Assignment #3: Simple Linear Regression Refresher (60 points)

Due: by 11:59 pm, Sunday, February 28th, 2021

** No assignments will be accepted late without written documentation of an unavoidable or unforeseeable event.*

Why this assignment?

This assignment is intended to achieve the goal of our Module 6 learning objective, "conduct a simple linear regression analysis to investigate and evaluate the relationship between two variables." Students will need to conduct a simple linear regression analysis to design a pay structure, which is the main activity in Module 7.

Instruction

Watch the video *Simple linear regression refresher* in Module 6 and answer the questions on pages 3 to 7 (on this document). For this assignment, you'll need to analyze the attached data (*Dataset for Assignment 3*) using the Microsoft Excel software.

Important Note

1. In Excel, 'E' is often used to indicate a very small decimal number. For example, 2.37E-20 is not 2.37. 2.37E-20 is 0.00000000000000000000237. How to interpret 'E' was explained in the video (in a yellow box from runtime 23:14 to 23:42). Many students interpret 2.37E-20 as 2.37, and gets the p-value wrong on this assignment.

2. In writing your answers, **NEVER copy-and-paste any part of the textbook or any other sources.** Write the answers in your own words. If more than seven consecutive words are identical to a writing from another source (including the textbook, your past works, and any other sources), it will be considered as plagiarism, and you will receive a failing grade "F" in this course. You will also be referred to Students Affairs.

3. I change the numbers on the dataset every year. So, **do not copy-and-paste the answers from the assignments that were given out in this course in the past**. In the case of your answers being the same as the answers in the past, it will be regarded as plagiarism and behavior that violates the '4. Ethics' category of the professional etiquette outlined on page 5 of our syllabus. In this case, you will receive a failing grade "F" in this course and be referred to Students Affairs.

Grading Rubric

There is no rubric for this assignment. If you answer a question right, you'll receive a full credit assigned to that question. If you do not answer a question right, you'll receive 0 points for that question. The answers for the 'Example Question' will be provided on page 2 of this document. How you phrase your answers for Questions 1 to 5 should be similar. The cases that will be graded as wrong answers are also shown in the 'Example Question.' Please read this carefully before you start your assignment.

Example Question

For this question, use the data on the sheet *Example Question* in the attached Excel file (*Dataset for Assignment 3*).

A researcher wanted to examine the relationship between autonomy and job satisfaction.
As in the datasheet, the researcher has collected data from 99 participants. Column B represents autonomy and column C represents job satisfaction of the participants.

Conduct a simple linear regression analysis with *autonomy* as the independent variable (X) and *job satisfaction* as the dependent variable (Y).

Example Question 1. What is the estimated linear regression equation?

***Answer*: Job satisfaction = 0.30 * Autonomy + 1.86**

Example Question 2. What is the p-value of the slope in the simple linear regression equation?

***Answer*: 0.002**

Example Question 3. Explain the relationship between *job satisfaction* and *job security perception*.

***Answer*: Job satisfaction level is expected to increase by 0.30 points as the autonomy level increases by one point.**

** Cases that will be graded as wrong answers.*
- *The relationship is explained but not specifically in numbers. (e.g., "There is a positive relationship between job satisfaction and autonomy.", "Job satisfaction level increases as the autonomy level increases.")*
- *The relationship is explained but not in an easy-to-understand English format. (e.g., when the independent variable is gender, using a phrase like "one unit increase in gender (or female)")*

Question 1 (Total of 12 points)

For this question, use the data on the sheet *Question 1* in the attached Excel file (*Dataset for Assignment 3*).

A researcher wanted to examine the relationship between sleep hours and overall happiness level. As in the datasheet, the researcher has collected data from 100 participants. Column B shows the average sleep hours per day of the participants. Column C shows the overall happiness level of the participants.

Conduct a simple linear regression analysis with the *sleep hours* as the independent variable and the *overall happiness level* as the dependent variable.

1-1. What is the estimated simple linear regression equation? **(3 points)**

1-2. What is the p-value of the slope in the simple linear regression equation? **(3 points)**

(Hint: 2.37E-20 is not 2.37. 2.37E-20 is 0.000000000000000000237.)

1-3. Explain the relationship between *sleep hours* and *overall happiness level*. **(3 points)**

(Hint: Your answer should be "As a person sleeps one more hour, his or her expected overall happiness level … …")

1-4. According to the simple linear regression equation in 1-1, what is the estimated level of overall happiness for a person who sleeps 6 hours per day on average? **(3 point)**

Question 2 (Total of 12 points)

For this question, use the data on the sheet *Question 2* in the attached Excel file (*Dataset for Assignment 3*).

A researcher wanted to examine the relationship between job security perception and job satisfaction. As in the datasheet, the researcher has collected data from 99 participants. Column B shows the job security perception level of the participants. Column C shows the job satisfaction level of participants.

Conduct a simple linear regression analysis with the *job security perception level* as the independent variable and the *job satisfaction level* as the dependent variable.

2-1. What is the estimated simple linear regression equation? **(4 points)**

2-2. What is the p-value of the slope in the simple linear regression equation? **(4 points)**

2-3. Explain the relationship between *job security perception level* and *job satisfaction level*. **(4 points)**

Question 3 (Total of 12 points)

For this question, use the data on the sheet *Question 3* in the attached Excel file (*Dataset for Assignment 3*).

A researcher wanted to see if there is a gender difference in job satisfaction.
As in the datasheet, the researcher has collected data from 100 participants. Column B shows the gender of the participants (male = 0, female = 1). Column C shows the job satisfaction level of the participants.

Conduct a simple linear regression analysis with *gender* as the independent variable and *job satisfaction level* as the dependent variable.

3-1. What is the estimated simple linear regression equation? **(4 points)**

3-2. What is the p-value of the slope in the simple linear regression equation? **(4 points)**

3-3. Explain the relationship between *gender* and *job satisfaction level*. Is there a gender difference in job satisfaction? **(4 points)**

Question 4 (Total of 12 points)

For this question, use the data on the sheet *Question 4* in the attached Excel file (*Dataset for Assignment 3*).

Jake, the owner of the Jake Automotive, wanted to examine the relationship between the number of cars that his sales associates sold and their job satisfaction.
As in the datasheet, Jake has collected data from 49 sales associates. Column B shows the average number of cars per month a sales associate has sold. Column C shows the job satisfaction level of a given sales associate.

Conduct a simple linear regression analysis with *# of cars sold* as the independent variable and *job satisfaction* as the dependent variable.

4-1. What is the estimated simple linear regression equation? **(4 points)**

4-2. What is the p-value of the slope in the simple linear regression equation? **(4 points)**

4-3. Explain the relationship between *# of cars sold* and *job satisfaction level*. **(4 points)**

Question 5 (Total of 12 points)

For this question, use the data on the sheet *Question 5* in the attached Excel file (*Dataset for Assignment 3*).

Kevin, the mayor of Wonderland, wanted to see if there is a gender pay gap in Wonderland.
As in the datasheet, Kevin has collected data from 70 working people in the area. Column B shows the gender of participants (male = 0, female = 1). Column C shows the annual pay of a given participant.

Conduct a simple linear regression analysis with *gender* as the independent variable and *annual pay* as the dependent variable.

5-1. What is the estimated simple linear regression equation? **(4 points)**

5-2. What is the p-value of the slope in the simple linear regression equation? **(4 points)**

5-3. Explain the relationship between *gender* and *annual pay*. Is there a gender pay gap in Wonderland? If there is a gender pay gap, what is the estimated gap amount and who (male or female workers) earns more? **(4 points)**