

Description:

You are a Teaching Assistant working for a professor. She just administered a midterm, and half of her class is complaining about the test being unfair. She wants you to look at the exam scores and do an analysis the data by answering the questions about the data. You are responsible for completing a full report containing and explaining the related data analytics. This report is to be returned to the professor, so he/she can make an informed decision about the exam created based on your data analysis.

Students will be able to:

- Understand the statistical significance of mean, median, range, and mode
- Analyze a set of data using mean, median, range, and mode
- Make inferences from a data set
- Draw conclusions when given a graphical representation of a graph

Students will understand:

This lesson is designed to teach the students about the importance of mean, median, mode and range within statistics. They will learn about how the connections between these concepts can determine the strength and validity of a data set. Students will work with two sets of data. The first is both skewed and polarized. The majority of the values in the data set lie close to the extremes. Students will learn that they need to check all components (mean, median, mode and range) to ensure that the data is valid. The second set of data is a normal distribution of valid data. Students will explore this and compare it to the polarized data. Students will learn how to make these inferences through analyzing a physical data set and through analyzing a graphical representation of a data set. These statistical concepts are important for students to understand because it gives them the skills in determine the validity of statistics, data sets, and help prepare them to make informed decisions and carry out research of their own. Through this lesson, students will achieve these understandings through inquiry and discovery.

Key Definitions & Concepts:

- **Central tendency:** the central or typical value for a probability distribution
- **Mean:** the average of the data set
- **Median:** the number that is the midpoint of the data set
- **Mode:** the number that occurs most frequently in the set
- **Range:** the difference between the lowest and highest values in a data set

Standards [Copied from: 1]:

CC.2.4.HS.B.2 Summarize, represent, and interpret data on two categorical and quantitative variables.
CC.2.4.HS.B.5 Make inferences and justify conclusions based on sample surveys, experiments, and observational studies.

	Background Information	
--	------------------------	--

Prior Knowledge:

- Calculating mean, median, range, mode
- Definitions from previous lesson on Descriptive and Inferential Statistics

Mathematics Practices [Copied from: 2]:

- Model with Mathematics
- Use Appropriate Tools Strategically

Core Ideas [Copied from: 3]:

- Information Processing
- Optimizing the Design Solutions

Cross Cutting Concepts [Copied from: 4]:

- Scale, Proportion, and Quantity

Possible Preconceptions/Misconceptions:

Some students may struggle with understanding the difference between median and mean. Often times, they are similar numbers but not always. They are found using two different processes. Students may also not understand the utilization of mean, median, range, and mode in terms of their statistical significance. Since students are not typically asked about the importance of mathematical concepts, expect them to have an initial struggle when completing the analysis questions throughout this lesson. Most students will have an idea of what each concept is and how to find it, but they may not understand how they can be utilized in data analytics.

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

Engage [5]:

Students will take a brief quiz, *Review Quiz*, on the terms given from the concept map in the previous lesson. This allows students to review and recall important statistical definitions that will carry out throughout the *Measurement and Data Analysis* module. Students should first try to fill out the quiz without any resources. When completing the corrections, students should be given access to their concept maps from the previous lesson, so they can correct the *Review Quiz* in their own words. This serves as a way to reinforce understanding of the prior concepts in a recursive manner. This should take students 10 minutes to complete.

Explore:

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. 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: Data Analysis

Students will work through the *Data Analysis* worksheet in small groups of two or three. This activity has the students acting as Teaching Assistants for a professor. The professor has a list of test scores that need to be analyzed to analyze the statistical significance of the spread of data. This gives the students the opportunity to review and practice finding the mean, median, range and mode. Most importantly, the students will be taken through a sequence of conceptual questions designed for them to analyze the meaning behind these concepts with respect to the given data set. At the end of the guided analysis, students will summarize their conclusions by answering whether or not the test was fair. During this activity, the teacher should be circulating the room to observe student interactions and listen to their discussions, so he/she can address any misunderstandings or misconceptions as necessary. This worksheet should take 10 minutes for the students to complete and 5 minutes for the teacher to correct as a whole class.

