Math 1530 Capstone Project Fall 2014 100 points

**Directions:**

1. **Type your Name, E number, and Section number in the header (double click the header and then click the toolbar “Close header and footer” in the menu list).**
2. **DO YOUR OWN WORK! It is academic misconduct to copy or seek assistance from other people, or to share your work with other students. Any academic misconduct on this project results in a grade of 0 and a written report to dean’s office.**
3. **This part of the project counts for 10% of the total grade.**
4. **The project is due by on . No late projects will be accepted. Don’t wait till the last minute to start working—you know how computer technology can fail at ETSU or at home without even a moment’s notice.**
5. Start each problem on a new page.
6. Insert any graphs in the appropriate places (not attached as an addendum at the back or even at the end of the problem.)
7. Only insert the relevant portions of a Minitab display used to answer a question, not everything Minitab gives you in hoping the right information is somewhere in what you copied into the document.
8. **Do not hand in these 1st two pages—just the problems**, please.
9. Please make it easier for your instructor to find your answers/discussions. You might use a different font for your answers or make them bold print. If you are using a color printer (with fresh ink cartridges) you could highlight in yellow (other colors will obscure your typing in the printed version) or use a different color of ink for your responses.

**Here are the questions that were asked on the survey:**

1. **GENDER:** Are you male or female? (Male, Female)
2. **CHILDREN:** What do you think is the ideal number of children for a family to have?
3. **MARRIED:** What is your opinion about a married person having sexual relations with someone other than the marriage partner? (Always wrong, Almost always wrong, Wrong only sometimes, Not wrong at all)
4. **SEX:** There's been a lot of discussion about the way morals and attitudes about sex are changing in this country. What is your opinion about two people having sexual relations before marriage? (Always wrong, Almost always wrong, Wrong only sometimes, Not wrong at all)
5. **PARTY:** What political party do you identify with? (Democratic, Republican, Independent, Other)
6. **INCOME\_TAX:** Should Tennessee implement a state income tax? (Yes, No)
7. **SPANKING:** Do you strongly agree, agree, disagree, or strongly disagree that it is sometimes necessary to discipline a child with a good, hand spanking? (Strongly agree, Agree, Disagree, Strongly disagree)
8. **DEATH\_PENALTY:** Do you favor or oppose the death penalty for persons convicted of murder? (Favor, Oppose)
9. **TV:** How many hours do you personally watch television including Netflix, Hulu Plus, Amazon Prime, etc... in a day?
10. **DEVICES:** How many devices do you have that will allow you watch a TV show or movie (live, streaming, pre-recorded, online)?
11. **PREPARING:** How many hours per 7-day week do you spend preparing for class (studying, reading, writing, doing homework or lab work, analyzing data, rehearsing, and other academic activities)?
12. **SLEEP:** Usually, how many hours sleep do you get a night?
13. **SHOES:** How many pairs of shoes do you own?

 **A total of 1220 students responded to the MATH1530 class survey. The data for 1184 students were recorded in the Minitab worksheet MATH1530FALL2014CapProjData.MTW.

Two questions were taken from a MATH 1530 survey from a previous semester. A total of 791 students responded to this survey. The data for those 791 students were also recorded in the Minitab worksheet MATH1530FALL2014CapProjData.MTW. Below is the two questions taken from the survey.**14. **SHOE\_SIZE:** What is your U.S. shoe size?
15. **HEIGHT:** What is your height in inches?

The Minitab worksheet is set up as follows:

|  |
| --- |
| C1: **ID** |
| C2: **GENDER** |
| C3: **CHILDREN** |
| C4: **MARRIED** |
| C5: **SEX** |
| C6: **PARTY** |
| C7: **INCOME\_TAX** |
| C8: **SPANKING** |
| C9: **DEATH\_PENALTY** |
| C10: **TV** |
| C11: **DEVICES** |
| C12: **PREPARING** |
| C13: **SLEEP** |
| C14: **SHOES** |
| C15: **SHOE\_SIZE** |
| C16: **HEIGHT** |

1. **Identify Variable Type.** Which of these questions from the class survey produced variables that are categorical and which are quantitative? Use your word processor to underline the best option (or you may highlight in yellow if you are using a color printer).

