Student Performance

A Multiple Linear Regression Analysis

Ben Lawrence

Michael Xie

Introduction

In this paper we investigate the question of which student attributes affect student academic performance. The adage "Skip class you won't pass?" presents us with what might appear to be an obvious relationship. If a student does not show up to learn their class material, they will inevitably struggle when it comes time to evaluate their learning. Many state and national governments have adopted policies to ensure that students "stay in school", and parents who allow their children to be absent for enough days are punished.

Some educational researchers believe that there is a strong effect that failing a class has on subsequent student performance. If a student fails one class, they may fall into a pattern where they are psychologically influenced to perceive their previous effort as wasted. In the pressure to catch up in the following grading period, students may lose the persistence and self-efficacy needed to succeed. In this way, failure can compound and become a predictor of student performance.

The last major predictor of student academic performance is prior academic performance. Students are often judged by how they performed previously. Some hold to the assumption that how one performed before will not differ significantly from how one performs in the future. While the logic of this assumption may seem flawed, it has not stopped universities from using placement tests and secondary school grades of applicants as measures of how they might perform as university students. We will investigate these and other predictors of student success as we proceed in our regression analysis in this paper.

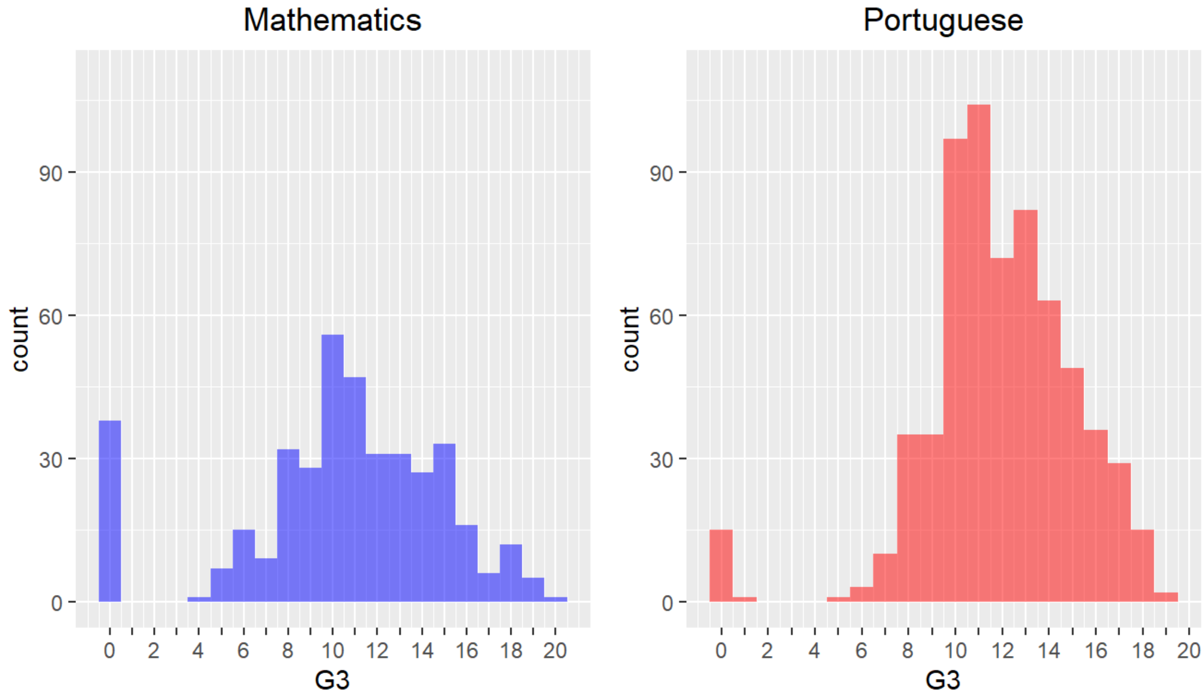
Data Description

We obtained the two datasets from the UC Irvine Machine Learning Repository (UCIMLR). They were donated by authors, Paulo Cortez and Alice Silva, of the article “Using Data Mining to Predict Secondary School Student Performance”. The authors belong to the University of Minho in Portugal and were interested in model efficacy in the prediction of the academic performance of Portuguese secondary students.

Each dataset aligns with a single subject, Portuguese and Mathematics, collected from two Portuguese secondary schools. The Mathematics dataset contains 395 observations and the Portuguese contains 649. They both have 33 columns.

Three of the columns are trimester grades of students with the third trimester grade **G3** taken as the student’s final grade. We will take this as our response and the previous two trimester grades as predictors. The remaining predictors consist of demographic survey data acquired for each student observation. These include number of absences, number of previous class failures, occupation of mother/father, etc. Most of the variables are categorical, taking ordinal and nominative values. There are also variables that take on binary and integer values. We will begin our analysis by visualizing the distribution of **G3** for the Mathematics and Portuguese data set.

