Student: **Karla M. Kassey**

THIS FORM MUST BE COMPLETELY FILLED IN

Follow these procedures: If requested by your instructor, please include an assignment cover sheet. This will become the first page of your assignment. In addition, your assignment header should include your last name, first initial, course code, dash, and assignment number. This should be left justified, with the page number right justified. For example:

DoeJXXX0000-1 1

Formatted Table

Save a copy of your assignments: You may need to re-submit an assignment at your instructor's request. Make sure you save your files in accessible location.

Academic integrity: All work submitted in each course must be your own original work. This includes all assignments, exams, term papers, and other projects required by your instructor. Knowingly submitting another person's work as your own, without properly citing the source of the work, is considered plagiarism. This will result in an unsatisfactory grade for the work submitted or for the entire course. It may also result in academic dismissal from the University.

BTM7104-8 **Dr. Abeer Yasin**

Statistics I **Paper—Fictitious Statistical Study**

Great work Karla. Wow and excellent, excellent study. Wonderful work on the discussion of your study, data collection, data analysis, data presentation, hypothesis statement and hypothesis testing. Excellent work on the use of correlational tests, regression tests and t-testing. Wonderful discussion of the results and conclusion to your study. Very well done! Congratulations on a wonderful work done in this course.

<Faculty Name>Dr. Abeer Yasin <Grade Earned>7/7 <Writing Score>3/3 total 100% A

<Date Graded>Sept 11th 2013

Fictitious Statistical Study—Signature Assignment

Do SAT Scores Reflect Income Level Ten Years after Graduation from College for Males versus Females?

Commented [A1]: Good question! ☺
Let's see what you have for us.

Dr. Abeer Yasin

Karla M. Kassey

Northcentral University

Do SAT Scores Reflect Income Level Ten Years after Graduation from College for Males versus Females?

Description of Research Study

This research study will attempt to determine whether there is a significant difference between males and females with regard to SAT scores versus the level of salary attained or income ten years after graduation from college with an undergraduate degree. In addition, this research study will be undertaken by focusing on data from its survey participants from Kent State University (2012), which is a Midwestern university that offers a Master of Business Administration degree program to students with a wide range of abilities and varied levels of professional work experience. Many of Kent State's participants in the graduate business program also received their undergraduate degree from this same university. Also, according to *U.S. News* (2013), 87.5% of applicants are accepted into the graduate business program. What this means is that these same students have a wide array of scores on their SAT tests, which makes this an ideal setting from which to conduct a research project to compare males and females who are ten years out from receiving their undergraduate degree and their income level.

Students in the MBA program who also work full or part time will be asked to complete a short questionnaire that determines the number of years since graduation from college as well as current income level. The survey will be drawn from a population of 636 students, many of whom have ten years of work experience or are ten years out from receiving their undergraduate degree. The strategy will be to include an envelope in which the survey can be placed and sealed upon completion of the questionnaire. This will ensure the anonymity of these students/career

professionals in order to be certain that honest responses are provided with regard to income, etc. The strategy is also to get responses from 20 males and 20 females, many of whom participate in the part time and executive level MBA programs at Kent State University that fit the above criteria of ten years out from graduating from an undergraduate program and currently employed.

Commented [A2]: Good work

Description of the SAT and its relevance to this study

According to the College Board (2013, para 1), "The SAT is a globally recognized college admission test that lets you show colleges what you know and how well you can apply that knowledge. It tests your knowledge of reading and math. Most students take the SAT during their junior or senior years of high school, and almost all colleges and universities use the SAT to make admission decisions." For the SAT, scores fall between 200 and 800 for the Verbal Section and between 200 and 800 for the Math Section. This results in a total possible score of 1,600 on the SAT test. The writing score for the SAT would not have become available as part of the test until 2007 and the students completing the questionnaire for the research study would have graduated from high school at least seven years before the institution of the writing portion of the exam. The relevance of the SAT to this study is that it might be presumed that students with higher SATs would attend colleges with a more rigorous curriculum and would, therefore, be able to attain higher salaries that would persist ten years after graduation from college. It also will be presumed that salaries ten years after college differ for males versus females. This presumption, however, is not absolute and that is the reason that the researcher is leaning toward a two-tailed rather than a one-tailed test.

