

Statistics Competency Exam Spring 2019 Part II

Name _____

R and/or Minitab along with a calculator may be used for this entire section. Once you have completed this section, e-mail Part II to Dr. Price at pricejr@etsu.edu, Dr. Lewis at lewiscn2@etsu.edu, and Dr. Hendrickson at hendricksonj@etsu.edu. Please make sure you e-mail it to all three professors.

The data labeled SP16 (that I have emailed you) is needed to answer all questions.

The data in the csv file is based on a student survey that asked the following questions:

- **GENDER:** What gender do you identify with? (Female, Male, Neither, Both)
- **AGE:** What is your age (in years)?
- **EVENING_TIME:** What is your favorite way of spending an evening? (Staying at home, Watching TV, Resting/Relaxing, Reading, Visiting with friends, Other)
- **SMOKING:** Have you, yourself, ever smoked cigarettes in the past week? (Yes, No)
- **PRAYER_CLASSROOM:** Do you favor or oppose daily prayer in the classroom? (Favor, Oppose)
- **POLITICAL_AFFILIATION:** What political party do you identify with? (Republican, Democrat, Independent, Other)
- **BOOKS:** How many books (not including textbooks or other books for class) did you read in 2015?
- **MIN_WAGE:** How much do you believe minimum wage should be? (in US dollars)
- **TEXT:** How many different people do you text on a normal day?
- **RELIGION_ID:** What is your religious identification? (Christian Religion, Non-Christian Religion, None)
- **FASTEST_SPEED:** What is the fastest you have ever driven a car? (in mph)
- **SPEED_TICKETS:** How many speeding tickets have you got since you started driving?
- **CAR_AGE:** How old is your car? (in years)
- **MPG:** How many miles per gallon (mpg) does your car get? (in mpg)

1. **MIN_WAGE:** Question 8 from the survey asked students, “How much do you believe minimum wage should be? (in US dollars)”
 - (A) Create an appropriate display for this variable and insert it here.
 - (B) Describe the shape of the distribution.
 - (C) Calculate numerical measures 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?
 - ii. What is the value of that statistic?
 - 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)?
 - (D) Are there any outliers in this distribution? Justify your answer using the IQR rule or an appropriate plot.
 - (E) Create a side-by-side boxplot to display **MIN_WAGE** for the different levels of **POLITICAL_AFFILIATION**. Insert your graph here.
 - (F) Describe the distributions of **MIN_WAGE** for the different levels of **POLITICAL_AFFILIATION** and compare them.

2. **POLITICAL_AFFILIATION:** Question 6 from the survey asked students “What political party do you identify with? (Republican, Democrat, Independent, Other)” As of October 2017, Gallup polling found that 31% of Americans identified as Democrat. Is the same true for the population of all U.S. college/university students?
 - (A) Create an appropriate graph to display the distribution of **POLITICAL_AFFILIATION** and insert it here.
 - (B) What is the mode of this distribution?

- (C) How many students surveyed said “Democrat” was the political party they identified with?
- (D) What proportion of our sample said “Democrat” was the political party they identified with?
- (E) 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. Calculate a 95% confidence interval for the proportion of students in the population who identify with the “Democrat” party (based on our sample data). Copy and paste the output here.
- (F) Interpret the confidence interval you reported above.
- (G) What do you think? Is the percentage of Americans that identify as Democrat, 31%, the same for the population of all U.S. college/university students? Explain.
3. **AGE and MIN_WAGE:** Question 2 from the survey asked students, “What is your age (in years)?” and question 8 from the survey asked students, “How much do you believe minimum wage should be? (in US dollars)” We are interested in seeing whether we can use age to predict how much a student believes minimum wage should be. Assume the respondents are an SRS of all ETSU students.
- (A) Create an appropriate plot to display the relationship between **AGE** and **MIN_WAGE** and insert it here.
- (B) What is the correlation between the pair of variables?
- (C) Based on the information displayed in the graph and the correlation you reported, how would you describe the association between **AGE** and **MIN_WAGE**?
- (D) Obtain the least squares regression equation for the pair of variables. Insert it here.
- (E) Interpret the value of the slope in the least squares regression equation you found from part (c).
- (F) What percent of the variation in **MIN_WAGE** is accounted for by its linear relationship with **AGE**?
- (G) What is the predicted minimum wage for a student who is 20 years old?
4. **BOOKS:** Question 7 from the survey asked students, “How many books (not including textbooks or other books for class) did you read in 2015?” According to Pew Research, Americans read an average of 12 books per year. On average, is the number of books read by ETSU students the same as reported by the research?
- (A) Create an appropriate graph to display the distribution of number of books read in 2015. Insert it here.
- (B) Calculate a 99% confidence interval for the mean number of books read by ETSU students and interpret this interval.
- (C) Using $\alpha = 0.01$, perform a test of significance to see if, on average, the number of books read by ETSU students in 2015 is the same as the average reported by Pew Research. Insert the output here.
- (D) Write the correct null and alternative hypotheses for the test.
- (E) Perform an appropriate hypothesis test and include the output here.
- (F) What is the P-value for this test?
- (G) State your decision and conclusion for the test.
- (H) What assumptions are you making in order to carry out the test? Is the P-value valid in this case?
5. **POLITICAL_AFFILIATION AND EVENING_TIME:** Question 6 from the survey asked students “What political party do you identify with? (Republican, Democrat, Independent, Other)” and Question 3 from the survey asked students “What is your favorite way of spending an evening? (Staying at home, Watching TV, Resting/Relaxing, Reading, Visiting with friends, Other)” We want to check if there is a relationship between political affiliation and evening time. Assume the students who took the survey are from an SRS of ETSU students.
- (A) Create an appropriate two-way table to summarize the data between **POLITICAL_AFFILIATION** and **EVENING_TIME** and insert it here.
- (B) Find the probability that a randomly selected student identifies with the Republican party and whose favorite way of spending an evening is staying at home.
- (C) Find the probability of a randomly selected student whose favorite way of spending an evening is reading or identifies as an Independent for political affiliation.

- (D) Find the probability that a randomly selected student identifies with the Democratic party given their favorite way of spending an evening is watching TV.
- (E) Carry out a test for the hypothesis that there is no relationship between **POLITICAL_AFFILIATION** and **EVENING_TIME** for ETSU students. Use a significance level of $\alpha = 0.05$.
- State the null and alternative hypothesis.
 - Perform the test and include any output here.
 - Which test statistic are you using and what is its value?
 - State your decision and conclusion for the test.
 - Examine the data. Are the conditions for inference in part (ii) violated? Explain.