Figure 1: Histogram of 3rd Trimester Grades by Subject

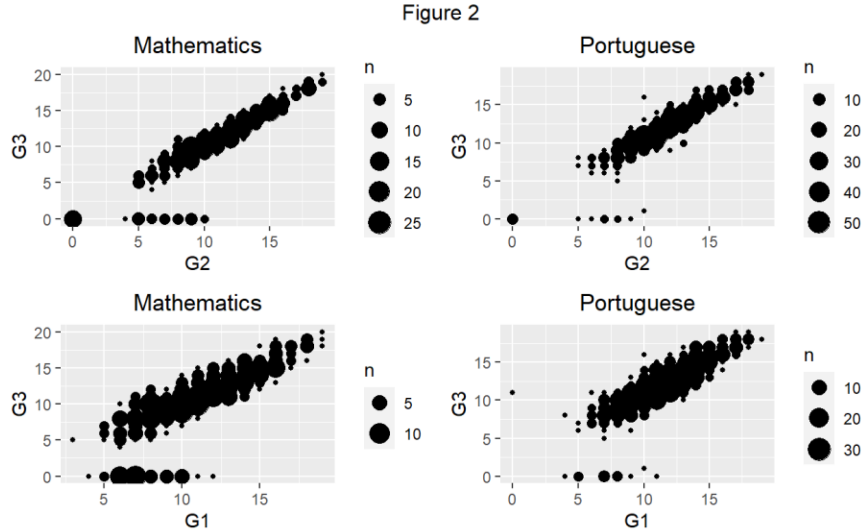


In **Figure 1** we can observe that the grades are skewed to the right. This is reasonable since we would expect a higher proportion of students passing (scoring over 10) than failing. There is a notable set of outlying points where a much higher proportion of students received 0s than we would reasonably expect. This is especially true in the mathematics class.

Figure 2 below provides examples of the predictors available in our dataset. (Full list available HERE) We will use all of these in our full linear model. Before creating such a model however, we will take time to analyze the relationship between G1, G2, absences, and failures with the response G3 to see if transformations of these variables are warranted. The other variables are categorical and do not offer obvious visual correlations with G3 that can inform our transformation choices.

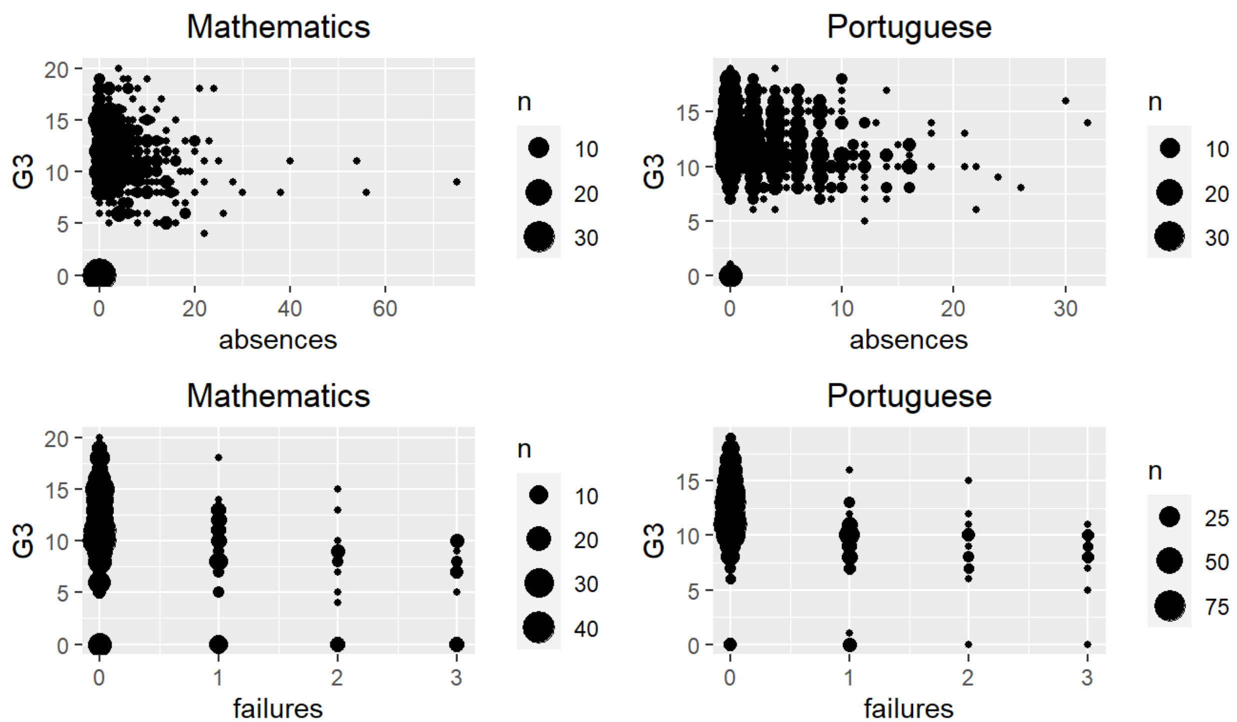
First, we examine scatterplots of G3 vs. G1 and G2. We can observe a clear linear correlation between the response and both predictors in **Figure 3**. We can also observe that outliers exist where students who received varying grades for G2 and G3 suddenly received 0 grades for G3, contrasting the linear trend.

