Final Project Paper

Ekaterina Hofrenning, Lauren Lazaro, Arthur Starodynov, Yuki Low

## Abstract (condenses a brief introduction, brief description of methods, and main results into a one-paragraph summary)

## Introduction (brief context and background of the problem)

It is well-documented that educational success is influenced by a multitude of factors that extend beyond one’s academic preparation. Our study aims to predict math, reading and writing scores based on personal and socio-economic variables. The dataset used for our study provides us with a diverse range of variables and gives us the opportunity to uncover patterns that influence a student’s test outcome.

## Methods (data description and statistical methods)

The dataset is a sample of 948 individuals from a public school with math, reading, and writing test scores (ranging from 0-100) and the following variables of interest: gender, ethnic group, parental education, lunch type, test prep, parent’s marital status, sports participation, first child, number of siblings, means of transportation to school, number of weekly study hours. Due to significant missing data, we imputed the missing values with the average or most common value per variable.

Math, reading and writing tests scores were the response variables and were scored from 0 to 100. Math test scores ranged from 0 to 100 with a mean score of 65.9821 and a median score of 66. Reading test scores ranged from 17 to 100 with a mean score of 68.8418 and a median score of 69.5. Writing test scores ranged from 10 to 100 with a mean score of 67.9293 and a median score of 68. The histograms of the distributions of math, reading and writing tests scores are slightly skewed to the left. In order to make the distributions more normal, we attempted both logarithmic transformations and square root transformations in which both of these types of transformations resulted in more severely skewed distributions. In the end, we decided to proceed with using no transformation on the test score variables.

We conducted a data description (Table 1), computing the averages and standard deviations for continuous variables and counts and percentages for categorical variables. Our sample is well-balanced across sex. The sample is composed of majority first born children, children on standard fee lunch, children with married parents, school bus riders, and children who study 5-10 hours. The average number of siblings is 2.

| **Characteristic** | **N = 948**1 |
| --- | --- |
| Gender |  |
| female | 488 / 948 (51%) |
| male | 460 / 948 (49%) |
| EthnicGroup |  |
| group A | 80 / 948 (8.4%) |
| group B | 171 / 948 (18%) |
| group C | 336 / 948 (35%) |
| group D | 237 / 948 (25%) |
| group E | 124 / 948 (13%) |
| ParentEduc |  |
| some high school | 163 / 948 (17%) |
| high school | 176 / 948 (19%) |
| associate's degree | 198 / 948 (21%) |
| some college | 252 / 948 (27%) |
| bachelor's degree | 104 / 948 (11%) |
| master's degree | 55 / 948 (5.8%) |
| LunchType |  |
| free/reduced | 331 / 948 (35%) |
| standard | 617 / 948 (65%) |
| TestPrep |  |
| completed | 322 / 948 (34%) |
| none | 626 / 948 (66%) |
| ParentMaritalStatus |  |
| divorced | 146 / 948 (15%) |
| married | 565 / 948 (60%) |
| single | 213 / 948 (22%) |
| widowed | 24 / 948 (2.5%) |
| PracticeSport |  |
| never | 112 / 948 (12%) |
| sometimes | 493 / 948 (52%) |
| regularly | 343 / 948 (36%) |
| IsFirstChild | 634 / 948 (67%) |
| NrSiblings | 2 (1) |
| TransportMeans |  |
| private | 337 / 948 (36%) |
| school\_bus | 611 / 948 (64%) |
| WklyStudyHours |  |
| < 5 | 253 / 948 (27%) |
| 5-10 | 545 / 948 (57%) |
| > 10 | 150 / 948 (16%) |
| MathScore | 66 (16) |
| ReadingScore | 69 (15) |
| WritingScore | 68 (15) |
| 1n / N (%); Mean (SD) | |

After looking at the histograms of the continuous outcomes, we decided that the variables looked fairly normal so we did not use log transformations.

We employed a multi-faceted approach to model building. Initially, we fit simple linear regression models for each of the three scores, utilizing all available covariates in our dataset. Subsequently, we conducted a comprehensive exploration by implementing backward, forward, and stepwise regression techniques. This allowed us to assess how variables varied when adjusting for all covariates through multiple linear regression modeling. The combination of these methods provided a nuanced understanding of the relationships between predictors and scores, capturing both individual and collective effects.

The chosen model was then tested through model diagnostics and influential observation diagnostics to make sure all assumptions are met and no significant outliers were influencing the model.

## Results

From the SLR model, we found that the linear association between MathScore and the predictors LunchType, Gender, TestPrep and WklyStudyHours were significant. The average increase in MathScore was 8.7216 (p=3e-09) for students of ethnic group B and 3.4986 (p=0.00759) for students of ethnic group C compared to students of ethnic group A without adjusting for all other covariates. The average increase in MathScore between students with parents with some high school education and students with parents who completed their high school education was 6.0952 (p=0.000139) without adjusting for all other covariates. The average decrease in MathScore between students with parents with some high school education and students with parents who have their Master’s degree was 2.8793 (p=0.009353) without adjusting for all other covariates. The average increase in MathScore between students with divorced parents and students with married parents was 4.142 (p=0.040) without adjusting for all other covariates. The average decrease in MathScore was 3.6763 (p=0.041) for students with 1 sibling and 3.8757 (p=0.0390) for students with 2 siblings compared to only children without adjusting for all other covariates. Note that the linear association between MathScore and PracticeSport, IsFirstChild, and TransportMeans were not significant.

After fitting a MLR using forwards, backwards and stepwise regression model building methods, we noticed that the linear association between MathScore and EthnicGroup, ParentEduc, ParentMaritalStatus, PracticeSport and IsFirstChild became significant after adjusting for all other covariates. Note that the linear association between MathScore and LunchType, Gender, TestPrep and WklyStudyHours were still significant after adjusting for all other covariates. Therefore, for Mathscore as the outcome, the optimal predictors are Gender, EthnicGroup, ParentEduc, LunchType, TestPrep, ParentMaritalStatus, PracticeSport, IsFirstChild, and WklyStudyHours. Our estimates for coefficients for each of our significant covariates associated with MathScore are located in our Appendix in Table 2, 3 and 4. Note that each of the three methods selected the same results.

Table 2. Math score stepwise model

| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 55.976 | <.001 |
| Gendermale | 4.981 | <.001 |
| EthnicGroup.L | 7.401 | <.001 |
| EthnicGroup.Q | 2.940 | 0.011 |
| EthnicGroup.C | 0.737 | 0.482 |
| EthnicGroup^4 | -0.985 | 0.251 |
| ParentEduc.L | 6.642 | <.001 |
| ParentEduc.Q | -0.045 | 0.972 |
| ParentEduc.C | -0.300 | 0.806 |
| ParentEduc^4 | 1.402 | 0.213 |
| ParentEduc^5 | -2.644 | 0.006 |
| LunchTypestandard | 11.155 | <.001 |
| TestPrepnone | -5.582 | <.001 |
| ParentMaritalStatusmarried | 3.876 | 0.002 |
| ParentMaritalStatussingle | 1.102 | 0.444 |
| ParentMaritalStatuswidowed | 5.066 | 0.087 |
| PracticeSport.L | 2.514 | 0.015 |
| PracticeSport.Q | -0.518 | 0.503 |
| IsFirstChildyes | 2.368 | 0.011 |
| WklyStudyHours.L | 2.621 | 0.008 |
| WklyStudyHours.Q | -1.032 | 0.158 |

From the SLR model, we found that the linear association between ReadingScore and the predictors LunchType, Gender, TestPrep and IsFirstChild were significant. The average increase in ReadingScore between students of ethnic group A and ethnic group B was 5.95373 (p=2.64e-05) without adjusting for all other covariates. The average increase in ReadingScore between students with parents with some high school education and students with parents who completed their high school education was 8.2312 (p=5.85e-08) without adjusting for all other covariates. The average decrease in ReadingScore between students with parents with some high school education and students with parents who have their Master’s degree was 3.5466 (p=0.000717) without adjusting for all other covariates. The average increase in ReadingScore between students with divorced parents and students with married parents was 3.965 (p=0.039) without adjusting for all other covariates. The average increase in ReadingScore between students who studied less than 5 hours and students who studied 10-May was 2.6298 (p=0.0194) without adjusting for all other covariates. The average difference in ReadingScore between students who studied less than 5 hours and students who studied greater than 10 hours was insignificant (p=0.0541) without adjusting for all other covariates. Note that the linear association between ReadingScore and PracticeSport, NrSiblings, and TransportMeans were not significant.

