Predicting Student Overall Course Grades Using Sex, Midterm Exam Scores, and the Semester and Year the Students Took the Course

[redacted]

Abstract

This study aims to identify what factors are predictive of college student's final course grade. A dataset published by OpenIntro was used to create a multiple linear regression predicting student's final course grades in a statistics course. Student average scores of exams 1 and 2, the year the course was taken, the sex of the student, and interactions between exam 3 scores and year were used to predict overall course grades. The results suggest that measures of academic performance (e.g., midterms) and individual characteristics (e.g., sex) are predictive of overall course grades among college students. Furthermore, tracking midterm grades could also be useful in identifying students who are at risk of failing coursework. Future research should examine how a student's major, year in school, and content of the course predict college student performance in their coursework.

Background and Significance

Tracking and predicting student success in college has important implications for attrition rates, early detection of the risk of dropout, and the identification of students who may need extra academic support (Ashenafi et al., 2015). Many institutions often rely on standardized tests, such as the ACT, SAT, and GRE to predict college performance**.** However, it is important that research and institutions continue to track student's academic performance to ensure their success and support in the course. Jensen and Barron (2014) found that exam grades and final grades were highly correlated, and that exam grades stayed fairly consistent throughout the semester. This research suggests that students struggling on the first exam may struggle throughout the rest of the semester, which will reflect on the student's final course grade.

Past research has also demonstrated gender differences in college coursework (e.g., exam performance, participation, and homework; Kost-Smith et al., 2010). Therefore, it is important that researchers consider the impact of gender when examining student performance in coursework. Furthermore, past research has examined the difference in course grades across years and semesters. Some studies have chalked this up to a difference in grading systems, teachers, or the format of coursework (Reardon et al., 2007). However, Reardon et al. (2007) tracked undergraduate performance across 26 years and found that students performed relatively consistently across quarters, semesters, and years until the more consistent use of the internet in the education system.

Research in this area is inconclusive, and leaves questions such as, "Does a student's midterm exam scores, sex, semester, and year the course was taken predict overall course grades among college students?" This study aims to contribute to this gap in the literature by using sex, the average of exam one and two scores, exam three scores, semester the course was taken, and the year the course was taken to predict overall student course grades in a statistics course.

Data Collection and Method Analysis

This study utilized a dataset of statistics students obtained from OpenIntro (https://www.openintro.org/data/index.php?data=exam_grades). Data was collected from (n = 233) students from an undergraduate statistics course at a university. It is important to note that the intended population is all college students. However, this was not a random sample, and students in this sample may have unique or different characteristics than the general population of college students. For example, some students may have more experience or interest in statistics than others. Our response variable is the overall course grade in percentage students received for their course grade.

Descriptive statistics were examined and recorded for each of the variables. The independent variables recorded were the average score of exam 1 and exam 2, exam 3 scores, sex (i.e., man or woman), the year the student took the statistics course (i.e., 2000, 2001, 2002, and 2003), and the semester they took the course (i.e., semester 1 or semester 2). The mean, standard deviation, and the minimum and maximum observations for all of the quantitative variables were recorded (Table 1). The count and proportions were recorded for the categorical variables (Table 2). We also calculated the correlations between each of the quantitative variables (Table 3). Based on the correlations, overall course grade and exam 3 scores had strong positive relationships, as well as overall course grade and the average scores of exams 1 and 2. Exam 3 scores and the average score of exams 1 and 2 had a moderately strong relationship.

To better understand these relationships, we looked at the slopes for the quantitative variables to check for a quadratic term (Figure 1). We concluded there was insufficient evidence of a concave up or concave down relationship amongst our variables to incorporate a quadratic term. Based on the slopes, overall course grade and exam 3 scores have a moderately strong linear relationship with no outliers. Overall course grade and the average score of exams 1 and 2 have a strong positive linear relationship with no outliers. Lastly, exam 3 scores and the average score of exams 1 and 2 had a strong positive linear relationship with no outliers.

There were also three different interactions we examined to best predict the overall course grade. We created one plot with an interaction between the average score of exams 1 and 2 and the year the student took the statistics course (Figure 2), and there was an indication of a slight interaction. We also examined the interaction between the average score of exams 1 and 2 and exam 3 scores (Figure 3), but there was also no interaction between the slopes. The last interaction we looked at was between exam 3 scores and the years the data was collected (Figure 4), and there was a clear interaction between the slopes.