Part III: Investigation Lesson: Reading Data on a Graph

For this activity, students will work through the *Reading Data on a Graph* worksheet in small groups of two or three. Students are provided with a graphical representation of data that is meant to show students' results after completing a retake exam that the professor gave to all of his/her classes. Students will approximate the mean, median, range, and mode by using the graph. The purpose of this activity is two-fold. It gives the students the opportunity to practice interpreting data in different forms, and it provides instructors with an introduction for the subsequent lesson covering distributions. Students are asked guided questions to analyze the information that they pulled from the graph in a similar format to the previous lesson's worksheet, *Data Analysis*. Students then have to determine whether or not the data spread is statistically valid and to explain why or why not. They also have to determine which test was most effective based on the information and analysis from both activities. During this activity, the teacher should be circulating the room to observe student interactions and listen to their discussions, so he/she can address any misunderstandings or misconceptions as necessary. This worksheet should take 10 minutes for the students to complete and 5 minutes for the teacher to correct as a whole class.

Explain:

During each activity in the exploration of this lesson, students are expected to explain their reasoning, choices, and/or answers (depending the worksheet). In the introduction part of the exploration, students must explain why they chose to connect two words. In the benchmark and investigation parts of the exploration, students must explain the purpose and meaning behind mean, median, range, and mode in terms of their significance in the data set at-hand. They must also explain what each concept tells them about the data set from an analytic standpoint. From this, the students must draw conclusions and support such via reasoning and sound explanations. The teacher should also be

asking probing questions during completion of the worksheets that prompt the students to explain what they are doing and why. This allows for students to learn through inquiry and discovery, which ensures that they gain deeper understanding of the learning goals of this lesson.

Elaborate:

Students can connect to the real-world example provided in this lesson - observing and analyzing data sets that represent exam scores. All students take exams, and, sometimes, they can be frustrated when exam scores are polarized like in the *Data Analysis* worksheet. This allows students to see why polarized test scores are statistically proven to be unfavorable because exam scores should result in a normalized distribution, not skewed. It also gives students the tools to provide these conclusions and explanations in a similar fashion to current data analysts.

Evaluate:

Students are evaluated both informally and formally throughout the entirety of this lesson. The informal evaluations occur during the open class discussions, and especially during the *Concept Map* activity. When the teacher is circulating the classroom, they are 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. The instructor gauge student progress through observing whether or not they are struggling to complete the activities. The formal assessments of this lesson are the worksheets, by checking for correctness, and the *Review Quiz*. The *Review Quiz* is an excellent assessment to determine if students retained information from the previous lesson.

Enrich:

This lesson can be differentiated into a business class. Statistics and data analytics are incredibly important in all aspects of business. Students can practice analyzing data sets and performing statistics to determine the meaning and significance behind the data given. Students can also infer the reasoning the explains why a data set shows unfavorable conditions based on their analysis and learn to present these findings to the organization or business that they completed data analytics for.

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

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

Name: _____ Date: _____

Review Quiz

In your own words, define each of the following keywords.