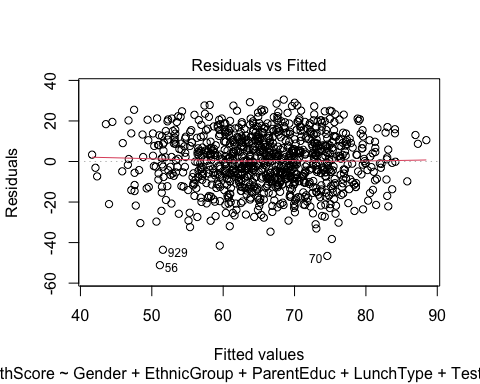
After fitting a MLR using forwards, backwards and stepwise regression model building methods, we noticed that the linear association between ReadingScore and EthnicGroup, ParentEduc, ParentMaritalStatus and WklyStudyHours became significant after adjusting for all other covariates. Note that the linear association between ReadingScore and Gender, LunchType, TestPrep and IsFirstChild were still significant after adjusting for all other covariates. Note that each of the three methods selected the same results. Therefore, for ReadingScore as the outcome, the optimal predictors are Gender, EthnicGroup, ParentEduc, LunchType, TestPrep, ParentMaritalStatus, IsFirstChild, and WklyStudyHours.

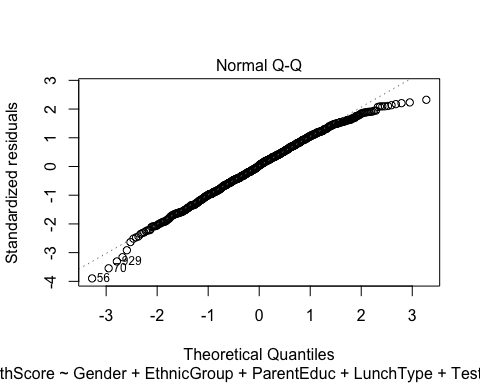
Table 3. Reading score stepwise model

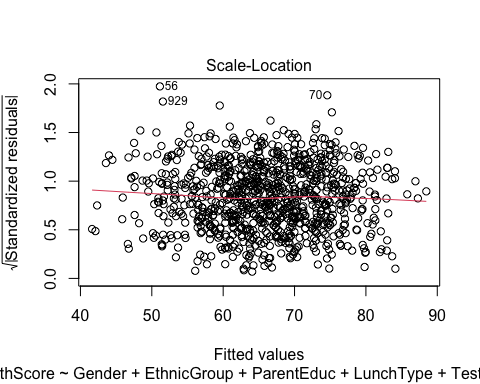
| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 68.495 | <.001 |
| Gendermale | -7.282 | <.001 |
| EthnicGroup.L | 4.149 | 0.001 |
| EthnicGroup.Q | 1.438 | 0.203 |
| EthnicGroup.C | -0.471 | 0.645 |
| EthnicGroup^4 | -0.942 | 0.261 |
| ParentEduc.L | 7.638 | <.001 |
| ParentEduc.Q | 1.535 | 0.22 |
| ParentEduc.C | 0.566 | 0.635 |
| ParentEduc^4 | 1.547 | 0.158 |
| ParentEduc^5 | -3.033 | 0.001 |
| LunchTypestandard | 7.494 | <.001 |
| TestPrepnone | -6.972 | <.001 |
| ParentMaritalStatusmarried | 4.113 | 0.001 |
| ParentMaritalStatussingle | 1.275 | 0.363 |
| ParentMaritalStatuswidowed | 4.645 | 0.106 |
| IsFirstChildyes | 2.446 | 0.007 |
| WklyStudyHours.L | 1.431 | 0.135 |
| WklyStudyHours.Q | -0.933 | 0.191 |

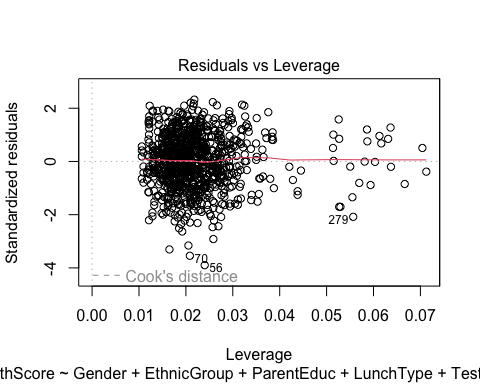
For writing score as the outcome, the optimal predictors are Gender, EthnicGroup, ParentEduc, LunchType, TestPrep, ParentMaritalStatus, PracticeSport, IsFirstChild, and WklyStudyHours.

Model diagnosis is an essential step in the modeling process. The models need to be checked through visual and statistical tests to make sure all assumptions are met. Through this process, we could evaluate if the models we fit well represented our study.









Start: AIC=4830.6  
WritingScore ~ Gender + EthnicGroup + ParentEduc + LunchType +   
 TestPrep + ParentMaritalStatus + PracticeSport + IsFirstChild +   
 NrSiblings + TransportMeans + WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
- NrSiblings 1 75.5 147549 4829.1  
- TransportMeans 1 150.6 147624 4829.6  
<none> 147473 4830.6  
- WklyStudyHours 2 769.3 148243 4831.5  
- PracticeSport 2 840.3 148314 4832.0  
- IsFirstChild 1 1057.3 148531 4835.4  
- ParentMaritalStatus 3 2701.0 150174 4841.8  
- EthnicGroup 4 5141.7 152615 4855.1  
- ParentEduc 5 11724.7 159198 4893.1  
- LunchType 1 14984.1 162457 4920.3  
- TestPrep 1 19289.0 166762 4945.1  
- Gender 1 19644.7 167118 4947.2  
  
Step: AIC=4829.09  
WritingScore ~ Gender + EthnicGroup + ParentEduc + LunchType +   
 TestPrep + ParentMaritalStatus + PracticeSport + IsFirstChild +   
 TransportMeans + WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
- TransportMeans 1 148.6 147698 4828.0  
<none> 147549 4829.1  
- WklyStudyHours 2 791.7 148341 4830.2  
- PracticeSport 2 837.9 148387 4830.5  
- IsFirstChild 1 1000.7 148550 4833.5  
- ParentMaritalStatus 3 2719.9 150269 4840.4  
- EthnicGroup 4 5128.8 152678 4853.5  
- ParentEduc 5 11676.1 159225 4891.3  
- LunchType 1 14922.6 162472 4918.4  
- TestPrep 1 19298.9 166848 4943.6  
- Gender 1 19731.7 167281 4946.1  
  
Step: AIC=4828.04  
WritingScore ~ Gender + EthnicGroup + ParentEduc + LunchType +   
 TestPrep + ParentMaritalStatus + PracticeSport + IsFirstChild +   
 WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
<none> 147698 4828.0  
- WklyStudyHours 2 784.4 148482 4829.1  
- PracticeSport 2 837.9 148535 4829.4  
- IsFirstChild 1 1006.8 148704 4832.5  
- ParentMaritalStatus 3 2765.8 150463 4839.6  
- EthnicGroup 4 5112.3 152810 4852.3  
- ParentEduc 5 11814.5 159512 4891.0  
- LunchType 1 14933.9 162631 4917.4  
- TestPrep 1 19221.8 166919 4942.0  
- Gender 1 19738.0 167436 4945.0

Call:  
lm(formula = WritingScore ~ Gender + EthnicGroup + ParentEduc +   
 LunchType + TestPrep + ParentMaritalStatus + PracticeSport +   
 IsFirstChild + WklyStudyHours, data = data\_cleaned %>% select(-c(MathScore,   
 ReadingScore)))  
  
Coefficients:  
 (Intercept) Gendermale   
 69.3473 -9.2094   
 EthnicGroup.L EthnicGroup.Q   
 4.6731 0.6274   
 EthnicGroup.C EthnicGroup^4   
 -1.8905 -1.6492   
 ParentEduc.L ParentEduc.Q   
 9.9826 1.3651   
 ParentEduc.C ParentEduc^4   
 0.2894 1.7154   
 ParentEduc^5 LunchTypestandard   
 -3.1061 8.3876   
 TestPrepnone ParentMaritalStatusmarried   
 -9.6288 4.1346   
 ParentMaritalStatussingle ParentMaritalStatuswidowed   
 1.0557 3.9501   
 PracticeSport.L PracticeSport.Q   
 2.2506 -0.7075   
 IsFirstChildyes WklyStudyHours.L   
 2.2077 1.3382   
 WklyStudyHours.Q   
 -0.9604

Writing score stepwise model

| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 69.347 | <.001 |
| Gendermale | -9.209 | <.001 |
| EthnicGroup.L | 4.673 | <.001 |
| EthnicGroup.Q | 0.627 | 0.569 |
| EthnicGroup.C | -1.891 | 0.058 |
| EthnicGroup^4 | -1.649 | 0.043 |
| ParentEduc.L | 9.983 | <.001 |
| ParentEduc.Q | 1.365 | 0.265 |
| ParentEduc.C | 0.289 | 0.803 |
| ParentEduc^4 | 1.715 | 0.109 |
| ParentEduc^5 | -3.106 | 0.001 |
| LunchTypestandard | 8.388 | <.001 |
| TestPrepnone | -9.629 | <.001 |
| ParentMaritalStatusmarried | 4.135 | 0.001 |
| ParentMaritalStatussingle | 1.056 | 0.44 |
| ParentMaritalStatuswidowed | 3.950 | 0.16 |
| PracticeSport.L | 2.251 | 0.022 |
| PracticeSport.Q | -0.708 | 0.335 |
| IsFirstChildyes | 2.208 | 0.012 |
| WklyStudyHours.L | 1.338 | 0.152 |
| WklyStudyHours.Q | -0.960 | 0.167 |