Models and Results

Four different multiple regression models were utilized to examine how certain independent variables and our chosen interaction are associated with the overall course grade (Table 4). To improve the fit of the data, we used mean centering on each of our quantitative independent variables. Mean centering allowed for the interpretation of the intercept in our equation to be meaningful. Initially, we included all independent variables in our model. T-tests revealed that the semester the student was taking the course was not predictive of a student's overall course grade. Therefore, we removed the variable semester from our model. Furthermore, we added the interaction between exam 3 and year because the interaction plot indicated a clear interaction that needed to be accounted for between these variables. We chose this model because it included the interaction between exam 3 and year and removed the variable semester, which provided an excellent fit to the data. For example, our chosen multiple regression model using the average score of exams 1 and 2, exam 3 scores, and year explained 84.47% of the variance in overall course grades, after adjusting for the complexity of the model (adjusted-$R^2 = .8447$).

We then tested the assumptions for our multiple regression model (Figure 5). There were no clear curves within the scatterplot matrix and the residual plot, which suggests the linearity assumption is reasonably met. Furthermore, the histogram of residuals was slightly skewed left. There was also a slight curve in the QQ-plot, and the Shapiro-Wilkes test indicated a slight violation ($p = 0.028$). Although there may be some slight concerns of the normality assumption, this assumption is reasonably met because the curves and skewness was minimal. The residual plot also indicated no clear signs of fanning, and the results of the Levene's test ($p = 0.3741$) suggest the homoskedasticity assumption is met. Therefore, all assumptions were reasonably met. Furthermore, we compared the average exam grades for exam 1, exam 2, and exam 3 for both genders between the years 2000 and 2003 (Figure 6). The student's overall course exam grades seem heavily dependent on the specific class on the time that the student took the course (e.g., semester 1 of the year 2000). For example, we see that students who took the course in semester two of the year 2000 seemed to struggle in comparison to students in semester one of the year 2001 who performed higher. Our proposed multiple regression model is:

Predicted Overall Course Grade = 73.11 + 0.596(X1-76.72) + 0.246(X2-75.48) + -0.373(X3) + -5.395(X4) + -1.224(X5) + 2.226 (X6) + 0.156((X2-75.48)*(X3) + 0.048((X2-75.48)*(X4)) + 0.079((X2-75.48)*(X5)).

Note: X1 = average score of exams 1 & 2; X2 = Exam 3 scores; X3 = Year 2001; X4 = Year 2002; X5 = Year 2003

The intercept in our equation is 73.11. For students who are male and took the course in 2000, the predicted overall course grade is 73.11, when the average score of exams 1 and 2 is 76.72 and exam 3 scores is 75.48%.

The slope for the average score of exams 1 and 2 is 0.60. This suggests that when the average score of exams 1 and 2 increases by 1%, the predicted overall course grade will increase by 0.60, when exam 3 scores, sex, and year are held constant.

The slope for sex(woman) is 2.23. This suggests the predicted overall course grade for students who are female is 2.23 points above students who are male, when exam 1 and 2 scores, exam 3 scores, and year are held constant.

The slope for exam 3 is 0.08. For students who took the statistics course in 2000, when exam 3 scores increase by 1%, we predict overall course grades to increase by 0.08%, when the average of exam 1 and 2 scores and sex are held constant.

The slope for the interaction between exam 3 scores and the year 2001 is 0.156. This means that the estimated slope between exam 3 scores and overall course grade for students who took the course in 2001 is 0.156 percent higher than students who took the course in 2000.

We cannot draw meaningful interpretations from the slope of the interaction between exam 3 scores and the year 2002. For example, consider the following 95% confidence interval that contains zero. We are 95% confident the population mean slope between overall course grade and exam 3 scores that took the course in 2002 is between 0.12 points lower and 0.22 points higher than students who took the course in 2000, while holding the student's sex and average score on exams 1 and 2 constant.

Additionally, we cannot draw meaningful interpretations from the slope of the interaction between exam 3 scores and the year 2003. For example, consider the following 95% confidence interval that contains zero. We are 95% confident the population mean slope between overall course grade and exam 3 scores that took the course in 2003 is between 0.014 points lower and 0.171 points higher than students who took the course in 2000, while holding the student's sex and average score on exams 1 and 2 constant.

