

**Description:**

Officials in the city of Philadelphia are investigating whether global trends apply to the city's population. As a recent hire in the demographics department, you are tasked with compiling and analyzing data. To perform this task, you must understand the difference between the two categories of statistics that you will be working with: descriptive and inferential. First, you will learn the definitions of both descriptive and inferential statistics and how they differ in terms of sampling and of the purpose of the data collected. Next, you will collect data from your classmates and compile it. Last, you will explore the GapMinder website and evaluate data from the website and draw trends from it and answer.

**Leading Question**

How is descriptive statistics different from inferential statistics?

**Students will be able to:**

- Differentiate between descriptive and inferential statistics
- Make statistical inferences
- Make distinctions between sampling types
- Determine if a chosen sample is proper or not for the given example
- Choose their own sample set

**Students will understand:**

Students will investigate descriptive and inferential statistics. They will use real world examples to explore sampling, how to choose a sample, and make inferences based on a data set. Through this lesson, students will discover the difference between descriptive and inferential statistics by exploring the purpose and uses of each through examples. They will understand the importance of choosing a proper sample, of different types of sampling and of how sampling methods can affect statistics.

**Key Definitions & Concepts [1]:**

- **Descriptive Statistics:** Numbers that are used to summarize and describe data. They do not involve generalizing beyond the data at hand.
- **Inferential Statistics:** Data from a sample that is used to draw inferences about a population
- **Random Sampling:** A sample that is collected in such a way that some members of the intended population are less likely to be included than others.
- **Stratified Sampling:** This method can be used if the population has a number of distinct "strata" or groups. In stratified sampling, you first identify members of your sample who belong to each group. Then you randomly sample from each of those subgroups in such a way that the sizes of the subgroups in the sample are proportional to their sizes in the population.
- **Population:** The larger set from which a sample may be drawn.
- **Biased Sample:** A sample that is collected in such a way that some members of the intended population are less likely to be included than others.

**Standards[Copied from: 2]:**

CC.2.4.HS.B.4 Recognize and evaluate random processes underlying statistical experiments.  
CC.2.4.HS.B.5 Make inferences and justify conclusions based on sample surveys, experiments, and observational studies.

**Background Information**

**Prior Knowledge:**

- Basic arithmetic
- Ability to read tables and graphs
- Understanding how to collect data

**Math Practices [Copied from: 3]:**

- Use appropriate tools strategically.
- Reason abstractly and quantitatively.
- Construct viable arguments and critique the reasoning of others.

**Core Ideas [Copied from: 4]:**

- Information Processing
- Defining and Delimiting Problems

**Cross Cutting Concepts [Copied from: 5]:**

- Patterns
- Cause and Effect
- Scale, Proportion, and Quantity

**Possible Preconceptions/Misconceptions:**

Students may automatically assume that random sampling is the best way to choose a sample but there are potential caveats to that. For example, random sampling does not always guarantee an accurate representation of the population as whole (i.e. a population may have 50% male and 50% female but a random sample may cause a sample that is 40% male and 60% female, which is not an accurate representation of the population). They may also have trouble distinguishing that descriptive statistics are strictly descriptive of an entire population, whereas inferential statistics require interpretation. Descriptive statistics are designed to describe a population to get a better understanding of it. One great example of this is the US Census. Inferential statistics pertain more to what we think of when we conduct research. An example of this is by collecting and analyzing or interpreting data to make inferences on that data that is taken from a sample that is representative of a population. Students may also struggle with why a seemingly proper sample is actually not beneficial for the experiment in question.

**Lesson Plan - 5E(+) Model**

**Engage [6]:**

Students will first write down their answer to the first question on the *Will Saving Poor Children Lead to Overpopulation* half sheet. Then the teacher will then play the video through the link: [Will Saving Poor Children Lead to Overpopulation?](#) After watching the video, the teacher should allot 5 minutes for

the students to write their answers for numbers 2 through 5 on the half sheet. The instructor should then facilitate an open class discussion to talk about how the students answered all of the questions on the half sheet. From this discussion, students should understand that analyzing statistical data is key to understanding trends rather than going off bias or preconceived notions. Also, this discussion should get students interested in gathering and looking at data. This activity will serve as a preview to Gapminder, which is where the students will navigate the web during the elaboration section. This section of the lesson will take less than 10 minutes.