\

Start: AIC=5201.34  
MathScore ~ 1  
  
 Df Sum of Sq RSS AIC  
+ LunchType 1 28689.3 199719 5076.1  
+ EthnicGroup 4 11794.2 216614 5159.1  
+ TestPrep 1 7994.3 220414 5169.6  
+ Gender 1 6835.6 221573 5174.5  
+ ParentEduc 5 6656.3 221752 5183.3  
+ WklyStudyHours 2 3204.8 225204 5191.9  
+ ParentMaritalStatus 3 2855.2 225553 5195.4  
+ IsFirstChild 1 902.6 227506 5199.6  
<none> 228409 5201.3  
+ TransportMeans 1 86.4 228322 5203.0  
+ NrSiblings 1 18.2 228391 5203.3  
+ PracticeSport 2 306.1 228103 5204.1  
  
Step: AIC=5076.1  
MathScore ~ LunchType  
  
 Df Sum of Sq RSS AIC  
+ EthnicGroup 4 10168.9 189550 5034.6  
+ TestPrep 1 8075.5 191644 5039.0  
+ Gender 1 5817.8 193902 5050.1  
+ ParentEduc 5 7142.6 192577 5051.6  
+ WklyStudyHours 2 3096.1 196623 5065.3  
+ ParentMaritalStatus 3 2220.4 197499 5071.5  
+ IsFirstChild 1 1033.3 198686 5073.2  
<none> 199719 5076.1  
+ PracticeSport 2 633.7 199086 5077.1  
+ NrSiblings 1 120.2 199599 5077.5  
+ TransportMeans 1 81.6 199638 5077.7  
  
Step: AIC=5034.56  
MathScore ~ LunchType + EthnicGroup  
  
 Df Sum of Sq RSS AIC  
+ TestPrep 1 7552.6 181998 4998.0  
+ Gender 1 5722.6 183828 5007.5  
+ ParentEduc 5 6136.8 183414 5013.4  
+ ParentMaritalStatus 3 2881.3 186669 5026.0  
+ WklyStudyHours 2 2483.7 187067 5026.1  
+ IsFirstChild 1 1145.7 188405 5030.8  
+ PracticeSport 2 799.9 188751 5034.5  
<none> 189550 5034.6  
+ NrSiblings 1 111.5 189439 5036.0  
+ TransportMeans 1 99.6 189451 5036.1  
  
Step: AIC=4998.01  
MathScore ~ LunchType + EthnicGroup + TestPrep  
  
 Df Sum of Sq RSS AIC  
+ Gender 1 5498.6 176499 4970.9  
+ ParentEduc 5 6022.9 175975 4976.1  
+ ParentMaritalStatus 3 3121.7 178876 4987.6  
+ WklyStudyHours 2 1984.0 180014 4991.6  
+ IsFirstChild 1 871.2 181127 4995.5  
+ PracticeSport 2 769.8 181228 4998.0  
<none> 181998 4998.0  
+ TransportMeans 1 139.8 181858 4999.3  
+ NrSiblings 1 116.2 181882 4999.4  
  
Step: AIC=4970.93  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender  
  
 Df Sum of Sq RSS AIC  
+ ParentEduc 5 6377.1 170122 4946.0  
+ ParentMaritalStatus 3 2781.5 173718 4961.9  
+ WklyStudyHours 2 2007.7 174492 4964.1  
+ IsFirstChild 1 1024.6 175475 4967.4  
+ PracticeSport 2 772.3 175727 4970.8  
<none> 176499 4970.9  
+ NrSiblings 1 149.6 176350 4972.1  
+ TransportMeans 1 141.5 176358 4972.2  
  
Step: AIC=4946.04  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc  
  
 Df Sum of Sq RSS AIC  
+ WklyStudyHours 2 2137.6 167985 4938.1  
+ ParentMaritalStatus 3 2404.8 167717 4938.5  
+ IsFirstChild 1 1037.2 169085 4942.2  
+ PracticeSport 2 1019.4 169103 4944.3  
<none> 170122 4946.0  
+ NrSiblings 1 201.5 169921 4946.9  
+ TransportMeans 1 57.1 170065 4947.7  
  
Step: AIC=4938.05  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc +   
 WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
+ ParentMaritalStatus 3 2320.08 165664 4930.9  
+ IsFirstChild 1 960.71 167024 4934.6  
+ PracticeSport 2 988.62 166996 4936.5  
<none> 167985 4938.1  
+ NrSiblings 1 136.00 167849 4939.3  
+ TransportMeans 1 62.26 167922 4939.7  
  
Step: AIC=4930.87  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc +   
 WklyStudyHours + ParentMaritalStatus  
  
 Df Sum of Sq RSS AIC  
+ IsFirstChild 1 1097.67 164567 4926.6  
+ PracticeSport 2 999.86 164665 4929.1  
<none> 165664 4930.9  
+ NrSiblings 1 99.66 165565 4932.3  
+ TransportMeans 1 34.82 165630 4932.7  
  
Step: AIC=4926.57  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc +   
 WklyStudyHours + ParentMaritalStatus + IsFirstChild  
  
 Df Sum of Sq RSS AIC  
+ PracticeSport 2 1060.53 163506 4924.4  
<none> 164567 4926.6  
+ NrSiblings 1 213.32 164353 4927.3  
+ TransportMeans 1 32.10 164535 4928.4  
  
Step: AIC=4924.44  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc +   
 WklyStudyHours + ParentMaritalStatus + IsFirstChild + PracticeSport  
  
 Df Sum of Sq RSS AIC  
<none> 163506 4924.4  
+ NrSiblings 1 220.911 163285 4925.2  
+ TransportMeans 1 33.331 163473 4926.2

Call:  
lm(formula = MathScore ~ LunchType + EthnicGroup + TestPrep +   
 Gender + ParentEduc + WklyStudyHours + ParentMaritalStatus +   
 IsFirstChild + PracticeSport, data = data\_cleaned %>% select(-c(ReadingScore,   
 WritingScore)))  
  
Coefficients:  
 (Intercept) LunchTypestandard   
 55.97610 11.15459   
 EthnicGroup.L EthnicGroup.Q   
 7.40062 2.94026   
 EthnicGroup.C EthnicGroup^4   
 0.73671 -0.98454   
 TestPrepnone Gendermale   
 -5.58185 4.98086   
 ParentEduc.L ParentEduc.Q   
 6.64225 -0.04451   
 ParentEduc.C ParentEduc^4   
 -0.29953 1.40239   
 ParentEduc^5 WklyStudyHours.L   
 -2.64422 2.62125   
 WklyStudyHours.Q ParentMaritalStatusmarried   
 -1.03208 3.87581   
 ParentMaritalStatussingle ParentMaritalStatuswidowed   
 1.10195 5.06618   
 IsFirstChildyes PracticeSport.L   
 2.36796 2.51364   
 PracticeSport.Q   
 -0.51783

Math score forwards stepwise model

| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 55.976 | <.001 |
| Gendermale | 4.981 | <.001 |
| EthnicGroup.L | 7.401 | <.001 |
| EthnicGroup.Q | 2.940 | 0.011 |
| EthnicGroup.C | 0.737 | 0.482 |
| EthnicGroup^4 | -0.985 | 0.251 |
| ParentEduc.L | 6.642 | <.001 |
| ParentEduc.Q | -0.045 | 0.972 |
| ParentEduc.C | -0.300 | 0.806 |
| ParentEduc^4 | 1.402 | 0.213 |
| ParentEduc^5 | -2.644 | 0.006 |
| LunchTypestandard | 11.155 | <.001 |
| TestPrepnone | -5.582 | <.001 |
| ParentMaritalStatusmarried | 3.876 | 0.002 |
| ParentMaritalStatussingle | 1.102 | 0.444 |
| ParentMaritalStatuswidowed | 5.066 | 0.087 |
| PracticeSport.L | 2.514 | 0.015 |
| PracticeSport.Q | -0.518 | 0.503 |
| IsFirstChildyes | 2.368 | 0.011 |
| WklyStudyHours.L | 2.621 | 0.008 |
| WklyStudyHours.Q | -1.032 | 0.158 |

Start: AIC=5109.96  
ReadingScore ~ 1  
  
 Df Sum of Sq RSS AIC  
+ Gender 1 11752.6 195668 5056.7  
+ LunchType 1 11567.9 195852 5057.6  
+ TestPrep 1 10808.1 196612 5061.2  
+ ParentEduc 5 9368.1 198052 5076.2  
+ EthnicGroup 4 4456.5 202964 5097.4  
+ IsFirstChild 1 1637.0 205783 5104.5  
+ ParentMaritalStatus 3 2223.6 205197 5105.7  
+ WklyStudyHours 2 1358.6 206062 5107.7  
<none> 207420 5110.0  
+ TransportMeans 1 349.9 207070 5110.4  
+ NrSiblings 1 29.9 207390 5111.8  
+ PracticeSport 2 123.3 207297 5113.4  
  