Descriptive statistics-

Data-

Sample-

Population-

Inferential Statistics-

Biased sample-

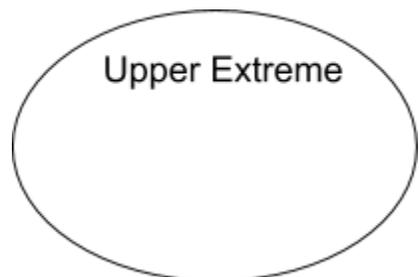
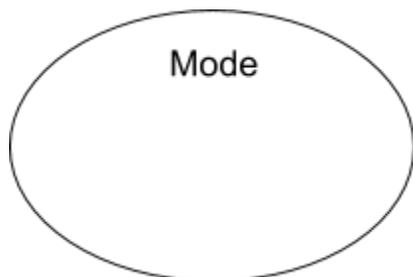
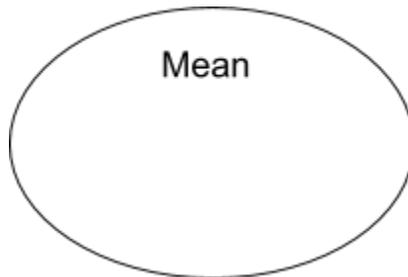
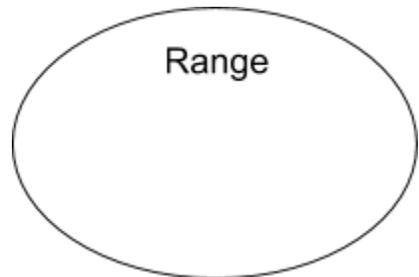
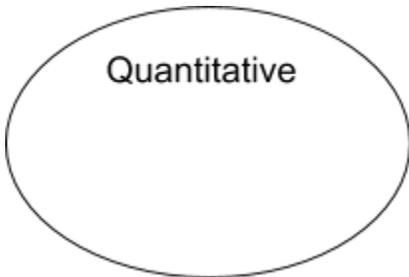
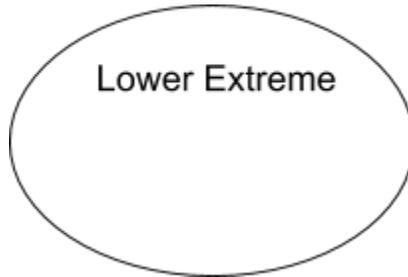
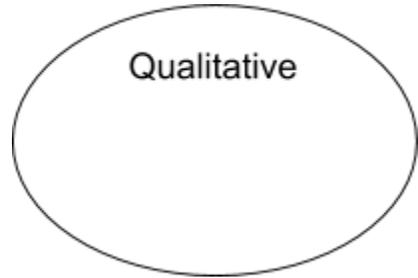
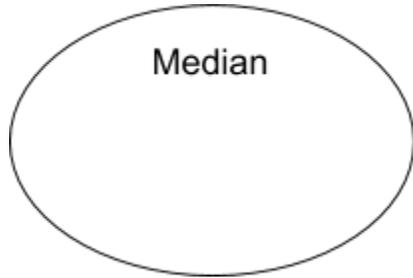
Random sampling-

Stratified sampling-

Name: _____ Date: _____

Statistics Concept Map

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.



Definition List

A: The average of the data set. Found by finding the sum of the data points then dividing by the number of data points

B: The number that is the midpoint of the data set

C: The number that occurs most frequently in the set

D: The difference between the lowest and highest values in a data set

E: The largest number in the data set

F: The smaller number in the data set

G: Scientific method of observation to gather non-numerical data. This type of research "refers to the meanings, concepts, definitions, characteristics, metaphors, symbols, and description of things" and not to their "counts or measures".

H: Information about quantities; that is, information that can be measured and written down with numbers.

Name: _____ Date: _____

Data Analysis

You are a Teaching Assistant working for a professor. She just administered a midterm, and half of her class is complaining about the test being unfair. She wants you to look at the exam scores and do an analysis on the data by answering the following questions.

Exam Scores: 95, 58, 90, 42, 57, 98, 62, 31, 92, 61, 97, 93, 68, 100, 48, 94, 50, 62

Calculate the mean test score.

What does this number tell you about the data qualitatively? (Hint: compare the mean score to the rest of the data)

What is the median score?

How is this different from the mean?

What are the extremes for data?

Find the range.

What does each range tell you about the data qualitatively?

What is the mode?

What does the mode tell you about the data qualitatively?

Why do we analyze statistics and data using mean, median, range, and mode?

