# MATH 1530 CAPSTONE TECHNOLOGY PROJECT SPRING 2017

Problem 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/highlight the best option.

### **a.** **CHILDREN** Categorical Quantitative Neither

State an appropriate plot for this variable: **\_\_\_\_\_\_\_\_\_\_\_\_**

**b.** **RELATIONS** Categorical Quantitative Neither

State an appropriate plot for this variable: **\_\_\_\_\_\_\_\_\_\_\_\_**

**c.** **CLASS** Categorical Quantitative Neither

State an appropriate plot for this variable: **\_\_\_\_\_\_\_\_\_\_\_\_**

Problem 2: Sampling. In the survey data, the variable “**NUMBER**”is the favorite number between 0 and 50 for each student.

**a.** Type the last 10 observations from the column representing the variable **NUMBER** into the table below,and use this as your sample data for part (b). Then calculate the mean favorite number of these last 10 observations and report the value below.

| **N** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **NUMBER** |  |  |  |  |  |  |  |  |  |  |

**b.** The mean favorite of the last 10 students is **\_\_\_\_\_\_**. (Type the value into the space provided.)

**c.** 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 “**NUMBER**” 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** |  |  |  |  |  |  |  |  |  |  |
| **NUMBER** |  |  |  |  |  |  |  |  |  |  |

**d.** Calculate and report the mean favorite number for your random sample of 10 students. The sample mean favorite number is **\_\_\_\_\_\_\_**.

e. 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* favorite number is 16.454. Discuss (two or more complete sentences) the **differences and similarities** between 16.454 and the answers you got in (b) and (d).

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 1 of the SPRING 2017 survey asked students, “What is your age (in years)?”

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

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

 Skewed left Uniform Skewed right Bimodal Symmetric

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

Choose statistics that are appropriate for the shape of the distribution to describe the center and spread of **AGE**.

**d.** Which statistic will you use to describe the center of the distribution? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**e.** In one or two sentences, describe why this statistic was chosen.

**f.** What is the value of that statistic? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**g.** Which statistic(s) will you use to describe the spread of the distribution? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**h.** What is (are) the value(s) of that (those) statistic(s)? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**i.** Are there any outliers in this distribution? Justify your answer using the IQR rule or an appropriate plot.

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 4 of the SPRING 2017 survey asked students, “Approximately, how many phone calls do you make per day?”

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

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

 Skewed left Uniform Skewed right Bimodal Symmetric

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

Choose statistics that are appropriate for the shape of the distribution to describe the center and spread of **CALLS**.

**d.** Which statistic will you use to describe the center of the distribution? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**e.** In one or two sentences, describe why this statistic was chosen.

**f.** What is the value of that statistic? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**g.** Which statistic(s) will you use to describe the spread of the distribution? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**h.** What is (are) the value(s) of that (those) statistic(s)? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**i.** Are there any outliers in this distribution? Justify your answer using the IQR rule or an appropriate plot.

Problem 4: CLASS versus AGE**.** Question 1 of the survey asked students, “What is your age (in years)?” Question 2 of the survey asked students, “What is your classification in college? (Freshman/first-year, Sophomore, Junior, Senior)”

**a.** Create a suitable graph to display the *distribution* of **CLASS** and insert it here.

**b.** What is the mode of this distribution? (Please underline one option.)

 Freshman/first-year Sophomore Junior Senior

**c.** Create a side-by-side boxplot to display the age of students for the different levels of **CLASS**. (Go to Graph > Boxplot > One Y with Groups > OK. Select **AGE** for the “Graph variables” slot and **CLASS** for the “Categorical variables for grouping” slot.) Insert your graph here.

Use the side-by-side boxplot found in part (c) to answer the following questions.

**d.** Which class has the oldest student? **\_\_\_\_\_\_\_\_\_\_\_\_**

**e.** Which class has the youngest student? **\_\_\_\_\_\_\_\_\_\_\_\_**

**f**. Which age-group has the largest IQR? **\_\_\_\_\_\_\_\_\_\_\_\_**

Problem 5: MILES vs. GAS. On the SPRING 2017 Math 1530 survey, question 12 asked students, “Approximately, how many miles do you live from campus? (Enter 0 if you live on campus)” and question 13 asked students, “Approximately, how do you spend on gas (in U.S. dollars) a week?” We are interested in seeing whether we can use the number of miles to predict the amount of dollars spent on gas for a week.

**a.** Create an appropriate graph to display the relationship between **MILES** and **GAS**. 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 **MILES** and **GAS.**

**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 **MILES** on **GAS**. 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 amount of dollars spent on gas for a week for a student that lives 5 miles from campus. (Show your math.)

**Predicted amount of dollars =**

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

Problem 6 (T): **Flip a fair coin. If it lands on tails do this problem (Omit this page/problem AND DO PROBLEM 6(H) if it lands on heads.)**

**LGBT AND AGE GROUP** Question 9 from the SPRING 2017 Math 1530 survey asked students “In the U.S., more Americans are identifying as LGBT. Do you, personally, identify as lesbian, gay, bisexual, or transgender? (Yes, No)” and Question 1 of the survey asked students, “What is your age (in years)?” This variable was divided into three age groups: Ages “12 to 20”, “21 to 25”, and “Over 25”. We named this variable **AGE GROUP**. We want to check if there is a relationship between **LGBT AND AGE GROUP** among ETSU students. Assume the students who took the (SPRING 2017 Math 1530) class survey are from an SRS of ETSU students.

