Are You Above Average?

Activity to enhance understanding of random variation and using probability distributions as models for natural phenomena

At least 1 week before using this activity, notify students that daily screen time usage on cell phones will be used in a class activity. Suggest that students become familiar with how to get that information from their phones (it may require downloading a free app for the Android platform). **Do not** have students monitor their actual screen time usage; just make sure the data is being collected by their phones.

If needed, a 30-day free trial of Minitab software is available for a one-time use per computer at [www.minitab.com](http://www.minitab.com). The videos that accompany this module were produce with Minitab 18 software. Minitab has just released Minitab 19 software; there may be slight variations in the screens between Minitab 18 and Minitab 19.

Learning Outcomes:

1. Identifies population of interest with specific enumeration of who is included in the population, determines correct sample size needed to estimate a mean with a specified confidence interval width, and selects best sampling collection methodology (simple random sample, stratified sample, cluster sample, 1 in k systematic random sampling) drawing on knowledge about sample collection schemes learned earlier. Correctly implements the collection scheme. (Connections – Draws upon knowledge acquired earlier. This skillset is needed whenever data collection is warranted)
2. Effectively works as part of a team dividing the workload as measured by peer review process at end of module (Jigsaw presentation of needed techniques)
3. Correctly uses Minitab software to explore collected cell phone usage data to answer questions such as what percent of students spend 30 minutes or less using their phones? More than 3 hours? Less than you? More than you? (Curiosity)
4. Compares their prior estimate of screen time usage for yesterday to the actual data collected for their screen time yesterday. This can illustrate the need for collecting data to make decisions instead of using assumptions (Curiosity – how good of an estimate did you make?) Was your estimate lower than actual or higher than actual? Was your estimate influenced by observation of screen time values seen when checking to see that your phone was set up to collect the data? Provides reflection on the possible pitfalls of having people estimate data.
5. Communicates results in plain language with graphs for understanding using a poster presentation and gallery walks. The poster is clear, visually appealing, and the analysis is accurate.
6. Compares outcomes from various teams to see the impact of sampling variation on analysis direction – although the same models may not be used, how different are the percentage estimates of less than 30 minutes or more than 3 hours? Student reflects on the impact that having different teams draw different samples for analysis has on the essential outcome of predicting what percentage of students are on for less than 30 minutes or more than 3 hours.

(Recommend instructor collect the summarized results (see activity 13) from each team before the gallery walk class to create a comparison poster/table for all students)

Activities:

1. Predict how many minutes screen time you had yesterday before looking at the data from your phone.
2. Record the actual number of minutes screen time you had yesterday on your phone.
3. Form a team of exactly 4 members based on the people sitting near you in class.
4. As a team, identify the population of interest. *Write a formal description of the population definition*.
5. As a team, *predict how many samples you think you will need to create a confidence interval on the mean* which has a +/- magnitude equal to 1/3 of the standard deviation of values
6. Have one team member go to each of the 4 stations in the room to watch the video or videos with members from other teams. If you have your laptop, take it to the station as at least 1 computer with access to MyCourses will be needed in each station.

When the video is finished, discuss the topics with the other people who saw the video. Ask the instructor any questions you have about the material. After watching the video you will be the expert on the topic or topics for your team and will be expected to teach the other members of your team about the topic and how to get Minitab to do the analysis

1. Sample Size for Confidence Intervals on the mean
2. Graphical and numerical summaries of data
3. Distribution identification
4. Probability plots / Probability distribution plots and wrap up
5. As a team, led by the member who attended the sample size for confidence intervals, *determine sample size needed* to estimate the population mean to within +/- 1/3 of a standard deviation. The student who attended the video group is teaching the rest of the team about this topic.
6. Using the sampling collection methodology best suited to your population of interest, collect the needed amount of data. *Write a 200 – 300 word explanation of why you chose the sampling scheme you did and how your data collection methods supported this scheme.*
7. Enter the collected data into Minitab. Have the student who attended the graphical and numerical summaries of data create a summary of your data. The student who attended the video group is teaching the rest of the team about this topic. What is the average screen time. What is the standard deviation of screen time measurements. Based on the graphical summary, can the data be assumed to be normally distributed? What is the 95% confidence interval on the mean? *Compare the +/- width (upper bound minus the average or average minus the lower bound) of the interval to the standard deviation of the data. Is it near one third?*
8. Have the student who attended the distribution identification group find an appropriate probability model for your data (normal is a possibility but may not be appropriate) The student who attended the video group is teaching the rest of the team about this topic. *In 200 words or less, explain why your team made this specific choice of model, especially if more than one model could be appropriate.* For the model chosen, record the estimated parameters such as scale and location.
9. Have the student who attended the probability plots group make an appropriate probability plot of your data (normal is a possibility but may not be appropriate) The student who attended the video group is teaching the rest of the team about this topic. Find what percentage of the observed data was less than 30 minutes and what percentage of the observed data was more than 3 hours. Use the probability model to predict the percentage of the population described by this model that is less than 30 minutes and more than 3 hours. Also have the student who attended the probability distribution plots and wrap up video generate probability distribution plots using the assumed model and parameters to show the percentage of the population expected to be less than 30 minutes and more than 3 hours.
10. Provide the following summary information to the instructor **before midnight** of the day of the gallery walk presentations:
    1. Team member names
    2. Sample Size
    3. Mean of collected data
    4. Standard Deviation of collected data
    5. Assumed probability model selected by team
    6. Parameters of model such as location and scale
    7. Percent of observed data less than 30 minutes
    8. Percent of predicted population less than 30 minutes
    9. Percent of observed data greater than 3 hours
    10. Percent of predicted population greater than 3 hours
    11. Number of team members who underestimated the amount of screen time
    12. Number of team members on your team
11. As a team, *create a poster presentation* summarizing what your team did for this activity. Make sure there are graphs for understanding. Post a copy of your poster to the Assignments folder in MyCourses for grading.
12. Prepare to present your findings to other teams during a gallery walk
13. Compare the outcomes from various teams provided by the instructor to observe the impact of sampling variation on analysis direction – although the same models may not be used, how different are the percentage estimates of less than 30 minutes or more than 3 hours?
14. *Individually complete the project summary* form and submit the form to the Assignments folder in MyCourses for grading