List and explanation of the variables upon which information will be collected in this study

The variables will include the following:

1. Male/Female

2. SAT score
3. Salary/Income

The first variable is that of male/female and will be used to determine whether there is a significant difference in salary based on one's SAT score and dependent upon the sex of the respondent to the survey. The second variable is the SAT score and a test will be conducted to determine whether there is a correlation between SAT score and income and whether this correlation is due to correlation versus causation. The first null hypothesis is that there is no significant difference between men and women with regard to income ten years after completing an undergraduate degree. The alternative hypothesis is that there is a difference in salary for men compared to women for the income received ten years after college. This difference is also presumed to be linked to one's SAT score. In this case, it is presumed that SAT scores influence salary, so the dependent variable which is a variable that is caused by or influenced by another is that of salary/income. The independent variable, therefore, is the SAT score as it is presumed that SAT scores influence income ten years after graduation and that the level of influence is dependent upon whether the subject is male or female.

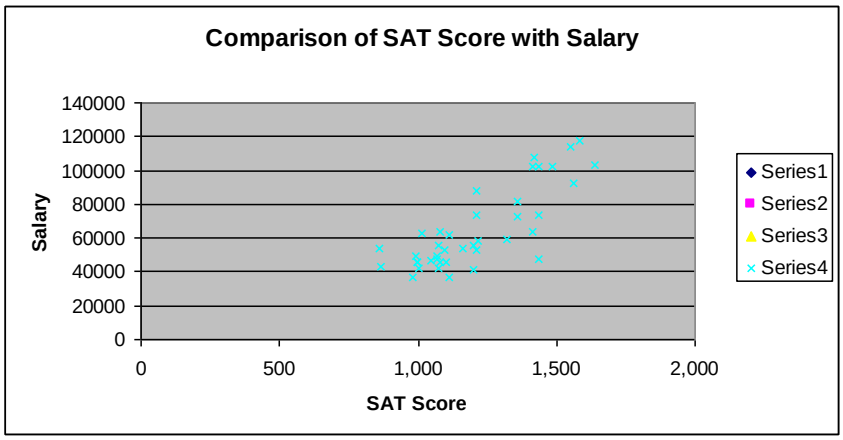
Commented [A3]: Good work on the discussion of variables

Creation of a data set that will be analyzed. The data set includes data from 20 males and 20 females who participated in an anonymous survey

1.0 males and 2.0 females	SAT Score	Salary
1.0	1,049	47,000
1.0	1,200	56,000
1.0	1,433	74,000
1.0	982	37,000
1.0	864	43,000
1.0	1,322	59,000
1.0	1,433	48,000
1.0	1,212	53,000

1.0	**1,068**	**48,000**
1.0	**1,485**	**102,000**
1.0	**1,102**	**46,000**
1.0	**1,000**	**42,000**
1.0	**996**	**46,000**
1.0	**1,358**	**82,000**
1.0	**1,412**	**64,000**
1.0	**1,113**	**37,000**
1.0	**1,201**	**41,000**
1.0	**1,080**	**46,000**
1.0	**1,074**	**42,000**
1.0	**1,582**	**118,000**
2.0	**1,082**	**64,000**
2.0	**1,213**	**88,000**
2.0	**1,436**	**102,000**
2.0	**1,360**	**73,000**
2.0	**998**	**46,000**
2.0	**863**	**54,000**
2.0	**1,418**	**108,000**
2.0	**1,563**	**92,000**
2.0	**1,098**	**53,000**
2.0	**1,416**	**102,000**
2.0	**1,212**	**74,000**
2.0	**1,069**	**49,000**
2.0	**1,218**	**58,000**
2.0	**1,549**	**114,000**
2.0	**1,112**	**62,000**
2.0	**1,600**	**103,000**
2.0	**1,012**	**63,000**
2.0	**1,163**	**54,000**
2.0	**993**	**49,000**
2.0	**1,074**	**56,000**

Descriptive Data Analysis



What this graph means is that there is a positive relationship between income and SAT score with some degree of dispersion. The mean for the SAT based on the sample is 1,211 compared with a mean of 1,016 in 1999 (which accounts for 14 years or four years of college and ten years after college) according to the College Board (2011) Total Group Profile Report.

