

Data Driven

Overview

Description

Students will represent, compare, and analyze sales data from two businesses. In the process, students will calculate measures of central tendency and variability and interpret the results. Students will evaluate their data to make decisions and determine which company has the greatest income potential for a prospective employee.

Final Product: Students will prepare a written report detailing their job recommendation. The report will contain strong statistical evidence that supports their data-driven decision.

Subject

Mathematical Models with Applications

Task Level

Grades 11–12

Objectives

Students will:

- Analyze a situation involving career choices and identify information relevant to the situation.
- Understand how changes made to the data set will affect both measures of center.
- Create and use graphical representations of numerical data, including bar graphs and box-and-whisker plots, to organize data, describe patterns, and analyze the data.
- Calculate, apply, and interpret measures of central tendency and variability.
- Compare and interpret data through the use of statistical measures and draw conclusions from the data.
- Consider additional information as presented throughout the task and re-evaluate their decisions based on this new data.
- Analyze data and interpret the results to formulate conclusions.

Preparation

- Make a copy of the Student Notes for each student, and the handouts for each partner group. Make copies of grid paper to be used as needed or requested.
- Gather at least one calculator for every two students (calculators with graphing capabilities preferred).
- Review graphing data with technology if needed (i.e. entering data into lists, graphing the data as a bar graph and box-and-whisker plot, setting appropriate window values, etc.).
- Plan questions you will use to facilitate the Getting Started class discussion.
- Compile resources students may use for reference if review of vocabulary or calculations is needed (e.g., box-and-whisker plot, standard deviation, etc.).

Prior Knowledge

Students should be able to draw and interpret data in graphical representations, specifically bar graphs and box-and-whisker plots. They should be able to calculate measures of center and variability from a discrete set of data, and compare and interpret these measures. Students should be familiar with the meaning and use of standard deviation and have working knowledge of the process for calculating it.

Key Concepts and Terms

- Bar graphs
- Box-and-whisker plot
- Commission
- Data distribution
- Five-number summary (minimum value, 1st quartile, median, 3rd quartile, maximum value)
- Measures of center (mean, median)
- Range
- Scale
- Standard deviation
- Variability

Time Frame

This assignment can be done in about three 50-minute class periods, depending on the use of technology and the number of oral presentations given.

Instructional Plan

Getting Started

Learning Objectives

- Analyze a situation involving career choices and identify information relevant to the situation.
- Compute, compare, and interpret measures of central tendency.
- Understand how changes made to the data set will affect both measures of center.
- Compute, compare, and interpret measures of variability.
- Compare and interpret data through the use of statistical measures and draw conclusions from the data.

Procedure

1. Arrange students in groups of two and distribute the Student Notes.
2. Direct students to read the introduction found on page S-1.
3. Discuss the two career choices given, making sure students understand the concept of a 100 percent commission-based salary.
4. Direct students to discuss the Getting Started items #1-4 found on pages S-1 and S-2 with their partner.
5. Facilitate a whole-class discussion regarding the ideas raised in the Getting Started section. Include questions such as:
 - a. What is a typical, or representative, sales figure based on the data given?
 - b. Which measure of center, mean or median, best describes this set of data and why?
 - c. Is it possible to change one of the third-quarter monthly sales values without changing the median? Explain.
 - d. What is the greatest possible change that can be made without affecting the median value? Explain.
 - e. Is it possible to change one of the third-quarter monthly sales values without changing the mean? Explain.
 - f. Is it possible to change one of the third-quarter monthly sales values so that the means for both Harpoon Boats and Angler Haven are the same? Explain.

Investigating

Learning Objectives

- Create and use graphical representations of numerical data, including bar graphs and box-and-whisker plots, to organize data, describe patterns, and analyze the data.
- Compute, compare, and interpret measures of central tendency.
- Understand how changes made to the data set will affect both measures of center.
- Compute, compare, and interpret measures of variability.
- Compare and interpret data through the use of statistical measures and draw conclusions from the data.
- Consider additional information as presented throughout the task and re-evaluate their decisions based on this new data.