### **Explore[1]:**

#### **Part I: Introduction**

Students will use the *Statistics Concept Map* worksheet to fill out a concept map of the important concepts in this lesson. Students will be given a list of definitions, and they need to match the correct definition to the words in the bubbles. Students should draw lines between the words that relate to each other; then, on the connecting line, students will write their reasoning to explain why these concepts are connected (e.g. “sample” connects to “population” because a sample is a subset of a population). This helps students draw connections between words, develop a stronger understanding of the concepts, and bridge any misunderstanding gaps that students may have had prior to this lesson. The teacher should facilitate a discussion of some of the connections that students made in their maps as well as go over the correct answers. This should take a total of 20 minutes - 10 minutes for the students to complete the concept map and 10 minutes to review the answers as a class.

#### \*\*\*\*\*NOTES FOR THE TEACHER\*\*\*\*\*

Throughout the *Measurement and Data Analysis* module, students will make concept maps similar to this one for each lesson. It is recommended to have the students turn in their concept maps after each lesson but be able to reference each of their concept maps during the subsequent lessons. Through these worksheets, students can then compile all of their concept maps together to continue to grow their mathematical connections and develop a deeper understanding of statistics. Also, students will be able to visualize how all of the concepts within the *Measurement and Data Analysis* module connect.

#### **Part II: Benchmark Lesson: Descriptive Statistics and Choosing Samples [1]**

Students will work on the *Descriptive Statistics* worksheet. They will first collect class data to create their own set of descriptive statistics about the class. They will answer questions about the data and descriptive statistics. This will allow students to see how descriptive work is completed and how descriptive data are collected. The associated questions ensure that the students think deeper about the purpose descriptive statistics. It is highly recommended to allow the students choose how they want to collect the information. This may lead to them choosing a chaotic method, such as each student individually survey each student, but it is an important step in learning how to collect data efficiently. Some students may elect a more efficient way of collecting data, such as a show of hands. By allowing for this individuality, students will learn how data can be collected in different ways, and why some methods of data collection are more efficient than others through a hands-on approach. This worksheet should take about 10 minutes to complete.

Students will then move to the *Choosing Samples* worksheet. Here, students will learn the importance of sampling in inferential statistics. Students will examine sample sets and why they may be good or bad. The worksheet provides students with a sample and information about the data collected. It then asks students guided questions about each specific sample to help lead them to make proper

conclusions about the sample and what may be wrong with it and how it could be fixed. This will help them develop a stronger idea of what makes a sample good. They will then choose their own sample and make their own inference based on a set of data. This will help familiarize students with the functionality of inferential statistics as they will begin to understand the importance and foundation of having a good sample to collect data from. They will have the opportunity to both see and explain why a sample set may be bad and also create their own sample and defend why the sample is a good sample to use. This worksheet should take no more than 10 minutes.

**Part III: Investigation Lesson: GapMinder [7]**

Students will work in groups of 2 or 3 on the *Exploring GapMinder and Making Inferences* worksheet. Students will have the opportunity to explore GapMinder and use it to analyze statistics and make inferences. This is a useful tool because it provides students with chance to see how statistics can be used on a larger scale and answer real world questions and solve problems. They also get to explore statistics and practice making inferences with real data that has been collected world-wide. They will use the interactive graphs and statistics to draw conclusions about the presented data and answer questions about it. They will also see the difference between single and multivariable statistics from a conceptual standpoint and visual level before being introduced to the concept later in the curriculum. This worksheet should take about 10 minutes for students to complete the worksheet.

**Explain:**

Throughout the lesson, students will be asked to explain their answers and discuss the reasoning behind their choices and decisions. In the *Statistics Concept Map* worksheet, students must explain the connections between the key concepts and definitions. In the *Descriptive Statistics, Choosing Samples*, and *Exploring GapMinder and Making Inferences* worksheets, students must explain their choices, describe the data sets and trends, and provide reasoning for how they arrived at a particular conclusion. Throughout the lesson, the teacher should be circulating the classroom to ensure that the students are focused and to address any misconceptions that the students might have.