Figure 2



In **Figure 3**, we examine the relationship between G3 and the predictors absences and failures. We can see a correlation between absences and G3, but it is not linear. This suggests a transformation is needed to appropriately fit absences into a linear model. We will use a square root transformation as this is normal practice when dealing with a count variable such as number of absences. We can observe that failures only take on values 0-3 so although it may seem similar to absences in terms of being count data we will treat it as a categorical variable and thus not apply a transformation.

Figure 3



Methods and Results

Now we will construct our full model containing all predictors in our dataset, appropriately classifying all categorical variables as factors in R.

```
Figure 4: Summary for Mathematics Model
Significant Coefficients:
            Estimate Std. Error   t value     Pr(>|t|)
age       -0.2167846 0.10062393 -2.154403 3.194768e-02
failures1 -0.9389238 0.32994127 -2.845730 4.714093e-03
absences   0.4292284 0.06921107  6.201731 1.706608e-09
G1         0.2064483 0.06139562  3.362589 8.649267e-04
G2         0.9228330 0.05252456 17.569552 5.701784e-49

R-squared:  0.8726161
Adjusted R-squared:   0.8446153
F-statistic:  31.16396  on  71  and  323  DF, p-value:  8.483994e-109

GVIF for provided predictors:
G1 :  5.016762
G2 :  4.71554
```

```
Figure 5: Summary for Portuguese Model
Significant Coefficients:
              Estimate Std. Error   t value     Pr(>|t|)
reasonother -0.36219563 0.17468951 -2.073368 3.858122e-02
traveltime4  0.78846906 0.34980198  2.254044 2.456777e-02
failures1   -0.53967035 0.18770903 -2.875037 4.188834e-03
Dalc4       -0.84274543 0.35158917 -2.396961 1.684879e-02
absences     0.09847597 0.04307662  2.286065 2.261203e-02
G1           0.13739308 0.03868280  3.551788 4.138774e-04
G2           0.85282116 0.03632078 23.480255 4.447773e-86

R-squared:  0.867014
Adjusted R-squared:   0.8506501
F-statistic:  52.98315  on  71  and  577  DF, p-value:  4.063962e-209

GVIF for provided predictors:
G1 :  4.688056
G2 :  4.655541
```

In **Figure 4** and **Figure 5** we observe that our models explain more than 80% of the variance for their respective classes. G1 and G2 are very significant in both with low p-values. Absences is significant and having 1 previous class failure is also significant. The variables reason, traveltime, age, and Dalc (Weekday alcohol consumption) disagree between the classes on their significance in predicting G3. This could be a true population difference between academic subjects, but the relatively high p-values of these variables leads us to hypothesize these variables are not actually significant.

We also can see the generalized variance inflation factors (GVIFs) for G1 and G2 are high, showing a likely issue with multicollinearity. This is reasonable since our hypothesis is that prior grades predict future grades, so it would naturally follow that G1 and G2 are highly correlated.

Next we will examine our model diagnostics for the full model.

