-01 26.405 < 2e-16 \*\*\*  
## sqrt(nexis\_sum\_3) 3.639e-02 1.724e-02 2.111 0.034889 \*   
## Percent.Catholic 3.633e-05 3.573e-03 0.010 0.991890   
## his\_prop 5.392e-03 3.606e-01 0.015 0.988072   
## pop\_dens -1.087e+02 2.418e+01 -4.497 7.18e-06 \*\*\*  
## income\_per\_capita -1.192e-05 3.184e-06 -3.744 0.000185 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.304 on 2857 degrees of freedom  
## (2362 observations deleted due to missingness)  
## Multiple R-squared: 0.0129, Adjusted R-squared: 0.01117   
## F-statistic: 7.467 on 5 and 2857 DF, p-value: 5.685e-07

### GRADUATE ENROLLMENT - Square Root Transformation

enroll\_grad <- lm(sqrt(EFTOTLT\_3\_change) ~ sqrt(nexis\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])

## Warning in sqrt(EFTOTLT\_3\_change): NaNs produced

summary(enroll\_grad)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT\_3\_change) ~ sqrt(nexis\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.257 -3.489 -0.738 1.914 49.640   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.799e+00 2.436e-01 19.702 < 2e-16 \*\*\*  
## sqrt(nexis\_sum\_3) -1.484e-02 3.146e-02 -0.472 0.6372   
## Percent.Catholic -2.914e-03 5.746e-03 -0.507 0.6120   
## his\_prop -2.944e-01 5.891e-01 -0.500 0.6173   
## pop\_dens 1.173e+02 4.949e+01 2.369 0.0179 \*   
## income\_per\_capita -2.493e-05 5.570e-06 -4.476 7.82e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.331 on 3597 degrees of freedom  
## (1622 observations deleted due to missingness)  
## Multiple R-squared: 0.008845, Adjusted R-squared: 0.007467   
## F-statistic: 6.42 on 5 and 3597 DF, p-value: 6.023e-06

### UNDERGRADUATE ENROLLMENT - Square Root Transformation

enroll\_undergrad <- lm(sqrt(EFTOTLT\_3\_change) ~ sqrt(nexis\_sum\_3) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])

## Warning in sqrt(EFTOTLT\_3\_change): NaNs produced

summary(enroll\_undergrad)

##   
## Call:  
## lm(formula = sqrt(EFTOTLT\_3\_change) ~ sqrt(nexis\_sum\_3) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.6287 -1.3964 -0.3912 0.8114 27.7469   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.784e+00 1.433e-01 26.405 < 2e-16 \*\*\*  
## sqrt(nexis\_sum\_3) 3.639e-02 1.724e-02 2.111 0.034889 \*   
## Percent.Catholic 3.633e-05 3.573e-03 0.010 0.991890   
## his\_prop 5.392e-03 3.606e-01 0.015 0.988072   
## pop\_dens -1.087e+02 2.418e+01 -4.497 7.18e-06 \*\*\*  
## income\_per\_capita -1.192e-05 3.184e-06 -3.744 0.000185 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.304 on 2857 degrees of freedom  
## (2362 observations deleted due to missingness)  
## Multiple R-squared: 0.0129, Adjusted R-squared: 0.01117   
## F-statistic: 7.467 on 5 and 2857 DF, p-value: 5.685e-07

### TOTAL ENROLLMENT - Log

enroll\_total <- lm(log(EFTOTLT\_3\_change+1) ~ log(nexis\_sum\_3+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])

## Warning in log(EFTOTLT\_3\_change + 1): NaNs produced

summary(enroll\_total)

##   
## Call:  
## lm(formula = log(EFTOTLT\_3\_change + 1) ~ log(nexis\_sum\_3 + 1) +   
## Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita,   
## data = data[data$LINE == "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.0880 -0.6769 0.1325 0.8139 4.7906   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.296e+00 7.923e-02 28.985 < 2e-16 \*\*\*  
## log(nexis\_sum\_3 + 1) 4.515e-02 1.956e-02 2.308 0.02106 \*   
## Percent.Catholic -1.143e-03 1.968e-03 -0.581 0.56150   
## his\_prop 1.351e-01 1.992e-01 0.678 0.49773   
## pop\_dens -4.566e+01 1.353e+01 -3.374 0.00075 \*\*\*  
## income\_per\_capita -7.367e-06 1.801e-06 -4.090 4.42e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.315 on 3056 degrees of freedom  
## (2163 observations deleted due to missingness)  
## Multiple R-squared: 0.009225, Adjusted R-squared: 0.007604   
## F-statistic: 5.691 on 5 and 3056 DF, p-value: 3.128e-05