Correlation between income and SAT score

Correlations

Gender			Gender	Income	SAT
1.00	Gender	Pearson Correlation	.[a]	.[a]	.[a]
		Sig. (2-tailed)		.	.
		N	20	20	20
	Income	Pearson Correlation	.[a]	1	.808**
		Sig. (2-tailed)	.		.000
		N	20	20	20
	SAT	Pearson Correlation	.[a]	.808**	1
		Sig. (2-tailed)	.	.000	
		N	20	20	20
2.00	Gender	Pearson Correlation	.[a]	.[a]	.[a]

		Sig. (2-tailed)		.	.
		N	20	20	20
		Pearson Correlation	.[a]	1	.883**
	Income	Sig. (2-tailed)	.		.000
		N	20	20	20
		Pearson Correlation	.[a]	.883**	1
	SAT	Sig. (2-tailed)	.	.000	
		N	20	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

a. Cannot be computed because at least one of the variables is constant.

There is a positive correlation for both income and SAT scores divided up by female (1.00) and male (2.00) at the .01 level of significance. A positive score indicates a correlation in the positive direction which means as SAT score increases so does income for both males and females as r = .808 for women and r= .883 for men. According to Pallant (2014), the r value can range from -1 to 1 with -1 being perfect negative correlation and positive 1 being perfect positive correlation. The results here indicate a correlation in the positive direction that is fairly strong for both males and females. Using the causation versus correlation principle, it can be argued that there is an additional variable that influences income ten years after college for women versus men and that variable could be the fact that there is discrimination in the workplace with regard to the lack of equal pay for equal work for women. Employers argue that women lag based on income in the workplace due to taking time off for maternity that will affect their pay. Employers also argue that women will work part-time due to having families and that perhaps that is another one of the reasons that women lag in pay ten years after graduating from college.

Commented [A4]: Excellent work

Regression Line Excel Table and Formula

	x	y	x^2	y^2	x*y
1	**1,049**	**47,000**	1,100,401	2,209,000,000	49303000
1	**1,200**	**56,000**	1,440,000	3,136,000,000	67200000
1	**1,433**	**74,000**	2,053,489	5,476,000,000	106042000

1	**982**	**37,000**	964,324	1,369,000,000	36334000
1	**864**	**43,000**	746,496	1,849,000,000	37152000
1	**1,322**	**59,000**	1,747,684	3,481,000,000	77998000
1	**1,433**	**48,000**	2,053,489	2,304,000,000	68784000
1	**1,212**	**53,000**	1,468,944	2,809,000,000	64236000
1	**1,068**	**48,000**	1,140,624	2,304,000,000	51264000
1	**1,485**	**102,000**	2,205,225	10,404,000,000	151470000
1	**1,102**	**46,000**	1,214,404	2,116,000,000	50692000
1	**1,000**	**42,000**	1,000,000	1,764,000,000	42000000
1	**996**	**46,000**	992,016	2,116,000,000	45816000
1	**1,358**	**82,000**	1,844,164	6,724,000,000	111356000
1	**1,412**	**64,000**	1,993,744	4,096,000,000	90368000
1	**1,113**	**37,000**	1,238,769	1,369,000,000	41181000
1	**1,201**	**41,000**	1,442,401	1,681,000,000	49241000
1	**1,080**	**46,000**	1,166,400	2,116,000,000	49680000
1	**1,074**	**42,000**	1,153,476	1,764,000,000	45108000
1	**1,582**	**118,000**	2,502,724	13,924,000,000	186676000
2	**1,082**	**64,000**	1,170,724	4,096,000,000	69248000
2	**1,213**	**88,000**	1,471,369	7,744,000,000	106744000
2	**1,436**	**102,000**	2,062,096	10,404,000,000	146472000
2	**1,360**	**73,000**	1,849,600	5,329,000,000	99280000
2	**998**	**46,000**	996,004	2,116,000,000	45908000
2	**863**	**54,000**	744,769	2,916,000,000	46602000
2	**1,418**	**108,000**	2,010,724	11,664,000,000	153144000
2	**1,563**	**92,000**	2,442,969	8,464,000,000	143796000
2	**1,098**	**53,000**	1,205,604	2,809,000,000	58194000
2	**1,416**	**102,000**	2,005,056	10,404,000,000	144432000
2	**1,212**	**74,000**	1,468,944	5,476,000,000	89688000
2	**1,069**	**49,000**	1,142,761	2,401,000,000	52381000
2	**1,218**	**58,000**	1,483,524	3,364,000,000	70644000
2	**1,549**	**114,000**	2,399,401	12,996,000,000	176586000
2	**1,112**	**62,000**	1,236,544	3,844,000,000	68944000
2	**1,600**	**103,000**	2,679,769	10,609,000,000	168611000
2	**1,012**	**63,000**	1,024,144	3,969,000,000	63756000
2	**1,163**	**54,000**	1,352,569	2,916,000,000	62802000
2	**993**	**49,000**	986,049	2,401,000,000	48657000
2	**1,074**	**56,000**	1,153,476	3,136,000,000	60144000
	48,452	2,595,000	60,354,870	190,069,000,000	3,297,934,000
	1211.3	64875			