Step: AIC=5056.67  
ReadingScore ~ Gender  
  
 Df Sum of Sq RSS AIC  
+ LunchType 1 12492.3 183175 4996.1  
+ TestPrep 1 11191.7 184476 5002.8  
+ ParentEduc 5 8383.7 187284 5025.2  
+ EthnicGroup 4 4418.9 191249 5043.0  
+ ParentMaritalStatus 3 2580.1 193088 5050.1  
+ IsFirstChild 1 1351.9 194316 5052.1  
+ WklyStudyHours 2 1392.8 194275 5053.9  
<none> 195668 5056.7  
+ TransportMeans 1 359.3 195308 5056.9  
+ NrSiblings 1 57.7 195610 5058.4  
+ PracticeSport 2 127.3 195540 5060.1  
  
Step: AIC=4996.12  
ReadingScore ~ Gender + LunchType  
  
 Df Sum of Sq RSS AIC  
+ TestPrep 1 11270.3 171905 4937.9  
+ ParentEduc 5 8809.6 174366 4959.4  
+ EthnicGroup 4 3778.7 179397 4984.4  
+ IsFirstChild 1 1444.9 181730 4990.6  
+ ParentMaritalStatus 3 2168.1 181007 4990.8  
+ WklyStudyHours 2 1302.7 181873 4993.4  
<none> 183175 4996.1  
+ TransportMeans 1 353.2 182822 4996.3  
+ NrSiblings 1 10.7 183165 4998.1  
+ PracticeSport 2 102.3 183073 4999.6  
  
Step: AIC=4937.93  
ReadingScore ~ Gender + LunchType + TestPrep  
  
 Df Sum of Sq RSS AIC  
+ ParentEduc 5 8374.9 163530 4900.6  
+ EthnicGroup 4 3491.4 168414 4926.5  
+ ParentMaritalStatus 3 2462.8 169442 4930.2  
+ IsFirstChild 1 1065.4 170840 4934.0  
+ WklyStudyHours 2 892.1 171013 4937.0  
+ TransportMeans 1 442.7 171462 4937.5  
<none> 171905 4937.9  
+ NrSiblings 1 9.8 171895 4939.9  
+ PracticeSport 2 89.6 171815 4941.4  
  
Step: AIC=4900.58  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc  
  
 Df Sum of Sq RSS AIC  
+ EthnicGroup 4 2838.31 160692 4892.0  
+ ParentMaritalStatus 3 2129.11 161401 4894.2  
+ IsFirstChild 1 1058.56 162472 4896.4  
+ WklyStudyHours 2 1062.99 162467 4898.4  
<none> 163530 4900.6  
+ TransportMeans 1 298.08 163232 4900.8  
+ NrSiblings 1 0.04 163530 4902.6  
+ PracticeSport 2 80.83 163449 4904.1  
  
Step: AIC=4891.98  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc + EthnicGroup  
  
 Df Sum of Sq RSS AIC  
+ ParentMaritalStatus 3 2520.47 158171 4883.0  
+ IsFirstChild 1 1138.50 159553 4887.2  
+ WklyStudyHours 2 911.62 159780 4890.6  
<none> 160692 4892.0  
+ TransportMeans 1 319.62 160372 4892.1  
+ NrSiblings 1 0.02 160692 4894.0  
+ PracticeSport 2 113.68 160578 4895.3  
  
Step: AIC=4882.99  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc + EthnicGroup +   
 ParentMaritalStatus  
  
 Df Sum of Sq RSS AIC  
+ IsFirstChild 1 1287.60 156884 4877.2  
+ WklyStudyHours 2 869.71 157302 4881.8  
<none> 158171 4883.0  
+ TransportMeans 1 253.86 157917 4883.5  
+ NrSiblings 1 2.24 158169 4885.0  
+ PracticeSport 2 103.76 158068 4886.4  
  
Step: AIC=4877.24  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild  
  
 Df Sum of Sq RSS AIC  
+ WklyStudyHours 2 821.26 156062 4876.3  
<none> 156884 4877.2  
+ TransportMeans 1 246.11 156638 4877.8  
+ NrSiblings 1 11.09 156873 4879.2  
+ PracticeSport 2 139.87 156744 4880.4  
  
Step: AIC=4876.27  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild + WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
<none> 156062 4876.3  
+ TransportMeans 1 255.415 155807 4876.7  
+ NrSiblings 1 3.985 156058 4878.2  
+ PracticeSport 2 116.550 155946 4879.6

Call:  
lm(formula = ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc +   
 EthnicGroup + ParentMaritalStatus + IsFirstChild + WklyStudyHours,   
 data = data\_cleaned %>% select(-c(MathScore, WritingScore)))  
  
Coefficients:  
 (Intercept) Gendermale   
 68.4954 -7.2819   
 LunchTypestandard TestPrepnone   
 7.4941 -6.9723   
 ParentEduc.L ParentEduc.Q   
 7.6375 1.5354   
 ParentEduc.C ParentEduc^4   
 0.5659 1.5467   
 ParentEduc^5 EthnicGroup.L   
 -3.0328 4.1491   
 EthnicGroup.Q EthnicGroup.C   
 1.4383 -0.4714   
 EthnicGroup^4 ParentMaritalStatusmarried   
 -0.9416 4.1130   
 ParentMaritalStatussingle ParentMaritalStatuswidowed   
 1.2752 4.6451   
 IsFirstChildyes WklyStudyHours.L   
 2.4465 1.4306   
 WklyStudyHours.Q   
 -0.9328

Reading score forwards stepwise model

| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 68.495 | <.001 |
| Gendermale | -7.282 | <.001 |
| EthnicGroup.L | 4.149 | 0.001 |
| EthnicGroup.Q | 1.438 | 0.203 |
| EthnicGroup.C | -0.471 | 0.645 |
| EthnicGroup^4 | -0.942 | 0.261 |
| ParentEduc.L | 7.638 | <.001 |
| ParentEduc.Q | 1.535 | 0.22 |
| ParentEduc.C | 0.566 | 0.635 |
| ParentEduc^4 | 1.547 | 0.158 |
| ParentEduc^5 | -3.033 | 0.001 |
| LunchTypestandard | 7.494 | <.001 |
| TestPrepnone | -6.972 | <.001 |
| ParentMaritalStatusmarried | 4.113 | 0.001 |
| ParentMaritalStatussingle | 1.275 | 0.363 |
| ParentMaritalStatuswidowed | 4.645 | 0.106 |
| IsFirstChildyes | 2.446 | 0.007 |
| WklyStudyHours.L | 1.431 | 0.135 |
| WklyStudyHours.Q | -0.933 | 0.191 |

Start: AIC=5186.91  
WritingScore ~ 1  
  
 Df Sum of Sq RSS AIC  
+ TestPrep 1 19817.7 205141 5101.5  
+ Gender 1 18846.0 206112 5106.0  
+ LunchType 1 13742.3 211216 5129.2  
+ ParentEduc 5 14395.9 210562 5134.2  
+ EthnicGroup 4 6413.2 218545 5167.5  
+ IsFirstChild 1 1481.7 223477 5182.6  
+ ParentMaritalStatus 3 2228.1 222730 5183.5  
+ WklyStudyHours 2 1494.5 223464 5184.6  
<none> 224958 5186.9  
+ TransportMeans 1 197.6 224761 5188.1  
+ NrSiblings 1 0.9 224957 5188.9  
+ PracticeSport 2 244.0 224714 5189.9  
  
Step: AIC=5101.49  
WritingScore ~ TestPrep  
  
 Df Sum of Sq RSS AIC  
+ Gender 1 19503.9 185637 5008.8  
+ LunchType 1 13830.7 191310 5037.3  
+ ParentEduc 5 13691.9 191449 5046.0  
+ EthnicGroup 4 6170.5 198970 5080.5  
+ ParentMaritalStatus 3 2609.5 202531 5095.4  
+ IsFirstChild 1 988.2 204152 5098.9  
+ WklyStudyHours 2 995.9 204145 5100.9  
<none> 205141 5101.5  
+ TransportMeans 1 290.1 204850 5102.1  
+ NrSiblings 1 1.3 205139 5103.5  
+ PracticeSport 2 216.4 204924 5104.5  
  
Step: AIC=5008.78  
WritingScore ~ TestPrep + Gender  
  
 Df Sum of Sq RSS AIC  
+ LunchType 1 15135.3 170501 4930.2  
+ ParentEduc 5 12275.4 173361 4953.9  
+ EthnicGroup 4 6174.8 179462 4984.7  
+ ParentMaritalStatus 3 3148.0 182489 4998.6  
+ IsFirstChild 1 704.6 184932 5007.2  
+ WklyStudyHours 2 1015.7 184621 5007.6  
<none> 185637 5008.8  
+ TransportMeans 1 302.9 185334 5009.2  
+ NrSiblings 1 2.5 185634 5010.8  
+ PracticeSport 2 198.5 185438 5011.8  
  
Step: AIC=4930.15  
WritingScore ~ TestPrep + Gender + LunchType  
  
 Df Sum of Sq RSS AIC  
+ ParentEduc 5 12883.8 157618 4865.7  
+ EthnicGroup 4 5581.8 164920 4906.6  
+ ParentMaritalStatus 3 2642.1 167859 4921.3  
+ IsFirstChild 1 777.9 169723 4927.8  
+ WklyStudyHours 2 872.2 169629 4929.3  
<none> 170501 4930.2  
+ TransportMeans 1 297.0 170204 4930.5  
+ PracticeSport 2 409.6 170092 4931.9  
+ NrSiblings 1 10.1 170491 4932.1  
  
