

STAT 250 Fall 2018 Data Analysis Assignment 3

Your submitted document should include the following items. Points will be deducted if the following are not included.

1. Type your **Name, STAT 250** with your correct section number (e.g. STAT 250-xxx) right justified and **Data Analysis Assignment #3** centered on the top of page 1 of your document.
2. Number your pages across your entire solutions document.
3. Your document should include the **ANSWERS ONLY** with each answer labeled by its corresponding number and subpart. Keep the answers in order. Do NOT include the questions in your submitted document.
4. Generate all requested graphs and tables using **StatCrunch**.
5. Upload your document onto Blackboard as a Word or pdf document using the link provided by your instructor.
6. You may not work with other individuals on this assignment. It is an honor code violation if you do.

Elements of good technical writing:

Use complete and coherent sentences to answer the questions.

Graphs must be appropriately titled and should refer to the context of the question.

Graphical displays must include labels with units if appropriate for each axis.

Units should always be included when referring to numerical values.

When making a comparison you must use comparative language, such as “greater than”, “less than”, or “about the same as.”

Ensure that all graphs and tables appear on one page and are not split across two pages.

Type all mathematical calculations when directed to compute an answer ‘by-hand.’

Pictures of actual handwritten work are not accepted.

When writing mathematical expressions into your document you may use either an equation editor or common shortcuts such as: \sqrt{x} can be written as sqrt(x), \hat{p} can be written as p-hat, \bar{x} can be written as x-bar.

Problem 1: Confidence Interval for Percentage of B's.

The data set “**STAT 250 Final Exam Scores**” contains a random sample of 269 STAT 250 students’ final exam scores (maximum of 80) collected over the past two years. Answer the following questions using this data set.

- a) What proportion of students in our sample earned B's on the final exam? A letter grade of B is obtained with a score of between 64 and 71 inclusive. Hint: You can do this many ways, but in StatCrunch, go to Data → Row Selection → Interactive Tools. In the slider selectors box, click the variable “Scores” into the variable box. Then click compute. Use the slider to obtain the count by looking at the “# rows selected” presented in the first line of the box. Show your work (i.e. describe the method you used to obtain the number of B's) and express this value as a proportion rounded to four decimal places.
- b) Using the sample proportion obtained in (a), construct a 90% confidence interval to estimate the population proportion of students who earned a B on the final exam. Please do this “by hand” using the formula and showing your work (please type your work, no images accepted here). Assume all Central Limit Theorem conditions hold. Round your confidence limits to four decimal places.
- c) Verify your result from part (b) using Stat → Proportions Stats → One Sample → With Summary. Inside the box, select confidence interval and click Compute! Copy and paste your StatCrunch result in your document.
- d) Interpret the StatCrunch confidence interval in part (c) in one sentence using the context of the question.
- e) Did this confidence interval capture the true population proportion $p = 0.21$ given in Problem 4 of Data Analysis 2. Answer this question in one sentence.
- f) Use the Confidence Interval applet (for a Proportion) in StatCrunch to simulate constructing one thousand 90% confidence intervals using $p = 0.21$ and $n = 269$. Once the window is open, **click reset and select (or click) 1000 intervals**. Copy and paste your image into your document.

Box 1: Enter the given population proportion, 0.21

Box 2: Enter the given confidence level 0.90

Box 3: Enter the given sample size, n=269

- g) Compare the “Prop. contained” value from part (f) to the confidence level associated with the simulation.
- h) Write a long-run interpretation for your confidence interval method in one sentence.

Problem 2: Opinion on Sports Betting

About one year ago, polling numbers began to show that public opinion about legalizing sports betting had changed. For the first time, a majority of Americans supported making wagering on professional sports legal. A researcher in the state of Virginia is currently interested in this topic. She wants to test the claim that more than 54% of Virginians would support this legalization. To test this claim she collected a random sample of 382 Virginia adults and then asked whether they support or oppose legalized professional sports betting. The responses (0 = Oppose and 1 = Support) are found in StatCrunch in a data set called “**Virginia Sports Betting Survey.**”

- a) Obtain the sample proportion of individuals who said “Support” using Stat → Tables → Frequency in StatCrunch. Only the value of the sample proportion is needed in your answer. Present this sample proportion as a fraction or a decimal rounded to 4 decimal places.
- b) Using $\alpha = 0.05$, is there sufficient evidence to conclude that more than 54% of Virginia adults support legalized sports betting? Conduct a full hypothesis test by following the steps below.
- i. Define the population parameter in one sentence.
 - ii. State the null and alternative hypotheses using correct notation.
 - iii. State the significance level for this problem.
 - iv. Check the three conditions of the Central Limit Theorem that allow you to use the One-Proportion z-Test using one complete sentence for each condition. Show work for the numerical calculation.