a = y-bar - b * x-bar
x-bar = the average value of x = 1,211.3
y-bar = the average value of y =
64,875

b = n * sum of xy - (Sum of x) * (Sum of y) / n * sum of x^2 - (sum of x) ^2
b = 40 * 3,297,934,000 - 48,452 * 2,595,000 / 40 * 60,354,870 - (48,452) ^2
131,917,360,000 - 48,452 * 2,595,000 / 2,414,194,800 - 2,347,596,304
131,917,360,000 - 125,732,940,000 / 2,414,194,800 - 2,347,596,304
6,184,420,000 / 66,598,496 = 92.86
b = 92.86
a = y-bar - b* x-bar
a = 64,875 - 92.86 * 1,211.3
a = 64,875 - 112,481 = -47606

y-hat = a + b (x)
y-hat = -47,606 + 92.86 * x

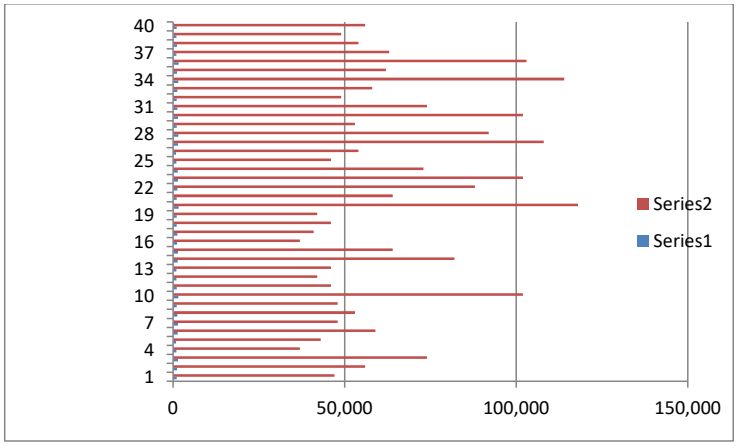
Using the (y-hat = -47,606 + 92.86 * x) equation which was derived from the data listed above through the use of an Excel spreadsheet, it is possible to determine what an individual should earn on a yearly basis based on his/her SAT score. Some of the participants in the survey process may be earning more or less than the amount determined by this equation. The premise is that those individuals that scored higher on their SATs will attend universities or colleges, on average, with more rigorous curriculums. Attending a college such as this will allow for a higher than average salary upon completion of one's degree program and it has been determined that this difference in pay will persist ten years after receiving one's bachelor's degree. The downside to earning a higher than average salary that is beyond the salary determined by the equation above based on one's SAT score, however, is that these individuals may be targeted first for layoffs if the company is forced into a situation in which it must downsize. Individuals who recognize this situation may be able to be proactive in taking on additional duties at work in order to be less attractive as a target when the company must lay off workers in order to cut costs. It is also the job as the researcher to determine the validity of whether there is a

correlation between SAT score and income which appears likely and would validate the above data on the usefulness of the regression equation. Also, higher SAT scores are, again, linked with colleges whose graduates generally command a higher salary but, as discussed above, these may be the first individuals targeted for layoffs due to their higher salaries unless the employer sees a clear reason to keep them on the payroll.