Figure 6

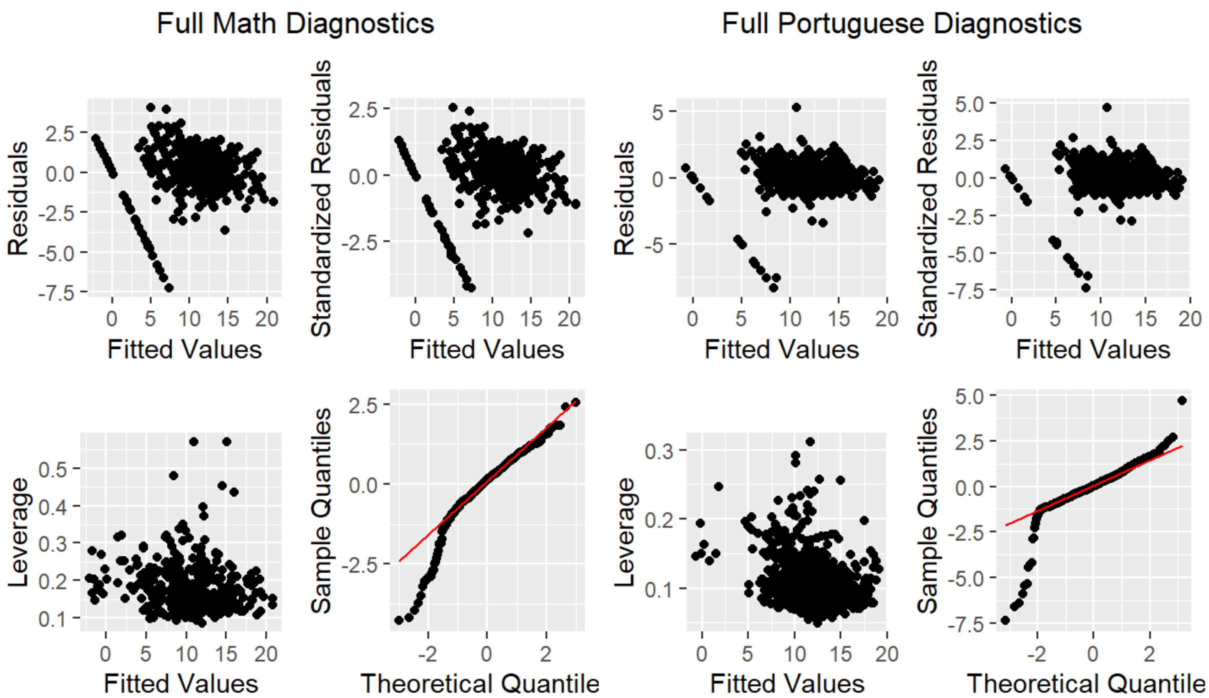


Figure 6 shows a clear division in the distribution of points in the plots of standardized residuals vs. fitted values. One set demonstrates relatively constant variance, and the other shows a clear linear pattern in its variance. Additionally, we can see that the QQ plot shows drop off from rest of the points on the lower end of Theoretical Quantiles, meaning the sample quantiles are much lower than what we would expect from normally distributed data.

To further examine our assumption of normality we will apply the Shapiro-Wilk test to both classes.

Figure 7

```
        Shapiro-Wilk normality test

data:  rstandard(lm_full_math)
W = 0.91826, p-value = 8.046e-14
```

```
        Shapiro-Wilk normality test

data:  rstandard(lm_full_port)
W = 0.81994, p-value < 2.2e-16
```

Figure 7 gives p-values for both classes that allow us to reject the null hypothesis that are residuals are normally distributed. Although the W values may allow us to claim that the full mathematics model may satisfy our assumptions enough to draw conclusions, we see a lower W value for the full Portuguese model which was constructed from more observations than the mathematics model.

Our diagnostics of the full multiple linear regression models showed multiple issues. These issues as discussed above, alongside non-normal results from Shapiro-Wilk tests and relatively high GVIF values provides evidence that our assumptions for MLR Regression in the full model are not satisfied. This leads us to attempt to reduce and adjust our full model into a model that will improve our diagnostics.

Investigation into the points of the deviant standardized residual line for both models revealed that the 22 of 24 (Portuguese: 8/10, Mathematics: 14/14) had drop offs of their G3 grade to 0, and all 24 had no absences recorded. There is not information about what these 0s entail in the data set, nor in the research article where the data set was obtained. It is possible that these outliers were data entry errors, a non-pass being entered as a 0 instead of the actual G3 grade. However, this is only speculation, and, per the article, students were "dropped due to missing data".

Because of these facts, we decided to remove the points that exhibited the drop off of the G3 grade from the data sets due to the large bias caused by outliers and the lack of clarity on these points' validity to the study.