- v. Calculate the test statistic “by-hand.” Show the work necessary to obtain the value by typing your work and provide the resulting test statistic. Do not round while doing the calculation. Then, round the test statistic to two decimal places after you complete the calculation.
- vi. Calculate the p -value using the standard Normal table and provide the answer. Use four decimal places for the p -value.
- vii. State whether you reject or do not reject the null hypothesis and the reason for your decision in one sentence (compare your p -value to the significance level to do this).
- viii. State your conclusion in context of the problem (i.e. interpret your results and/or answer the question being posed) in one or two complete sentences.
- ix. Use StatCrunch (Stat \rightarrow Proportion Stats \rightarrow One Sample \rightarrow with Data) to verify your test statistic and p -value. Copy and paste this box into your document.

Problem 3: Electric Cars

According to both a Consumer Report’s and a AAA survey, about 1 in 5 Americans will buy an electric car as their next vehicle. To test this claim, an independent surveyor obtained records for 2018 car sales. The surveyor generated a random sample of 175 car sales and found that 28 of these new car purchases were of electric cars.

- a) Check the three conditions of the Central Limit Theorem that allow you to use the One-Proportion Confidence Interval using one complete sentence for each condition. Show work for the numerical calculation.
- b) Construct a 99% confidence interval to estimate the population proportion of Americans new car purchases that were electric. Calculate this “by hand” using the formula and showing your work (please type your work, no images accepted here). Round your confidence limits to four decimals.
- c) Verify your result in part (a) using Stat \rightarrow Proportions Stats \rightarrow One Sample \rightarrow With Summary. Copy and paste your StatCrunch result in your document as well.
- d) Using $\alpha = 0.01$, is there sufficient evidence to conclude that the proportion of Americans who purchased an electric car is different from 0.2? Conduct a full hypothesis test by following the steps below. Enter an answer for each of these steps in your document.
 - i. Define the population parameter in one sentence.
 - ii. State the null and alternative hypotheses using correct notation.
 - iii. State the significance level for this problem.
 - iv. Calculate the test statistic “by-hand.” Show the work necessary to obtain the value by typing your work and provide the resulting test statistic. Do not round during the calculation. Then, round the test statistic to two decimal places after you complete the calculation.

- v. Calculate the p -value using the standard Normal table and provide the answer. Use four decimal places for the p -value.
 - vi. State whether you reject or do not reject the null hypothesis and the reason for your decision in one sentence (compare your p -value to the significance level to do this).
 - vii. State your conclusion in context of the problem (i.e. interpret your results and/or answer the question being posed) in one or two complete sentences.
 - viii. Use StatCrunch (Stat \rightarrow Proportion Stats \rightarrow One Sample \rightarrow with Summary) to verify your test statistic and p -value. Copy and paste this box into your document.
- e) Explain the connection between the confidence interval and the hypothesis test in this problem (discuss this in relation to the decision made from your hypothesis test and connect it to the confidence interval you constructed in part (b)). Answer this question in one to two sentences.

Problem 4: House Prices

Use the “**Fairfax City Home Sales**” dataset for parts of this problem.

- a) Use StatCrunch to construct an appropriately titled and labeled relative frequency histogram of Fairfax home closing prices stored in the “Price” variable. Copy your histogram into your document.
- b) What is the shape of this distribution? Answer this question in one complete sentence.
- c) Assuming the population has a similar shape as the sample with population mean \$510,000 and population standard deviation \$145,000; calculate the probability that in a random sample of size 10, the mean of the sample will be greater than \$600,000. You may assume a random sample was taken and the sample came from a big population. However, be sure to check the central limit theorem condition of a large sample size before completing this problem using one complete sentence. If this condition is not met, you cannot complete the problem.
- d) Assuming the population has a similar shape as the sample with population mean \$510,000 and population standard deviation \$145,000; calculate the probability that in a random sample of size 36, the mean of the sample will be greater than \$600,000. You may assume a random sample was taken and the sample came from a big population. However, be sure to check the central limit theorem condition of a large sample size before completing this problem using one complete sentence. If this condition is not met, you cannot complete the problem.