

Gender and graduate record exam scores



**ASSIGN
BUSTER**

Student performance at any academic level is paramount to the success of all academic institutes due to the fact all academic establishments are rated by said student performance, especially at the college level. Commonly, educational organizations use Grade Point Average (GPA), Graduate Record Exam (GRE) scores, Scholastic Aptitude Test (SAT), American College Testing (ACT), etc. as measuring tools for current and future student success (Cole & Gonyea, 2010). While many studies have been done which show the GRE to be an unreliable tool for admissions, there are many institutions which still uses it for the admission process. For one hypothetical institution in particular, it is the GPA and GRE which are being called into question as to whether or not these success measuring tools are appropriate for the school's needs in determining the success of current and future students. The dean of the School of Education has employed the new director of institutional research to analyze whether or not these measurements should continue to be used, or if other means are on the rise for the institution. The director has many things to look into before the analysis can take place.

A Hypothetical Analysis

Several variables must be analyzed and researched in order to help the dean decide whether or not to continue using the GRE as a measurement of success or eliminate it completely from graduate school admissions requirements. In order to give the dean a well-rounded analysis, four areas (variables) must be researched: 1) student gender, 2) the GPA of current and future students, 3) the GRE scores for current and future students, and 4) the graduate degree completion frequency. Each of these variables must be analyzed through specific statistical algorithms in order for the institution to

have a proper reading of the desired outcome. Once it is calculated, a proposal must be written and presented to the dean. The proposal's analysis will allow the dean to see the relationships between GPA and GRE scores, the relationship between gender, GPA, and GRE scores, and the relationship involving gender, GRE scores, and frequency of completion. Also, a final analysis will show recommendations to the dean based on the findings of this analysis based on the information gathered by the new director. As the director begins their investigative study, they must start with the gathering of data pertaining to the four variables previously mentioned, but more importantly, understanding the variables themselves.

Understanding Variables

Before being able to analyze the different parts the dean wishes the director to look into, one must first understand each and every one of the variables which will be used in the coming calculations, the GPE, GRE scores, gender, and completion frequency. A student's GPA shows the success of a student in the class based on their ability to do well on assessments. This scale is based on a 0 to 4-point scale. It is a cumulative score over the student's learning career. Through honors or Advance Placement (AP) courses, students can obtain a 5.0 or higher GPA (Caskie, Sutton, & Eckhardt, 2014). Conversely, the GRE is a cumulative test which shows how well the students retain and use learned information. It tests critical thinking, as well as analytical, verbal, and quantitative reasoning skills (Wao, Ries, Flood, Lavy, & Ozbek, 2016). Next, gender is usually labeled by whether the subject is a male or female. And finally, the completion frequency shows if a student was able to complete their degree successfully or if they were unable to

<https://assignbuster.com/gender-and-graduate-record-exam-scores/>

complete the degree. Each of these variables will play a part in the calculations. GRE and GPA will usually be dependent variables while gender and completion frequency will be the independent variables. With the understanding of what each variable means, it is now time for the director to begin collecting data.

Collecting Data

It is important to collect the data for each of the variables in the study by asking the appropriate research questions and/or hypotheses. In order to check the hypotheses, one needs to collect the data for gender, GPA, GRE scores, and completion frequency. These variables should be collected from the records of the institution itself. Any bias should be strictly prohibited, and only the unfiltered and raw data should be analyzed throughout the entirety of the study in order to have a complete, unbiased view of the results of the study (Campbell, 2014). Participants chosen for the study should be comprised of a random samples which represent the entire student population in order for the test results to be a valid means of analysis. That is, students included should have an equal range of genders, broad range of GPA and GRE scores, and a well-balanced number of participants who completed their degree. Once the data is collected, it will be imperative to use the proper statistical tools and techniques to calculate the correct statistics for analysis or else the proper results will not be calculated and the data inventories nor the statistical prediction will not be valid (Garb, Wood, & Baker, 2018) and therefore, the results given to the dean will have little to no real validity to whether or not the GRE and GPA are

good tools for measuring the academic success of current and future students of the institution.

Analyzing Data

Once the data collection is complete, it will be important to analyze all data using the appropriate calculations, values, formulas, and techniques. In order to test the different variables as valid measuring factors, one would use descriptive statistics for the data which has been collected. These descriptives can be graphical analyses and inferential statistics. Any graphical analyses would be used to interpret the facts properly and visually. Inferential statistics would be used to test each of the hypothesis for all four variables in question. Of course, one would also check the research questions using the inferential data or while testing the different kinds of hypotheses. Each test would require both null and alternative hypotheses to begin with. Then the variables must be properly tested.

