MATH 1530 CAPSTONE TECHNOLOGY PROJECT SPRING 2016

**Problem 1: Identify Variable Type.** One of these is a variable that is categorical and one is quantitative. Consider the different graphs that correspond to each variable type. Use Minitab to create ***two*** *different* graphs appropriate for **each** variable’s type. That is, there is a total of four graphs. EXTRA CREDIT if you can resize to fit all of the graphs on one page.

**Evening\_Time Books**

**Problem 2: Sampling.** In the survey data, the variable “AGE”is the current age reported by each student.   
  
a. Type the first 10 observations from the column representing the variable AGEinto the table below,and use this as your sample data for part (b). Then calculate the mean age of these first 10 observations and report the value below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **AGE (yrs)** |  |  |  |  |  |  |  |  |  |  |

b. The mean age of the first 10 students isyears. (Type the value into the space provided.)

c. Identify the type of sampling method you have just used:

d. Next, select a random sample of size n = 10 (Go to Calc > Random Data > Sample from Columns). Type the number 10 in the “Number of rows to Sample” slot. Enter the variable “ID” and “AGE” into the “From columns” slot. Enter C17-C18 into the “Store samples in” slot. Record the data for your sample in the table below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **ID** |  |  |  |  |  |  |  |  |  |  |
| **AGE (yrs)** |  |  |  |  |  |  |  |  |  |  |

e. Calculate and report the mean age for your random sample of 10 students. The sample mean age is

years.

f. Identify the type of sampling method you have just used:

g. REPEAT the random sample selection process three more times. Calculate and report the mean age for each random sample of 10 students.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **ID** |  |  |  |  |  |  |  |  |  |  |
| **AGE (yrs)** |  |  |  |  |  |  |  |  |  |  |

ii) The sample mean age is **\_\_\_\_\_\_** years.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **ID** |  |  |  |  |  |  |  |  |  |  |
| **AGE (yrs)** |  |  |  |  |  |  |  |  |  |  |

iii) The sample mean age is **­­­­\_\_\_\_\_\_\_\_** years.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **ID** |  |  |  |  |  |  |  |  |  |  |
| **AGE (yrs)** |  |  |  |  |  |  |  |  |  |  |

iv) The sample mean age is **­­­\_\_\_\_\_\_\_\_** years.

h. Suppose we think of *all* the students who responded to the survey as a *population* for the purposes of this problem. In that case, the *population mean* age is 21.25. Discuss (two or more complete sentences) the **differences and similarities** between 21.25 and the answers you got in (b), (e), and ii), iii), and iv).

**Problem 3(e): 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.)**

Question 11 of the SPRING 2016 survey asked students, “What is the fastest you have ever driven a car? (in mph)”

a. Create an appropriate graph to display the *distribution* of the variable called **Fastest\_Speed** and insert it here.

b. Which of the following best describes the shape of the distribution? Underline (or highlight) your answer.

Skewed left Symmetric Skewed right

c. Using Minitab, calculate the basic statistics for the data collected on **Fastest\_Speed**. Copy and paste all of the Minitab output here.

d. Choose statistics that are appropriate for the shape of the distribution to describe the center and spread of **Fastest\_Speed**.

Which statistic will you use to describe the center of the distribution? (Type name of the statistic here.)

e. What is the value of that statistic? (Type value here.)

f. Which statistic(s) will you use to describe the spread of the distribution?

g. What is (are) the value(s) of that (those) statistic(s)?

h. Look up the IQR rule on page 55 in our textbook. Are there any outliers in this distribution? If so, what are their values? How many are there? Justify your answer.

**Problem 3(o): 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.)**

Question 8 of the SPRING 2016 survey asked students, “How much do you believe minimum wage should be? (in US dollars)”

a. Create an appropriate graph to display the *distribution* of the variable called **Min\_Wage** and insert it here.

b. Which of the following best describes the shape of the distribution? Underline (or highlight) your answer.

Skewed left Symmetric Skewed right

c. Using Minitab, calculate the basic statistics for the data collected on **Min\_Wage** and copy & paste the Minitab output here.

d. Choose statistics that are appropriate for the shape of the distribution to describe the center and spread of **Min\_Wage**.

i) Which statistic will you use to describe the center of the distribution? (Type name of the statistic here.)

ii) What is the value of that statistic? (Type value here.)

iii) Which statistic(s) will you use to describe the spread of the distribution?

iv) What is (are) the value(s) of that (those) statistic(s)?

v) Look up the IQR rule on page 55 in our textbook. Are there any outliers in this distribution? If so, what are their values? How many are there? Justify your answer.

**Problem 4: CAR\_AGE vs. MPG.** On the SPRING 2016 Math 1530 survey, questions 13 and 14 asked students to state the age of their car (in years) and the miles per gallon (mpg) for that car. We are interested in seeing whether we can use the age of the car to predict the mpg of the car.

a. Create an appropriate graph to display the relationship between **CAR\_AGE** and **MPG**. Insert it here.

b. 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.

c. Describe the *form* of the relationship between **CAR\_AGE** and **MPG.**

d. Report the value of the correlation between this pair of variables? **r =**

e. Based on the information displayed in the graph and the correlation you just reported, how would you describe the *strength* of the association?

f. Using Minitab, obtain the equation for the least squares regression of MPG on CAR\_AGE. Copy & paste the output here.  
  
g. Interpret the value of the slope in the least squares regression equation you found in part (f).

h. Use the regression equation in part (f) to predict the mpg for a car that is 10 years old. (Show your math.)