Discussion/Conclusions

The study sought to determine if a student's midterm exam scores, sex, semester, and year the course was taken predict overall course grades among college students. Our findings suggest that the sex of the student, their average score on exams 1 and 2, exam 3 scores, and the year they took the statistics course are statistically significant in predicting overall course grades. This finding is important because it suggests that tracking student performance (e.g., exam one and two scores) and providing extra support for students who are struggling earlier in the semester may reduce the rate of students who fail their coursework at the end of the semester. This study encourages universities to take preventative measures by tracking student performance in an attempt to support struggling students.

Limitations include not accounting for confounding variables, such as a student's major, year in school, or hours studied. For example, a student who was a statistics major may have scored better on their midterm exams and ended with a better overall course grade than students who were not statistics majors. Furthermore, students who studied more may have scored better overall in the course. Thus, future studies should take these variables into account. In addition, there was little to no demographic information about the sample. We cannot be certain these findings can be generalized to the general population of college students, especially because this data was likely collected via convenience sampling. Future studies should attempt to use a more representative sample and continue to identify risk factors for dropout among college students.

References

Ashenafi, M. M., Riccardi, G., & Ronchetti, M. (2015). Predicting students' final exam scores from their course activities. In *2015 IEEE Frontiers in education conference (FIE)* (pp. 1-9). IEEE. https://doi.org/10.1109/FIE.2015.7344081.

Jensen, P. A., & Barron, J. N. (2014). Midterm and first-exam grades predict final grades in biology courses. *Journal of College Science Teaching*, *44*(2), 82-89. http://www.jstor.org/stable/43631961.

Kost-Smith, L. E., Pollock, S. J., & Finkelstein, N. D. (2010). Gender disparities in second-semester college physics: The incremental effects of a "smog of bias". *Physical Review Special Topics-Physics Education Research*, *6*(2), 020112. https://doi.org/10.1103/PhysRevSTPER.6.020112.

OpenIntro. (n.d.). *Exam and course grades for statistics students*. Data Sets. https://www.openintro.org/data/index.php?data=exam_grades

Reardon, R. C., Leierer, S. J., & Lee, D. (2007). Charting grades over 26 years to evaluate a career course. *Journal of Career Assessment*, *15*(4), 483-498. https://doi.org/10.1177/1069072707305767

Appendix

Descriptive Statistics

Table 1. Chart of the sample size, mean, standard deviation, minimum observation, and maximum observation for each quantitative variable

Variable	*n*	*M*	*SD*	*Min*	*Max*
Average Score of Exam 1&2	232	76.72	9.92	50.50	98.25
Score on Exam 3	233	75.48	14.71	28.00	98.89
Overall Course Grade	233	72.24	9.81	43.27	97.57

Table 2. Chart of counts and percentages of each of our categorical variables

Variable	*Counts*	*Proportions*
Sex Female	45	.1931
Sex Male	188	.8069
Semester 1	161	.6910
Semester 2	72	.3090
Year 2000	86	.3691
Year 2001	75	.3219
Year 2002	36	.1545
Year 2003	36	.1545

Table 3. Chart depicts correlations between all the quantitative variables

Variable	Average score of exams 1 and 2	Overall course grade
Exam 3	0.4663319	0.69126577