**a.** Create an appropriate **graph** to display the relationship between **LGBT** and **AGE GROUP**. Insert your graph here.

**b.** Create an appropriate two-way table to summarize the data. Insert your table here. (**IN MINITAB: STAT 🡪 TABLES 🡪 CROSS TABULATION AND CHI-SQUARE. Make sure to select “Options” and click “No variables” under the *Display missing values for*”).**

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

**c.** What is the probability that this student identifies as LGBT *and* is aged 12 to 20?

**P = \_\_\_\_\_\_\_\_\_\_\_\_**

**d.** What is the probability that this student identifies as LGBT *or* is aged 12 to 20?

**P = \_\_\_\_\_\_\_\_\_\_\_\_**

**e.** What is the probability that this student does not identify as LGBT given that the student is aged over 25?

**P = \_\_\_\_\_\_\_\_\_\_\_\_**

**f.** What is the probability that this student is aged over 25 given that the student does not identify as LGBT?

**P = \_\_\_\_\_\_\_\_\_\_\_\_**

Problem 6 (H): **Flip a fair coin. If it lands on heads do this problem (Omit this page/problem AND DO PROBLEM 6(T) if it lands on tails.)**

**GUNS AND TERRORISM** Question 10 from the SPRING 2017 Math 1530 survey asked students “Are you satisfied with America’s law or policies on guns? (Yes, No)” and Question 11 from the SPRING 2017 Math 1530 survey asked students “Are you satisfied with America’s security with terrorism? (Yes, No)” We want to check if there is a relationship between **GUNS** and **TERRORISM** among ETSU students. Assume the students who took the (SPRING 2017 Math 1530) class survey are from an SRS of ETSU students.

**a.** Create an appropriate **graph** to display the relationship between **GUNS** and **TERRORISM**. Insert your graph here.

**b.** Create an appropriate two-way table to summarize the data. Insert your table here. (**IN MINITAB: STAT 🡪 TABLES 🡪 CROSS TABULATION AND CHI-SQUARE. Make sure to select “Options” and click “No variables” under the *Display missing values for*”).**

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

**c.** What is the probability that this student is satisfied with America’s law or policies on guns *and* the student is satisfied with America’s security with terrorism?

**P = \_\_\_\_\_\_\_\_\_\_\_\_**

**d.** What is the probability that this student is satisfied with America’s law or policies on guns *or* the student is satisfied with America’s security with terrorism?

**P = \_\_\_\_\_\_\_\_\_\_\_\_**

**e.** What is the probability that this student is satisfied with America’s law or policies on guns given that the student is not satisfied with America’s security with terrorism?

**P = \_\_\_\_\_\_\_\_\_\_\_\_**

**f.** What is the probability that this student is not satisfied with America’s security with terrorism given that this student is satisfied with America’s law or policies on guns?

**P = \_\_\_\_\_\_\_\_\_\_\_\_**

Problem 7: In 2013, Gallop found that the ideal number of children Americans want is 2.6 children per family. (<http://www.gallup.com/poll/164618/desire-children-norm.aspx>). Question 3 of the survey asked students, “What is your ideal number of children? ” A professor feels that this number may be lower for college students. Is ETSU student’s ideal number of children, on average, less than 2.6 children?

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

Perform a test of significance to see if ETSU college student’s ideal number of children, on average, is lower than the 2.6 children reported by Gallop using *α* = 0.05.

**b.** Write the correct null and alternative hypothesis for the test: **\_\_\_\_\_\_\_**

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

**d.** What is the name of your test statistic and what is its value? **\_\_\_\_\_\_\_**

**e.** What is the P-value for the test? **P = \_\_\_\_\_\_\_**

**f.** State your decision regarding the hypotheses being tested.

**g.** State your conclusion. USE COMPLETE SENTENCES.

**h.** Is the P-value valid in this case? **\_\_\_\_\_\_\_**

**i.** What assumptions are you making in order to carry out this test?

Bonus Problem:Question 13 on the SPRING 2017 Math 1530 asked, “What is your religious identification? (Protestant/Other Christian, Catholic, Mormon, Jewish, Muslim, Other Non-Christian religion, None/Atheist/Agnostic)” The Gallup took a survey of U.S. adults in December 2016 and reported that and reported that 18.2% of U.S. adults said their religion identification was None/Atheist/Agnostic (<http://www.gallup.com/poll/200186/five-key-findings-religion.aspx?g_source=Religion&g_medium=newsfeed&g_campaign=tiles>). Is the same true for the population of all U.S. college/university students?

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

**b.** How many of the students surveyed said “None/Atheist/Agnostic?”

**c.** What proportion of our sample said “None/Atheist/Agnostic?”

**d.** Assume (for the purpose of this problem) that we may treat the SPRING 2017 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 “None/Atheist/Agnostic” 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 survey by Gallup or do they appear to agree with it? EXPLAIN.