Math 1530 Capstone Project Spring 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 three 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. **AGE:** What is your age?
3. **AREA\_CODE:** What is your area code?
4. **ZODIAC:** What is your zodiac sign? (Aries, Taurus, Gemini, Cancer, Leo, Virgo, Libra, Scorpio, Sagittarius, Capricorn, Aquarius, Pisces)
5. **COMMUTE:** What is your average one-way commute time to campus in minutes? Enter 0 if you live on campus.
6. **WORK:** What is the average number of hours you plan to work per week at a paid job during the school semester? Enter 0 if you do not have a paid job during the school semester.
7. **PHONES\_OWNED:** How many cell phones have you owned?
8. **FIRST\_PHONE:** At what age did you get your first cell phone?
9. **TIME\_PHONE:** On average, how much time do you usually spend on your phone per day? (Round to the nearest minute)
10. **TEETH:** Have you ever had your teeth examined by a dentist? (Yes, No)
11. **CAVITY:** Approximately how many dental cavities have you ever had (either active or filled)?
12. **BRUSH:** How many times do you brush your teeth a day? (Zero, Once, Twice, Three or more)
13. **WATER\_TYPE:** What type of water do you drink? (Bottled, Tap, Filtered, Other)
14. **SOFT\_DRINK:** What type of carbonated soft drink do you prefer? (Regular, Diet, Zero, No preference, Do not drink carbonated soft drinks)
15. **GENERATION:** Are you a first generation college student (i.e., first generation in your family to attend college)? (Yes, No)
16. **OBSTACLES:** What obstacles have you experienced or do you anticipate that you many encounter in pursuit of your college education? Check all that may apply
17. **HOGWARTS:** If you could ask the Sorting Hat to place you in any House at Hogwarts, which House would you choose? (Gryffindor, Hufflepuff, Ravenclaw, Slytherin)
18. **MARIJUANA:** Do you think the use of marijuana should be legal in the USA?
19. **MARRIAGE:** Do you think same-sex marriage should be legal in the USA?

Overpopulation is one of the world’s impending problems. In the new Dan Brown (author of the Da Vinci Code) book Inferno, the following question was asked: “What if you were told that if you flipped a switch, you would kill half the world’s population?”

1. **FIRST\_SWITCH:** Would you flip the switch? (Yes, No)
2. **SECOND\_SWITCH:** “What if you were told that if you didn’t flip the switch right then, the human race would be extinct in the next hundred years. Would you flip the switch then?” (Yes, No)

Thefollowing questions were included in the Research Assignment for MATH1530 students at the beginning of the semester. The Research Assignment asked students to input some information of their plan for a summer sea cruise.

1. **DEPARTURE:** In what city is the port of departure for the cruise? Enter city only.
2. **TRANSPORTATION:** What means of transportation will you use to get to the port of departure? (Airplane, Car, Train, Bus, Other)
3. **TRANSPORTATION\_COST:** What is the cost of the air tickets(s), fares, or estimated cost of gas to arrive at the port of departure? (Round to nearest dollar)
4. **CRUISE\_LINE:** Which cruise line would you choose? (Royal Caribbean, Carnival Cruises, Princess Cruises, Disney Cruise Line, Celebrity Cruises, Norwegian Cruise Line, Other)
5. **DESTINATION:** What is your cruise destination? (Alaska, Asia & Africa, Australia & New Zealand, Canada & New England, Caribbean, Europe, Mexico, Hawaii & California Coastal, Panama Canal, South America, Tahiti & South Pacific, Other)
6. **COMPANION:** Who will travel with you on your cruise? (Alone, Significant Other, Friend(s), Family)
7. **TOTAL\_COMPANION:** How many people, including you, will go on the cruise together?
8. **DAYS:** How many days will the cruise last?
9. **CABIN:** What type of cabin will you choose? (Inside, Outside, Balcony, Suite)
10. **CRUISE\_COST:** What is the total cost of the cruise for your entire party, not including shore excursions and travel to port of departure? (Round to the nearest dollar.)
11. **ACTIVITY:** Which on-ship activity are you looking forward to the most? (Entertainment, Eating & Drinking, Meeting new people, Gambling, Games & Activities, Hanging out on the deck, Other)
12. **PORTS:** How many ports are included on your cruise ship’s itinerary?
13. **ACTIVITY\_COST:** What is your total estimated cost of planned activities during shore excursions? (Round to the nearest dollar)