Overall course grade	0.81271037	-------

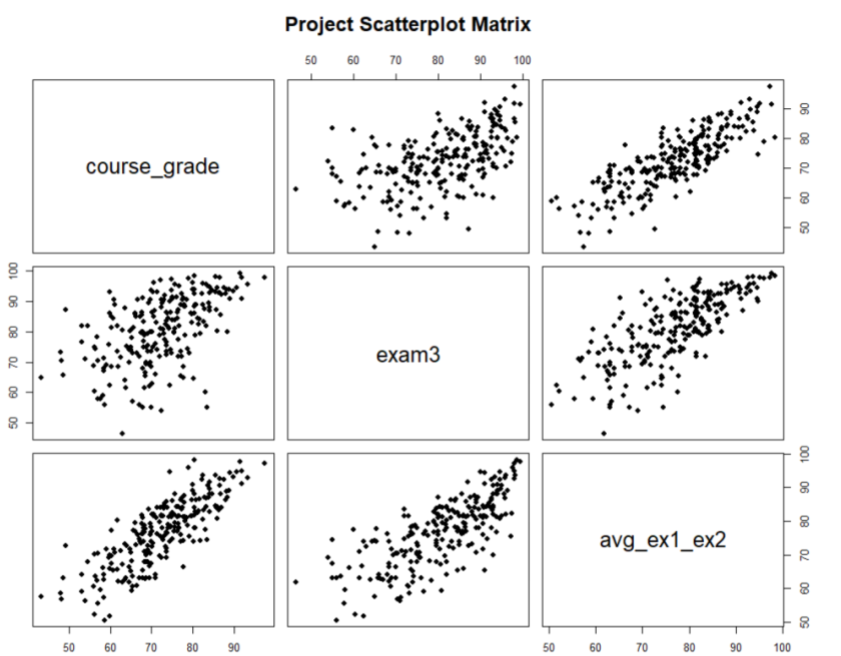

Figure 1. Scatterplot matrix depicts the relationships between the quantitative variables

Table 4. Results of Fitting a Taxonomy of OLS Regression Models Predicting Overall Course Grade of Students in their Statistics Course a Sample of 233 People

	Model A	**Model B**	**Model C**	**Model D**
Average score of exams 1 and 2	0.596 (0.538, 0.655) < 2e-16	0.596 (0.538, 0.655) < 2e-16	0.64 (0.539, 0.669) <2e-16	0.615 (0.556, 0.674) < 2e-16
Exam 3 scores	0.253 (0.199, 0.308) < 2e-16	0.246 (0.192, 0.299) < 2e-16	0.273 (0.229, 0.316) <2e-16	0.308 (0.266, 0.350 < 2e-16
Semester	0.885 (-0.382, 2.151) 0.172473			0.650 (-0.621, 1.921 0.31744
Year 2001	-0.429 (-1.661, 0.802) 0.494981	-0.373 (-1.604, 0.858) 0.553383		-0.900 (-2.136, 0.336 0.15504
Year 2002	-5.056 (-7.240, -2.871) 9.39e-06	-5.395 (-7.529, -3.261) 1.43e-06		-5.505 (-7.132, -3.879) 2.44e-10
Year 2003	-0.880 (-2.468, 0.708) 0.278369	-1.224 (-2.737, 0.288) 0.114081		-1.116 (-2.730, 0.498) 0.17668

Sex (woman)	2.203 (0.947, 3.459) 0.000702	2.226 (0.967, 3.484) 0.000633		2.111 (0.828, 3.395) 0.00145
Exam 3*Year 2001	0.162 (0.077, 0.247) 0.000232	0.156 (0.071, 0.240) 0.000375		
Exam 3*Year 2002	0.041 (-0.128, 0.209) 0.636367	0.048 (-0.120, 0.216) 0.575963		
Exam 3*Year 2003	0.071 (-0.022, 0.164) 0.136538	0.079 (-0.014, 0.171) 0.097608		
Intercept	71.888 (69.936, 73.840) < 2e-16	73.111 (72.248, 73.974) < 2e-16	72.352 (71.786, 72.918) <2e-16	72.405 (70.444, 74.366) < 2e-16
Adjusted R^2	0.8453	0.8447	0.7936	0.8376
F statistic (p-value)	126.2 (< 2.2e-16)	140.6 (< 2.2e-16)	445.1 (< 2.2e-16)	171.2 (< 2.2e-16)

Note 1. The quantitative independent variables were centered at their means of 76.72 (average score of exams 1 and 2) and 75.48 (exam 3 scores)

Note 2. The confidence intervals for each variable are presented in parentheses.