	1. **CHILDREN** Categorical Quantitative Neither
	2. **PARTY** Categorical Quantitative Neither
	3. **DEATH\_PENALTY** Categorical Quantitative Neither
	4. **TV** Categorical Quantitative Neither
	5. **SHOES** Categorical Quantitative Neither

1. **Sampling.** In the survey data, the variable “**SLEEP”** is the number of hours you sleep in a night.

a. Type in the first 10 observations from the variable “**SLEEP”** and use this as your sample data. Record the values in the table below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **SLEEP** |  |  |  |  |  |  |  |  |  |  |

Obtain the mean number of hours that one sleeps in a night for the first 10 observations.

The mean number of hours of sleep in a night is \_\_\_\_\_\_ hours.

Identify the type of sampling method you have just used: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

b. Now, generate a random sample of size n=10 (Calc > Random Data > Sample from Columns). Enter 10 in the “Number of rows to Sample” box. Enter the variable “ID” and “**SLEEP**” into the “From columns” box. Enter C17-C18 into the “Store samples in” box. Record the information in the table below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ID |  |  |  |  |  |  |  |  |  |  |
| **SLEEP** |  |  |  |  |  |  |  |  |  |  |

Obtain the mean number of hours that one sleeps in a night for the first 10 observations.

The mean number of hours of sleep in a night is \_\_\_\_\_\_ hours.

Identify the type of sampling method you have just used: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

c. Obtain the mean number of hours that one sleeps in a night for all 1184 observations.

 The population mean number of hours of sleep in a night is \_\_\_\_\_\_ hours.

d. Compare the population mean you found in Part (c) to the means you found in Parts (a) and (b). Which sampling method provides a better estimate of the population mean number of hours one sleeps at night?

1. **If you are female then do this question. (Omit this page/problem if you are male.) SHOES.** Question 13 from the survey asked, “How many pairs of shoes do you own?”
	1. Create an appropriate display for this variable and insert it here.
	2. Which of the following best describes the shape of the distribution? Circle your answer.

Skewed left Symmetric Skewed right

* 1. Calculate numerical measures appropriate for the shape of the distribution to describe the center and spread of **shoes**. Include appropriate output from Minitab here.

		1. Which statistic will you use to describe the center of the distribution? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
		2. What is the value of that statistic? \_\_\_\_\_\_\_\_\_\_\_\_
		3. Which statistic(s) will you use to describe the spread of the distribution?

		\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
		4. What is(are) the value(s) of that statistic? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
	2. Create a side-by-side boxplot to compare the distributions of **shoes** for **genders**. Insert the graph below.
	3. Describe the distributions of **shoes** for the two groups and compare them.
	4. Are there any outliers in each group? Identify them and justify your answers.

1. **If you are male then do this question. (Omit this page/problem if you are female.) CHILDREN.** Question 2 from the survey asked, “What do you think is the ideal number of children for a family to have?”
	1. Create an appropriate display for this variable and insert it here.
	2. Which of the following best describes the shape of the distribution? Circle your answer.

Skewed left Symmetric Skewed right

* 1. Calculate numerical measures appropriate for the shape of the distribution to describe the center and spread of **children**. Include appropriate output from Minitab here.

		1. Which statistic will you use to describe the center of the distribution? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
		2. What is the value of that statistic? \_\_\_\_\_\_\_\_\_\_\_\_
		3. Which statistic(s) will you use to describe the spread of the distribution?

		\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
		4. What is(are) the value(s) of that statistic? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
	2. Create a side-by-side boxplot to compare the distributions of **children** for **gender**. Insert the graph below.
	3. Describe the distributions of **children** for the two groups and compare them.
	4. Are there any outliers in each group? Justify your answer.

1. **Shoe size versus height.** Someone’s shoe size depends on several variables and one of them could be someone’s height. A previous MATH1530 class survey asked students to state their shoe size and height. Questions 14 and 15 asked students to input their U.S. shoe size (SHOE\_SIZE) and their height (HEIGHT) in inches. Assume the respondents are an SRS of all ETSU students. We are interested in studying the relationship between a student’s height and their shoe size. That is, we are interested in seeing whether knowing one’s height can explain one’s shoe size.

	1. Create an appropriate plot to display the relationship between **SHOE\_SIZE** and **HEIGHT**. Insert the plot here.

Does the plot show a positive association, a negative association, or no association between these two variables? EXPLAIN what this means with respect to the variables being studied.