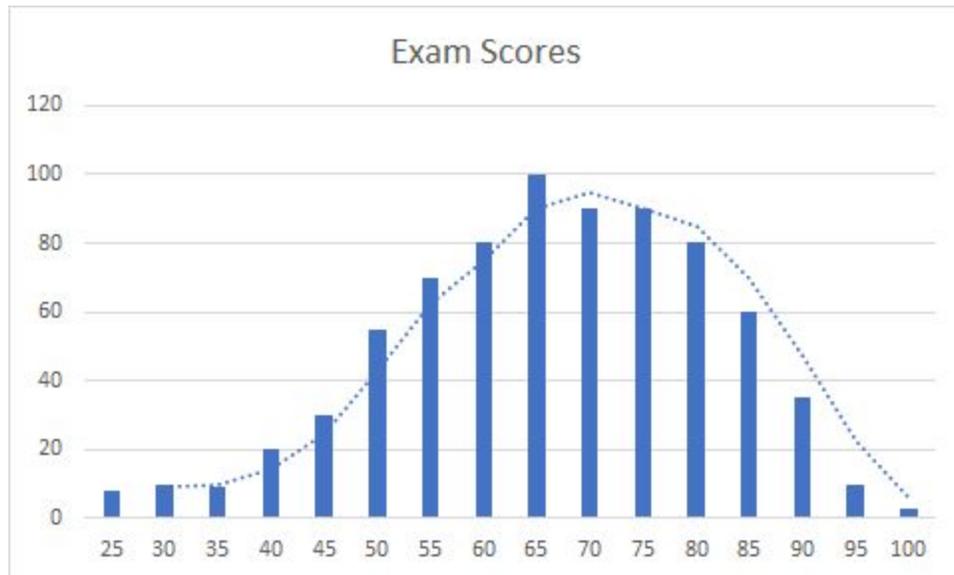
Based on the analysis of the data, do you think the data spread is good? Why or why not?

Do you think this teacher's test was an accurate or fair way to test the students? Why or why not?

Name: _____ Date: _____

Reading Data on a Graph

The professor has issued a retake for all of her sections of the class. Analyze this set data by answering the following questions.



What is the mode?

Approximately, what is the median exam score?

Approximately what is the mean exam score?

Are these numbers similar and near each other, or are they very different?

What is the significance of the data's proximity to one another?

What are the extremes of the data?

What is the range?

How is this data spread different from the previous?

Based on the analysis of the data do you think the data spread is good? Why or why not?

Do you think this teacher's new exam was an accurate way to test the students? Why or why not?

Which exam should the professor use in her class next year?

What can you say about the shape of the graph?

Name: _____ ANSWER KEY _____ Date: _____

Review Quiz

In your own words, define each of the following keywords.

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

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

Sample- a small subset of a larger set of data

Population- the larger set from which the sample is drawn

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

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.

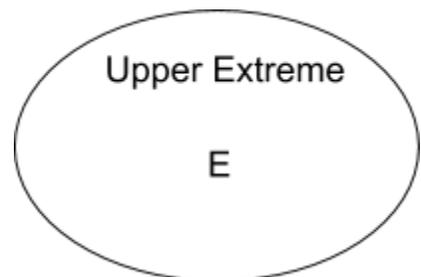
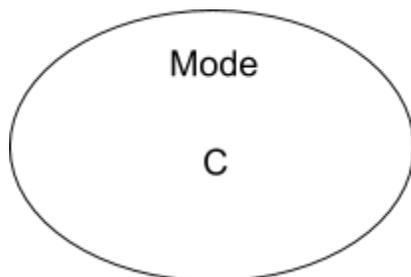
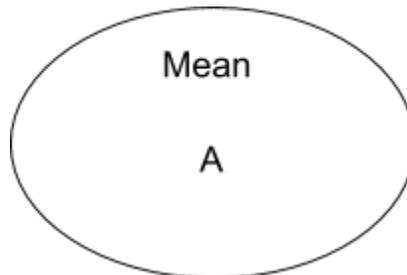
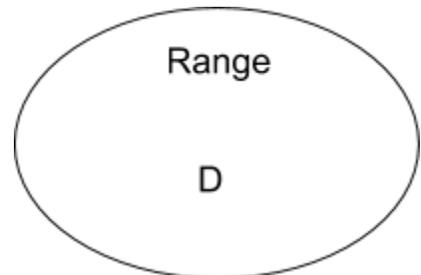
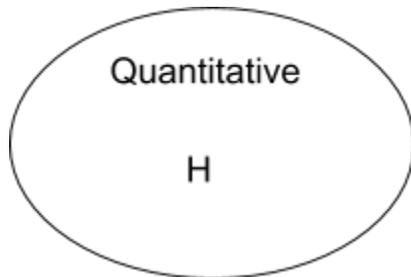
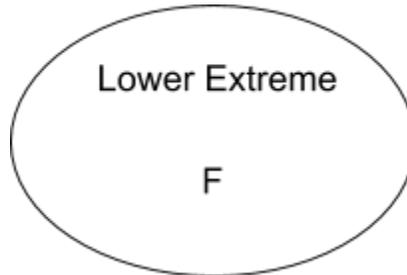
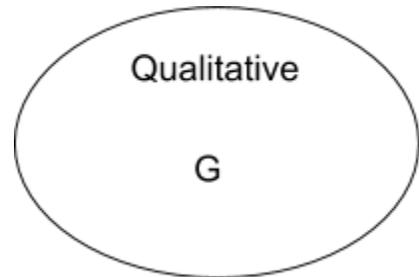
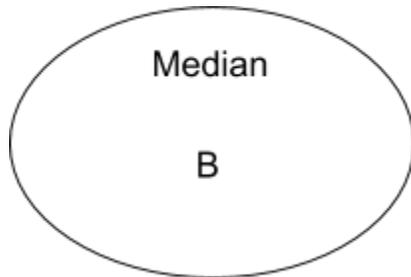
Random sampling- requires every member of the population to have an equal chance of being selected into the sample. In addition, the selection of one member must be independent of the selection of every other member. (Random samples, especially if the sample size is small are not necessarily representative of the entire population)