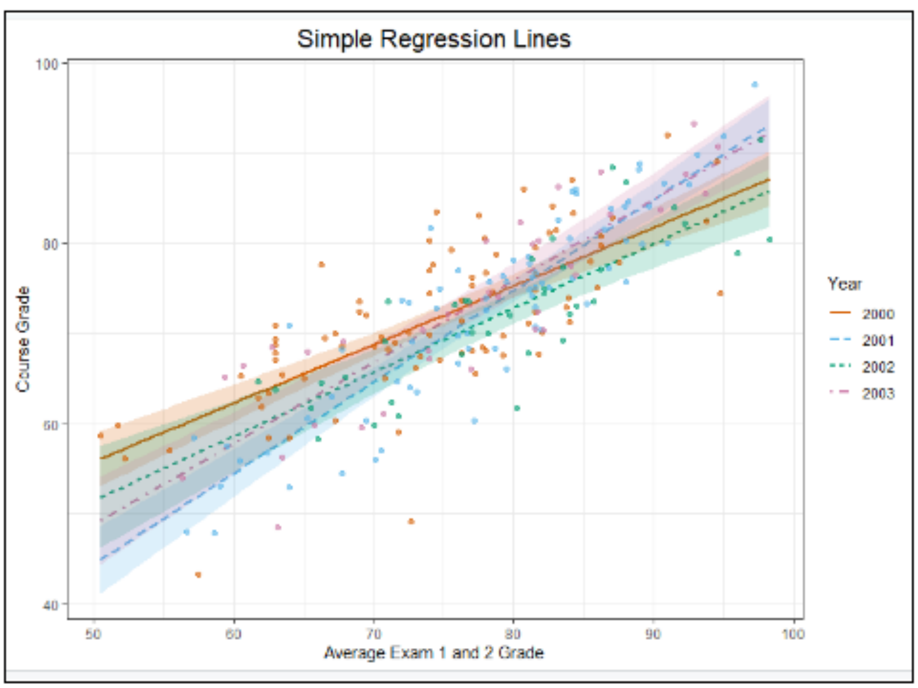

Figure 2. Interactive plot depicts the interaction in between the average score of exams 1 and 2 and the years the data was collected from to predict the overall course grade

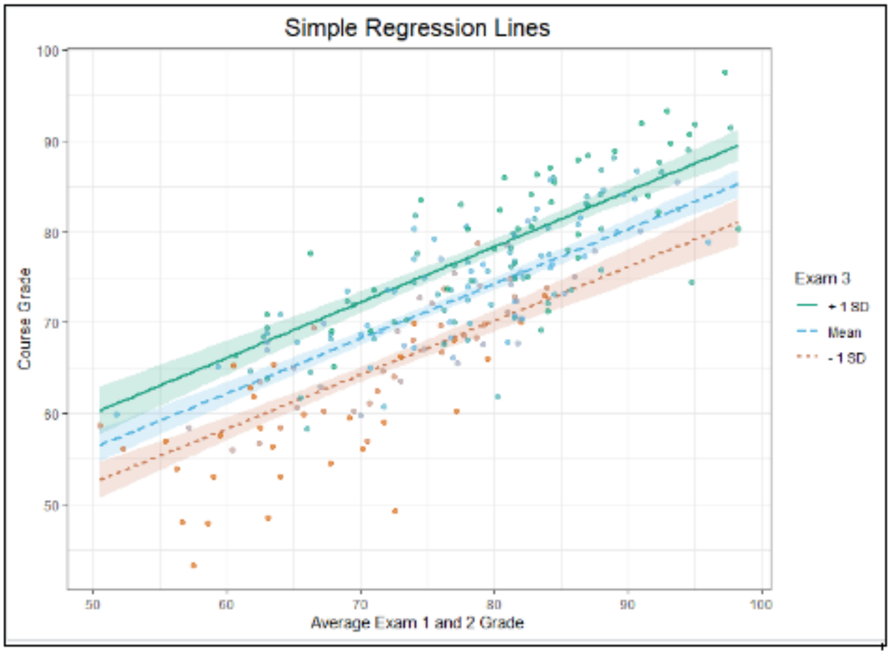

Figure 3. Interactive plot depicts an interaction between the average score of exams 1 and 2 and exam 3 scores to predict the overall course grade