### GRADUATE ENROLLMENT - Log

enroll\_grad <- lm(log(EFTOTLT+1) ~ log(EFTOTLT\_3\_change+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Graduate", ])

## Warning in log(EFTOTLT\_3\_change + 1): NaNs produced

summary(enroll\_grad)

##   
## Call:  
## lm(formula = log(EFTOTLT + 1) ~ log(EFTOTLT\_3\_change + 1) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Graduate", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7.420 -2.038 -0.066 1.706 11.828   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.434e+00 1.389e-01 10.318 < 2e-16 \*\*\*  
## log(EFTOTLT\_3\_change + 1) 9.508e-01 2.193e-02 43.353 < 2e-16 \*\*\*  
## Percent.Catholic -1.327e-02 3.111e-03 -4.265 2.04e-05 \*\*\*  
## his\_prop 2.203e+00 3.168e-01 6.953 4.20e-12 \*\*\*  
## pop\_dens 3.953e+02 2.653e+01 14.897 < 2e-16 \*\*\*  
## income\_per\_capita 2.608e-05 2.832e-06 9.208 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.378 on 3676 degrees of freedom  
## (1543 observations deleted due to missingness)  
## Multiple R-squared: 0.4012, Adjusted R-squared: 0.4004   
## F-statistic: 492.5 on 5 and 3676 DF, p-value: < 2.2e-16

### UNDERGRADUATE ENROLLMENT - Log

enroll\_undergrad <- lm(log(EFTOTLT+1) ~ log(EFTOTLT\_3\_change+1) + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Undergraduates", ])

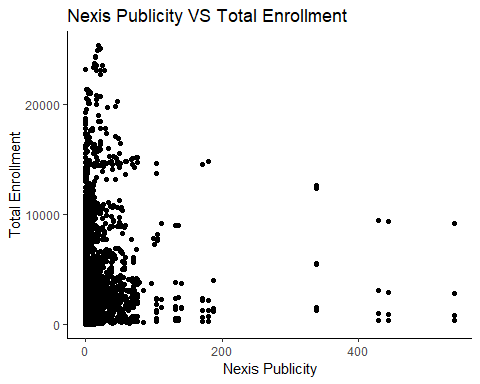
## Warning in log(EFTOTLT\_3\_change + 1): NaNs produced

summary(enroll\_undergrad)

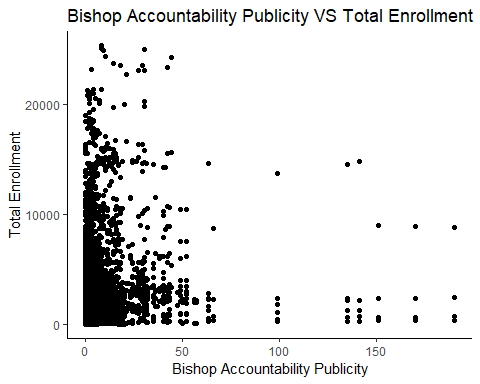
##   
## Call:  
## lm(formula = log(EFTOTLT + 1) ~ log(EFTOTLT\_3\_change + 1) + Percent.Catholic +   
## his\_prop + pop\_dens + income\_per\_capita, data = data[data$LINE ==   
## "Undergraduates", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.2078 -0.4020 0.2272 0.6845 2.7229   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.432e+00 7.472e-02 99.470 < 2e-16 \*\*\*  
## log(EFTOTLT\_3\_change + 1) -2.730e-01 1.535e-02 -17.785 < 2e-16 \*\*\*  
## Percent.Catholic 4.410e-03 1.663e-03 2.652 0.00803 \*\*   
## his\_prop 4.941e-01 1.687e-01 2.929 0.00343 \*\*   
## pop\_dens 1.138e+02 1.122e+01 10.140 < 2e-16 \*\*\*  
## income\_per\_capita 3.805e-06 1.417e-06 2.686 0.00727 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.117 on 3056 degrees of freedom  
## (2163 observations deleted due to missingness)  
## Multiple R-squared: 0.1595, Adjusted R-squared: 0.1581   
## F-statistic: 115.9 on 5 and 3056 DF, p-value: < 2.2e-16