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: _____

Statistics Concept Map

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.



Definition List

A: Mean- The average of the data set. Found by finding the sum of the data points then dividing by the number of data points

B: Median- The number that is the midpoint of the data set

C: Mode- The number that occurs most frequently in the set

D: Range- The difference between the lowest and highest values in a data set

E: Upper Extreme- The largest number in the data set

F: Lower Extreme- The smaller number in the data set

G: Qualitative- Scientific method of observation to gather non-numerical data. This type of research "refers to the meanings, concepts, definitions, characteristics, metaphors, symbols, and description of things" and not to their "counts or measures".

H: Quantitative- Information about quantities; that is, information that can be measured and written down with numbers.

Name: _____ ANSWER KEY _____ Date: _____

Data Analysis

You are a Teaching Assistant working for a professor. She just administered a midterm, and half of her class is complaining about the test being unfair. She wants you to look at the exam scores and do an analysis on the data by answering the following questions.

Exam Scores: 95, 58, 99, 42, 57, 98, 62, 62, 92, 61, 97, 93, 68, 100, 48, 94, 50, 62

Calculate the mean test score.

$$\frac{(95+58+99+42+57+98+62+62+92+61+97+93+68+100+48+94+50+62)}{18} = 74.333$$

What does this number tell you about the data qualitatively? (Hint: compare the mean score to the rest of the data)

Nobody in the class got a score that either was the average or was near the average. Everybody either scored a lot higher (in the 90's) or lower (below a 68). The mean here is not a good representation of how students performed overall

What is the median score?

62

How is this different from the mean?

The median is the midpoint of the data and the mean is the sum of the scores divided by the total number of scores. Here, the median score is a lower score than the average.

What are the extremes for data?

Lower: 42 Upper:100

Find the range.

$$\underline{100 - 42 = 58}$$

What does each range tell you about the data qualitatively?

The range is pretty big meaning there is a large spread of data. The lowest score is very low while the highest is very high.

What is the mode?

62

What does the mode tell you about the data qualitatively?

The most frequent score is not close to the average, but it is the same number as the median which is also different from the mean. This shows a problem in the data set.

Why do we analyze statistics and data using mean, median, range, and mode?

It allows us to see the quality of data as well as see if something is wrong with it. If something is off we can start to try to make inferences as to why. In this case students only scored very high or low.

Based on the analysis of the data do you think the data spread is good? Why or why not?