**Elaborate:**

Students are able to explore real world statistics and their applications through GapMinder. They will see the importance of statistics and how it can be used to answer questions, to understand the world, and to solve problems. Through GapMinder, students will develop and enhance their skills in interpreting statistics and in making inferences by utilizing real data sets. All data, graphs and statistics on GapMinder is real data from across the globe and is representative of real populations from both the past and present. Gapminder is an interactive tool that allows students to explore what they are interested about in the world. Students can discover how statistics can answer questions and correct their misconceptions. Students may think or feel one way about a particular topic or think that they know the answer to a question, but this may change when they are properly informed and have information supported by statistics and data. They can also see how statistics can be used to make predictions and see where the world will be in 5, 10, or more years. Through the GapMinder activity, students can truly see the power of statistics.

**Evaluate:**

During whole class discussions and while the students are completing the worksheets, the teacher is able to check for surface level understanding and make sure that the class is all on the same page by listening to students' discussions and observing students' responses. These serve as the informal evaluations within this lesson. Formal evaluation can be done by checking for correctness in the students' worksheets.

**Enrich:**

This lesson can easily be extended into science classes such as biology, where data and statistics is collected. When dealing with sample sizes and collected data, it can be useful for students to understand the basics of statistics, how to choose samples, and how to make inferences from data. In a biology lab, students may be collecting data on fruit fly populations and genetics for example. Students can collect data that describes the population of fruit flies they have, but they can also collect inferential statistics where they need to draw conclusions and make assumptions about the population. This may include observing and determining which genes are dominant versus recessive.

**\*\*All associated documents are attached below\*\***

**\*\*Reference *Annotated Bibliography* on the very last page of this packet\*\***

Name: \_\_\_\_\_ Date: \_\_\_\_\_

Will Saving Poor Children Lead to Overpopulation? [6]

1. Use the space below to write your answer to the title question.
2. After watching the video, how did your answer change? If your answer did not change, explain why.
3. Why are statistics important?
4. What would happen if more of the world made statistical inference instead of assumptions?

Name: \_\_\_\_\_ Date: \_\_\_\_\_

Will Saving Poor Children Lead to Overpopulation? [6]

1. Use the space below to write your answer to the title question.
2. After watching the video, how did your answer change? If your answer did not change, explain why.
3. Why are statistics important?
4. What would happen if more of the world made statistical inference instead of assumptions?

Name: \_\_\_\_\_ Date: \_\_\_\_\_

### Statistics Concept Map [1]

Use the definition list on the next page to match the word to its definition by writing the corresponding letter inside the bubble. Then draw a line between words that connect with each other. On the line, write a brief reason of why or how they connect.

Random Sampling

Descriptive Statistics

Population

Biased Sample

Sample

Inferential Statistics

Data

Stratified Sampling

Name: \_\_\_\_\_ Date: \_\_\_\_\_

### Definition List [1]

**A-** are numbers that are used to summarize and describe data

**B-** refers to the information that has been collected from an experiment, a survey, a historical record, etc.

**C-** a small subset of a larger set of data

**D-** the larger set from which a sample may be drawn

**E-** using information from a sample to draw inferences about a population

**F-** a sample that is collected in such a way that some members of the intended population are less likely to be included than others.

**G-** requires every member of the population to have an equal chance of being selected into the sample. In addition, the selection of one member must be independent of the selection of every other member.

**H-** this method can be used if the population has a number of distinct “strata” or groups. In stratified sampling, you first identify members of your sample who belong to each group. Then you randomly sample from each of those subgroups in such a way that the sizes of the subgroups in the sample are proportional to their sizes in the population.



Name: \_\_\_\_\_ Date: \_\_\_\_\_

## Descriptive Statistics

### **Introduction:**

Officials in the city of Philadelphia are investigating whether global trends apply to the city's population. As a recent hire in the demographics department, you are tasked with compiling and analyzing data. To perform this task, you must understand the difference between the two categories of statistics that you will be working with: descriptive and inferential.