Figure 8

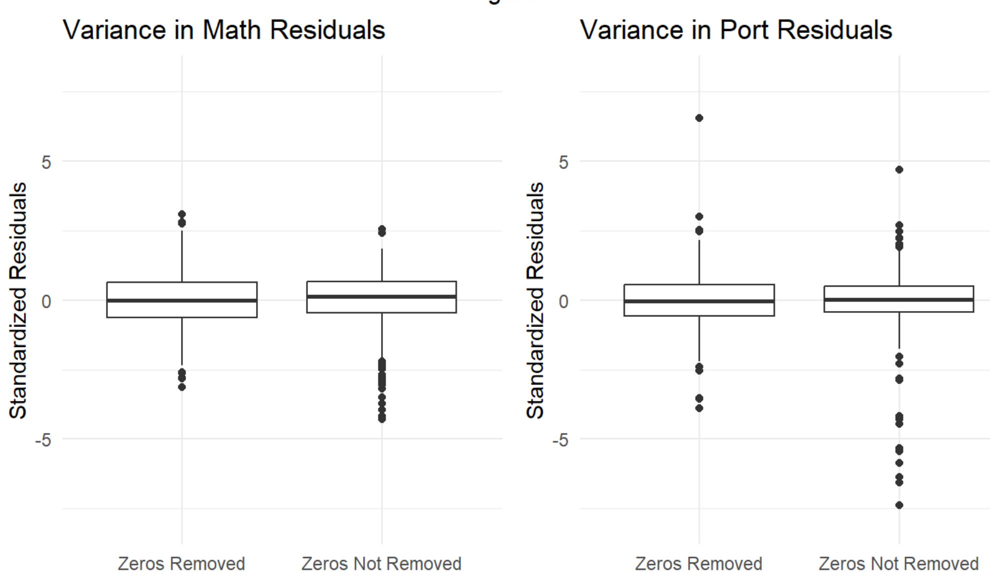
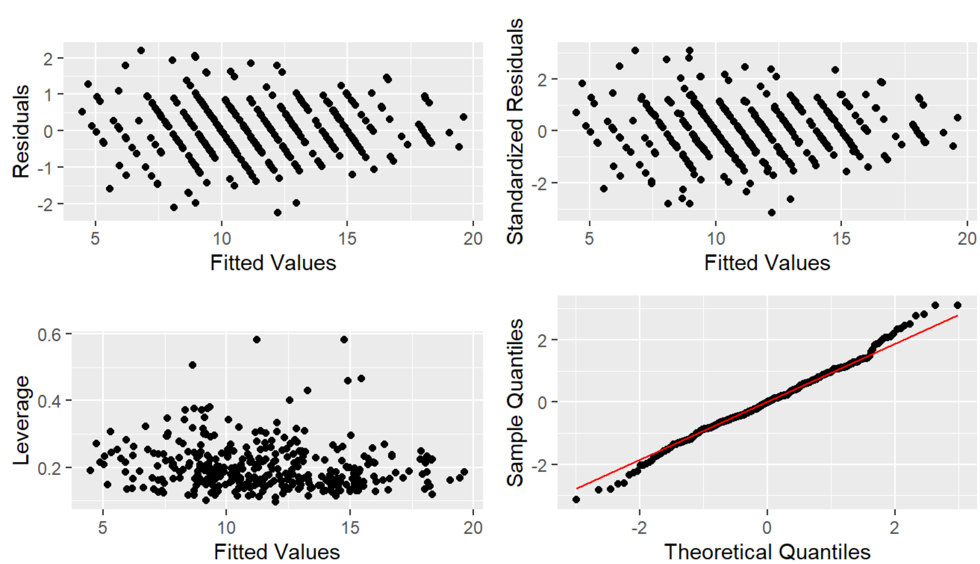


Figure 9



From the box plots in **Figure 8** we can see that there is a clear improvement in how uniformly dispersed the standardized residuals are once we remove the rows where G3 = 0. In **Figure 9** we also see improvement in our diagnostic plots for the full math model with the zeros removed. Apart from a few outlying points, most of the residuals seem randomly distributed with constant variance. However there appears to be a fundamental pattern in the residuals where they are divided into discrete groups that are arranged diagonally. This is actually a reasonable outcome since our response variable is a discrete integer. These results are also present in the Portuguese full model with removed zeroes. Thus, it follows that each increment in G3 will have its own group of residual points distributed around it.

Nonetheless this is a fundamental problem with our approach because a linear model assumes a continuous relationship between response and predictors. One way of solving this is to change our response variable into an average of the three grade variables (G1, G2, and G3).

After creating full models with GPA as a predictor with G1, G2, and G3 dropped, we can see an improvement in our model diagnostics, where the residuals are no longer separated into discrete sets of diagonal patterns. We then used a Backwards Stepwise variable selection procedure with AIC as our selection criterion.

```
Figure 10: Summary for Reduced Mathematics Model
Significant Coefficients:
              Estimate Std. Error   t value     Pr(>|t|)
(Intercept)  14.3787365  1.0802483 13.310585 1.336609e-32
sexM          0.6538912  0.3145877  2.078566 3.843670e-02
Mjobhealth    1.7742185  0.6286054  2.822468 5.057305e-03
Mjobservices  1.2366072  0.4856457  2.546316 1.134482e-02
studytime3    1.1121153  0.4943670  2.249574 2.514078e-02
studytime4    1.4192375  0.6444916  2.202104 2.835566e-02
failures1    -1.2763254  0.4976691 -2.564606 1.077532e-02
failures2    -2.4552254  0.8530506 -2.878171 4.263105e-03
failures3    -3.4877791  0.8570681 -4.069431 5.913990e-05
schoolsupyes -2.0909554  0.4291309 -4.872535 1.720352e-06
famsupyes    -0.7231220  0.3049584 -2.371215 1.830875e-02
health3      -1.6037258  0.5095535 -3.147316 1.799429e-03
health4      -1.0925124  0.5464915 -1.999139 4.642055e-02
health5      -1.3146756  0.4730484 -2.779156 5.765207e-03
absences     -0.3168997  0.1027885 -3.083027 2.223480e-03

R-squared:  0.347129
Adjusted R-squared:  0.2892291
F-statistic:  5.995331  on  29  and  327  DF, p-value:  2.314663e-17
```

```
Figure 11: Summary for Reduced Portuguese Model
Significant Coefficients:
                Estimate Std. Error   t value     Pr(>|t|)
(Intercept)    6.4453534 1.67247052  3.853792 1.288831e-04
schoolMS      -0.9448490 0.19681446 -4.800709 2.000097e-06
sexM          -0.6019816 0.17984094 -3.347300 8.672519e-04
age            0.2480849 0.07849408  3.160555 1.654291e-03
Fjobteacher    1.2793288 0.49557161  2.581522 1.007346e-02
guardianmother -0.4110824 0.19829248 -2.073111 3.858961e-02
traveltime4   -1.2853433 0.53593022 -2.398341 1.677493e-02
studytime3     0.6854127 0.27563764  2.486644 1.316715e-02
studytime4     0.9970648 0.38685651  2.577351 1.019446e-02
failures1     -1.8830489 0.29520891 -6.378699 3.577825e-10
failures2     -2.7206480 0.56727597 -4.795987 2.046013e-06
failures3     -2.4942695 0.59429103 -4.197051 3.115044e-05
schoolsupyes  -1.1818250 0.27687669 -4.268416 2.288935e-05
higheryes      1.6447484 0.29916989  5.497707 5.704702e-08
goout2         0.9004455 0.35731720  2.520017 1.199418e-02
health5       -0.5180421 0.25963638 -1.995260 4.646778e-02
absences      -0.3198224 0.06776225 -4.719773 2.943331e-06

R-squared:  0.4154338
Adjusted R-squared:  0.38122
F-statistic:  12.14231  on  35  and  598  DF, p-value:  1.396738e-49
```