Step: AIC=4865.67  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc  
  
 Df Sum of Sq RSS AIC  
+ EthnicGroup 4 4636.0 152982 4845.4  
+ ParentMaritalStatus 3 2136.7 155481 4858.7  
+ IsFirstChild 1 757.3 156860 4863.1  
+ WklyStudyHours 2 1069.8 156548 4863.2  
+ PracticeSport 2 755.4 156862 4865.1  
<none> 157618 4865.7  
+ TransportMeans 1 168.3 157449 4866.7  
+ NrSiblings 1 43.8 157574 4867.4  
  
Step: AIC=4845.37  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup  
  
 Df Sum of Sq RSS AIC  
+ ParentMaritalStatus 3 2649.91 150332 4834.8  
+ IsFirstChild 1 855.16 152126 4842.1  
+ WklyStudyHours 2 884.81 152097 4843.9  
+ PracticeSport 2 781.38 152200 4844.5  
<none> 152982 4845.4  
+ TransportMeans 1 195.29 152786 4846.2  
+ NrSiblings 1 47.79 152934 4847.1  
  
Step: AIC=4834.8  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup +   
 ParentMaritalStatus  
  
 Df Sum of Sq RSS AIC  
+ IsFirstChild 1 988.67 149343 4830.5  
+ WklyStudyHours 2 849.42 149482 4833.4  
+ PracticeSport 2 797.97 149534 4833.8  
<none> 150332 4834.8  
+ TransportMeans 1 146.78 150185 4835.9  
+ NrSiblings 1 29.46 150302 4836.6  
  
Step: AIC=4830.55  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild  
  
 Df Sum of Sq RSS AIC  
+ PracticeSport 2 861.11 148482 4829.1  
+ WklyStudyHours 2 807.65 148535 4829.4  
<none> 149343 4830.5  
+ TransportMeans 1 141.62 149201 4831.6  
+ NrSiblings 1 94.43 149249 4831.9  
  
Step: AIC=4829.06  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild + PracticeSport  
  
 Df Sum of Sq RSS AIC  
+ WklyStudyHours 2 784.42 147698 4828.0  
<none> 148482 4829.1  
+ TransportMeans 1 141.31 148341 4830.2  
+ NrSiblings 1 95.46 148386 4830.5  
  
Step: AIC=4828.04  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild + PracticeSport + WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
<none> 147698 4828.0  
+ TransportMeans 1 148.608 147549 4829.1  
+ NrSiblings 1 73.458 147624 4829.6

Call:  
lm(formula = WritingScore ~ TestPrep + Gender + LunchType + ParentEduc +   
 EthnicGroup + ParentMaritalStatus + IsFirstChild + PracticeSport +   
 WklyStudyHours, data = data\_cleaned %>% select(-c(MathScore,   
 ReadingScore)))  
  
Coefficients:  
 (Intercept) TestPrepnone   
 69.3473 -9.6288   
 Gendermale LunchTypestandard   
 -9.2094 8.3876   
 ParentEduc.L ParentEduc.Q   
 9.9826 1.3651   
 ParentEduc.C ParentEduc^4   
 0.2894 1.7154   
 ParentEduc^5 EthnicGroup.L   
 -3.1061 4.6731   
 EthnicGroup.Q EthnicGroup.C   
 0.6274 -1.8905   
 EthnicGroup^4 ParentMaritalStatusmarried   
 -1.6492 4.1346   
 ParentMaritalStatussingle ParentMaritalStatuswidowed   
 1.0557 3.9501   
 IsFirstChildyes PracticeSport.L   
 2.2077 2.2506   
 PracticeSport.Q WklyStudyHours.L   
 -0.7075 1.3382   
 WklyStudyHours.Q   
 -0.9604

Writing score forwards stepwise model

| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 69.347 | <.001 |
| Gendermale | -9.209 | <.001 |
| EthnicGroup.L | 4.673 | <.001 |
| EthnicGroup.Q | 0.627 | 0.569 |
| EthnicGroup.C | -1.891 | 0.058 |
| EthnicGroup^4 | -1.649 | 0.043 |
| ParentEduc.L | 9.983 | <.001 |
| ParentEduc.Q | 1.365 | 0.265 |
| ParentEduc.C | 0.289 | 0.803 |
| ParentEduc^4 | 1.715 | 0.109 |
| ParentEduc^5 | -3.106 | 0.001 |
| LunchTypestandard | 8.388 | <.001 |
| TestPrepnone | -9.629 | <.001 |
| ParentMaritalStatusmarried | 4.135 | 0.001 |
| ParentMaritalStatussingle | 1.056 | 0.44 |
| ParentMaritalStatuswidowed | 3.950 | 0.16 |
| PracticeSport.L | 2.251 | 0.022 |
| PracticeSport.Q | -0.708 | 0.335 |
| IsFirstChildyes | 2.208 | 0.012 |
| WklyStudyHours.L | 1.338 | 0.152 |
| WklyStudyHours.Q | -0.960 | 0.167 |

\

Start: AIC=5201.34  
MathScore ~ 1  
  
 Df Sum of Sq RSS AIC  
+ LunchType 1 28689.3 199719 5076.1  
+ EthnicGroup 4 11794.2 216614 5159.1  
+ TestPrep 1 7994.3 220414 5169.6  
+ Gender 1 6835.6 221573 5174.5  
+ ParentEduc 5 6656.3 221752 5183.3  
+ WklyStudyHours 2 3204.8 225204 5191.9  
+ ParentMaritalStatus 3 2855.2 225553 5195.4  
+ IsFirstChild 1 902.6 227506 5199.6  
<none> 228409 5201.3  
+ TransportMeans 1 86.4 228322 5203.0  
+ NrSiblings 1 18.2 228391 5203.3  
+ PracticeSport 2 306.1 228103 5204.1  
  
Step: AIC=5076.1  
MathScore ~ LunchType  
  
 Df Sum of Sq RSS AIC  
+ EthnicGroup 4 10168.9 189550 5034.6  
+ TestPrep 1 8075.5 191644 5039.0  
+ Gender 1 5817.8 193902 5050.1  
+ ParentEduc 5 7142.6 192577 5051.6  
+ WklyStudyHours 2 3096.1 196623 5065.3  
+ ParentMaritalStatus 3 2220.4 197499 5071.5  
+ IsFirstChild 1 1033.3 198686 5073.2  
<none> 199719 5076.1  
+ PracticeSport 2 633.7 199086 5077.1  
+ NrSiblings 1 120.2 199599 5077.5  
+ TransportMeans 1 81.6 199638 5077.7  
- LunchType 1 28689.3 228409 5201.3  
  
Step: AIC=5034.56  
MathScore ~ LunchType + EthnicGroup  
  
 Df Sum of Sq RSS AIC  
+ TestPrep 1 7552.6 181998 4998.0  
+ Gender 1 5722.6 183828 5007.5  
+ ParentEduc 5 6136.8 183414 5013.4  
+ ParentMaritalStatus 3 2881.3 186669 5026.0  
+ WklyStudyHours 2 2483.7 187067 5026.1  
+ IsFirstChild 1 1145.7 188405 5030.8  
+ PracticeSport 2 799.9 188751 5034.5  
<none> 189550 5034.6  
+ NrSiblings 1 111.5 189439 5036.0  
+ TransportMeans 1 99.6 189451 5036.1  
- EthnicGroup 4 10168.9 199719 5076.1  
- LunchType 1 27064.0 216614 5159.1  
  
Step: AIC=4998.01  
MathScore ~ LunchType + EthnicGroup + TestPrep  
  
 Df Sum of Sq RSS AIC  
+ Gender 1 5498.6 176499 4970.9  
+ ParentEduc 5 6022.9 175975 4976.1  
+ ParentMaritalStatus 3 3121.7 178876 4987.6  
+ WklyStudyHours 2 1984.0 180014 4991.6  
+ IsFirstChild 1 871.2 181127 4995.5  
+ PracticeSport 2 769.8 181228 4998.0  
<none> 181998 4998.0  
+ TransportMeans 1 139.8 181858 4999.3  
+ NrSiblings 1 116.2 181882 4999.4  
- TestPrep 1 7552.6 189550 5034.6  
- EthnicGroup 4 9646.0 191644 5039.0  
- LunchType 1 27205.5 209203 5128.1  
  
Step: AIC=4970.93  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender  
  
 Df Sum of Sq RSS AIC  
+ ParentEduc 5 6377.1 170122 4946.0  
+ ParentMaritalStatus 3 2781.5 173718 4961.9  
+ WklyStudyHours 2 2007.7 174492 4964.1  
+ IsFirstChild 1 1024.6 175475 4967.4  
+ PracticeSport 2 772.3 175727 4970.8  
<none> 176499 4970.9  
+ NrSiblings 1 149.6 176350 4972.1  
+ TransportMeans 1 141.5 176358 4972.2  
- Gender 1 5498.6 181998 4998.0  
- TestPrep 1 7328.6 183828 5007.5  
- EthnicGroup 4 9553.8 186053 5012.9  
- LunchType 1 26159.4 202659 5099.9  
  