Commented [A5]: Excellent work

Right Skewed data

The graph below demonstrates the fact that the data is Right-skewed. What this means is that the data is more spread out on the right side. Most individuals, according to Bennett, Briggs & Triola (2014), are middle income earners so the mode of this distribution would likely be a middle-class income. A few higher income individuals, however, can push the distribution to the right which causes a right-skewed distribution as illustrated by the graph below. This graph is important from the standpoint that it illustrates that there are higher than average earners that skew the data and make it more difficult to determine whether there is actually a correlation between income and male versus female subjects. To further validate the data, we can use a confidence interval to determine where most of the incomes are located on the continuum.



Margin of error

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Income	40	37000.00	118000.00	64875.0000	23598.33490
SAT	40	863.00	1600.00	1211.3000	204.73844
Valid N (listwise)	40				

To specify the confidence interval, we can calculate the margin of error with a t distribution, and the formula we would use is that of E = t * s / sqrt n. To do this, we must first determine the degrees of freedom, which in this case is n -1 or 40 – 1 = 39. Next, based on the degrees of freedom, we must find the t-value that corresponds to a 95% confidence interval, so we must look in the column labeled .05 for a two-tailed test. In this case, the t value is 2.0227. What this means is that 2.0227 * 23598 / sqrt 40 = 2.027 * 23598 / 6.32 = 2.0227 * 3733.86 = 7,552. In this case, the confidence interval is an estimation of the true value of the population mean according to Bennett, Briggs & Triola (2014). Using the data above, the following equation 64,875 – 7,522 < u < 64,875 + 7,522 or 57,353 < u < 72,387 contains the confidence interval for the population mean. What this also means is that 95% of the incomes will fall within this range based on the total population of individuals ten years out from college who are working professionals in their fields. For the SAT scores, the margin of error is 66 and the confidence interval is 1145 < u < 1277. This data is helpful from the standpoint that it tells us where the majority of the participants in the study lie with regard to income and SAT scores. To actually determine what the connection is between males and females with regard to a difference based on income, we can use the following diagnostic tool.

Hypothesis for this research study

First, we will state the hypothesis and alternative hypothesis:

Ho: There is no difference in the level of salary attained ten years after college for males versus females who are working professionals in their field

Ha: There is a difference in the level of salary attained by males versus females ten years out from graduation from an undergraduate degree program who are working professionals in their field

This is a two-tailed test as according to Voelker, Orton, & Adams (2001, p. 66), "Some hypothesis tests predict only that one value will be different from another, without additionally predicting which will be higher. The test of such a hypothesis is non-directional or two-tailed because an extreme test statistic in either tail of the distribution (positive or negative) will lead to the rejection of the null hypothesis of no difference." According to this same source, Voelker et al. (2001, p. 67), "The decision of whether to use a one- or a two-tailed test is important because a test statistic that falls in the region of rejection in a one-tailed test may not do so in a two-tailed test, even though both tests use the same probability level."

The challenge here will be to determine whether gender versus income is significant to the point that we can reject the null hypothesis in favor of the alternative hypothesis. In other words, we must determine whether males and females are divergent with regard to income and whether this divergence is or is not responsible for a significant difference in income.

Independent Sample t-test

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Income	1.00	20	56550.0000	21828.21689	4880.93768
	2.00	20	73200.0000	22818.73653	5102.42460

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	9
Income	Equal variances assumed	.892	.351	-2.358	38	.024	-16650.00000	7061.04025	-
	Equal variances not assumed			-2.358	37.925	.024	-16650.00000	7061.04025	-

