S.IC.A.2: Analysis of Data
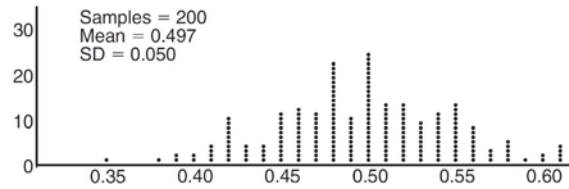
1   Anne has a coin. She does not know if it is a fair coin. She flipped the coin 100 times and obtained 73 heads and 27 tails. She ran a computer simulation of 200 samples of 100 fair coin flips. The output of the proportion of heads is shown below.

Given the results of her coin flips and of her computer simulation, which statement is most accurate?
1)   73 of the computer's next 100 coin flips will be heads.
2)   50 of her next 100 coin flips will be heads.
3)   Her coin is not fair.
4)   Her coin is fair.

2   The results of simulating tossing a coin 10 times, recording the number of heads, and repeating this 50 times are shown in the graph below.
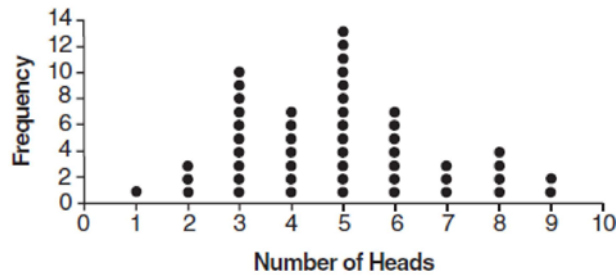
Based on the results of the simulation, which statement is *false*?
1)   Five heads occurred most often, which is consistent with the theoretical probability of obtaining a heads.
2)   Eight heads is unusual, as it falls outside the middle 95% of the data.
3)   Obtaining three heads or fewer occurred 28% of the time.
4)   Seven heads is not unusual, as it falls within the middle 95% of the data.

3   Robin flips a coin 100 times. It lands heads up 43 times, and she wonders if the coin is unfair. She runs a computer simulation of 750 samples of 100 fair coin flips. The output of the proportion of heads is shown below.

Mean = 0.499
SD = 0.049

Frequency

Proportion of Heads
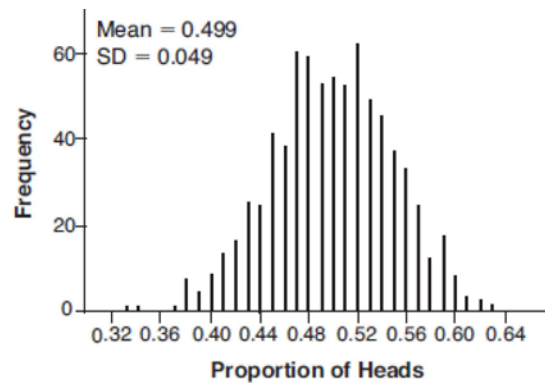
Do the results of the simulation provide strong evidence that Robin's coin is unfair? Explain your answer.

4   A game spinner is divided into 6 equally sized regions, as shown in the diagram below.
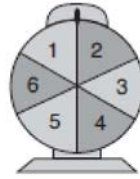

For Miles to win, the spinner must land on the number 6. After spinning the spinner 10 times, and losing all 10 times, Miles complained that the spinner is unfair. At home, his dad ran 100 simulations of spinning the spinner 10 times, assuming the probability of winning each spin is $\frac{1}{6}$. The output of the simulation is shown in the diagram below.
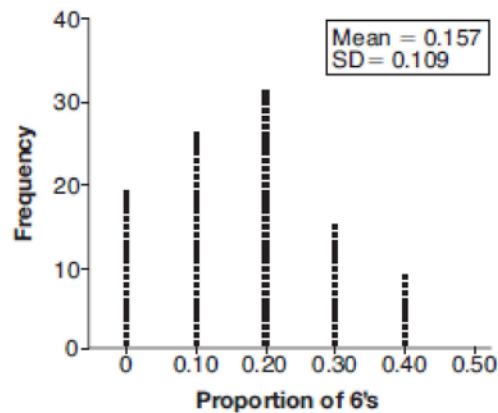

Which explanation is appropriate for Miles and his dad to make?
1) The spinner was likely unfair, since the number 6 failed to occur in about 20% of the simulations.
2) The spinner was likely unfair, since the spinner should have landed on the number 6 by the sixth spin.
3) The spinner was likely not unfair, since the number 6 failed to occur in about 20% of the simulations.
4) The spinner was likely not unfair, since in the output the player wins once or twice in the majority of the simulations.

