

THAT'S SO RANDOM

DOUG TYSON

PCTM ANNUAL MEETING 2018

HARRISBURG, PA

ACTIVITY 1 – SHOW ME THE MONEY

ACTIVITY 2 – DOES SPONGEBOB MAKE US DUMBER?

COPIES OF THIS HANDOUT CAN BE FOUND AT MRTYSONSTATS.COM

RANDOM SAMPLING - SHOW ME THE MONEY!

INTRODUCTION

1. Name your favorite movie of all time.
2. On average, how much money do you think a movie grosses (earns) in theaters?
3. Take a guess at the title of the top-grossing movie of 2017.
4. What do you think was the maximum amount grossed by a movie in 2017?

SAMPLING METHODS

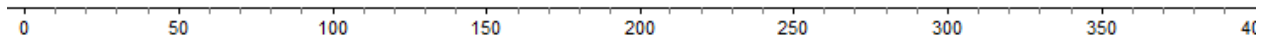
5. Take a look at the list of the 200 top-grossing movies of 2017 and select 10 that you saw (or wanted to see) in theaters. For purposes of this activity, we will consider these 200 movies as a small population. In practice, populations are often much larger than 200 individuals, sometimes reaching the hundreds of millions or more. Write the titles of those movies in the table below, along with the amount they grossed in 2017. Notice that the listing of movies gives the gross income rounded to the nearest tenth of a million, so that a movie listed as earning \$264.6 really grossed \$264,600,000. The order in which you write the movies in the table below does not matter.

Movie Title	Gross Income (Millions)

6. Compute and record the mean gross income for the 10 movies you selected. This number is called the **sample mean**.
7. Is your sample mean the same as all the other sample means computed by the other students in your class?

It's probably not surprising to you that your sample mean differs from those of other students because you have most likely chosen different samples. The fact that different samples yield different statistics (in this case different sample means) is called **sampling variability**.

8. Combine your results with those of your classmates by creating a dotplot of sample means on the board. Then, record this dotplot on the number line below, carefully labeling the axis.



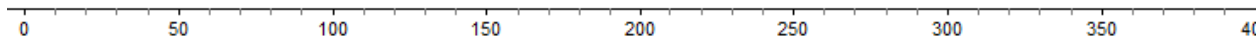
9. Based on the previous dotplot, without any calculations, what do you suppose the mean gross income for the population of all 200 movies might be?

The method you used to take samples from the population is based on your experience and interest in movies. It turns out that this is not a particularly good way to sample if you wish to generate samples that are representative (good images) of the population. Instead of using human experience, judgment, or interest to choose samples, statisticians use chance to select samples from large populations. Samples selected by a chance process are called **random samples**.

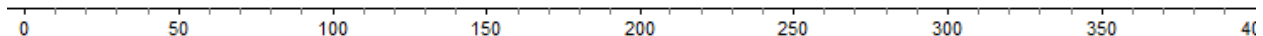
10. Your instructor will give you instructions on how to use random chance to select 10 movies from the population of 200 movies. You will draw chips from a container, use a table of random digits, or use technology to generate 10 random numbers from 1-200. Find the ID numbers in the table below that match your randomly chosen numbers. Record the ID number, title of the movie, and gross income in the following table.

ID Number	Movie Title	Gross Income (Millions)

11. Compute and record the mean gross income for the 10 movies you selected. This number is called the **sample mean**, but this time the sample mean has been generated by a **random sample**. Is your sample mean the same as all the other sample means computed by the other students in your class?
12. Combine your results with those of your classmates by creating a dotplot of sample means on the board. Then, record this dotplot on the number line below, carefully labeling the axis.



13. Based on the previous dotplot, without any calculations, what do you suppose the mean gross income for the population of all 200 movies might be?
14. Is your guess for the mean gross income from a random sample somewhat different from your guess when you chose your own sample?
15. The population's mean gross income for the population of all 200 movies is \$54.98 million. Go back to your dotplots in #8 and #12 and draw a vertical line on your number line at 54.98. Did the sample means from you and your classmates do a pretty good job of estimating the population mean when you chose your sample by thinking of movies you saw or wished to see? How about when you obtained your samples through chance? What have you learned about the use of random samples from populations?
16. With enough time, you and your classmates could continue drawing *random samples* of 10 movies, computing the sample mean for each sample, and building the dotplot in #12 above. In order to save time, your instructor will show you a simulation of this process using Fathom software. When the simulation is finished, create a rough sketch of the resulting dotplot below. Label the horizontal axis and draw two vertical lines: one at the center of all the dots and one at the population mean.



It should be clear that rather than producing haphazard results, random sampling actually follows regular patterns that can be predicted with advanced mathematics and statistics. These patterns can then be used to make inferences (conclusions) about the population from random samples.

4.2 DOES SPONGEBOB IMPAIR THINKING? – RANDOM ASSIGNMENT

INTRODUCTION

1. What is your favorite cartoon?
2. Just your best guess, do you think that cartoons can impair children's thinking?

THE SPONGEBOB SQUAREPANTS STUDY

As reported in the journal *Pediatrics*, researchers Angeline Lillard and Jennifer Peterson conducted a study to determine whether cartoons could produce an impact on the thinking of 4-year-olds. Sixty 4-year-olds were randomly placed into one of three different groups. One group of twenty watched a fast-paced cartoon (*SpongeBob SquarePants*), a different group of twenty watched an educational cartoon (*Caillou*), and the final group of twenty was allowed to draw pictures. At the end of the study, all sixty children were asked to perform four tasks and their mental function was measured.