 **A total of 661 students responded to the MATH1530 class survey. The data for 610 students were recorded in the Minitab worksheet MATH1530SPRING2014CapProjData.MTW. Note that there are some missing values, denoted by an asterisk (\*), in the data set.**

The Minitab worksheet is set up as follows:

|  |
| --- |
| C1: **ID** |
| C2: **GENDER** |
| C3: **AREA\_CODE** |
| C4: **ZODIAC** |
| C5: **PHONES\_OWNED** |
| C6: **FIRST\_PHONE** |
| C7: **HOGWARTS** |
| C8: **MARIJUANA** |
| C9: **MARRIAGE** |
| C10: **FIRST\_SWITCH** |
| C11: **SECOND\_SWITCH** |
| C12: **TRANSPORTATION** |
| C13: **TRANSPORTATION\_COST** |
| C14: **COMPANION** |
| C15**: DAYS** |
| C16: **CRUISE\_COST** |
| C17: **PORTS** |

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. **AREA\_CODE** Categorical Quantitative Neither
	2. **HOGWARTS** Categorical Quantitative Neither
	3. **FIRST\_PHONE** Categorical Quantitative Neither
	4. **DAYS** Categorical Quantitative Neither
	5. **COMPANION** Categorical Quantitative Neither

1. **Sampling.** In the survey data, the variable “**FIRST\_PHONE”** is the age that one received their first cell phone.

a. Type in the first 10 observations from the variable “**FIRST\_PHONE”** 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 |
| **FIRST\_PHONE** |  |  |  |  |  |  |  |  |  |  |

Obtain the mean age that one received their first cell phone for the first 10 observations.

The mean age is \_\_\_\_\_\_ years.

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 “**FIRST\_PHONE**” into the “From columns” box. Enter C18-C19 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 |  |  |  |  |  |  |  |  |  |  |
| **FIRST\_PHONE** |  |  |  |  |  |  |  |  |  |  |

 Obtain the mean age that one received their first cell phone of your random sample.

 The mean age is \_\_\_\_\_\_ years.

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

c. Obtain the mean age that one received their first cell phone for all 610 observations.

 The population mean age is \_\_\_\_\_\_ years.

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 age?

1. **If you are female then do this question. (Omit this page/problem if you are male.) TRANSPORTATION\_COST.** Question 24 from the survey asked, “What is the cost of the air tickets(s), fares, or estimated cost of gas to arrive at the port of departure?”
	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 **transportation cost**. 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 **transportation cost** for different modes of **transportation**. Insert the graph below.
	3. Describe the distributions of **transportation cost** for the five 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.) CRUISE\_COST.** Question 31 from the survey asked, “What is the total cost of the cruise for your entire party, not including shore excursions and travel to port of departure?”
	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 **cruise cost**. 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 **cruise cost** for **companion**. Insert the graph below.
	3. Describe the distributions of **cruise cost** for the four groups and compare them.
	4. Are there any outliers in each group? Justify your answer.

1. **Number of Ports versus Length of Cruise.** The number of ports visited during a sea cruise depends on several variables and one of them could be the length of the cruise. The MATH1530 class survey asked students to plan a sea cruise for the summer of 2014. Questions 29 and 33 asked students to input the number of days that the cruise would last (DAYS) and the number of ports that are included on the cruise ship’s itinerary (PORTS). Assume the respondents are an SRS of all ETSU students. We are interested in studying the relationship between the days of the cruise and the number of ports visited and whether knowing the length of the cruise would explain the number of ports visited on the cruise.

