Lesson 1.1
The Statistical Analysis Process

INTRODUCTION

What is statistics? Why do we study statistics?

Statistics is about using data to answer questions. Data is information that we collect from our world. Data involves facts and observations that we make. Before scientists discovered statistics and before people used data, they would use opinions and hunches to explain how the world worked. A lot of times these explanations were wrong.

For example, people once believed that the earth was the center of the universe. When people started making observations and using measures they discovered that this was incorrect. The earth was not the center of the universe. Data helps us make better conclusions. In this example, data helped us see that the earth goes around the sun, not the other way around.

Data can help us answer many types of questions.
- Students can use data to help pick a college that is best for them.
- Teachers use data to find the best ways to educate their students.
- Medical professionals use data to learn if new treatments actually work.
- Voters need data about their society and planet to create a better democracy.
- Politicians use data to better represent the people who elect them.

In statistics, we gather, summarize, and analyze data to search for answers to our questions.

Let’s begin with an idea that will help us think about how statistics can help us answer a question that we may have.

1 LeBron James is an NBA basketball super star who, in 2008, played for the Cleveland Cavaliers. His team made it to the playoffs that year and faced off against the Boston Celtics. However, sports announcers and fans alike claimed that he did not perform up to his abilities in the playoffs, claiming that he “choked”. Imagine you want to determine if this criticism is justified. You know that data involves facts, observations, and measurements about a particular topic or idea. If you wanted to gather data about LeBron’s performance, what information might you look for? How could you use data to answer this question?
Statistical analysis is the process of looking at data to learn about something bigger. We can think of the statistical analysis process in 4 steps.

Steps in a Statistical Analysis

1. Ask a question that can be answered by collecting data.
2. Decide what to measure and then collect data.
3. Summarize and analyze the data.
4. Draw a conclusion and communicate the results.

Did LeBron James choke in the 2008 playoffs?

We will now do an activity that will help us learn about the statistical analysis process. In this activity we use the statistical analysis process to investigate a question about whether LeBron choked in the 2008 playoffs.

As you discovered in answering question number 1 above, there are many ways to judge the performance of a basketball player. We will choose to look at his 3-point shooting percentage and use this to see if his ability to make 3-point shots decreased in the playoffs. Note, we are using his 3-point shooting percentage to answer the same question we saw above: Did LeBron James choke in the playoffs?

In order to answer our question, we need to collect the necessary data. Specifically we need to know what LeBron’s 3-point shooting percentages were in the regular season of 2007-2008 and in the 2008 playoffs.
Lesson 1.1
The Statistical Analysis Process

The data shown below is from [www.basketballreference.com](http://www.basketballreference.com).

### Regular Season Stats

| Year | Age | Team | Lg | G     | Min | Pts | PPG | FGM | FGA | FGP | FTM | FTA | FTP | 3PM | 3PA | 3PP | ORB | DRB | TRB | RPG | AST | APG | STL | BLK | TO  |
|------|-----|------|----|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 2003-04 | 19  | CLE  | NBA| 79    | 1654| 20.9| 622 | 1492| 417 | 347 | 460 | 600 | 575 | 63  | 217 | 290 | 99  | 333 | 432 | 5.5 | 465 | 5.9 | 130 | 58  | 273 |
| 2004-05 | 20  | CLE  | NBA| 80    | 2175| 27.2| 765 | 1684| 472 | 477 | 636 | 750 | 108 | 308 | 351 | 111 | 477 | 588 | 7.4 | 577 | 7.2 | 177 | 52  | 262 |
| 2005-06 | 21  | CLE  | NBA| 79    | 3362| 2478| 31.4| 875 | 1823| 480 | 601 | 814 | 738 | 127 | 379 | 335 | 75  | 461 | 556 | 7.0 | 521 | 6.6 | 123 | 66  | 260 |
| 2006-07 | 22  | CLE  | NBA| 78    | 3195| 2132| 27.3| 772 | 1621| 478 | 489 | 701 | 698 | 99  | 310 | 319 | 83  | 443 | 528 | 6.7 | 470 | 6.0 | 125 | 55  | 250 |
| 2007-08 | 23  | CLE  | NBA| 75    | 3027| 2250| 30.0| 764 | 1642| 484 | 549 | 771 | 712 | 113 | 369 | 315 | 133 | 459 | 592 | 7.9 | 639 | 7.2 | 138 | 81  | 255 |
| 2008-09 | 24  | CLE  | NBA| 81    | 3057| 2304| 28.4| 769 | 1613| 489 | 594 | 762 | 790 | 132 | 364 | 344 | 106 | 507 | 613 | 7.6 | 597 | 7.2 | 137 | 93  | 241 |
| 2009-10 | 25  | CLE  | NBA| 76    | 2966| 2258| 29.7| 768 | 1528| 503 | 593 | 773 | 767 | 129 | 367 | 333 | 71  | 463 | 554 | 7.3 | 651 | 8.6 | 125 | 77  | 281 |
| 2010-11 | 26  | Mia  | NBA| 79    | 3063| 2111| 26.7| 768 | 1485| 510 | 503 | 663 | 759 | 92  | 279 | 330 | 80  | 510 | 590 | 7.5 | 554 | 7.0 | 124 | 50  | 284 |