5 Joette is playing a carnival game. To win a prize, one has to correctly guess which of five equally sized regions a spinner will land on, as shown in the diagram below.
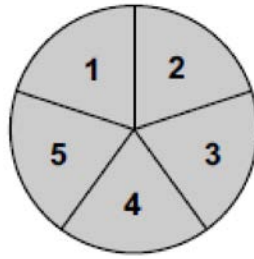
She complains that the game is unfair because her favorite number, 2, has only been spun once in ten times she played the game. State the proportion of 2's that were spun. State the theoretical probability of spinning a 2. The simulation output below shows the results of simulating ten spins of a fair spinner, repeated 100 times.
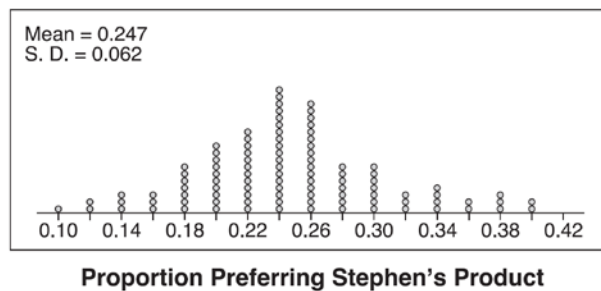

Proportion of 2's
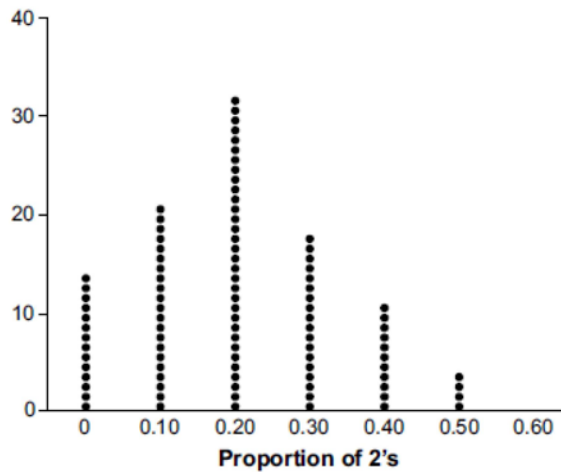
Does the output indicate that the carnival game was unfair? Explain your answer.

6 Some smart-phone applications contain "in-app" purchases, which allow users to purchase special content within the application. A random sample of 140 users found that 35 percent made in-app purchases. A simulation was conducted with 200 samples of 140 users assuming 35 percent of the samples make in-app purchases. The approximately normal results are shown below.
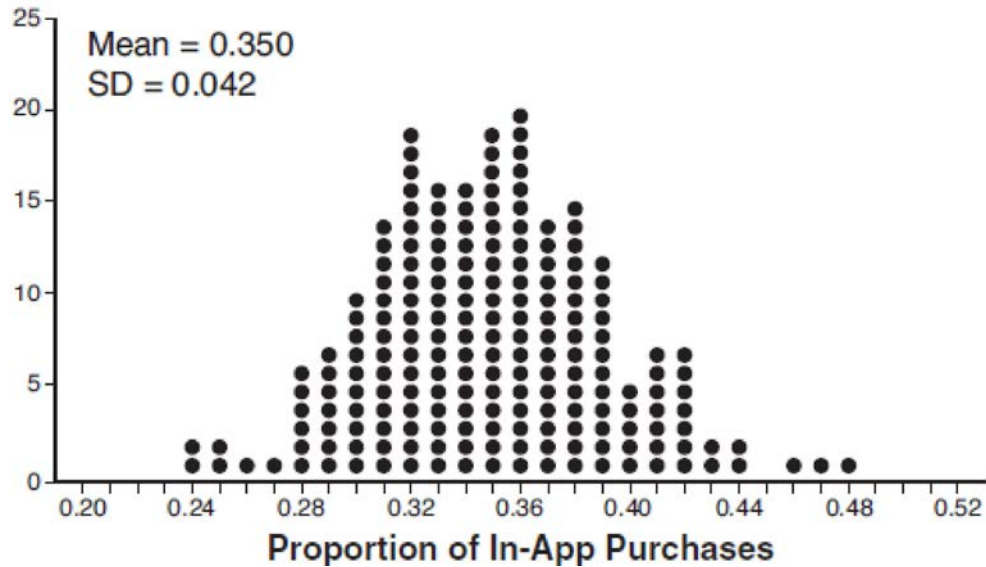