From **Figures 10 and 11** we can see there has been a large drop in the R^2 value of both models. Our reduced models do not explain as much of the variance in GPA as the prior models explained with G3.

Conclusion

For both classes, Mathematics and Portuguese, we saw the most significant predictors were G1 and G2 when in models where they were available. In our final reduced models, both classes share the following variables as significant: absences, failures (1,2,3), studytime (3,4), health(5), and being male. Interestingly being male has a positive coefficient in Mathematics and a negative coefficient in the Portuguese class, implying equity differences in the curriculums of the two subjects. The rest of the predictors are specific to each subject in their significance and will not be further investigated here as our goal is to generalize predictors of student performance.

On this question, our analyses lead us to several conclusions. First, in using a linear model, prior grades were the strongest predictor of G3 across both classes. This implies that students who perform well early on will generally continue to perform well and vice versa. It should kept in mind that the prior grades we were using as predictors come from the same class environment in the same academic year, so it is reasonable that students would perform similarly in the same class through the school year. The strength of prior performance may not be as strong when taken to predict performance in a different academic environment (high school vs. college, between subjects, etc.). However research on this subject does confirm prior academic performance generalizes as a strong predictor (Alyahyan, Düştegör).

Even so, prior academic performance is not a very useful predictor if the intention is to improve student performance. The factors that correlate with prior academic performance also correlate with future performance so no useful information is obtained if one wishes to design interventions to improve student performance.

The next significant variables in our final reduced models were absences and all levels of failures except 0. This is reasonable as it implies that students who miss class will do poorly (likely because of the additional effort of catching up). According to our results, students with no previous class failures do not possess a significant advantage but students with at least 1 previous failure experience a significant disadvantage in the magnitude of GPA. This is likely due to a psychological response where the experience of failure negatively affects student performance. (Ajjawi, Zacharias, et. al).

The most puzzling significant predictor is health5. This is the highest health rating, but for both Math and Portuguese it seems to be a significant negative predictor of GPA, implying students who self report as very healthy are at an academic disadvantage. In Mathematics there is also a negative relationship with other levels of health rating. This is likely not a truly significant predictor since it is unreasonable to claim that the better a student feels about their health, the worse they will perform.

To close, we are not confident that a linear model is ideal for predicting student performance in the context of our data. While some predictors may provide hints at possible interventions to improve student performance (Equity interventions, Delinquency prevention programs, Psychological/Academic Interventions to manage effects of failing a class), many predictors that could contribute to the improvement of student performance are left out of our model completely or produce unreasonable results. Because of this it is our judgement that nonlinear models, as pursued in wider research (Cortez, Silva) should be investigated to see how they may compensate for the problems previously discussed. They may be superior at detecting the significance of useful predictors and less sensitive to the strong linear relationship between prior grades and subsequent grades (in the context of our data), leading to more targeted interventions to help improve student performance.

References

Ajjawi, R., Dracup, M., Zacharias, N., Bennett, S., & Boud, D. (2020). Persisting students' explanations of and emotional responses to academic failure. *Higher Education Research & Development*, *39*(2), 185–199. https://doi.org/10.1080/07294360.2019.1664999

Alyahyan, E., Düştegör, D. Predicting academic success in higher education: literature review and best practices. *Int J Educ Technol High Educ* **17**, 3 (2020). https://doi.org/10.1186/s41239-020-0177-7

Cortez, P., & Silva, A.M. (2008). Using data mining to predict secondary school student performance.