8 Season Totals | 627 | 25160 | 17362 | 27.7 | 6173 | 12886 | 479 | 4153 | 5580 | 744 | 863 | 2623 | 329 | 758 | 3693 | 4451 | 7.1 | 4384 | 7.0 | 1079 | 532 | 2086 |

### Playoff Stats

| Year | Team | Lg | G     | Min | Pts | PPG | FGM | FGA | FGP | FTM | FTA | FTP | 3PM | 3PA | 3PP | ORB | DRB | TRB | RPG | AST | APG | STL | BLK | TO  |
|------|------|----|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 2005-06 | CLE  | NBA| 13    | 804 | 146 | 30.8| 257 | 516 | 457 | 115 | 737 | 21  | 210 | 105 | 23  | 76  | 5.8 | 18  | 9   | 65  |     |     |     |
| 2006-07 | CLE  | NBA| 20    | 893 | 166 | 25.1| 399 | 769 | 416 | 146 | 196 | 755 | 21  | 275 | 280 | 26  | 135 | 161 | 8.1 | 159 | 9.0 | 34  | 10  | 66  |
| 2007-08 | CLE  | NBA| 13    | 552 | 113 | 28.2| 275 | 611 | 412 | 122 | 187 | 731 | 18  | 70  | 257 | 16  | 86  | 102 | 7.8 | 99  | 7.6 | 23  | 17  | 54  |
| 2008-09 | CLE  | NBA| 14    | 580 | 159 | 35.3| 312 | 510 | 149 | 199 | 749 | 27  | 81  | 333 | 19  | 109 | 128 | 9.1 | 102 | 7.3 | 23  | 12  | 38  |
| 2009-10 | CLE  | NBA| 11    | 459 | 106 | 32.0| 211 | 502 | 88  | 120 | 733 | 20  | 60  | 400 | 15  | 87  | 102 | 9.3 | 84  | 7.6 | 19  | 20  | 42  |
| 2010-11 | Mia  | NBA| 21    | 922 | 174 | 23.7| 373 | 466 | 119 | 156 | 763 | 30  | 85  | 353 | 34  | 142 | 176 | 8.4 | 123 | 5.9 | 21  | 25  | 66  |

Totals | 60  | 2529 | 1761 | 29.4 | 584 | 1293 | 452 | 506 | 680 | 744 | 87  | 289 | 301 | 83  | 413 | 496 | 8.3 | 436 | 7.3 | 98  | 48  | 223 |

#### Key to Column Abbreviations

- G: Games Played
- Min: Minutes Played
- MPG: Minutes per Game
- Pts: Total Points
- PPG: Points per Game
- FGM: Field Goals Made
- FGA: Field Goals Attempted
- FGP: Field Goal Percentage
- FTM: Free Throws Made
- FTA: Free Throws Attempted
- FTP: Free Throw Percentage
- 3PM: Three Pointers Made
- 3PA: Three Pointers Attempted
- 3PP: Three Point Percentage
- REB: Total Rebounds
- RPG: Rebounds per Game
- AST: Total Assists
- APG: Assists per Game
- STL: Steals
- BLK: Blocks
- TO: Turnovers
TRY THESE

2 Answer each of the following questions about the four steps of the statistical analysis process.

Step 1: Ask a question that can be answered with data.

