Introduction

The purpose of this project is to test whether past student ratings can predict the difficulty of exams in undergraduate classes. For this project, I modeled course difficulty based on past ratings for professor giving the class, as well previous student ratings of the course being given .

Data

Most of the data for this project is from faculty evaluations on Brooklyn College's website which can be found @ http://websql.brooklyn.cuny.edu/facultyevaluations/. These evaluations are based on questionnaires that students fill out at the end of each semester. The site shows the students' responses by percentage for each of the questions on the questionnaire. The site supplies the percentages either by professor (for all courses given by that professor in a given semester), by course code (for all sections of a particular course given in a semester), or for each individual class.

For the response variable, I measured the difficulty of each class based on the responses to Question 15 of the questionnaire, which asks students to rate whether the examinations were "Very Difficult", "Difficult", "Somewhat Difficult" or "Not Difficult". The response variable that I used represents the percentage of students that found the examinations "Very Difficult". Since I needed to create multiple online queries for each data point, I only used only data from accounting classes given in Spring 2013 for the response variable.

For the explanatory variables, I experimented with three variables that are either a measure of the difficulty level of the professor teaching the class, or a measure of the difficulty of the course itself. The first explanatory variable (from here on referred to as **Variable A**) represents the percentage of students that found that particular professor's examinations "very difficult" during the previous (Fall 2012) semester. The second explanatory variable (**Variable B**) is the easiness rating for the professor from http://www.ratemyprofessors.com/. Ratemyprofessors.com allows visitors to rate professors "easiness level" on a scale of one to five. The Third variable (**Variable C**) equals the percentage of students that rated the examinations for that particular course as "very difficult" during the previous semester (irrespective of who the professor was).

Model # 1

My first model used all three explanatory variables for the regression. The results of the regression from excel are shown below.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.643054075
R Square	0.413518543
Adjusted R Square	0.360202047
Standard Error	0.119138507
Observations	37

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	0.330262262	0.110087421	7.755921208	0.000469871
Residual	33	0.468401469	0.014193984		
Total	36	0.798663731			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*
Intercept	0.265487658	0.149964614	1.770335351	0.085910113	-0.039617642	0.570592958	-0.039617642
Variable A	0.41578084	0.294925668	1.409781806	0.167959707	-0.184249939	1.015811619	-0.184249939
Variable B	-0.057412181	0.033057776	-1.73672242	0.091764126	-0.124668731	0.00984437	-0.124668731
Variable C	0.209368314	0.394420239	0.530825483	0.599094347	-0.593085692	1.01182232	-0.593085692

As can be seen from the P- value of .599 for Variable C (The percentage of students who found the course very difficult in the previous semester), Variable C is not significant when used in a regression together with both variables A & B. This may be due to the variables being highly correlated with each other. (Professors often teach the same course for multiple semesters in a row so professor ratings and course ratings should be highly correlated.) Scarcity of the data may also be an issue.

It should be noted that in the regression Variable B (ratemyprofessor rating) has a negative slope. That is because ratemyprofessor.com has an easiness rating on a scale of one to five. The higher ratings correspond to easier professors. Easier professors should have a smaller percentage of students that describe their examinations as “very difficult” on the BC questionnaires. As a result, the higher values for Variable B result in lower values for the response variable, which causes the negative slope.

Model 2

To determine which variable to remove for model 2, I created the following correlation matrix:

	Variable A	Variable B	Variable C
Variable A	1	-0.64795	0.7402996
Variable B	-0.64795	1	-0.608357
Variable C	0.7402996	-0.608357	1

I decided to remove Variable A since it has a correlation of .64 with variable B, and a correlation of -.65 with variable C, and those correlations are higher than the correlation between Variables B & C. This resulted in the following linear regression:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.614976939
R Square	0.378196635
Adjusted R Square	0.341619967
Standard Error	0.120856234
Observations	37