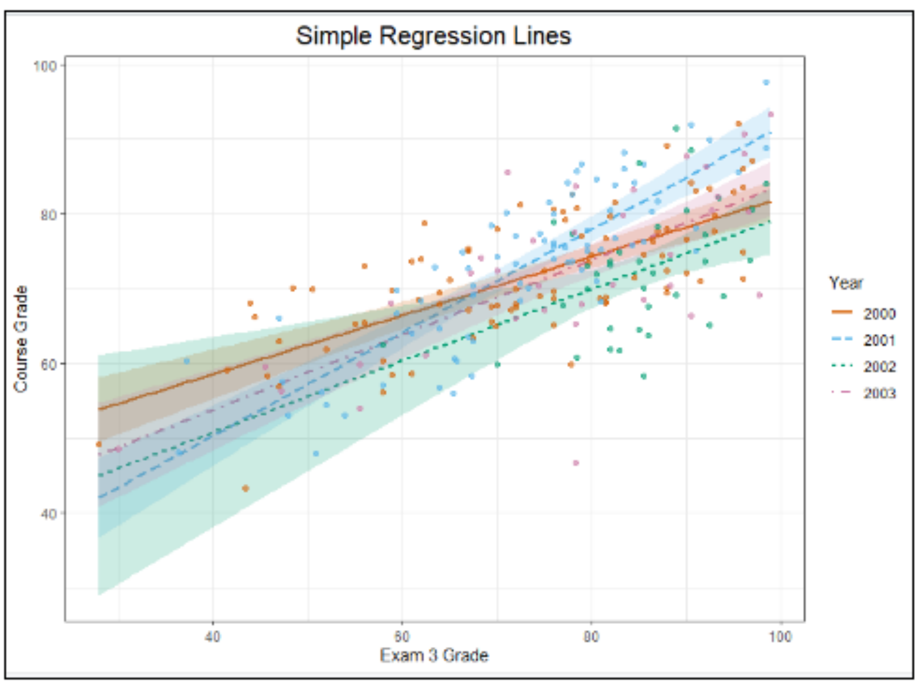

Figure 4. Interactive plot depicts the interaction of our independent variables (exam 3 scores and years the data was collected from) to predict the overall course grade

Model and Results

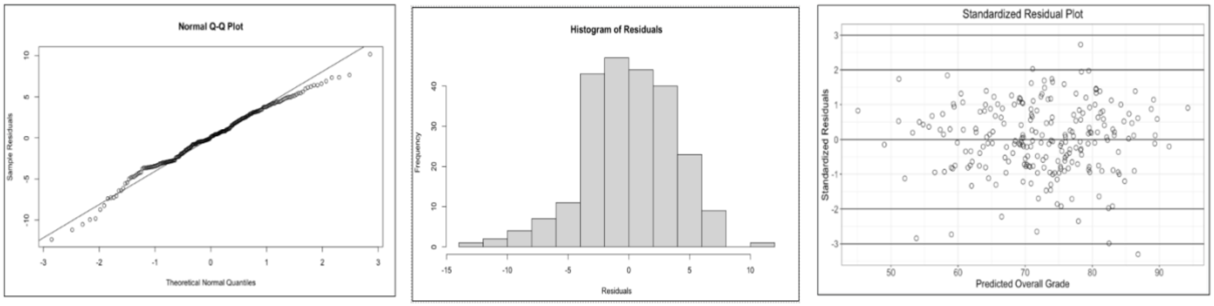

Figure 5. QQ-plot, Histogram of residual, and standardized residual plot for assumptions

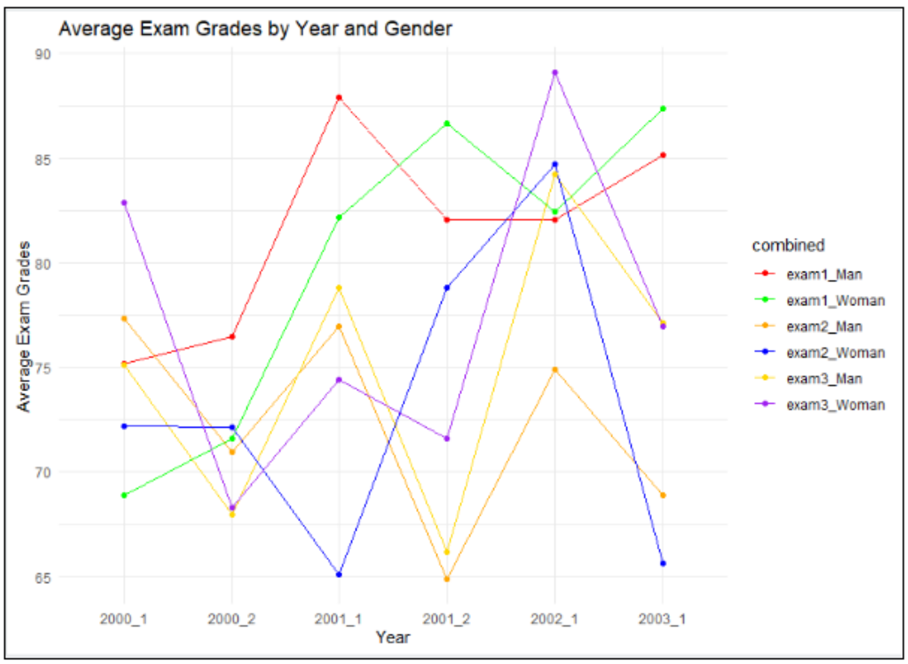

Figure 6. Graph depicts average exam grades for exam 1, exam 2, and exam 3 for both genders between the years 2000 and 2003