A What question are we trying to answer in our investigation of LeBron’s 3-point shooting percentages in the regular season and in the playoffs of 2007-2008?

Step 2: Decide what to measure and then collect data.

B What information did we get from the website? Are the data related to the question we are trying to answer?

Step 3: Summarize and Analyze Data

C Use the data shown above to summarize the necessary data in the table below. How might we use this data to decide if his ability to make 3-point shots in the playoffs decreased?

Record the appropriate data below:

<table>
<thead>
<tr>
<th>LeBron James 2007-2008</th>
<th>3 point shots made</th>
<th>3 point shots attempted</th>
<th>3 point percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular Season</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Playoffs</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 4: Draw a conclusion and communicate the results.

D Once the data are summarized and analyzed, how can we use this to answer the research question?
Lesson 1.1
The Statistical Analysis Process

NEXT STEPS

As we proceed, we will address the question of whether LeBron choked in the playoffs. We have completed the data collecting and summarizing process by researching on the internet and then recording the data we found in a table.

We now prepare to analyze our data. To do this, we need to think about how the data can be used to answer our question. We will use simulation to help in this. Talk about the following questions with your group to begin this process.

3 Suppose LeBron did not choke in the playoffs. Would you expect his 3-point percentage to be different than what it was in the regular season? Explain why you think this.

If LeBron did not choke in the playoffs, about what percentage of 3-point shots would you expect him to make?

4 Suppose LeBron did choke in the playoffs. Would you expect the percentage of 3-point shots he made in the playoffs to be greater or less than the percentage you wrote above?

How poorly would LeBron had to have performed in order to convince you that he choked in the playoffs, that is, that his 3-point shooting ability decreased in the playoffs? Give a specific percentage. Tell why you think this.

5 Imagine that LeBron’s 3-point percentage in the playoffs was 22%. Does this prove without a doubt that LeBron’s performance in the playoffs decreased in the playoffs? If not, give another explanation for why his percentage might have been so different than his percentage in the regular season.
Lesson 1.1  
The Statistical Analysis Process

NEXT STEPS

Using simulation to discover what may happen by chance
If LeBron did not choke in the playoffs, we would expect him to perform at close to his regular season percentage, 31.5%. But how far below 31.5% would convince us that he did choke?

We will answer this by simulating LeBron’s 3-point shooting ability in the regular season. From that, we’ll see the typical range of percentages to expect from LeBron when he’s performing at his usual level of ability, in other words, when he’s not choking. This range of percentages is what would likely occur just due to random (or chance) variation. We’ll also be able to see what percentages are most likely.

6 Your instructor has given you a spinner marked to match LeBron’s regular season 3 point percentage: 31.5% of the time he makes it, and 68.5% of the time he misses it. Spin the spinner 70 times (why 70?) and keep track of how many times the spinner lands on “Made It”. Complete the following information.
   ▪ Total Spins ___ 70 ___
   ▪ Number of “Made Its” : ____________
   ▪ Fraction of your spins that “Made It”: ____________
   ▪ Decimal proportion of your spins that “Made It”: ____________
   ▪ Percentage of your spins that “Made It”: ____________

In statistics, a proportion is a number between 0 and 1. It represents a portion out of the total. We usually give proportions as decimals or percents. We can calculate a decimal proportion by dividing the numerator of a fraction by the denominator. For example, if the fraction of “Made Its” is 21/70, then you divide 70 by 21, or 21 ÷ 70. The proportion would be 0.30. To change to a percent we multiply by 100 or move the decimal 2 places to the right. 0.30 = 30%.

7 Is your percentage of “Made Its” equal to 31.5%? If not, is it greater than or less than 31.5%?

8 Do you think everyone in the class got the same percentage of “Made Its”? Why or why not?
Lesson 1.1
The Statistical Analysis Process

We will gather the class percentages to help us understand what kinds of shooting percentages we could expect if LeBron was shooting with the same ability he displayed in the regular season.