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0.302051936	0.151025968	10.33983276	0.000310529
Residual	34	0.496611795	0.014606229		
Total	36	0.798663731			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.33347089	0.144047248	2.315010483	0.026781077	0.040731662	0.626210118	0.040731662	0.626210118
Variable B	-0.074670245	0.031150367	-2.397090355	0.022167196	-0.137975407	-0.011365082	-0.137975407	-0.011365082
Variable C	0.527729744	0.328037104	1.608750162	0.116918544	-0.138921856	1.194381344	-0.138921856	1.194381344

For this regression Variable B has a slope of -.075, and Variable C has a slope .528. The intercept is .333. Variable B has a P value of .0222, so it is considered significant. Variable C has a P value of .117. Although the P value isn't significant enough to conclude that Variable C helps explain the result, it probably does play a role in explaining the response variable. My data set only has 37 data points, and it's possible that with more data we would get lower p values.

To check to see if this is the best model to use, I tried three other models with only one explanatory variable.

Additional Models

Model 3 regresses the percentage of students that rate examinations as "very difficult" based on the prior semester's ratings for the professor (Variable A) The regression for Model 3 is below.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.584949386
R Square	0.342165784
Adjusted R Square	0.32337052
Standard Error	0.122519774
Observations	37

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.273275401	0.273275401	18.20489438	0.000143359
Residual	35	0.525388329	0.015011095		
Total	36	0.798663731			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.057343651	0.039746733	1.442726154	0.157987986	-0.023346506	0.138033808	-0.023346506	0.138033808
Variable A	0.808130767	0.189403308	4.266719393	0.000143359	0.423621613	1.192639922	0.423621613	1.192639922

Model 4 Regresses the response variable based on the professors ratemyprofessor.com ratings. The Regression is as follows.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.575208672
R Square	0.330865017
Adjusted R Square	0.311746874
Standard Error	0.123567661
Observations	37

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.264249889	0.264249889	17.30633709	0.000195983
Residual	35	0.534413842	0.015268967		
Total	36	0.798663731			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*
Intercept	0.52754029	0.080488006	6.554272045	1.44771E-07	0.364140951	0.690939629	0.364140951
Variable B	-0.105156936	0.025277571	-4.160088592	0.000195983	-0.156473134	-0.053840739	-0.156473134

Model 5 Regresses the response variable based on the Brooklyn College's ratings for that course from the previous semester (Variable C)

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.522600321
R Square	0.273111096
Adjusted R Square	0.252342841
Standard Error	0.128789951
Observations	37

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.218123926	0.218123926	13.15041169	0.000906093
Residual	35	0.580539804	0.016586852		
Total	36	0.798663731			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.012634048	0.056743192	0.222653101	0.825099775	-0.102560756	0.127828851	-0.102560756	0.127828851
Variable C	1.006101902	0.277442057	3.626349636	0.000906093	0.442864586	1.569339218	0.442864586	1.569339218

It's clear from the low P values for the explanatory variables in models 3, 4, and 5 (.00014, .00019 and .0009 respectively) that each of the explanatory variables when used alone can help predict the response variable. Model 3 is the best model from these three models, because it has the highest R Squared value (.342) and the lowest P value for the explanatory variable.

When comparing Model 2 to Model 3, I would recommend using model 2 because the adjusted R Squared (.3416) for model 2 is higher than the adjusted R Squared of model 3 (.3234)

Conclusion

It is clear from the low P values in each of the models 3, 4, and 5, that either of the explanatory variables are significant when the regression doesn't include any of the other variables. This means that past Brooklyn College ratings for either the professor, or the course, as well as professor ratings from ratemyprofessor.com, can help predict the percentage of Students that will rate the examinations as "Very Difficult"

The recommended model to use for Brooklyn College students who want to avoid hard exams is model 2, which regresses the exam difficulty based on previous BC student ratings for the course, as well as ratings of professors from myprofessor.com. Although I would recommend Model 2 as the best option, the p value for variable C in model 2 isn't significant enough to conclude that it is a better model than the models with only one explanatory variable.