# Other Things I Have Tried

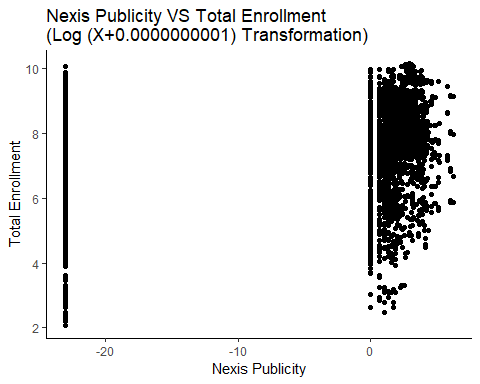
library(ggplot2)  
  
ggplot(data[data$LINE == "Total", ], aes(x= nexis\_sum\_3, y = EFTOTLT))+  
 geom\_point() +  
 theme\_classic() +  
 labs(title = "Nexis Publicity VS Total Enrollment", x = "Nexis Publicity", y = "Total Enrollment")



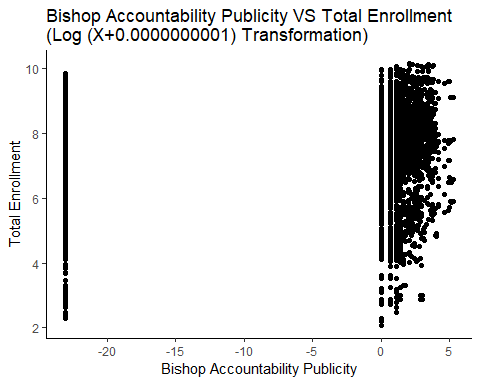
ggplot(data[data$LINE == "Total", ], aes(x= bishop\_pub\_sum\_3, y = EFTOTLT))+  
 geom\_point()+  
 theme\_classic()+  
 labs(title = "Bishop Accountability Publicity VS Total Enrollment", x = "Bishop Accountability Publicity", y = "Total Enrollment")



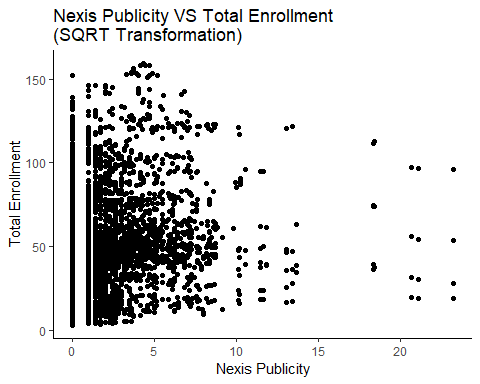
ggplot(data[data$LINE == "Total", ], aes(x= log(nexis\_sum\_3+0.0000000001), y = log(EFTOTLT+0.0000000001)))+  
 geom\_point() +  
 theme\_classic() +  
 labs(title = "Nexis Publicity VS Total Enrollment \n(Log (X+0.0000000001) Transformation)", x = "Nexis Publicity", y = "Total Enrollment")



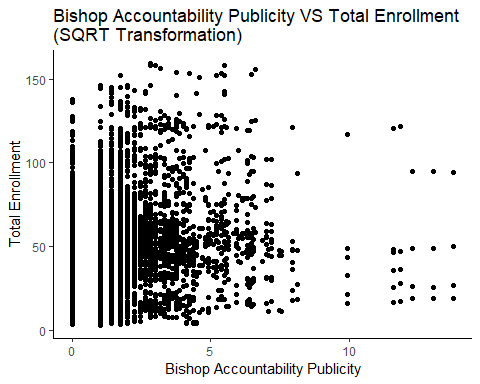
ggplot(data[data$LINE == "Total", ], aes(x= log(bishop\_pub\_sum\_3+0.0000000001), y = log(EFTOTLT+0.0000000001)))+  
 geom\_point()+  
 theme\_classic()+  
 labs(title = "Bishop Accountability Publicity VS Total Enrollment \n(Log (X+0.0000000001) Transformation)", x = "Bishop Accountability Publicity", y = "Total Enrollment")