### **Directions:**

First you will collect data from your classmates and compile it. Collect data on your classmates' favorite music genre, social media platform and clothing brand. As a class, choose which music, social media platforms and brands you are going to poll. It is recommended to have a poll size no greater than 5 per category.

1. How are you going to keep track of the data you collect?
  
2. How do you want to collect information?

### **Data Sets:**

## Questions

1. Is there a better way to collect data that you collected from above?
2. Do the data sets correlate? Why or why not?
3. What does the data tell us?
4. What are other examples of descriptive data? What are they used for?
5. What is an example of a big famous descriptive data collection in US?





Name: \_\_\_\_\_ Date: \_\_\_\_\_

### Exploring GapMinder and Making Inferences [7]

Looking at global trends is a big part of statistics. One website that looks at these global trends is called GapMinder.

Using the web-link [GapMinder1](#), answer the following questions.

1. What two things are the graph comparing?
2. What does the size of the circle represent?
3. What trends can you infer from this data set?

Using the web-link [GapMinder2](#), answer the following questions.

1. What is does the graph show?
2. How is it different from the bubbles?
3. When would you use each type of graph?
4. How is this data different from the class data we collected?
5. What is the difference between Descriptive and Inferential Statistics?

Name: \_\_\_\_\_ ANSWER KEY \_\_\_\_\_ Date: \_\_\_\_\_

Will Saving Poor Children Lead to Overpopulation? [6]

1. Use the space below to write your answer to the title question.

Students preconception, either yes or no.

2. After watching the video, how did your answer change? If your answer did not change, explain why.

Based on students' above answer

3. Why are statistics important?

Statistics are important because we may have bias, and they can correct this bias. Statistics answer questions and provide potential solutions or identify potential problems based in fact instead of from assumptions or misconceptions.

4. What would happen if more of the world made statistical inference instead of assumptions?

People would be more informed and understand the world problems on a deeper level and have a better understanding of how to fix them.

Name: \_\_\_\_\_ ANSWER KEY \_\_\_\_\_ Date: \_\_\_\_\_

### Statistics Concept Map [1]

Use the definition list on the next page to match the word to its definition by writing the corresponding letter inside the bubble. Then draw a line between words that connect with each other. On the line, write a brief reason of why or how they connect.

Random Sampling  
G

Descriptive Statistics  
A

Population  
D

Biased Sample  
F

Sample  
C

Inferential Statistics  
E

Data  
B

Stratified Sampling  
H

Name: \_\_\_\_\_ ANSWER KEY \_\_\_\_\_ Date: \_\_\_\_\_

### Definition List [1]

**A-Descriptive statistics:** are numbers that are used to summarize and describe data

**B-Data:** refers to the information that has been collected from an experiment, a survey, a historical record, etc.

**C-Sample:** a small subset of a larger set of data

**D-Population:** the larger set from which a sample may be drawn

**E-Inferential Statistics:** using information from a sample to draw inferences about a population

**F-Biased sample:** a sample that is collected in such a way that some members of the intended population are less likely to be included than others.

**G-Random sampling:** requires every member of the population to have an equal chance of being selected into the sample. In addition, the selection of one member must be independent of the selection of every other member.

**H-Stratified sampling:** this method can be used if the population has a number of distinct "strata" or groups. In stratified sampling, you first identify members of your sample who belong to each group. Then you randomly sample from each of those subgroups in such a way that the sizes of the subgroups in the sample are proportional to their sizes in the population.



Name: \_\_\_\_\_ ANSWER KEY \_\_\_\_\_ Date: \_\_\_\_\_

## Descriptive Statistics

### Introduction:

Officials in the city of Philadelphia are investigating whether global trends apply to the city's population. As a recent hire in the demographics department, you are tasked with compiling and analyzing data. To perform this task, you must understand the difference between the two categories of statistics that you will be working with: descriptive and inferential.