Step: AIC=4946.04  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc  
  
 Df Sum of Sq RSS AIC  
+ WklyStudyHours 2 2137.6 167985 4938.1  
+ ParentMaritalStatus 3 2404.8 167717 4938.5  
+ IsFirstChild 1 1037.2 169085 4942.2  
+ PracticeSport 2 1019.4 169103 4944.3  
<none> 170122 4946.0  
+ NrSiblings 1 201.5 169921 4946.9  
+ TransportMeans 1 57.1 170065 4947.7  
- ParentEduc 5 6377.1 176499 4970.9  
- Gender 1 5852.9 175975 4976.1  
- TestPrep 1 7156.2 177278 4983.1  
- EthnicGroup 4 8496.3 178618 4984.2  
- LunchType 1 26736.8 196859 5082.4  
  
Step: AIC=4938.05  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc +   
 WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
+ ParentMaritalStatus 3 2320.1 165664 4930.9  
+ IsFirstChild 1 960.7 167024 4934.6  
+ PracticeSport 2 988.6 166996 4936.5  
<none> 167985 4938.1  
+ NrSiblings 1 136.0 167849 4939.3  
+ TransportMeans 1 62.3 167922 4939.7  
- WklyStudyHours 2 2137.6 170122 4946.0  
- ParentEduc 5 6507.1 174492 4964.1  
- Gender 1 5911.1 173896 4968.8  
- TestPrep 1 6624.6 174609 4972.7  
- EthnicGroup 4 8063.7 176048 4974.5  
- LunchType 1 26502.3 194487 5074.9  
  
Step: AIC=4930.87  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc +   
 WklyStudyHours + ParentMaritalStatus  
  
 Df Sum of Sq RSS AIC  
+ IsFirstChild 1 1097.7 164567 4926.6  
+ PracticeSport 2 999.9 164665 4929.1  
<none> 165664 4930.9  
+ NrSiblings 1 99.7 165565 4932.3  
+ TransportMeans 1 34.8 165630 4932.7  
- ParentMaritalStatus 3 2320.1 167985 4938.1  
- WklyStudyHours 2 2052.9 167717 4938.5  
- ParentEduc 5 6110.6 171775 4955.2  
- Gender 1 5583.4 171248 4960.3  
- TestPrep 1 6811.1 172476 4967.1  
- EthnicGroup 4 8601.5 174266 4970.9  
- LunchType 1 25882.2 191547 5066.5  
  
Step: AIC=4926.57  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc +   
 WklyStudyHours + ParentMaritalStatus + IsFirstChild  
  
 Df Sum of Sq RSS AIC  
+ PracticeSport 2 1060.5 163506 4924.4  
<none> 164567 4926.6  
+ NrSiblings 1 213.3 164353 4927.3  
+ TransportMeans 1 32.1 164535 4928.4  
- IsFirstChild 1 1097.7 165664 4930.9  
- WklyStudyHours 2 1979.1 166546 4933.9  
- ParentMaritalStatus 3 2457.0 167024 4934.6  
- ParentEduc 5 6107.0 170674 4951.1  
- Gender 1 5746.4 170313 4957.1  
- TestPrep 1 6505.5 171072 4961.3  
- EthnicGroup 4 8738.2 173305 4967.6  
- LunchType 1 25952.6 190519 5063.4  
  
Step: AIC=4924.44  
MathScore ~ LunchType + EthnicGroup + TestPrep + Gender + ParentEduc +   
 WklyStudyHours + ParentMaritalStatus + IsFirstChild + PracticeSport  
  
 Df Sum of Sq RSS AIC  
<none> 163506 4924.4  
+ NrSiblings 1 220.9 163285 4925.2  
+ TransportMeans 1 33.3 163473 4926.2  
- PracticeSport 2 1060.5 164567 4926.6  
- IsFirstChild 1 1158.3 164665 4929.1  
- WklyStudyHours 2 1940.4 165447 4931.6  
- ParentMaritalStatus 3 2466.0 165972 4932.6  
- ParentEduc 5 6377.7 169884 4950.7  
- Gender 1 5773.6 169280 4955.3  
- TestPrep 1 6459.7 169966 4959.2  
- EthnicGroup 4 8854.1 172360 4966.4  
- LunchType 1 26412.5 189919 5064.4

Call:  
lm(formula = MathScore ~ LunchType + EthnicGroup + TestPrep +   
 Gender + ParentEduc + WklyStudyHours + ParentMaritalStatus +   
 IsFirstChild + PracticeSport, data = data\_cleaned %>% select(-c(ReadingScore,   
 WritingScore)))  
  
Coefficients:  
 (Intercept) LunchTypestandard   
 55.97610 11.15459   
 EthnicGroup.L EthnicGroup.Q   
 7.40062 2.94026   
 EthnicGroup.C EthnicGroup^4   
 0.73671 -0.98454   
 TestPrepnone Gendermale   
 -5.58185 4.98086   
 ParentEduc.L ParentEduc.Q   
 6.64225 -0.04451   
 ParentEduc.C ParentEduc^4   
 -0.29953 1.40239   
 ParentEduc^5 WklyStudyHours.L   
 -2.64422 2.62125   
 WklyStudyHours.Q ParentMaritalStatusmarried   
 -1.03208 3.87581   
 ParentMaritalStatussingle ParentMaritalStatuswidowed   
 1.10195 5.06618   
 IsFirstChildyes PracticeSport.L   
 2.36796 2.51364   
 PracticeSport.Q   
 -0.51783

Math score stepwise model

| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 55.976 | <.001 |
| Gendermale | 4.981 | <.001 |
| EthnicGroup.L | 7.401 | <.001 |
| EthnicGroup.Q | 2.940 | 0.011 |
| EthnicGroup.C | 0.737 | 0.482 |
| EthnicGroup^4 | -0.985 | 0.251 |
| ParentEduc.L | 6.642 | <.001 |
| ParentEduc.Q | -0.045 | 0.972 |
| ParentEduc.C | -0.300 | 0.806 |
| ParentEduc^4 | 1.402 | 0.213 |
| ParentEduc^5 | -2.644 | 0.006 |
| LunchTypestandard | 11.155 | <.001 |
| TestPrepnone | -5.582 | <.001 |
| ParentMaritalStatusmarried | 3.876 | 0.002 |
| ParentMaritalStatussingle | 1.102 | 0.444 |
| ParentMaritalStatuswidowed | 5.066 | 0.087 |
| PracticeSport.L | 2.514 | 0.015 |
| PracticeSport.Q | -0.518 | 0.503 |
| IsFirstChildyes | 2.368 | 0.011 |
| WklyStudyHours.L | 2.621 | 0.008 |
| WklyStudyHours.Q | -1.032 | 0.158 |

Start: AIC=5109.96  
ReadingScore ~ 1  
  
 Df Sum of Sq RSS AIC  
+ Gender 1 11752.6 195668 5056.7  
+ LunchType 1 11567.9 195852 5057.6  
+ TestPrep 1 10808.1 196612 5061.2  
+ ParentEduc 5 9368.1 198052 5076.2  
+ EthnicGroup 4 4456.5 202964 5097.4  
+ IsFirstChild 1 1637.0 205783 5104.5  
+ ParentMaritalStatus 3 2223.6 205197 5105.7  
+ WklyStudyHours 2 1358.6 206062 5107.7  
<none> 207420 5110.0  
+ TransportMeans 1 349.9 207070 5110.4  
+ NrSiblings 1 29.9 207390 5111.8  
+ PracticeSport 2 123.3 207297 5113.4  
  
Step: AIC=5056.67  
ReadingScore ~ Gender  
  
 Df Sum of Sq RSS AIC  
+ LunchType 1 12492.3 183175 4996.1  
+ TestPrep 1 11191.7 184476 5002.8  
+ ParentEduc 5 8383.7 187284 5025.2  
+ EthnicGroup 4 4418.9 191249 5043.0  
+ ParentMaritalStatus 3 2580.1 193088 5050.1  
+ IsFirstChild 1 1351.9 194316 5052.1  
+ WklyStudyHours 2 1392.8 194275 5053.9  
<none> 195668 5056.7  
+ TransportMeans 1 359.3 195308 5056.9  
+ NrSiblings 1 57.7 195610 5058.4  
+ PracticeSport 2 127.3 195540 5060.1  
- Gender 1 11752.6 207420 5110.0  
  
Step: AIC=4996.12  
ReadingScore ~ Gender + LunchType  
  
 Df Sum of Sq RSS AIC  
+ TestPrep 1 11270.3 171905 4937.9  
+ ParentEduc 5 8809.6 174366 4959.4  
+ EthnicGroup 4 3778.7 179397 4984.4  
+ IsFirstChild 1 1444.9 181730 4990.6  
+ ParentMaritalStatus 3 2168.1 181007 4990.8  
+ WklyStudyHours 2 1302.7 181873 4993.4  
<none> 183175 4996.1  
+ TransportMeans 1 353.2 182822 4996.3  
+ NrSiblings 1 10.7 183165 4998.1  
+ PracticeSport 2 102.3 183073 4999.6  
- LunchType 1 12492.3 195668 5056.7  
- Gender 1 12677.0 195852 5057.6  
  