* 1. What is the correlation between the pair of variables? \_\_\_\_\_\_\_\_\_\_\_\_\_
	2. Obtain the least squares regression equation for the pair of variables. Insert it here.
	3. Interpret the value of the slope in the least squares regression equation you found in part (c).
	4. Use the regression equation in part (c) to predict a student’s shoe size for someone that is 73 inches tall.
	5. How well does the regression equation fit the data? Explain. Justify your answer with appropriate plot(s) and summary statistics.

1. **If your E number ends in an even number (0, 2, 4, 6, or 8) then do this question. (Omit this page/problem if your E# ends with an odd number.) SPANKING AND THE DEATH PENALTY** Question 7 from the survey asked students “Do you strongly agree, agree, disagree, or strongly disagree that it is sometimes necessary to discipline a child with a good, hand spanking?” and Question 8 from the survey asked students “Do you favor or oppose the death penalty for persons convicted of murder?” We want to check if there is a relationship between the belief about spanking and the belief about the death penalty. Assume the students who took the class survey are from an SRS of ETSU students.
2. Create an appropriate graph to display the data and insert it here.
3. Create an appropriate two-way table to summarize the data and insert it here.

**SUPPOSE WE SELECT ONE STUDENT AT RANDOM:**

1. Find the probability that the student strongly believes in spanking a child and favors the death penalty for persons convicted of murder.
2. Find the probability that a student strongly disagrees in spanking a child or they oppose the death penalty for persons convicted of murder.
3. Find the probability that a student agrees in spanking a child given they favor the death penalty for persons convicted of murder.
4. Find the probability that a student favors the death penalty for persons convicted of murder given they agree in spanking a child.
5. **BONUS**: Carry out a test for the hypothesis that there is no relationship between the belief about spanking a child and the belief about the death penalty of ETSU students. Use a significance level of α = 0.05.

i. State the null and alternative hypothesis.

ii. Perform the test and include any output from Minitab here.

iii. Which test statistic are you using and what is its value?

iv. State your decision and conclusion for the test.

v. Examine the data. Are the conditions for inference in part (ii) violated? Explain.
6. **If your E number ends in an odd number (1, 3, 5, 7, or 9) then do this question. (Omit this page/problem if your E# ends with an even number.) POLITICAL PARTY AND INCOME TAX** Question 5 from the survey asked students “What political party do you identify with?” and Question 6 from the survey asked students “Should Tennessee implement a state income tax?” We want to check if there is a relationship between political party and the belief about an income tax for Tennessee. Assume the students who took the class survey are from an SRS of ETSU students.
7. Create an appropriate graph to display the data and insert it here.
8. Create an appropriate two-way table to summarize the data and insert it here.

**SUPPOSE WE SELECT ONE STUDENT AT RANDOM:**

1. Find the probability that the student identifies with the Republican Party and agrees Tennessee should implement an income tax.
2. Find the probability that a student identifies with the Independent Party or they disagree Tennessee should implement an income tax.
3. Find the probability that a student identifies with the Democratic Party given they disagree Tennessee should implement an income tax.
4. Find the probability that a student disagrees Tennessee should implement an income tax given they identify with the Democratic Party.
5. **BONUS**: Carry out a test for the hypothesis that there is no relationship between one’s political party and the belief about an income tax in Tennessee. Use a significance level of α = 0.05.

i. State the null and alternative hypothesis.

ii. Perform the test and include any output from Minitab here.

iii. Which test statistic are you using and what is its value?

iv. State your decision and conclusion for the test.

v. Examine the data. Are the conditions for inference in part (ii) violated? Explain.
6. **TV.** A marketing analyst for a cable provider wishes to know if males or females watch more TV in order for the company’s advertisements to target that specific gender. After talking to the company’s current sales representatives across the US, he concludes that males watch more TV. Questions 9 from the survey asked students “How many hours do you personally watch television including Netflix, Hulu Plus, Amazon Prime, etc... in a day?” Assume that the students who responded the survey are a SRS of all ETSU students. Is there good evidence to support the idea that male students at ETSU watch more TV, on average, than female students?
7. Create an appropriate graph to display the distribution of number of hours of TV watched in day and insert it here.
8. Use Minitab to calculate a 95% confidence interval for the difference in the mean number of hours of TV watched between male and female students. Interpret the confidence interval.

1. Perform an appropriate hypothesis test and include the output from Minitab here.

1. What is the value of the test statistic?
2. What is the P-value for this test?
3. State your decision and conclusion for the test using a significance level of α = 0.05
4. What assumptions are we making about the samples for our interpretation to be valid?