### Directions:

First you will collect data from your classmates and compile it. Collect data on your classmates' favorite music genre, social media platform and clothing brand. As a class choose which music, social media platforms and brands you are going to poll. It's recommended no more than 5 categories per topic

They need you, the students to collect data:

1. How are you going to keep track of the data you collect?  
Students' choice: Students may choose to make a table, graph, make tallies, etc.
2. How do you want to collect information?  
Students' choice: Students may choose an effective method such as deciding as a class to take a show of hands, or they may choose to sample individually. They could also choose a method such as each group collects information then shares it with the class. There can be a lot of variation, some methods being more effective than others, but it's important to let students decide so they can practice going through the process and learn the importance of their decisions in data collection

### Data Sets:

(The following are examples; answers may vary based on how students choose to represent data)

Music

Pop	Country	HipHop	Indie/Folk	Alternative
10	3	5	4	3

Social Media

SnapChat	Instagram	Tumblr	TicTok	Facebook
8	10	2	5	0

## Brands

Nike	Adidas	American Eagle	Under Armour	Hollister
7	4	6	5	3

## Questions

1. Is there a better way to collect data you collected from above?

Dependent on classes choices on collecting data; generally, a class poll or show of hands would be the most effective with a data table.

2. Do the data sets correlate? Why or why not?

No, the sets do not relate to each other because they are used strictly to describe the class. One set of data does not imply another.

3. What does the data tell us?

It tells us about the class and their preferences. These preferences are dependent on the data collected by the class.

4. What are other types of descriptive data? What are they used for?

Age, height, hair color, are some other types. Descriptive data comprises of anything that can be used to describe characteristics of an entire population without numerical analysis.

5. What is an example of a big famous descriptive data collection in US?

The U.S. Census.

Name: \_\_\_\_\_ ANSWER KEY \_\_\_\_\_ Date: \_\_\_\_\_

### Choosing Samples

Now evaluate the samples below to determine whether they are good or bad. After practicing with the given sample, you will choose your own sample.

1. A coach is interested in how many cartwheels the average college freshmen at his university can do. Eight volunteers from the freshman class step forward. After observing their performance, the coach concludes that college freshmen can do an average of 16 cartwheels in a row without stopping.

- a. Is this a good or bad sample? Why?

This is a bad sample. There are only 8 students to represent the entire class. Also, students, who volunteer to do a cartwheel, most likely already know how to cartwheel and are confident enough to attempt it in front of the class.

- b. How can it be improved?

The coach should have students from all backgrounds and athletic abilities in the sample. The sample should be representative of the freshman class. For instance, if only 3% of the freshman class does cheerleading, then only 3% of the sample should be cheerleaders.

2. A substitute teacher wants to know how students in the class did on their last test. The teacher asks the 10 students sitting in the front row to state their latest test score. The teacher concludes from the students' reports that the class did extremely well.

- a. What is the sample? What is the population?

The sample is 10 students in the front row, and the population is the entire class.

- b. Can you identify any problems with choosing the sample in the way that the teacher did? How can the sample be improved?

This is a bad sample. Students who chose to sit in the front row tend to be more attentive and pay more attention, meaning they will most likely do better on a test than the average score. The teacher should have asked a random selection of students sitting in all different locations in the classroom.

3. You have been hired by the National Election Commission to examine how the American people feel about the fairness of the voting procedures in the U.S.

a. What is the population?

The population is all United States citizens who are eligible to vote.

b. Describe the ideal sample that you should take. Defend your decision.

The ideal sample would be fairly representative of all groups in the population. This means the percentages of these groups in the sample should be the same percentage as those in the entire population. Some examples of these group may be male, female, non-gender conforming, LGBTQIA+, race, religion, age etc. This allows the sample to truly representative of the population and take all types of opinions from all types of people into consideration.

(There may be some student variation)

c. How would you choose your ideal sample? Explain.

To choose participants in this sample, one can set aside how many people they need to have identifying with each group and then randomly select people from each group to participate in the survey.

(There may be some student variation in method)

d. What are the pitfalls of having to complete this study? Why? If none exist, explain why not.