Code Appendix

```
{r}
# Custom functions to avoid redundant coding

# Creates basic scatter plots using ggplot2 between two continuous variables
gg_basic <- function(data, x, y, title="") {
  plot <- ggplot(data = data, aes(x = {{x}}, y = {{y}})) +
          geom_point() +
          labs(title=title) +
          theme(plot.title = element_text(hjust = 0.5))
  return(plot)
}

# Similar to gg_basic but uses geom_count to plot counts of points at locations;
useful when data has overlapping points
gg_basic_count <- function(data, x, y, title="") {
  plot <- ggplot(data = data, aes(x = {{x}}, y = {{y}})) +
          geom_count() +
          labs(title=title) +
          theme(plot.title = element_text(hjust = 0.5))
  return(plot)
}
```

```
{r}
print_summary <- function(model, p_value = 0.05, predictors = NULL) {
  # Extract the summary of the model
  summary_model <- summary(model)

  # Filter significant coefficients
  significant_coefs <- summary_model$coefficients[summary_model$coefficients[, "Pr
(>|t|)"] < p_value, ]

  # Print significant coefficients
  cat("Significant Coefficients:\n")
  print(significant_coefs)

  # Print R-squared and Adjusted R-squared
  cat("\nR-squared: ", summary_model$r.squared, "\n")
  cat("Adjusted R-squared: ", summary_model$adj.r.squared, "\n")

  # Print F-statistic
  f_value <- summary_model$fstatistic[1]
  f_df1 <- summary_model$fstatistic[2]
  f_df2 <- summary_model$fstatistic[3]
  f_pvalue <- pf(f_value, f_df1, f_df2, lower.tail = FALSE)
  cat("F-statistic: ", f_value, " on ", f_df1, " and ", f_df2, " DF, p-value: ",
f_pvalue, "\n")

  # If predictors are provided, calculate and print the VIF for each predictor
  if (!is.null(predictors) && length(predictors) > 0) {
      vif_values <- vif(model)
      cat("\nGVIF for provided predictors:\n")
      # Loop through each predictor and print its VIF value
      for (predictor in predictors) {
          cat(predictor, ": ", vif_values[predictor, 1], "\n")
      }
  }
}
```

```r
{r}
#Creates a 2x2 grid of basic diagnostic plots of a lm object
lm_diag <- function(lm_model, title = "", limResid = NULL, limSresid = NULL, limLev
= NULL, limQQ = NULL) {
  # Helper function to apply y-axis limits if provided
  apply_limits <- function(plot, ylim) {
    if (!is.null(ylim)) {
      plot + coord_cartesian(ylim = ylim)
    } else {
      plot
    }
  }

  # Generate diagnostic plots
  p1 <- gg_basic(lm_model$model, fitted(lm_model), residuals(lm_model)) +
    labs(x="Fitted Values", y="Residuals")
  p1 <- apply_limits(p1, limResid)

  p2 <- gg_basic(lm_model$model, fitted(lm_model), rstandard(lm_model)) +
    labs(x="Fitted Values", y="Standardized Residuals")
  p2 <- apply_limits(p2, limSresid)

  p3 <- gg_basic(lm_model$model, fitted(lm_model), hatvalues(lm_model)) +
    labs(x="Fitted Values", y="Leverage")
  p3 <- apply_limits(p3, limLev)

  p4 <- ggplot(lm_model$model, aes(sample = rstandard(lm_model))) +
        stat_qq() +
        stat_qq_line(colour="red") +
        labs(x ="Theoretical Quantiles", y="Sample Quantiles")
  p4 <- apply_limits(p4, limQQ)

  # Arrange all plots into a grid and return the combined plot
  return(grid.arrange(p1, p2, p3, p4, ncol=2, nrow=2, top = title))
}
```

```r
{r}
# Histogram for math_perf data frame
p1 <- ggplot(data = math_perf, aes(x = G3)) +
      geom_histogram(fill = "blue", alpha = 0.5, binwidth = 1) +
      scale_x_continuous(breaks = 0:20, labels = ifelse((0:20) %% 2 == 0, as
.character(0:20), "")) +  # Label only even numbers
      ylim(0, 110) +
      labs(title = "Mathematics") +
      theme(plot.title = element_text(hjust = 0.5))  # Center the title under the
plot

# Histogram for port_perf data frame
p2 <- ggplot(data = port_perf, aes(x = G3)) +
      geom_histogram(fill = "red", alpha = 0.5, binwidth = 1) +
      scale_x_continuous(breaks = 0:20, labels = ifelse((0:20) %% 2 == 0, as
.character(0:20), "")) +  # Label only even numbers
      ylim(0, 110) +
      labs(title = "Portuguese") +
      theme(plot.title = element_text(hjust = 0.5))  # Center the title under the
plot