ggplot(data[data$LINE == "Total", ], aes(x= sqrt(nexis\_sum\_3), y = sqrt(EFTOTLT)))+  
 geom\_point() +  
 theme\_classic() +  
 labs(title = "Nexis Publicity VS Total Enrollment \n(SQRT Transformation)", x = "Nexis Publicity", y = "Total Enrollment")



ggplot(data[data$LINE == "Total", ], aes(x= sqrt(bishop\_pub\_sum\_3), y = sqrt(EFTOTLT)))+  
 geom\_point()+  
 theme\_classic()+  
 labs(title = "Bishop Accountability Publicity VS Total Enrollment \n(SQRT Transformation)", x = "Bishop Accountability Publicity", y = "Total Enrollment")



zip model

M1 <- glm(EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Total", ],  
 family = 'poisson')  
  
summary(M1)

##   
## Call:  
## glm(formula = EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop +   
## pop\_dens + income\_per\_capita, family = "poisson", data = data[data$LINE ==   
## "Total", ])  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -137.53 -39.78 -14.82 10.18 224.36   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 7.565e+00 8.323e-04 9089.6 <2e-16 \*\*\*  
## nexis\_sum\_3 2.920e-04 6.775e-06 43.1 <2e-16 \*\*\*  
## Percent.Catholic 2.402e-03 2.012e-05 119.4 <2e-16 \*\*\*  
## his\_prop 5.795e-01 1.894e-03 305.9 <2e-16 \*\*\*  
## pop\_dens 1.044e+02 1.035e-01 1008.3 <2e-16 \*\*\*  
## income\_per\_capita 6.203e-06 1.556e-08 398.7 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for poisson family taken to be 1)  
##   
## Null deviance: 15638012 on 5199 degrees of freedom  
## Residual deviance: 13601844 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## AIC: 13650257  
##   
## Number of Fisher Scoring iterations: 5

E2 <- resid(M1, type = "pearson")  
N <- nrow(data)  
p <- length(coef(M1))   
sum(E2^2) / (N - p)

## [1] 1064.226

#overdispersion

library(MASS)  
M2 <- glm.nb(EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita, data=data[data$LINE == "Total",])  
summary(M2)

##   
## Call:  
## glm.nb(formula = EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop +   
## pop\_dens + income\_per\_capita, data = data[data$LINE == "Total",   
## ], init.theta = 1.055981708, link = log)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -3.1065 -0.8475 -0.2518 0.2072 2.7355   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 7.469e+00 4.534e-02 164.718 < 2e-16 \*\*\*  
## nexis\_sum\_3 4.940e-04 4.323e-04 1.143 0.25311   
## Percent.Catholic 3.604e-03 1.080e-03 3.336 0.00085 \*\*\*  
## his\_prop 8.939e-01 1.116e-01 8.013 1.12e-15 \*\*\*  
## pop\_dens 1.695e+02 8.288e+00 20.447 < 2e-16 \*\*\*  
## income\_per\_capita 4.765e-06 9.524e-07 5.003 5.65e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(1.056) family taken to be 1)  
##   
## Null deviance: 6707.8 on 5199 degrees of freedom  
## Residual deviance: 5964.1 on 5194 degrees of freedom  
## (25 observations deleted due to missingness)  
## AIC: 93747  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 1.0560   
## Std. Err.: 0.0184   
##   
## 2 x log-likelihood: -93733.3000

E2 <- resid(M2, type = "pearson")  
N <- nrow(data)  
p <- length(coef(M2)) + 1 # '+1' is for variance parameter in NB  
sum(E2^2) / (N - p)

## [1] 0.3397552

#underdispersion?

#M3 <- zeroinfl(EFTOTLT ~ nexis\_sum\_3 + Percent.Catholic + his\_prop + pop\_dens + income\_per\_capita,  
 # dist = 'poisson',  
 # data=data[data$LINE == "Total",])  
  
#summary(M3)