One pitfall may be that the size of the population is huge. Hence, a large sample size would be needed to be accurately representative. There are also a lot of identity groups to consider as well as intersectionality of those groups.

Name: \_\_\_\_\_ ANSWER KEY \_\_\_\_\_ Date: \_\_\_\_\_

### Exploring GapMinder and Making Inferences [7]

Looking at global trends is a big part of statistics. One website that looks at these global trends is called GapMinder.

Using the web-link [GapMinder1](#), answer the following questions.

1. What two things are the graph comparing?  
Life expectancy and income
2. What does the size of the circle represent?  
Population size of the country
3. What trends can you infer from this data set?  
Higher income leads to longer life expectancy

Using the web-link [GapMinder2](#), answer the following questions.

1. What is does the graph show?  
Income/GDP growth over time
2. How is it different from the bubbles?  
It does not show as many factors or variables. It also shows a trend over time instead of comparing two things.
3. When would you use each type of graph?  
The first graph shows the viewer that multiple variables have a correlation to one-another. This can be used if you want to see cause and effect of two variables. The second graph is useful if you want to look at just one variable, and it can be used to see a trend within a specific statistical variable.
4. How is this data different from the class data we collected?  
Students' answers may vary, but expect any of the following: Information can be inferred from; different statistics can relate to each other; this is inferential statistics, etc.
5. What is the difference between Descriptive and Inferential Statistics?  
Descriptive statistics describe a population and collects information from the entire population while inferential statistics uses samples (subsets of a population) and makes inferences based on the information being analyzed.

## Annotated Bibliography

- [1] Lane, D. M. (n.d.). [2.0]. Retrieved from [http://onlinestatbook.com/Online\\_Statistics\\_Education.pdf](http://onlinestatbook.com/Online_Statistics_Education.pdf)  
This online textbook was used for excerpt within the Descriptive vs Inferential Statics lesson plan as part of the Measurements and Data Analysis module. This reference aided in the completion of providing definitions for the key concepts and definitions sections as well as for the concept map activity. It was used to complete the Choosing Samples worksheet. Examples from the textbook were used and modified to create questions about sampling for students to complete. This book was useful because of its layout and completeness. The lesson expands upon the material used form this book as it uses the material in the creation of worksheets and activities that are not provided in the textbook.
- [2] Standards Aligned System. (n.d.). Retrieved from <https://www.pdesas.org/>  
This website was used in each lesson in the Measurements and Data Analysis module to select proper Pennsylvania State standards, which are based in Common Core, that each lesson is centered around.
- [3] Standards for Mathematical Practice. (n.d.). Retrieved from <http://www.corestandards.org/Math/Practice/>  
This website used in every lesson in the Measurements and Data Analysis module to find Standards for Mathematical Practices that are applicable in each lesson.
- [4] Nsta. (n.d.). Disciplinary Core Ideas. Retrieved from <https://ngss.nsta.org/DisciplinaryCoreIdeasTop.aspx>  
This website was used in each lesson in the Measurements and Data Analysis module to select appropriate disciplinary core ideas set forth by the NSTA that are at the center of each lesson.
- [5] Nsta. (n.d.). Crosscutting Concepts. Retrieved from <https://ngss.nsta.org/CrosscuttingConceptsFull.aspx>  
This website was used in each lesson in the Measurements and Data Analysis module to selecting appropriate crosscutting concepts set forth by the NSTA that apply to each mathematics lesson.
- [6] Will saving poor children lead to overpopulation? (n.d.). Retrieved from <https://www.gapminder.org/answers/will-saving-poor-children-lead-to-overpopulation/>  
This video on GapMinder is used as an engagement in the Descriptive vs Inferential Statistics lesson in the Measurements and Data Analysis module. Questions were developed based on this video for students to answer.
- [7] GapMinder. (n.d.). Retrieved from <https://www.gapminder.org/>  
This is an online tool centered around data and statics that is used as an instructional aid and student exploration in the Descriptive and Inferential Statistics lesson in Measurements and Data Analysis module. A student activity was developed based on the tools provided by GapMinder.