For this activity, you will “conduct” a similar, but fictitious, study to analyze the reason that researchers randomly assigned students to groups. Your **subjects** (participants) will be a group of fourteen 4-year-olds whose parents have consented to their participation in your research. The goal of your research is to compare the effects of a *SpongeBob cartoon* to *drawing pictures* on the mental function of 4-year-olds. The *SpongeBob cartoon* and *drawing pictures* are called **treatments**.

1. Your instructor will provide you with fourteen index cards. Each card will represent one child from the table below. On one side of each card, write the name of the child, their IQ, and their gender. For example, the card for Albus would read “Albus, 130, Male.”

Name	IQ	Gender
Albus	130	Male
Bellatrix	92	Female
Draco	103	Male
Fred	100	Male
Ginny	106	Female
Harry	110	Male
Hermione	122	Female
Lily	103	Female
Luna	105	Female
Minerva	114	Female
Mundungus	97	Male
Ron	98	Male
Severus	119	Male
Voldemort	116	Male

We will simulate the process of randomly assigning seven subjects to the *SpongeBob cartoon* treatment and seven subjects to the *drawing pictures treatment*. Turn the cards upside down (so that you cannot see your writing) and shuffle them so that they are well mixed. Then, deal out the cards into two piles without looking at them. Choose one group to be the *SpongeBob cartoon* group and the other to be the *drawing pictures* group. Record your groups in the table below.

SPONGEBOB CARTOON GROUP			DRAWING PICTURES GROUP		
Name	IQ	Gender	Name	IQ	Gender

2. Compute the proportion of females within each group (round your answer to the nearest thousandth). Compute the mean IQ for each group. Record your answers below.

	SPONGEBOB CARTOON GROUP	DRAWING PICTURES GROUP
Proportion of Females		
Mean IQ		

3. Compute the difference in the proportion of females in each group (*SpongeBob cartoon group* – *drawing pictures group*). Compute the difference in the mean IQ in each group (*SpongeBob cartoon group* – *drawing pictures group*). Record your answers below.

Difference in proportion of females (*SB* – *Draw*):

Difference in mean IQ (*SB* – *Draw*):

4. Thinking of the difference in proportions of females you calculated above, what would a difference of 0 indicate? a difference that is positive? a difference that is negative?

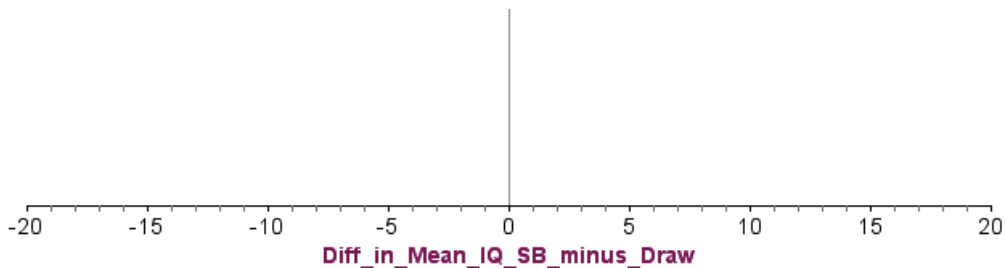
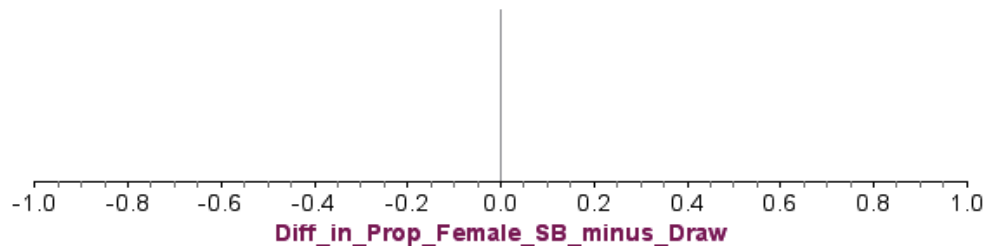
5. Thinking of the difference in mean IQs you calculated above, what would a difference of 0 indicate? a difference that is positive? a difference that is negative?

6. Combine your results with the rest of your class by creating dotplots of the differences on the board. Does a pattern appear to be emerging?

7. Does random assignment *always* balance gender between the two treatment groups? Does random assignment *tend* to balance gender between the two treatment groups?

8. Does random assignment *always* balance IQ between the two treatment groups? Does random assignment *tend* to balance IQ between the two treatment groups?

9. To see if there is a recognizable long-run pattern, you would need to keep shuffling, dealing, and calculating differences. However, to save time, your instructor will show you a demonstration with Fathom software. Below, recreate a rough sketch of the resulting dotplots after the demonstration is done.



10. Would you be surprised to see two treatment groups that were very different if random assignment was used to create them?

This investigation highlights an important property of random assignment: it tends to balance out variables among groups. This helps create similar groups so that when different treatments are applied to the groups, a large difference in their effects (difference in mental function) can reasonably be attributed to the treatments. Thus, if the *SpongeBob cartoon* group shows a significantly higher or lower mental function than the *drawing pictures* group, we could conclude that the *SpongeBob cartoon* caused this difference since the groups should have been roughly similar in mental function before we applied the treatments.

A statistical study in which researchers apply treatments to subjects is called an **experiment**. When random assignment is used in experiments, it has two important consequences:

1. **treatment groups tend to be roughly equal** for all variables that the researchers cannot directly control, and
2. **cause-and-effect inferences (conclusions) may be drawn.**