grid.arrange(p1, p2, ncol = 2, top = "Figure 1: Histogram of 3rd Trimester Grades
by Subject")
```

```r
{r}
p1 <- gg_basic_count(math_perf,G2,G3, title="Mathematics")
p2 <- gg_basic_count(port_perf,G2,G3, title="Portuguese")
p3 <- gg_basic_count(math_perf,G1,G3, title="Mathematics")
p4 <- gg_basic_count(port_perf,G1,G3, title="Portuguese")

grid.arrange(p1, p2,p3, p4, ncol=2, nrow=2, top="Figure 2")
```

```r
{r}
# Factor these data columns
fcols <- c('school', 'sex', 'address', 'famsize', 'Pstatus', 'Medu', 'Fedu', 'Mjob'
, 'Fjob', 'traveltime', 'studytime', 'reason', 'guardian', 'schoolsup', 'famsup',
'paid', 'failures', 'activities', 'nursery', 'higher', 'internet', 'romantic',
'famrel', 'freetime', 'goout', 'Dalc', 'Walc', 'health')

# Factorize columns in both dataframes
math_perf[fcols] <- lapply(math_perf[fcols], factor)
port_perf[fcols] <- lapply(port_perf[fcols], factor)  # Corrected this line

# Transform absences - Square Root
math_perf$absences <- math_perf$absences^(0.5)
port_perf$absences <- port_perf$absences^(0.5)
```

```r
{r}
cat("Figure 4: Summary for Mathematics Model \n")
lm_full_math <- lm(G3 ~ . - G3, data=math_perf)
print_summary(lm_full_math, predictors = c("G1","G2"))
```

```r
{r}
cat("Figure 5: Summary for Portuguese Model \n")
lm_full_port <- lm(G3 ~ . - G3, data=port_perf)
print_summary(lm_full_port, predictors = c("G1","G2"))
```

```r
{r}
p10 <- lm_diag(lm_full_math, title = "Full Math Diagnostics")
p11 <- lm_diag(lm_full_port, title = "Full Portuguese Diagnostics")

grid.arrange(p10, p11, ncol = 2, top = "Figure 6")
```

```r
{r echo = TRUE}
cat("Figure 7 \n")
shapiro.test(rstandard(lm_full_math))
shapiro.test(rstandard(lm_full_port))
```

```r
{r}
math_perf_remove <- subset(math_perf, G3 != 0)
port_perf_remove <- subset(port_perf, G3 != 0)

lm_full_math_remove <- lm(G3 ~ . - G3, data=math_perf_remove)
lm_full_port_remove <- lm(G3 ~ . - G3, data=port_perf_remove)
```

```{r}
# Extract standardized residuals
resid_math_remove <- rstandard(lm_full_math_remove)
resid_port_remove <- rstandard(lm_full_port_remove)
resid_math <- rstandard(lm_full_math)
resid_port <- rstandard(lm_full_port)

# Create a data frame for plotting
data_math <- data.frame(
  Residuals = c(resid_math_remove, resid_math),
  Condition = factor(rep(c("Zeros Removed", "Zeros Not Removed"),
                         c(length(resid_math_remove), length(resid_math))),
                     levels = c("Zeros Removed", "Zeros Not Removed"))
)

data_port <- data.frame(
  Residuals = c(resid_port_remove, resid_port),
  Condition = factor(rep(c("Zeros Removed", "Zeros Not Removed"),
                         c(length(resid_port_remove), length(resid_port))),
                     levels = c("Zeros Removed", "Zeros Not Removed"))
)

# Plot for Math models
p9 <- ggplot(data_math, aes(x = Condition, y = Residuals)) +
  geom_boxplot() +
  labs(title = "Variance in Math Residuals",
       x = "",
       y = "Standardized Residuals") +
  coord_cartesian(ylim = c(-8, 8)) +
  theme_minimal()

# Plot for Portuguese models
p10 <- ggplot(data_port, aes(x = Condition, y = Residuals)) +
  geom_boxplot() +
  labs(title = "Variance in Port Residuals",
       x = "",
       y = "Standardized Residuals") +
  coord_cartesian(ylim = c(-8, 8)) +
  theme_minimal()

grid.arrange(p9, p10, ncol = 2, top = "Figure 8")
```

```{r}
lm_diag(lm_full_math_remove, title = "Figure 9")
```

```{r}
math_perf_remove$GPA <- (math_perf_remove$G1 + math_perf_remove$G2 +
math_perf_remove$G3)/3

port_perf_remove$GPA <- (port_perf_remove$G1 + port_perf_remove$G2 +
port_perf_remove$G3)/3

lm_full_math_remove <- lm(GPA ~ . - G3 - G2 - G1, data=math_perf_remove)
lm_full_port_remove <- lm(GPA ~ . - G3 - G2 - G1, data=port_perf_remove)
```

```{r}
lm_math_reduced <- step(lm_full_math_remove, trace = 0)
lm_port_reduced <- step(lm_full_port_remove, trace = 0)
```

```{r}
cat("Figure 10: Summary for Reduced Mathematics Model \n")
print_summary(lm_math_reduced)
cat("\n")

cat("Figure 11: Summary for Reduced Portuguese Model \n")
print_summary(lm_port_reduced)
```