Considering the middle 95% of the data, determine the margin of error, to the *nearest hundredth*, for the simulated results. In the given context, explain what this value represents.

7 Stephen's Beverage Company is considering whether to produce a new brand of cola. The company will launch the product if at least 25% of cola drinkers will buy the product. Fifty cola drinkers are randomly selected to take a blind taste-test of products *A*, *B*, and the new product. Nine out of fifty participants preferred Stephen's new cola to products *A* and *B*. The company then devised a simulation based on the requirement that 25% of cola drinkers will buy the product. Each dot in the graph shown below represents the proportion of people who preferred Stephen's new product, each of sample size 50, simulated 100 times.

Proportion Preferring Stephen's Product

Assume the set of data is approximately normal and the company wants to be 95% confident of its results. Does the sample proportion obtained from the blind taste-test, nine out of fifty, fall within the margin of error developed from the simulation? Justify your answer. The company decides to continue developing the product even though only nine out of fifty participants preferred its brand of cola in the taste-test. Describe how the simulation data could be used to support this decision.

8   In a random sample of 250 men in the United States, age 21 or older, 139 are married. The graph below simulated samples of 250 men, 200 times, assuming that 139 of the men are married.
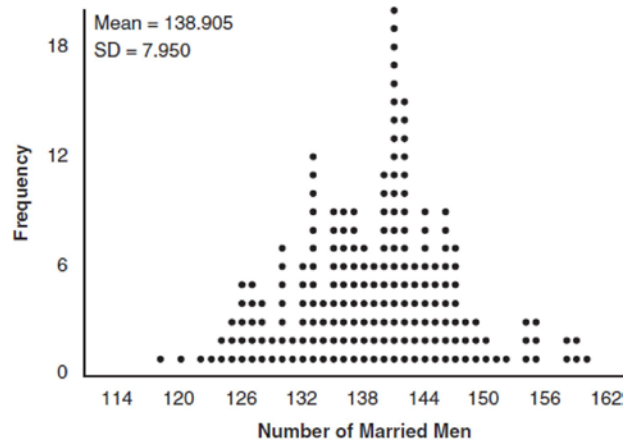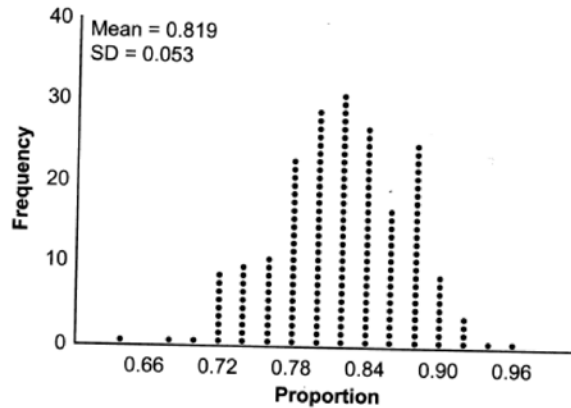

a) Based on the simulation, create an interval in which the middle 95% of the number of married men may fall. Round your answer to the *nearest integer*.
b) A study claims "50 percent of men 21 and older in the United States are married." Do your results from part a contradict this claim? Explain.