There are several different ways for the different relationships to be explored through statistics. The first would be through descriptive statistics to show distributions and variations between the desired variables (Knapp, 2017). The second would be to evaluate the relationship using predictive tools such a correlation and regression analysis to identify any linear relationships between the different variables. And finally, doing a trend analysis to see whether there is an increase, decrease, or constant trend over time which can explain the relationship between the different variables (Michaelson & Hardin, 2012). One would use the 5% level of significance in order to check the validity of the hypotheses and research questions. Since the dean's

main concern in the relationship between the GPA and GRE scores, it would behoove the director to begin with the analysis of this relationship.

GPA and GRE score relationship. The relationship to question through this analysis would be what the GRE scores are of current and future students, if statistically significant relationship exists between students GRE tests and GPA, what the GRE and GPA scores look like at the school itself, and what GRE and GPA scores look like at different universities. Before being able to analyze these questions, the first step in analyzing the validity and usefulness of the GRE scores as a way to measure student success must begin with a look at the relationship between GPA and GRE. To begin, one must first discuss the null and alternative hypotheses. For this study, the null hypothesis would be H_0 : The GRE score does not depend on GPA or vice versa. The alternate hypothesis would be H_a : The GRE score and the GPA are significantly different. That means the students who have a minimum or higher GRE score would also have a higher GPA. For this test, the best course of action would be to determine significance by applying a t-value test.

The t-value test will show if the test statistic does or does not equal the expected value minus the observed value. The t-test itself would look as such: $t = (m - \mu) / [S / \text{Sqrt } N]$. The m is the mean, with the standard deviation being S . The comparison m to the theoretical value μ will show whether or not the evaluation indicates a difference which is statistically significant or not. For there to be any statistical significance, the critical value must be above the significance level alpha of 5% (0.05) (Balakrishnan, Castilla, Martin, Pardo, 2019). Once the test was conducted, the scores

<https://assignbuster.com/gender-and-graduate-record-exam-scores/>

would need to be compared and analyzed. If the two values are significantly different, one would fail to reject the null hypothesis, thereby accepting it. Therefore, the GRE score would not be dependent on a student's GPA and therefore would not be a good predictor of current and future student success at the institution in question. It would then be recommended the administration question their use of GRE scores and GPA scores when trying to predict student performance. With this piece of information now completed, the next step in the analysis would be to see whether or not there is a relationship between gender, GPA, and GRE scoring.

Gender, GPA, and GRE scores. The relationship between GPA and GRE scores lies in whether or not there is a way to predict the GPA or GRE scoring for students who are enrolling in the institution. Once again, to begin the analysis, one would look for the null and alternative hypotheses for which testing will take place. Three potential null hypotheses for this study would be as follows: H_0 : The performance in the GRE is significantly different between genders, H_1 : The GPA is the same over the years, H_2 : The GRE scores have increased over the years. The alternate hypotheses would read as H_a : The performance in the GRE is not significantly different between genders, H_{a1} : The GPA is not constant over the years, and H_{a2} : The scores have remained the same over the years. For this study one will conduct another t-value test in order to test the statistics properly and “ calculate the necessary information to determine error and sample validity” (Ryan, 2016). As stated previously, the t-test would follow the $t = (m - \mu) / [S / \text{Sqrt } N]$ formula. Dependent variables would be the predictors of the GPA and GRE test scores and the gender would be the independent variable. Once the

calculations are done to test the outcome for both genders, if the critical value is below the significance level alpha of 5% (0.05), then there will be a significant statistical meaning, so the null hypothesis would be accepted. However, if there was a significant difference, say that of higher scores for women than men, then the null hypothesis would be rejected, and the alternative hypothesis would be accepted. Now the GPA and GRE score relationship has been viewed as well as the gender, GPA and GRE score relationships, the next part of the analysis process would be to see if gender affects GRE scores.