	1. Create an appropriate plot to display the relationship between **DAYS** and **PORTS**. 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 the number of ports that will be visited if a student would like to take a 14 day cruise.
	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.) Do males and females believe that same-sex marriage should be legal in the USA?** Question 19 from the survey asked students “Do you think same-sex marriage should be legal in the USA?” We want to check if there is a relationship between gender and the belief about same-sex marriage in the USA. 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 is a male and believes that same-sex marriage should be legal in the USA.
2. Find the probability that a student is female or they believe that same-sex marriage should be legal in the USA.
3. Find the probability that a student does not believe that same-sex marriage should be legal in the USA given that the student was a female.
4. Find the probability that a student was a female given that the student does not believe that same-sex marriage should be legal in the USA.
5. Carry out a test for the hypothesis that there is no relationship between gender and the belief about same-sex marriage in the USA 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.) Overpopulation.** Overpopulation is one of the world’s impending problems. In the new Dan Brown (author of the Da Vinci Code) book *Inferno*, the following question was asked: “What if you were told that if you flipped a switch, you would kill half the world’s population?” **Question 20** from the survey asked students “Would you flip the switch?” **Question 21** from the survey asked students ““What if you were told that if you didn’t flip the switch right then, the human race would be extinct in the next hundred years. Would you flip the switch then?” We want to check if there is a relationship between the choice of the flipping the switch based on question 20 and the choice of flipping the switch based on the secondary question 21. 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 chose to flip the switch based on question 20 and chose to flip the switch based on question 21.
2. Find the probability that a student does not choose to flip the switch based on question 20 or does not flip the switch based on question 21.
3. Find the probability that a student chooses to flip the switch based on question 20 given the chose to flip the switch based on question 21.
4. Find the probability that a student chooses to flip the switch based on question 21 given the chose to flip the switch based on question 20.
5. Carry out a test for the hypothesis that there is no relationship between the choice of the flipping the switch based on question 20 and the choice of flipping the switch based on the secondary question 21.ETSU students. Use 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. **Hogwarts School of Witchcraft and Wizardry.** Question 17 from the survey asked “If you could ask the Sorting Hat to place you in any House at Hogwarts, which House would you choose?” We suspect that the proportion of students that choose to be placed the Ravenclaw House at Hogwarts is different than 25%. Assume the students who took the class survey are from an SRS of ETSU students. Use the survey data to conduct a confidence interval and perform a hypothesis test. Answer the following questions.
7. Create pie graph of the variable Hogwarts and insert it here.
8. Based on our sample data, use Minitab to calculate a 90% confidence interval for the proportion of all ETSU students who would choose to be placed in Ravenclaw House at Hogwarts.
9. State the proper interpretation of the confidence interval you obtained in part (a).
10. Perform an appropriate hypothesis test to determine if the proportion of ETSU students that would choose Ravenclaw is not 25%. State the null and alternative hypotheses.

1. What is the sample proportion of the respondents who chose Ravenclaw?
2. Perform the test and include any output from Minitab here.
3. Which test statistic are you using? Report its value.
4. What is the P-value for the test?
5. State your decision and conclusion using a significance level of α=0.10.
6. Examine the data. What are the conditions for inference for this test? Are they met in this case? Explain.

1. **Number of Cell Phones Owned.** A marketing analyst wishes to know if males or females buy more cell phones 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 females buy more cell phones and thus, have owned more cell phones in their lifetime compared to males. Questions 7 from the survey asked students “How many cell phones have you owned?” Assume that the students who responded the survey are a SRS of all ETSU students. Is there good evidence to support the idea that female students at ETSU have owned more cell phones, on average, than male students?
2. Create an appropriate graph to display the distribution of number of cell phones owned and insert it here.
3. Use Minitab to calculate a 95% confidence interval for the difference in the mean number of cell phones owned between female and male 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?
5. **(BONUS PROBLEM)** **What’s your sign?** (For reasons unknown to anybody, this was thought to be a good pick-up line back in the 70’s.) For reasons known only to social scientists, the General Social Survey (GSS) regularly asks its subjects about their astrological sign. Question 4 from the survey asked, “What is your zodiac sign?” If births are spread uniformly across the year, we expect all 12 signs to be equally likely. Assume that the students who responded the survey are from an SRS of all ETSU students.
	1. Create an appropriate graph to display the distribution of astrological signs in our sample and comment on the results. Insert the graph here.
	2. Is there good evidence that the zodiac signs of ETSU students are not equally likely? Perform a test of hypotheses using a significance level of α = 0.05. Include the appropriate test output and accompanying graphs from Minitab here.
	What is your conclusion? If you concluded that not all signs are equally likely, which signs appear to be furthest from expectations based on the null hypothesis?