9   In order to decrease the percentage of its residents who drive to work, a large city launches a campaign to encourage people to use public transportation instead. Before starting the campaign, the city's Department of Transportation uses census data to estimate that 65% of its residents drive to work. The Department of Transportation conducts a simulation, shown below, run 400 times based on this estimate. Each dot represents the proportion of 200 randomly selected residents who drive to work.
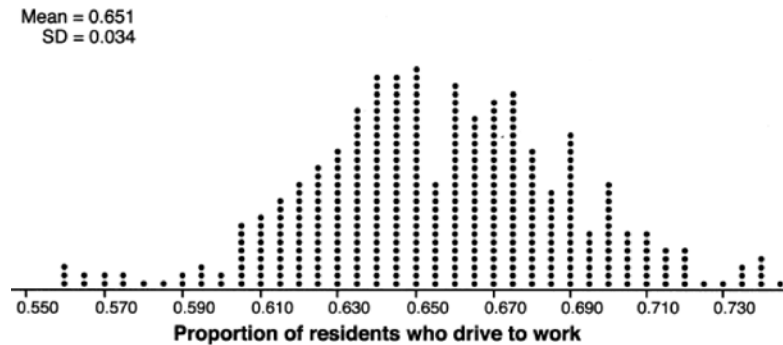
Use the simulation results to construct a plausible interval containing the middle 95% of the data. Round your answer to the *nearest hundredth*. One year after launching the campaign, the Department of Transportation conducts a survey of 200 randomly selected city residents and finds that 122 of them drive to work. Should the department conclude that the city's campaign was effective? Use statistical evidence from the simulation to explain your answer.

10 State officials claim 82% of a community want to repeal the 30 mph speed limit on an expressway. A community organization devises a simulation based on the claim that 82% of the community supports the repeal. Each dot on the graph below represents the proportion of community members who support the repeal. The graph shows 200 simulated surveys, each of sample size 60.

Based on the simulation, determine an interval containing the middle 95% of plausible proportions. Round your answer to the *nearest thousandth*. The community organization conducted its own sample survey of 60 people and found 70% supported the repeal. Based on the results of the simulation, explain why the organization should question the State officials' claim.

11 An app design company believes that the proportion of high school students who have purchased apps on their smartphones in the past 3 months is 0.85. A simulation of 500 samples of 150 students was run based on this proportion and the results are shown below.
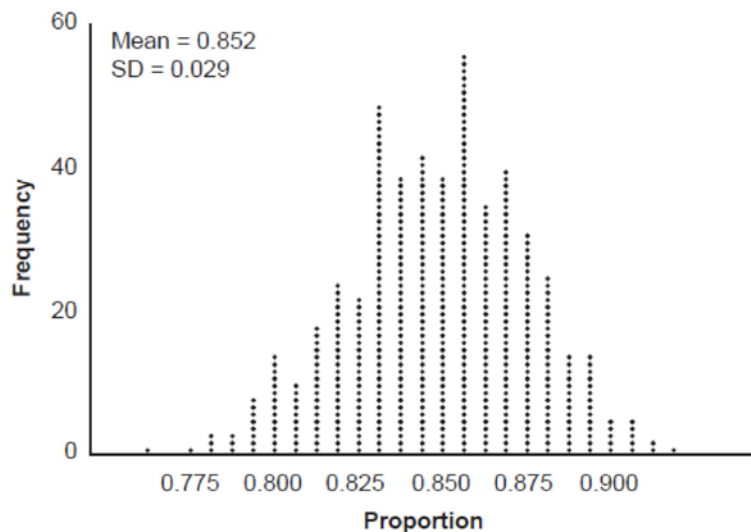

Suppose a sample of 150 students from your high school showed that 88% of students had purchased apps on their smartphones in the past 3 months. Based on the simulation, would the results from your high school give the app design company reason to believe their assumption is *incorrect*? Explain.