**Predicted mpg =**

i. How well does the regression equation fit the data? Explain. Justify your answer with appropriate plot(s) and summary statistics.

**Problem 5 (T): Flip a fair coin. If it lands on tails do this problem (Omit this page/problem AND DO PROBLEM 5(H) if it lands on heads.)**

**PRAYER\_CLASSROOM AND AGE\_GROUP** Question 5 from the SPRING 2016 Math 1530 survey asked students “Do you favor or oppose daily prayer in the classroom?” and Question 2 from the SPRING 2016 Math 1530 survey asked students “What is your age?” The variable **AGE** was split into two groups: ages “12 to 21” and “Over 21.” These two groups formed the variable **AGE\_GROUP.** We want to check if there is a relationship between PRAYER\_CLASSROOM and AGE\_GROUP among ETSU students. Assume the students who took the (SPRING 2016 Math 1530) class survey are from an SRS of ETSU students.

a. Create an appropriate **graph** to display the relationship between PRAYER\_CLASSROOM and AGE\_GROUP. Insert your graph here.

b. Create an appropriate two-way table to summarize the data. Insert your table here.

**SUPPOSE WE SELECT ONE STUDENT AT RANDOM:** (Calculate the following probabilities and show your work.)

c. What is the probability that this student is either “Over 21” *or* favors prayer in the classroom?

**P =**

d. What is the probability that this student is aged “12 to 21” given that the student opposes prayer in the classroom?

**P =**

e. What is the probability that this student opposes prayer in the classroom given that the student is aged “12 to 21”?

**P =**

f. Do you think there may be an association between PRAYER\_CLASSROOM and AGE\_GROUP? Why or why not? Explain your reasoning based on what you see in your graph.

**Problem 5 (H): Flip a fair coin. If it lands on heads do this problem (Omit this page/problem AND DO PROBLEM 5(T) if it lands on tails.)**

**POLITICAL\_AFFLILATION AND RELIGION\_ID** Question 6 from the SPRING 2016 Math 1530 survey asked students “What political party do you identify with?” and Question 10 from the SPRING 2016 Math 1530 survey asked students “What is your religious identification?” We want to check if there is a relationship between POLITICAL\_AFFLILATION AND RELIGION\_ID among ETSU students. Assume the students who took the (SPRING 2016 Math 1530) class survey are from an SRS of ETSU students.

a. Create an appropriate **graph** to display the relationship between POLITICAL\_AFFLILATION AND RELIGION\_ID. Insert your graph here.

b. Create an appropriate two-way table to summarize the data. Insert your table here.

**SUPPOSE WE SELECT ONE STUDENT AT RANDOM:** (Calculate the following probabilities and show your work.)

c. What is the probability that this student’s political affiliation is Independent or identifies with a Non-Christian religion?

**P =**

d. What is the probability that this student’s political affiliation is Republican given the student identifies with a Christian religion?

**P =**

e. What is the probability that this student identifies with a Christian religion given the student’s political affiliation is Republican?

**P =**

f. Do you think there may be an association between POLITICAL\_AFFLILATION AND RELIGION\_ID? Why or why not? Explain your reasoning based on what you see in your graph.

**Problem 6:** In 2015, the Pew Research Center found that American adults (aged 18+) read an average of 12 books per year (http://www.pewresearch.org/fact-tank/2015/10/19/slightly-fewer-americans-are-reading-print-books-new-survey-finds/ft\_15-10-09\_books\_averageread/). Do ETSU students read, on average, 12 books per year?

a. Create a suitable graph to display the distribution of BOOKS reported by our sample of college students and insert it here.

b. Describe the distribution shown in your graph.

c. Perform a test of significance to see if ETSU college students read, on average, the same amount of books in a year as American adults. If this is true, then the average BOOKS reported by ETSU students should be 12. Thus,

**Ho**: μ = 12 books

Write the correct alternative hypothesis for the test.

d. Use Minitab to perform the appropriate test. Copy and paste the output for the test here.

e. What is the name of your test statistic and what is its value?

f. What is the P-value for the test? **P =**

g. State your decision regarding the hypotheses being tested.

h. State your conclusion. USE COMPLETE SENTENCES.

i. Is the P-value valid in this case? What assumptions are you making in order to carry out this test?

**Bonus Problem:** Question 4 on the SPRING 2016 Math 1530 asked, “Have you, yourself, ever smoked cigarettes in the past week?” Gallup reported that the national proportion of adults ages 18 to 29 that have smoked cigarettes in the past week is 20% (<http://www.gallup.com/poll/187592/young-adults-cigarette-down-sharply.aspx?g_source=CATEGORY_WELLBEING&g_medium=topic&g_campaign=tiles>). Note the results are similar for ages 30 to 64. Is the same true for the population of all U.S. college/university students?

a. Create an appropriate graph to display the distribution of SMOKING and insert it here.

b. How many of the students surveyed said “yes?”

c. What proportion of our sample said “yes?”

d. Assume (for the purpose of this problem) that we may treat the SPRING 2016 sample of Math-1530 students as a simple random sample drawn from the population of all U.S. college/university students. Use Minitab to calculate a 95% confidence interval for the proportion of students in the population who chose “yes” to the survey question (based on our sample data). Copy and paste the Minitab output here.

e. Interpret the confidence interval you reported in part (d).

f. What do you think? Do our results contradict the results obtained from Gallup or do they appear to agree with it? EXPLAIN.