Analysis of data from Independent Sample t-test

Since the value in the Sig. (two-tailed) test is equal to or less than .05, there is a significant difference in the mean scores for the dependent variable, which in this case is the income of males versus females. Another way of interpreting this data is to conduct a test of the t-value which compares the test statistic t in the above table labeled Independent Sample Test to the value found in the T distribution critical values table. The table value, according to easycalculation.com (n.d.), is 2.0244 at the .05 significance level. The reason for a two-tailed test in this instance is that we are measuring whether there is a difference in the income of men versus women ten years after graduation from college from an undergraduate degree program versus whether one group's salary is greater than the other. According to Bennett, Briggs & Triola (2014), we can reject the null hypothesis if the computed test statistic is greater than or equal to the value found in the T distribution critical values table like the one found in table 10.1 in this textbook or the one by easycalculation.com (n.d.). The observed test statistic is 2.358. What this means is that since the observed t statistic is greater than the t statistic in the table, we can reject the null hypothesis in favor of the alternative hypothesis. The alternative hypothesis states that there is a difference between the incomes of men versus women ten years after receiving one's undergraduate degree.

The reason for choosing the independent sample t-test is that according to Pallant (2013 p. 247), that "this test is used when you want to compare the mean scores of two different groups of people or conditions." In this case, we want to compare gender differences based on income ten years after graduating from an undergraduate degree program. The difference between the means for men and women is 73,200 – 56,500 = 16,700, which is a substantial difference for men versus women with regard to income The difference in standard deviation is 22818.74 – 21828 = 990.74. What this means is that there is a substantial difference between the mean income for males versus females as well as a difference in the standard deviation from the mean.

We can also use this same test to determine whether there is a significant difference between males and females and SAT scores. The Null Hypothesis is that there is no significant difference between males and females with regard to SAT Scores and the Alternative Hypothesis is that there is a significant difference between males and females with regard to SAT scores.

Independent Sample t-test

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
SAT	1.00	20	1198.3000	198.72173	44.43553
	2.00	20	1224.3000	218.59434	48.87918

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95%
									Lo
SAT	Equal variances assumed	.197	.659	-.394	38	.696	-26.00000	66.05824	-
	Equal variances not assumed			-.394	37.660	.696	-26.00000	66.05824	-

Analysis of Data from Independent Sample t-test

In this case, the researcher conducted a second Independent Sample t-test to determine whether there is a significant difference among SAT scores between males and females. Based on the significance value of the Independent Sample Test, which is .696, it can be confirmed that the significance level is greater than .05 which means we cannot reject the null hypothesis that states that there is no significant difference between males and females with regard to SAT scores. Another method for verifying this is to compare the observed t score to that of the T Distribution Critical Values Table (n.d.), which states that the value based on this table is 2.0224. Since the observed t statistic of .394 is not greater than the table statistic of 2.0224, we cannot reject the null hypothesis based on this information.

Commented [A6]: Wonderful work

Conclusion

In conclusion, and based on the above data, we can conclude that even though there is a substantial difference between the incomes of males and females, there is no substantial difference between SAT scores. What this means is that even though women match their male counterparts with regard to level of ability academically, they are not paid the same income as men in the professional world. This may be simply due to discrimination against women in the workplace or it may be due to an external factor such as women working part-time due to family responsibilities, or it may be partly due to both factors. These are factors that should be considered by employers as not providing equal pay for equal work can lead to a loss of employee productivity, and this issue should be studied further in the corporate world to ensure that this theory is not the case. Otherwise, the salaries of women should be increased in order to increase productivity and morale in the workplace.

References

Bennett, Briggs, & Triola. (2014). *Statistical Reasoning for Everyday Life.* Boston: Pearson.

easycalculation.com. (n.d.). Retrieved July 25, 2013, from T Distribution Critical Values Table: http://easycalculation.com/statistics/t-distribution-critical-value-table.php

Kent State University. (2012). Retrieved August 25, 2013, from Flash Facts Fall 2012: www.kent.edu/rpie/data/upload/flash-facts-fall-2012.pdf

Pallant, J. (2013). *SPSS Survival Manual.* Berkshire: McGraw-Hill Companies.

The College Board. (2013). *SAT*. Retrieved August 25, 2013, from About the SAT: https://Sat.collegeboard.or/about-tests/sat

U.S. News . (2013). Retrieved August 25, 2013, from Kent State University: colleges.usnewsrankingsandreviews.com/best-colleges/Kent-State-University-3051

Voelker, Adams, & Orton. (2001). *Cliffs Quick Review--Statistics.* New York: Wiley Publishing, Inc.