12 A sporting goods manufacturer is trying to determine if they should continue to produce multiple types of hockey pucks. The company surveyed 50 randomly chosen customers and asked them if they purchased both game regulation pucks and lighter training pucks. Of those surveyed, 40 of them said that they purchase both types of pucks. A simulation that was run 100 times based on the survey results produced the approximately normal results below.
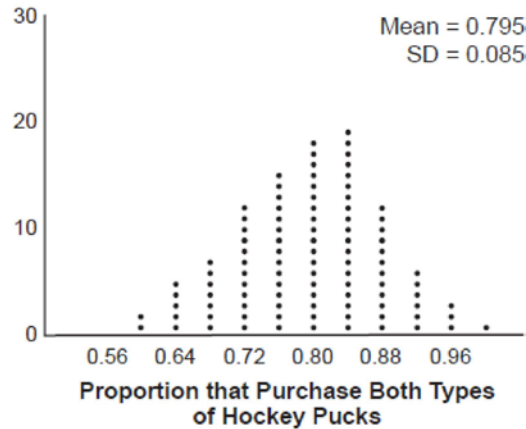
a) Determine an interval containing the middle 95% of plausible values that estimates the proportion of all customers who would purchase both types of pucks from the company.
b) The company will continue to manufacture both types of hockey pucks if it is reasonable to assume that the true proportion of customers who buy both types of hockey pucks is above 0.60. Using the interval from part a, explain whether or not the company should continue to produce both types of hockey pucks.

13 In a packaging plant, a machine packs boxes with jars. The machine's manufacturer states that a box is packed, on average, every 42 seconds. To test that claim, the packaging plant randomly selects a sample of 10 boxes and finds the sample mean to be 49.8 seconds. The company ran a simulation of 1000 trials based on the manufacturer's claim. The approximately normal results are shown below.
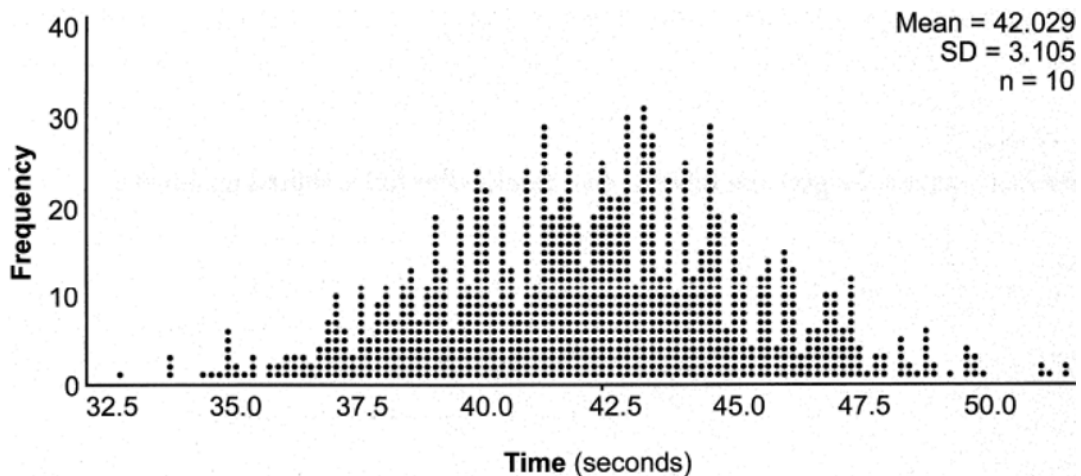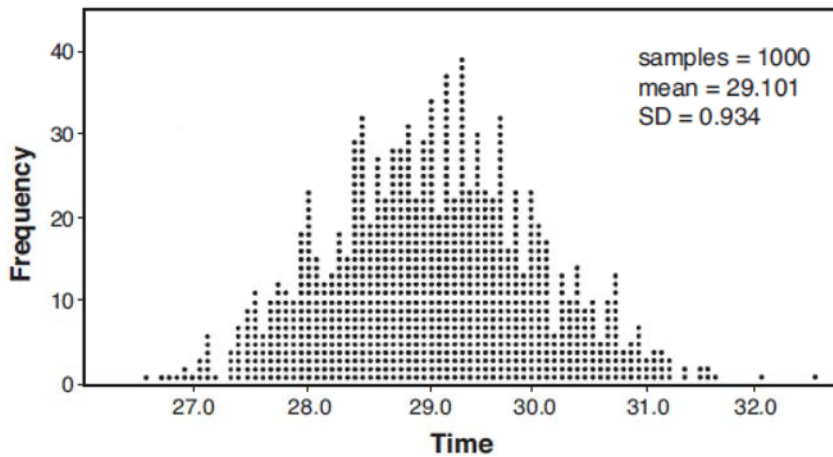

Based on the simulation, determine an interval containing the middle 95% of plausible mean times. Round your answer to the *nearest hundredth*. Is the time 49.8 seconds unusual? Use statistical evidence to justify your answer.