Procedure

1. Guide students to page SH-2, *Digging Deeper into the Data*. Direct students to examine the monthly sales data. Ask students to share their observations and solicit possible explanations for those observations.
2. Inform students that they are going to work with their partner to analyze the monthly sales data as directed on the *Data Exhibit Cards A-E*. Remind students that the goal of this task is to determine which job Rodney should accept and that the recommendation must be based on sound statistical evidence.
3. Distribute one set of *Data Exhibit Cards A-E* (SH-3 through SH-7) and a calculator to each pair of students. Note: Graphing calculators are recommended for efficiency due to the nature of the task and size of the data values.
4. Direct students to complete the *Data Exhibit Cards A-E* with their partner.
5. Provide graph paper if calculators do not have graphing capabilities or should students request it.
6. Circulate and monitor students' progress. Be prepared to direct students to additional resources (e.g. textbook, class notes, websites) if the student does not remember the vocabulary or recall how to calculate certain statistical measures.
7. Continually ask groups what job they will recommend and why. Ask them which representation and/or statistical measure best represents the data in this context and why.

8. Direct students to begin writing their report as outlined in the Drawing Conclusion section of this task found on page S-2 after completing the *Data Exhibit Cards*.

Drawing Conclusions

Learning Objectives

- Compare and interpret data through the use of statistical measures and draw conclusions from the data.
- Analyze data and interpret the results to formulate conclusions.

Procedure

1. Direct students to discuss with their partners items #1-3 from the Drawing Conclusions section on page S-2. When partners are in agreement with their responses, they should prepare a report that summarizes their conclusions.
2. Have students individually prepare a report that details their job recommendation for Rodney. Emphasize that the recommendation must be backed with strong statistical evidence and the significance of such evidence should be clearly articulated in the report.
3. Have some students present their recommendations to the class.
4. Following each presentation, facilitate class discussions that focus on the strengths of the data analysis. After all presentations, ask, *Whose report was most convincing to you and why?*
5. Follow up with a class discussion of other considerations that might influence Rodney's career choice (e.g., state income tax, cost of living).
6. Collect student work.

Scaffolding/Instructional Support

The goal of scaffolding is to provide support to encourage student success, independence, and self-management. Instructors can use these suggestions, in part or all together, to meet diverse student needs. The more skilled the student, however, the less scaffolding that he or she will need. Some examples of scaffolding that could apply to this assignment include:

- Review the vocabulary needed for this assignment.
- What is a typical, or representative, value that best portrays the sales data?
- How would you describe the shape of the monthly sales data? Do you see any trends in the data? Do you see any clusters or gaps in the data?
- Are the mean, median, and range easily obtained from your graphical representation?
- Can you reconstruct the exact data values from your graphical representation?
- Describe the algorithm used to calculate the mean. Why does this make sense? What does the mean monthly sales value represent in this context?
- If two data sets, with the same number of entries, have the same mean, what must be true of the values in the data set? (*The intent is for students to understand mean as an equal share or “balancing point” in the data. Therefore, students must connect that in order for the yearly mean to be the same for both dealerships, both dealerships must have the same total yearly sales.*)
- What is revealed about the shape of the distribution when the mean and median are very similar? Conversely, what would you expect to see in the shape of the distribution when the mean and median are vastly different?
- How will changing the scale of the graph affect the appearance of the data’s distribution? When comparing data representations, why should the same scale be used?
- For each dealership, about what percent of data is above the median? What percent of data is below the median? For each dealership, about how many months of the year fall below the median?
- In the box-and-whisker plot, about what percent of the sales year is represented in the 1st quartile? 2nd quartile? 3rd quartile? 4th quartile?
- Why does each quartile on the box plot have a different length when each quartile represents approximately the same number of monthly sales entries?
- For each dealership, which months fall within one standard deviation of the mean? How does that compare with a normal distribution?

Solutions

The solution provided in this section is intended to clarify the problem for instructors. This solution may not represent all possible strategies for approaching the problem or all possible solutions. It should be used for reference only.

Reports for the final job recommendations will vary. Some students may argue that Mississippi's Angler Haven is the best choice because the potential for monthly income will be more stable. Additionally, Angler Haven's mean is slightly higher. Others may argue that Rodney should plan to work at Harpoon Boat's during the tourist season so that he can make more money in a few short months. Regardless of the position taken, students must provide strong statistical evidence to support their recommendation. See the solutions to the *Data Exhibit Cards* for further analysis of the statistical evidence.