Gender and GRE scores. The research question for this particular part of the study should be what effect gender has on GRE scores. As with each analysis, the first step in the analysis is to determine the null and alternate hypotheses for this concern. H_1 : Gender and the GRE score are not dependent on one another. H_a : Gender and the GRE score are significantly different between genders. Once the hypotheses are created, it would be time to apply a p-value test because one wishes to see the marginal significance of the probability occurring between genders. The p-value shows the probability that the data is inconsistent with the hypothesis if the hypothesis itself is deemed true. In order to calculate this, one would use the following formula: $p\text{-value test} = (\bar{x} - \mu) / [\sigma / \sqrt{n}]$ because “tests that combine p-values, such as Fisher’s product test, are popular to test the global null hypothesis H_0 that each of n component null hypotheses, H_1, \dots, H_n , is true versus the alternative that at least one of H_1, \dots, H_n is false, since they are more powerful than classical multiple tests such as the Bonferroni test and the Simes tests” (Zhang, Chen, & Pfeiffer, 2013, p. 764). The \bar{x} is

the sample mean while μ is the population mean. The s shows the sample's standard deviation and then one takes the square root of n which is the sample size. Next, there would need to be an ANOVA test and a Chi-square test so the data between the GRE scores and the gender gap can be analyzed. The ANOVA test would show if there is a variation between the scores of the different genders. The formula for this would be $SS \text{ Total} = \sum x^2 - CF$. $SS \text{ Total}$ stands for the sum of squares total and estimates the sum of the estimate in question. This is calculated using the statistic times variable 2 minus the cumulative frequency. The Chi-square test will help assess whether or not the variables fit between the expected values and the theoretical expectations, that is it calculates the degrees of freedom. It is calculated with the following formula: $f(x) = \frac{2^{(1-n/2)} x^{(n-1)} e^{-(x^2)/2}}{\sqrt{n/2}}$. If women were to score consistently better than men on any of these tests, it would show there are other factors at play when looking at the GRE scoring such as backgrounds of the participants, the majors and studies of each of the students and so on. Based off of the outcome of this test for both genders, if the test value is less than the p-value, then the GRE scores are not significantly different between genders. Because of this, if the two values are not significantly different, one would fail to reject the null hypothesis.

Also, taking into consideration previous studies, one by Hirschfeld and Moore (1995) showed similar findings in women who also found the GRE had a fundamentally lower score than those of men due to the way women approached different questions. Men were better in the math and sciences than women and depending on the GRE criteria of the hypothetical college is <https://assignbuster.com/gender-and-graduate-record-exam-scores/>

looking into, could mean there might be a gender gap in the admissions process. The final test for the effectiveness of the GRE score would be to research whether the GRE score and the completion frequency have anything to do with gender at all.

Gender, GRE score, and completion frequency. This part of the study will require asking questions such as how the GRE scores can predict student success in academics and if the GRE is a valid predictor for student outcomes. For the first step in this analysis, it is the null and alternative hypotheses which must be recognized. H_0 : The GRE score, gender, and degree completion are not significantly different. H_a : The GRE score, gender, and degree completion are directly related to one another. For this analysis, it would be most appropriate to use another t-value test to identify the significant and non-significant effects related to the hypotheses as well as linear modeling, and correlation coefficients. There are other interpretive models which can also be used so the institution can use different criteria from the GRE and GPA scores in order to predict student success such as Kane's interpretive model (Deluca, 2011) if the school really wanted the best advice for how to proceed, but a simple t-value test would work also. If the tests show the values are more than the t-value, it would be concluded that gender, GRE score, and the frequency of degree completion are indeed linked to one another. If there are statistically significant differences in the linear modeling and correlation coefficients, then there would be a rejection of the null hypothesis. However, if not one would fail to reject the null hypothesis, and it would be accepted. Now all tests have been done on the

different parts of the requested analysis, it would be time to put together a finalized report for the dean.

Finalizing Information. Some educators and administrators believe the GRE to be a poor indicator of student success for undergraduate performance as well as graduate student work. For the final report, there are two potential outcomes. Either the values represented are less than the statistic, or they are greater than the statistic. Depending on the outcome of the different tests (test statistic, t-value, and p-value), the recommendation will be made. Using the p-value would help make the proper call for the decisions regarding any null hypotheses. This would help us decide whether the GRE is a proper tool for measuring student success, and if it is a statistically significant predictor for predicting the success of future students to the institution. The p-value would also help indicated which factor or predictor is more valuable for the outcome of the dependent variables.