Step: AIC=4937.93  
ReadingScore ~ Gender + LunchType + TestPrep  
  
 Df Sum of Sq RSS AIC  
+ ParentEduc 5 8374.9 163530 4900.6  
+ EthnicGroup 4 3491.4 168414 4926.5  
+ ParentMaritalStatus 3 2462.8 169442 4930.2  
+ IsFirstChild 1 1065.4 170840 4934.0  
+ WklyStudyHours 2 892.1 171013 4937.0  
+ TransportMeans 1 442.7 171462 4937.5  
<none> 171905 4937.9  
+ NrSiblings 1 9.8 171895 4939.9  
+ PracticeSport 2 89.6 171815 4941.4  
- TestPrep 1 11270.3 183175 4996.1  
- LunchType 1 12570.9 184476 5002.8  
- Gender 1 13079.4 184984 5005.4  
  
Step: AIC=4900.58  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc  
  
 Df Sum of Sq RSS AIC  
+ EthnicGroup 4 2838.3 160692 4892.0  
+ ParentMaritalStatus 3 2129.1 161401 4894.2  
+ IsFirstChild 1 1058.6 162472 4896.4  
+ WklyStudyHours 2 1063.0 162467 4898.4  
<none> 163530 4900.6  
+ TransportMeans 1 298.1 163232 4900.8  
+ NrSiblings 1 0.0 163530 4902.6  
+ PracticeSport 2 80.8 163449 4904.1  
- ParentEduc 5 8374.9 171905 4937.9  
- TestPrep 1 10835.7 174366 4959.4  
- Gender 1 12101.6 175632 4966.3  
- LunchType 1 13017.0 176547 4971.2  
  
Step: AIC=4891.98  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc + EthnicGroup  
  
 Df Sum of Sq RSS AIC  
+ ParentMaritalStatus 3 2520.5 158171 4883.0  
+ IsFirstChild 1 1138.5 159553 4887.2  
+ WklyStudyHours 2 911.6 159780 4890.6  
<none> 160692 4892.0  
+ TransportMeans 1 319.6 160372 4892.1  
+ NrSiblings 1 0.0 160692 4894.0  
+ PracticeSport 2 113.7 160578 4895.3  
- EthnicGroup 4 2838.3 163530 4900.6  
- ParentEduc 5 7721.9 168414 4926.5  
- TestPrep 1 10613.1 171305 4950.6  
- Gender 1 12243.8 172936 4959.6  
- LunchType 1 12500.5 173192 4961.0  
  
Step: AIC=4882.99  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc + EthnicGroup +   
 ParentMaritalStatus  
  
 Df Sum of Sq RSS AIC  
+ IsFirstChild 1 1287.6 156884 4877.2  
+ WklyStudyHours 2 869.7 157302 4881.8  
<none> 158171 4883.0  
+ TransportMeans 1 253.9 157917 4883.5  
+ NrSiblings 1 2.2 158169 4885.0  
+ PracticeSport 2 103.8 158068 4886.4  
- ParentMaritalStatus 3 2520.5 160692 4892.0  
- EthnicGroup 4 3229.7 161401 4894.2  
- ParentEduc 5 7359.8 165531 4916.1  
- TestPrep 1 10869.6 169041 4944.0  
- LunchType 1 12075.1 170246 4950.7  
- Gender 1 12697.5 170869 4954.2  
  
Step: AIC=4877.24  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild  
  
 Df Sum of Sq RSS AIC  
+ WklyStudyHours 2 821.3 156062 4876.3  
<none> 156884 4877.2  
+ TransportMeans 1 246.1 156638 4877.8  
+ NrSiblings 1 11.1 156873 4879.2  
+ PracticeSport 2 139.9 156744 4880.4  
- IsFirstChild 1 1287.6 158171 4883.0  
- ParentMaritalStatus 3 2669.6 159553 4887.2  
- EthnicGroup 4 3338.0 160222 4889.2  
- ParentEduc 5 7333.9 164218 4910.6  
- TestPrep 1 10431.3 167315 4936.3  
- LunchType 1 12126.0 169010 4945.8  
- Gender 1 12407.5 169291 4947.4  
  
Step: AIC=4876.27  
ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild + WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
<none> 156062 4876.3  
+ TransportMeans 1 255.4 155807 4876.7  
- WklyStudyHours 2 821.3 156884 4877.2  
+ NrSiblings 1 4.0 156058 4878.2  
+ PracticeSport 2 116.5 155946 4879.6  
- IsFirstChild 1 1239.1 157302 4881.8  
- ParentMaritalStatus 3 2631.1 158694 4886.1  
- EthnicGroup 4 3170.0 159232 4887.3  
- ParentEduc 5 7463.6 163526 4910.6  
- TestPrep 1 10081.3 166144 4933.6  
- LunchType 1 11954.0 168016 4944.2  
- Gender 1 12340.8 168403 4946.4

Call:  
lm(formula = ReadingScore ~ Gender + LunchType + TestPrep + ParentEduc +   
 EthnicGroup + ParentMaritalStatus + IsFirstChild + WklyStudyHours,   
 data = data\_cleaned %>% select(-c(MathScore, WritingScore)))  
  
Coefficients:  
 (Intercept) Gendermale   
 68.4954 -7.2819   
 LunchTypestandard TestPrepnone   
 7.4941 -6.9723   
 ParentEduc.L ParentEduc.Q   
 7.6375 1.5354   
 ParentEduc.C ParentEduc^4   
 0.5659 1.5467   
 ParentEduc^5 EthnicGroup.L   
 -3.0328 4.1491   
 EthnicGroup.Q EthnicGroup.C   
 1.4383 -0.4714   
 EthnicGroup^4 ParentMaritalStatusmarried   
 -0.9416 4.1130   
 ParentMaritalStatussingle ParentMaritalStatuswidowed   
 1.2752 4.6451   
 IsFirstChildyes WklyStudyHours.L   
 2.4465 1.4306   
 WklyStudyHours.Q   
 -0.9328

Reading score stepwise model

| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 68.495 | <.001 |
| Gendermale | -7.282 | <.001 |
| EthnicGroup.L | 4.149 | 0.001 |
| EthnicGroup.Q | 1.438 | 0.203 |
| EthnicGroup.C | -0.471 | 0.645 |
| EthnicGroup^4 | -0.942 | 0.261 |
| ParentEduc.L | 7.638 | <.001 |
| ParentEduc.Q | 1.535 | 0.22 |
| ParentEduc.C | 0.566 | 0.635 |
| ParentEduc^4 | 1.547 | 0.158 |
| ParentEduc^5 | -3.033 | 0.001 |
| LunchTypestandard | 7.494 | <.001 |
| TestPrepnone | -6.972 | <.001 |
| ParentMaritalStatusmarried | 4.113 | 0.001 |
| ParentMaritalStatussingle | 1.275 | 0.363 |
| ParentMaritalStatuswidowed | 4.645 | 0.106 |
| IsFirstChildyes | 2.446 | 0.007 |
| WklyStudyHours.L | 1.431 | 0.135 |
| WklyStudyHours.Q | -0.933 | 0.191 |

Start: AIC=5186.91  
WritingScore ~ 1  
  
 Df Sum of Sq RSS AIC  
+ TestPrep 1 19817.7 205141 5101.5  
+ Gender 1 18846.0 206112 5106.0  
+ LunchType 1 13742.3 211216 5129.2  
+ ParentEduc 5 14395.9 210562 5134.2  
+ EthnicGroup 4 6413.2 218545 5167.5  
+ IsFirstChild 1 1481.7 223477 5182.6  
+ ParentMaritalStatus 3 2228.1 222730 5183.5  
+ WklyStudyHours 2 1494.5 223464 5184.6  
<none> 224958 5186.9  
+ TransportMeans 1 197.6 224761 5188.1  
+ NrSiblings 1 0.9 224957 5188.9  
+ PracticeSport 2 244.0 224714 5189.9  
  
Step: AIC=5101.49  
WritingScore ~ TestPrep  
  
 Df Sum of Sq RSS AIC  
+ Gender 1 19503.9 185637 5008.8  
+ LunchType 1 13830.7 191310 5037.3  
+ ParentEduc 5 13691.9 191449 5046.0  
+ EthnicGroup 4 6170.5 198970 5080.5  
+ ParentMaritalStatus 3 2609.5 202531 5095.4  
+ IsFirstChild 1 988.2 204152 5098.9  
+ WklyStudyHours 2 995.9 204145 5100.9  
<none> 205141 5101.5  
+ TransportMeans 1 290.1 204850 5102.1  
+ NrSiblings 1 1.3 205139 5103.5  
+ PracticeSport 2 216.4 204924 5104.5  
- TestPrep 1 19817.7 224958 5186.9  
  
