Section 3.1 IRA Guide

Introduction

Screen 1: This problem focuses on the difference between a sample and a population. Go back to section 1.1 if you need a refresher.

Screen 2: This problem focuses on the difference between a statistic and a parameter. Go back to section 1.1 if you need a refresher.

Screen 3: This problem focuses on the difference between qualitative variables (categorical) and quantitative variables (numerical). Go back to section 1.1 if you need a refresher.

Screen 4: This problem focuses on the different types of sampling. Go back to sections 1.3 and 1.4 if you need a refresher.

Screen 5: List of Objectives
Also, an introduction to measures of central tendency.

Objective 1: Determine the Arithmetic Mean of a Variable from Raw Data

Screen 1: Definition of an arithmetic mean, which is what we have always called the "average" before this class. We use different symbols for a population mean (mu) and a sample mean (x-bar). Although the only difference for the mean is our labeling a value as coming from a population or a sample, in future sections this will affect the calculation. You must be able to determine whether you are working with a sample or a population from the context of the problem.

Screen 2: A quick explanation of the formulas for the population mean mu and the sample mean x-bar. Click the link and watch the video that explains the formulas.

Screen 3: Example 1 shows how to compute mu and x-bar from a set of data. Watch the StatCrunch video solution to learn how to do this on StatCrunch which really pays off when the size of the data set is large.

Screen 4: This problem is based on Example 1. You can open the data in StatCrunch to do the calculations there.

Screen 5: Slide the fulcrum in the picture to the left and to the right. Notice that the histogram is "balanced" when the fulcrum is placed at the location of the mean, which indicates that the mean is truly the "center" of the data.

Objective 2: Determine the Median of a Variable from Raw Data

Screen 1: Definition of the median *M*. We do not have different symbols for the median of a sample and the median of a population, and we often will label it as "median" instead of as *M*.

Screen 2: Scroll through the 3 steps for finding a median by hand. Typically we will use StatCrunch to do this. By hand, I like to circle the first half of the data and the second half of the data. If there is a value between those two halves, it is the median. Otherwise, the median is the mean of the two central values. You must have the data in ascending order to find the median (as well as the quartiles later in this chapter).

Screen 3: Example 2 explains how to find the median of a set of data when the number of observations is odd. If you watch the By Hand solution video you will see the process for breaking the data into two equal halves as I referred to in the notes for the previous screen.
Watch the StatCrunch solution video to learn how to find a median using StatCrunch.

Screen 4: This problem is based on Example 2 on the previous screen. You should open the data in StatCrunch to find the median.

Screen 5: Example 3 shows how to find the median of a data set with an even number of observations. Watch the By Hand solution video so you will understand the concept, as well as the StatCrunch solution video so you can practice finding the median using StatCrunch.

Screen 6: This problem is based on Example 3 on the previous screen. You should open the data in StatCrunch to find the median.

Objective 3: Explain What It Means for a Statistic to be Resistant

Screen 1: Open the activity by clicking the link, and walk through the 5 parts so you will understand how the mean and median relate to one another.

Screen 2: Watch the video for a summary of the same concepts you worked with in the activity on Screen 1.

Screen 3: Although the median is generally more reliable, there is a short explanation about why we use the mean.

Screen 4: This problem explores how the mean is affected by extreme values, as you saw in the activity on Screen 1.
You will only get one chance at this problem, so be sure of your answer before submitting it.

Screen 5: This problem explores how the median is affected by extreme values, as you saw in the activity on Screen 1.
You will only get one chance at this problem, so be sure of your answer before submitting it.

Screen 6: Definition of “resistant”, which is a very important concept going forward.

Screen 7: This multi-step problem will help you determine whether the mean or median are said to be resistant. Open the data in StatCrunch for all calculations, and be sure to use the correct columns.
You can compute the mean and median of all three columns at the same time through Stat > Summary Stats > Columns, selecting all 3 columns, and selecting only mean and median.
After answering the first 6 parts based on the original data set, go back and edit the first value in each column to be the value stated (In my example I had to change 106 to 160) and then redo the calculations.
Think about how the mean and median were affected by the new value. Also think about how the size of the sample affects the mean.

Screen 8: Work through the 3 steps of the activity to understand how the locations of the mean and median are related to the shape of a distribution.

Screens 9-11: These 3 problems are based on the observations of the activity on Screen 8. **You only get one chance at each of these 3 problems, so you need to be sure of your answers.** *You might want to look at the summary on Screen 12 before trying these 3 problems.*

Screen 12: A summary of how the relative positions of the mean and median are for differently shaped distributions.

Screen 13: Example 4 works through how to compute the mean and median of a distribution, as well as determine the shape. Open the data in StatCrunch by clicking on the icon so you can work through it while watching the StatCrunch solution video.

Screen 14: This problem is based on Example 4 on the previous screen. Open the data in StatCrunch to create your graphs and do the calculations.

Objective 4: Determine the Mode of a Variable from Raw Data

Screen 1: Definition of the mode of a variable. A set of data can have more than one mode, or it could have no mode. It is mostly used for categorical data.

Screen 2: Example 5 covers how to find the mode by hand.
By the way, StatCrunch now computes the mode. Stat > Summary Stats > Columns, you will find mode at the bottom of the list.

Screen 3: Example 6 shows a scenario where there is no mode.

Screen 4: This problem is based on Examples 5 & 6. You can use StatCrunch (Stat > Summary Stats > Columns, then select Mode from the bottom of the list), although it's quite easy to do by hand, too.

Screen 5: A quick discussion of scenarios in which there is more than one mode.

Screen 6: Example 7 walks through the process of finding the mode of qualitative data. You could create a frequency distribution, and then look for the category with the largest frequency.

Screen 7: This problem is based on Example 7 on the previous screen.

Screen 8: Summary of mean, median, and mode, including how to compute and interpret, as well as when to use.

Screen 9: End of Section