Less than value. If the outcome of the test statistic is less than the t-value and p-values, then the null hypotheses will be accepted. This means, in the case of the relationship between the GPA and GRE, the null hypothesis stating the GRE score does not depend on GPA would be accepted. In the case of gender, GPA, and GRE scores, the null hypotheses of the performance in the GRE being significantly different between the two genders, the grade point average being the same over the years, and the scores increasing over the years would be accepted. For gender and GRE scores, the approved hypothesis would be the GRE score does not depend on gender. And finally, if the effect research question involving gender, GRE score, and degree completion frequency is less than the statistic, the

approved hypothesis would be the GRE score, gender, and degree completion are not significantly different. In most cases, the null hypothesis would be accepted for the value being less than the statistic. Therefore, the GRE would not be recommended based on the results of the statistical analysis.

Greater than value. Should the t-value and p-value be less than the test statistic, then one would reject the null hypothesis and instead accept the alternate hypothesis. In the case of the relationship between GPA and GRE, the accepted hypothesis would be the GRE and GPA are significantly different from one another. For the gender, GPA, and GRE scores corresponding, the performance on the GRA is not significantly different between the two genders, the grade point average is not constant over the years, and the scores are the same over the years would all be accepted hypotheses though they are the alternative hypotheses. For the GRE scoring, if the value is greater than the statistic there would be a significant difference between the scoring of the genders so the alternate hypothesis would be accepted. Finally, in the last case of the gender, GRE score and degree completion frequency, the accepted hypothesis would be the GRE, gender, and degree completion are directly related to one another. Therefore, the GRE would be recommended based on the results of the statistical analysis.

Previous Studies

Some studies have shown GRE tests to have a designed bias against women and minorities (Hale, 2010) and therefore should not be used as a predictor

for student success, especially for women and minorities. Tock and Ericsson (2019) found there to be a significant difference in science related GRE questions due to the way men and women see the question's outcome differently. Another study showed the GRE was not a good predictor because students from different socioeconomic statuses did or did not have the resources to study or understand certain questions, and therefore the GRE did not give a proper reading for the potential students to other universities (Schwager, Hülshager, Bridgeman, & Lang, 2015). Therefore, the GRE and the GPA are not exclusive to one another, nor are they good indicators of what a student's potential is in the classroom. Other factors such as previous successes in the classroom, what the student's major focus in the STEM system was, as well as overall assessments from their previous experiences are all different ways in which a school could look into whether or not a student is a good fit for the institution and if they will be successful learners.

Expected Outcome

In most cases, it would be expected the GRE would not be a statistically significant or efficient predictor of undergraduate and graduate student performance. There would not be a statically significant relationship between the different variables included within the study. The results would not show any strong linear or non-linear relationship between GRE, GPA, gender, or completion frequency. So, it would be recommended that the GRE be removed as a predictor of the undergraduate and graduate school performance measuring test. Using other methods of predicting student success would be a much better idea for the school rather than using GRE

<https://assignbuster.com/gender-and-graduate-record-exam-scores/>

scores and GPAs to determine a student's current and future academic success. This along with the data of other studies which show the GRE is a low-ranking indicator of a student's ability to be successful at an undergraduate and graduate level, due to the fact that many studies show that " GRE scores failed to predict time to degree or to identify students who would leave during the first year of their programs" (Petersen, Erenrich, Levine, Vigoreaux, & Gile, 2018, p. 1).

As the new director for institutional research, it would be advised for the director to inform the dean it is time to look for other ways of admitting prospective students. Though the data is needed to say for sure which way to go, a majority of studies which have already been done show the GRE to be a poor predictor of student learning promise. Therefore, assuming the data were to show the GRE is a weak predictor, the fact gender does effect GRE scores, and GPA and GRE scores are not exclusive to one another, it would be time for the educational institution in question to find a new way to admit students and new criterion on which to base their admission process.

Conclusion

" In our society we put a huge premium on the kinds of analytical problems the GRE measures. So if you're a good abstract analytical thinker, you'll do well on these tests," says Robert J. Sternberg, a cognitive psychologist and professor of human development at Cornell University. " The GRE is like taking a cancer test that was invented in the 1940s, though. Most of us wouldn't have confidence in the results from a cancer test developed then.

We have more knowledge and a far better understanding of intelligence and ability now” (Sternberg, 2016).

Without the actual data to test these hypotheses, it is unclear as to which path the institution should take. However, it is clear the institution should keep the GRE and GPA as measuring tools of the statistical outcome of the different variables equal less than the statistic itself because this would show there is no statistically significant reason to change what is already in place. But, if there are any variables which are greater than the statistic, it might be time to look into alternatives for measuring current and future student success at the institution. Whatever the outcome would be, by testing the t-value, p-value, and statistical data from a random, yet well representative group of students given by the institution, the director will be able to give the dean a well-rounded, and close to definite answer as to whether or not they should keep the current measuring tools in place, or if it is time for alternate measuring methods. For the most part, according to past studies as well as the hypothetical outcomes of the different analyses of this study, it has been proven the GRE is not the most reliable way for educational institutions to predict whether or not students will be successful at their college or university.