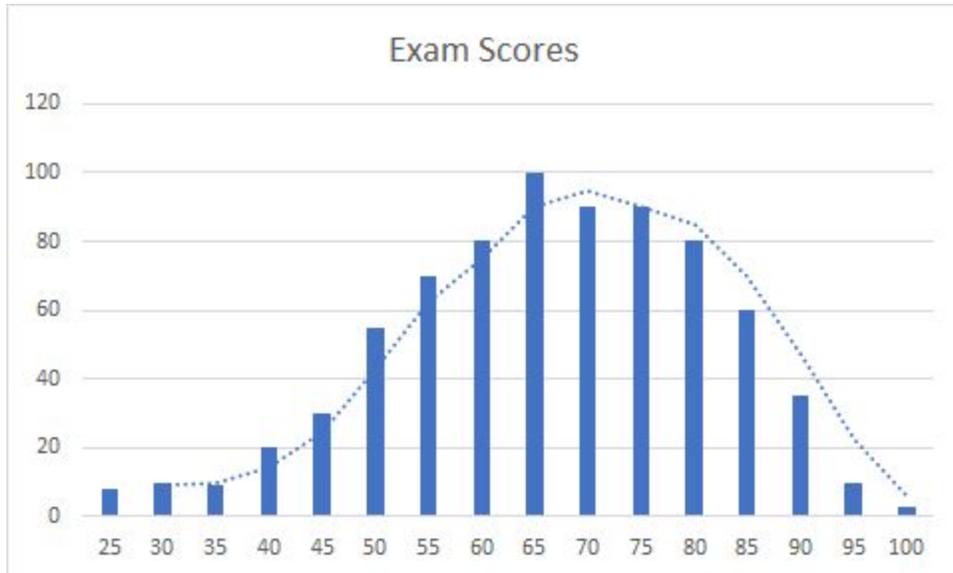
This data spread is inherently not good. This is because of the above analysis with the mean, median, and mode. The median and the average are not close to the same, and the most common exam score of 62 is also not close to the average. This shows that the data is polarized. If the spread was valid, then the mean, median, and mode would all be similar numbers or closer together.

Do you think this teacher's test was an accurate or fair way to test the students? Why or why not?

The test was either unfair in the fact that it favored one type of student to perform well and another poorly, or half the class was over prepared while the other was not. There could be a lot of reasons why this occurred.

Name: _____ ANSWER KEY _____ Date: _____

The professor has issued a retake for all of her sections of the class. Analyze this set data by answering the following questions.



What is the mode?

65: This can be seen on the graph as it is the score with the highest amount of students with that score.

Approximately, what is the median exam score?

Somewhere between 65 - 70: We can see, based on the shape of the graph and how many students scored high and low, that the middle score appears to be around 70. We can also find this more carefully by working our way from the outside of the graph in, moving in equal increments of students until we find the approximate representing the median.

Approximately, what is the mean exam score?

70: This can be read off of the graph because of the shape of the graph and because of how evenly the data is distributed. Hence, the average is going to be the peak of the graph. Students may come to this conclusion on their own, while others may run an approximate calculation based on the numbers in the graph.

Are these numbers similar and near each other, or are they very different?

They are closer to each other in value, and the mean and median are the same.

What is the significance of the data's proximity to one another?

It means that the data has a better spread and is less polarized.

What are the extremes of the data?

Lower: 25

Upper: 100

What is the range?

$75 = 100 - 25$: The upper extreme minus the lower extreme.

How is this data spread different from the previous?

There is a larger range but the mean, median, and mode are much more similar. Most people scored closer to the median/mean as opposed to the extremes.

Based on the analysis of the data do you think the data spread is good? Why or why not?

This data spread is valid and confirmed by the mean, median, and mode because they all have similar values. This means the most common score is similar to the middle score, which is also similar to the average score. This confirms that the data is not polarized.

Do you think this teacher's new exam was an accurate way to test the students? Why or why not?

The new exam showed better results statistically and seems to be the fairer exam.

Which exam should the professor use in her class next year?

This new exam is the one that the professor should continue to use.

What can you say about the shape of the graph?

It is highest in the middle and lowest at the extremes. Bonus: It is shaped like a bell curve

Annotated Bibliography

- [1] 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.
- [2] 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 Unit to find Standards for Mathematical Practices that are applicable in each lesson.
- [3] 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.
- [4] 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.
- [5] Lane, D. M. (n.d.). [2.0]. Retrieved from http://onlinestatbook.com/Online_Statistics_Education.pdf
This online textbook was used for excerpt within the Mean Median Range and Mode 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 in the previous lesson and the same definitions are used in answers keys of the engagement Review Quiz in this lesson. This book was useful because of its layout and completeness. The lesson expands upon the material used from this book as it uses the material in the creation of worksheets and activities that are not provided in the textbook.