Step: AIC=5008.78  
WritingScore ~ TestPrep + Gender  
  
 Df Sum of Sq RSS AIC  
+ LunchType 1 15135.3 170501 4930.2  
+ ParentEduc 5 12275.4 173361 4953.9  
+ EthnicGroup 4 6174.8 179462 4984.7  
+ ParentMaritalStatus 3 3148.0 182489 4998.6  
+ IsFirstChild 1 704.6 184932 5007.2  
+ WklyStudyHours 2 1015.7 184621 5007.6  
<none> 185637 5008.8  
+ TransportMeans 1 302.9 185334 5009.2  
+ NrSiblings 1 2.5 185634 5010.8  
+ PracticeSport 2 198.5 185438 5011.8  
- Gender 1 19503.9 205141 5101.5  
- TestPrep 1 20475.6 206112 5106.0  
  
Step: AIC=4930.15  
WritingScore ~ TestPrep + Gender + LunchType  
  
 Df Sum of Sq RSS AIC  
+ ParentEduc 5 12883.8 157618 4865.7  
+ EthnicGroup 4 5581.8 164920 4906.6  
+ ParentMaritalStatus 3 2642.1 167859 4921.3  
+ IsFirstChild 1 777.9 169723 4927.8  
+ WklyStudyHours 2 872.2 169629 4929.3  
<none> 170501 4930.2  
+ TransportMeans 1 297.0 170204 4930.5  
+ PracticeSport 2 409.6 170092 4931.9  
+ NrSiblings 1 10.1 170491 4932.1  
- LunchType 1 15135.3 185637 5008.8  
- TestPrep 1 20592.1 191094 5036.2  
- Gender 1 20808.5 191310 5037.3  
  
Step: AIC=4865.67  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc  
  
 Df Sum of Sq RSS AIC  
+ EthnicGroup 4 4636.0 152982 4845.4  
+ ParentMaritalStatus 3 2136.7 155481 4858.7  
+ IsFirstChild 1 757.3 156860 4863.1  
+ WklyStudyHours 2 1069.8 156548 4863.2  
+ PracticeSport 2 755.4 156862 4865.1  
<none> 157618 4865.7  
+ TransportMeans 1 168.3 157449 4866.7  
+ NrSiblings 1 43.8 157574 4867.4  
- ParentEduc 5 12883.8 170501 4930.2  
- LunchType 1 15743.7 173361 4953.9  
- Gender 1 19329.5 176947 4973.3  
- TestPrep 1 20021.2 177639 4977.0  
  
Step: AIC=4845.37  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup  
  
 Df Sum of Sq RSS AIC  
+ ParentMaritalStatus 3 2649.9 150332 4834.8  
+ IsFirstChild 1 855.2 152126 4842.1  
+ WklyStudyHours 2 884.8 152097 4843.9  
+ PracticeSport 2 781.4 152200 4844.5  
<none> 152982 4845.4  
+ TransportMeans 1 195.3 152786 4846.2  
+ NrSiblings 1 47.8 152934 4847.1  
- EthnicGroup 4 4636.0 157618 4865.7  
- ParentEduc 5 11937.9 164920 4906.6  
- LunchType 1 15250.4 168232 4933.5  
- Gender 1 19620.2 172602 4957.8  
- TestPrep 1 19985.4 172967 4959.8  
  
Step: AIC=4834.8  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup +   
 ParentMaritalStatus  
  
 Df Sum of Sq RSS AIC  
+ IsFirstChild 1 988.7 149343 4830.5  
+ WklyStudyHours 2 849.4 149482 4833.4  
+ PracticeSport 2 798.0 149534 4833.8  
<none> 150332 4834.8  
+ TransportMeans 1 146.8 150185 4835.9  
+ NrSiblings 1 29.5 150302 4836.6  
- ParentMaritalStatus 3 2649.9 152982 4845.4  
- EthnicGroup 4 5149.2 155481 4858.7  
- ParentEduc 5 11366.3 161698 4893.9  
- LunchType 1 14758.1 165090 4921.6  
- Gender 1 20199.9 170532 4952.3  
- TestPrep 1 20367.4 170699 4953.3  
  
Step: AIC=4830.55  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild  
  
 Df Sum of Sq RSS AIC  
+ PracticeSport 2 861.1 148482 4829.1  
+ WklyStudyHours 2 807.7 148535 4829.4  
<none> 149343 4830.5  
+ TransportMeans 1 141.6 149201 4831.6  
+ NrSiblings 1 94.4 149249 4831.9  
- IsFirstChild 1 988.7 150332 4834.8  
- ParentMaritalStatus 3 2783.4 152126 4842.1  
- EthnicGroup 4 5279.2 154622 4855.5  
- ParentEduc 5 11321.1 160664 4889.8  
- LunchType 1 14807.3 164150 4918.2  
- TestPrep 1 19818.0 169161 4946.7  
- Gender 1 19871.7 169215 4947.0  
  
Step: AIC=4829.06  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild + PracticeSport  
  
 Df Sum of Sq RSS AIC  
+ WklyStudyHours 2 784.4 147698 4828.0  
<none> 148482 4829.1  
+ TransportMeans 1 141.3 148341 4830.2  
+ NrSiblings 1 95.5 148386 4830.5  
- PracticeSport 2 861.1 149343 4830.5  
- IsFirstChild 1 1051.8 149534 4833.8  
- ParentMaritalStatus 3 2798.3 151280 4840.8  
- EthnicGroup 4 5308.9 153791 4854.4  
- ParentEduc 5 11666.6 160149 4890.8  
- LunchType 1 15142.9 163625 4919.1  
- TestPrep 1 19702.3 168184 4945.2  
- Gender 1 19820.5 168302 4945.8  
  
Step: AIC=4828.04  
WritingScore ~ TestPrep + Gender + LunchType + ParentEduc + EthnicGroup +   
 ParentMaritalStatus + IsFirstChild + PracticeSport + WklyStudyHours  
  
 Df Sum of Sq RSS AIC  
<none> 147698 4828.0  
- WklyStudyHours 2 784.4 148482 4829.1  
+ TransportMeans 1 148.6 147549 4829.1  
- PracticeSport 2 837.9 148535 4829.4  
+ NrSiblings 1 73.5 147624 4829.6  
- IsFirstChild 1 1006.8 148704 4832.5  
- ParentMaritalStatus 3 2765.8 150463 4839.6  
- EthnicGroup 4 5112.3 152810 4852.3  
- ParentEduc 5 11814.5 159512 4891.0  
- LunchType 1 14933.9 162631 4917.4  
- TestPrep 1 19221.8 166919 4942.0  
- Gender 1 19738.0 167436 4945.0

Call:  
lm(formula = WritingScore ~ TestPrep + Gender + LunchType + ParentEduc +   
 EthnicGroup + ParentMaritalStatus + IsFirstChild + PracticeSport +   
 WklyStudyHours, data = data\_cleaned %>% select(-c(MathScore,   
 ReadingScore)))  
  
Coefficients:  
 (Intercept) TestPrepnone   
 69.3473 -9.6288   
 Gendermale LunchTypestandard   
 -9.2094 8.3876   
 ParentEduc.L ParentEduc.Q   
 9.9826 1.3651   
 ParentEduc.C ParentEduc^4   
 0.2894 1.7154   
 ParentEduc^5 EthnicGroup.L   
 -3.1061 4.6731   
 EthnicGroup.Q EthnicGroup.C   
 0.6274 -1.8905   
 EthnicGroup^4 ParentMaritalStatusmarried   
 -1.6492 4.1346   
 ParentMaritalStatussingle ParentMaritalStatuswidowed   
 1.0557 3.9501   
 IsFirstChildyes PracticeSport.L   
 2.2077 2.2506   
 PracticeSport.Q WklyStudyHours.L   
 -0.7075 1.3382   
 WklyStudyHours.Q   
 -0.9604

Writing score stepwise model

| term | estimate | p.value |
| --- | --- | --- |
| (Intercept) | 69.347 | <.001 |
| Gendermale | -9.209 | <.001 |
| EthnicGroup.L | 4.673 | <.001 |
| EthnicGroup.Q | 0.627 | 0.569 |
| EthnicGroup.C | -1.891 | 0.058 |
| EthnicGroup^4 | -1.649 | 0.043 |
| ParentEduc.L | 9.983 | <.001 |
| ParentEduc.Q | 1.365 | 0.265 |
| ParentEduc.C | 0.289 | 0.803 |
| ParentEduc^4 | 1.715 | 0.109 |
| ParentEduc^5 | -3.106 | 0.001 |
| LunchTypestandard | 8.388 | <.001 |
| TestPrepnone | -9.629 | <.001 |
| ParentMaritalStatusmarried | 4.135 | 0.001 |
| ParentMaritalStatussingle | 1.056 | 0.44 |
| ParentMaritalStatuswidowed | 3.950 | 0.16 |
| PracticeSport.L | 2.251 | 0.022 |
| PracticeSport.Q | -0.708 | 0.335 |
| IsFirstChildyes | 2.208 | 0.012 |
| WklyStudyHours.L | 1.338 | 0.152 |
| WklyStudyHours.Q | -0.960 | 0.167 |

## Conclusions/Discussion

## A brief summary on each group member’s contribution (method, data analysis, writing, etc).