References

- Campbell, M. (2014). The importance of statistics. *African Journal of Midwifery & Women's Health* , 8 (3), 113-114.
- Caskie, G. I., Sutton, M. C., & Eckhardt, A. G. (2014). Accuracy of self-reported college gpa: gender-moderated differences by achievement

level and academic self-efficacy. *Journal of College Student Development*, 55 (4), 385-390. doi: 10. 1353/csd. 2014. 0038

- Cole, J., & Gonyea, R. (2010). Accuracy of self-reported SAT and ACT test scores: implications for research. *Research in Higher Education*, 51 (4), 305-319.
- Balakrishnan, N., Castilla, E., Martin, N., & Pardo, L. (2019). Robust estimators and test statistics for one-shot device testing under the exponential distribution. *IEEE Transactions on Information Theory*, *Information Theory*, *IEEE Transactions on*, *IEEE Trans. Inform. Theory*, (5), 3080.
- Deluca, C. (2011). Interpretive validity theory: Mapping a methodology for validating educational assessments. *Educational Research*, 53 (3), 303-320. doi: 10. 1080/00131881. 2011. 598659
- Garb, H. N., Wood, J. M., & Baker, M. (2018). The lackland behavioral questionnaire: the use of biographical data and statistical prediction rules for public safety screening. *Psychological Assessment*, 30 (8), 1039-1048.
- Hale, J. (2010). Does GRE measure anything related to graduate school? Retrieved from <https://psychcentral.com/blog/does-the-gre-measure-anything-related-to-graduate-school/>
- Hirschfeld, M., & Moore, R. L. (1995). Exploring the gender gap on the GRE subject test in economics. *Journal of Economic Education*, 26 (1), 3. <https://doi-org.proxy1.ncu.edu/10.2307/1183461>
- Knapp, H. (2017). *An introduction to the paired t-test* [Video file]
London: SAGE

- Michaelson, G., & Hardin, M. (2012). *Significance, statistical*. In N. J. Salkind (Ed.), *Encyclopedia of research design* (p. 1362-1366). Thousand Oak.
- Petersen, S. L., Erenrich, E. S., Levine, D. L., Vigoreaux, J., & Gile, K. (2018). Multi-institutional study of GRE scores as predictors of STEM PhD degree completion: GRE gets a low mark. *Plos One* , 13 (10), e0206570. <https://doi.org.proxy1.ncu.edu/10.1371/journal.pone.0206570>
- Ryan, C. (2016). One-Sample & Paired Samples t-tests: A Model Comparison Approach. <https://doi-org.proxy1.ncu.edu/10.4135/9781473969858>
- Schwager, I. T. L., Hülshager, U. R., Bridgeman, B., & Lang, J. W. B. (2015). Graduate student selection: graduate record examination, socioeconomic status, and undergraduate grade point average as predictors of study success in a western European University. *International Journal of Selection & Assessment* , 23 (1), 71-79. <https://doi-org.proxy1.ncu.edu/10.1111/ijsa.12096>
- Sternberg, R. J. (2016). The Gale Encyclopedia of Psychology. (p. 1143, 1145). Retrieved from <https://search-ebshost.com.proxy1.ncu.edu/login.aspx?direct=true&db=edsgvr&AN=edsgcl.3631000747&site=eds-live>
- Tock, J. L., & Ericsson, K. A. (2019). Effects of curricular emphasis in college on the GRE and its impact on the gender gap in performance. *Contemporary Educational Psychology* , 56 , 40-54. <https://doi-org.proxy1.ncu.edu/10.1016/j.cedpsych.2018.11.003>

- Wao, J. O., Ries, R., Flood, I., Lavy, S., & Ozbek, M. E. (2015). Relationship between admission GRE scores and graduation GPA scores of construction management graduate students. *International Journal of Construction Education and Research*, 12 (1), 37-53. doi: 10.1080/15578771.2015.1050562
- Zhang, S., Chen, H.-S., & Pfeiffer, R. M. (2013). A combined p-value test for multiple hypothesis testing. *Journal of Statistical Planning and Inference* , 143 (4), 764-770.