14 A radio station claims to its advertisers that the mean number of minutes commuters listen to the station is 30. The station conducted a survey of 500 of their listeners who commute. The sample statistics are shown below.

$\bar{x}$	29.11
s_x	20.718

A simulation was run 1000 times based upon the results of the survey. The results of the simulation appear below.

Based on the simulation results, is the claim that commuters listen to the station on average 30 minutes plausible? Explain your response including an interval containing the middle 95% of the data, rounded to the *nearest hundredth*.

15 Mrs. Jones had hundreds of jelly beans in a bag that contained equal numbers of six different flavors. Her student randomly selected four jelly beans and they were all black licorice. Her student complained and said "What are the odds I got all of that kind?" Mrs. Jones replied, "simulate rolling a die 250 times and tell me if four black licorice jelly beans is unusual." Explain how this simulation could be used to solve the problem.

S.IC.A.2: Analysis of Data
Answer Section

1 ANS: 3   REF: 061607aii
2 ANS: 2   REF: 011820aii
3 ANS:
No. $0.499 \pm 2(0.049) \rightarrow 0.401 - 0.597$. Since 0.43 falls within this interval, Robin's coin is likely not unfair.

REF: 061932aii
4 ANS: 3   REF: 061710aii
5 ANS:
$\frac{1}{10}, \frac{1}{5}$, and no, since 0.10 clearly falls within 95% of 0.20.

REF: 012334aii
6 ANS:
$2(0.042) = 0.084 \approx 0.08$ The percent of users making in-app purchases will be within 8% of 35%.

REF: 081832aii
7 ANS:
Yes. The margin of error from this simulation indicates that 95% of the observations fall within ± 0.12 of the simulated proportion, 0.25. The margin of error can be estimated by multiplying the standard deviation, shown to be 0.06 in the dotplot, by 2, or applying the estimated standard error formula, $\left(\sqrt{\frac{p(1-p)}{n}} \right) or \left(\sqrt{\frac{(0.25)(0.75)}{50}} \right)$

and multiplying by 2. The interval 0.25 ± 0.12 includes plausible values for the true proportion of people who prefer Stephen's new product. The company has evidence that the population proportion could be at least 25%. As seen in the dotplot, it can be expected to obtain a sample proportion of 0.18 (9 out of 50) or less several times, even when the population proportion is 0.25, due to sampling variability. Given this information, the results of the survey do not provide enough evidence to suggest that the true proportion is not at least 0.25, so the development of the product should continue at this time.

REF: spr1512aii
8 ANS:
$138.905 \pm 2 \cdot 7.95 = 123 - 155$. No, since 125 (50% of 250) falls within the 95% interval.

REF: 011835aii
9 ANS:
$.651 \pm 2 \cdot .034 = .58 - .72$. No, since .61 (122/200) falls within the 95% interval.

REF: 062235aii
10 ANS:
$.819 \pm 2 \cdot .053 = .713 - .925$. Since .70 does not fall within the 95% interval.

REF: 082236aii

11 ANS:
No. $0.852 \pm 2(0.029) \rightarrow 0.794 - 0.91$. 0.88 falls within this interval.

REF: 062332aii

12 ANS:
$.795 \pm 2 \cdot .085 = .625 - .965$. Yes, as it is plausible at least .625 of the customers will purchase both.

REF: 062435aii

13 ANS:
$42.029 \pm 2 \cdot 3.105 \approx 35.82 - 48.24$. Yes, since 49.8 falls outside the 95% interval.

REF: 082434aii

14 ANS:
$29.101 \pm 2 \cdot 0.934 = 27.23 - 30.97$. Yes, since 30 falls within the 95% interval.

REF: 011935aii

15 ANS:
Since there are six flavors, each flavor can be assigned a number, 1-6. Use the simulation to see the number of times the same number is rolled 4 times in a row.

REF: 081728aii