9 First, write the class shooting percentages in the following table. Add to the table as necessary to accommodate the number of trials your class produced. Each Trial represents 70 shots or 70 spins. Observed Percentage means the percentage of “Made Its” in each round of 70 spins produced.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Observed Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Second, copy the dotplot your class has made using the class data below:
Answer these questions using the dotplot you constructed above which shows the observed percentage of “Made Its” for 70 spins, assuming LeBron shot with the same percentage as he did in the regular season, 31.5%.

A What was the smallest shooting percentage observed?

B What was the largest shooting percentage observed?

C Did the shooting percentage differ much from trial to trial?

Use the dotplot produced in the movie to answer the following questions.

11 What does each dot represent in the dotplot?

12 Where is the dotplot centered? Why do you think this is?

13 What range of percentages is most likely?

14 Look at the dotplot. How can this graph help you decide whether LeBron choked or not?
Lesson 1.1
The Statistical Analysis Process

15 Recall that LeBron made 25.7% of his 3-point shots in the 2008 playoffs. Use the dotplot to answer the following question.

If LeBron was shooting in the playoffs with the same percentage as he did in the regular season, would it be unusual for him to have made 25.7% of his 3-point shots? Hint: Look at your answer to number 13.

Draw a Conclusion and Communicate the Results.

16 So, is there enough evidence for us to conclude that LeBron choked? That is, is there enough evidence for us to conclude that if LeBron was shooting in the playoffs at the same percentage as he was in the regular season, his shooting percentage in the playoffs was sufficiently low enough for us to conclude that his lower percentage was not just due to chance?
### NEXT STEPS

The task you have just completed with the LeBron example illustrates the **statistical analysis process**, which we have described in four steps. These are given again below.

**Steps in a Statistical Investigation**

1. Ask a question that can be answered by collecting data.
2. Decide what to measure and then collect data.
3. Summarize and analyze the data.
4. Draw a conclusion and communicate the results.

17 Identify each step of the Statistical Analysis Process for the LeBron investigation below.

<table>
<thead>
<tr>
<th>Steps in Statistical Analysis</th>
<th>For the LeBron Investigation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ask a question that can be answered by collecting data.</td>
<td></td>
</tr>
<tr>
<td>2. Decide what to measure and then collect data.</td>
<td></td>
</tr>
<tr>
<td>3. Summarize and analyze the data.</td>
<td></td>
</tr>
<tr>
<td>4. Draw a conclusion and communicate the results.</td>
<td></td>
</tr>
</tbody>
</table>
Lesson 1.1
The Statistical Analysis Process

STUDENT NAME ___________________________ DATE ______________________

TAKE IT HOME

1. Considering your answer to question 13, give an example of a 3 – point percentage that would convince you that LeBron DID choke in the playoffs.

2. Suppose we wanted to investigate LeBron’s Field Goal percentages, instead of his 3 – point percentages, to decide if he choked in the playoffs. Looking back at our data table, we see his regular season Field Goal percentage is 48.4%. We’ll proceed as before, and simulate this percentage in order to decide what we can expect just from random variation (luck). Note that he attempted 275 field goals in the playoffs. There are two dotplots shown below.

A. Which of the two dotplots would we expect to see if LeBron’s field goal percentage for the regular season was 48.4%? Explain why you made this choice.

B. Give an example of a percentage that would convince you LeBron choked in the playoffs, using field goal percentages as our indicator. Explain why you made this choice.

C. Look back in the playoffs data table, given at the beginning of this lesson, to find the number of field goals he made and the number he attempted in the 2008 playoffs. Record them below.

   Number of Field Goals made__________   Number of Field Goals attempted__________
Lesson 1.1
The Statistical Analysis Process

D  Use these values to calculate the percentage of field goals LeBron made in the 2008 playoffs. Show below what you divided. Check your answer in the data table given.

E  Looking at the dotplot above and the percentage you calculated in part D, what can we decide? If LeBron was on trial for choking in the playoffs, using Field Goal percentages as our measurement, could we convict LeBron of choking?