Getting Started

2. Student answers will vary. Examples of possible student answers include: state income tax, cost of housing, cost of food, cost of transportation, and other general costs of living for each of the two areas.
3. Student answers will vary. Students may notice that sales are increasing year after year in Mississippi and that the average of the sales for the years given is higher in Alaska than in Mississippi. They may also comment on the difference in variability between sales in the two states.
4. Student answers will vary. Students should justify their answers by identifying measures of central tendency, trends, and variance.

Data Exhibit A - Organizing the Data

Draw a graphical representation of the monthly sales data for both boat dealerships.

Representations will vary. Representations may include line graphs, bar graphs, circle graphs, histograms, scatterplots, line plots, stem-and-leaf plots, and box-and-whisker-plots. The intent of this card is for students to select their own representation based on what makes sense to them. Their choice of representation is a chance to assess their understanding of a representation's strengths and weaknesses. The next question will ask students to analyze the effectiveness of their representation. Later in the task, students will be specifically asked to create a bar graph and box-and-whisker plot.

What information does the representation provide? What information does the representation not provide?

Answers will vary depending on representation above. Press for responses to include some of the following:

- *Can the exact mean, median, mode, and range can be determined the representation?*
- *Can the exact data values be identified?*

- Can the shape of the data be readily observed?
- Can the number of data entries be retrieved?
- Can the appearance of your representation be manipulated if a different scale is used?

How do you think monthly sales data from previous years would compare to the data shown above? Explain.

Possible answer: It would be reasonable to expect similar sales figures in subsequent years. The weather patterns should remain fairly consistent; therefore, the tourism should continue to follow the same seasonal trends. There would be differences in the data but the differences should not vary greatly from year to year, assuming the economy remains relatively stable as well.

Data Exhibit B - Measures of Central Tendency

Calculate the mean and median monthly sales for each boat dealership.

	Alaska's Harpoon Boats Monthly Sales	Mississippi's Angler Haven Monthly Sales
Median	\$ 1,080.00	\$ 30,000.00
Mean	\$ 29,867.00	\$ 29,912.50

Compare each measure above to the mean and median sales figures for the 3rd quarter months only. In your opinion, which statistic better represents a typical monthly sales figure for each dealership? Why?

	Alaska's Harpoon Boats		Mississippi's Angler Haven	
	3rd Quarter Monthly Sales (July-Sept)	Monthly Sales (Jan-Dec)	3rd Quarter Sales (July-Sept)	Monthly Sales (Jan-Dec)
Median	\$ 61,000	\$ 1,080.00	\$ 54,000	\$ 30,000.00
Mean	\$ 58,500	\$ 29,867.00	\$ 55,000	\$ 29,912.50

Answers will vary. Possible answer: For Mississippi's Angler Haven, the median and mean are extremely close. This suggests a symmetrical and consistent distribution of the sales figures. Either measure, mean or median, would describe a typical sales figure for any given month. In contrast, Alaska's Harpoon Boats has very different measures of center. This suggests that the sales figures are not as consistent and clustering of the data values may occur. In the case of Harpoon Boats the data clusters at extremely low values and again at extremely high values. Perhaps the median value of \$1,080 is the most representative of the monthly sales data because, with the exception of a few months, the sales figures are extremely low. At no time during the year do the monthly sales figures gravitate anywhere near the actual mean (\$29,000-\$30,000).

Compare the mean values for both dealerships. What do you observe? Explain why this occurs?

The mean values for both companies are relatively similar in value. This is true for both the 3rd quarter sales and the monthly sales data. The reason is because despite the very different distribution of data, both companies had very similar total annual sales figures. For example, Alaska's Harpoon Boats had a total of \$358,410 for the year compared to Mississippi's Angler Haven, which had an annual sales figure of \$358,950. Overall, Angler Haven only outsold Harpoon Boats by a nominal \$540. When these annual sales figures are divided out or averaged over the same 12-month period, the means will be very close in value.

Describe how to change one value so the means for the two boat dealerships are the same.

Answers will vary as any data entry could be manipulated so that both companies had the same total annual sales. Possible answer:

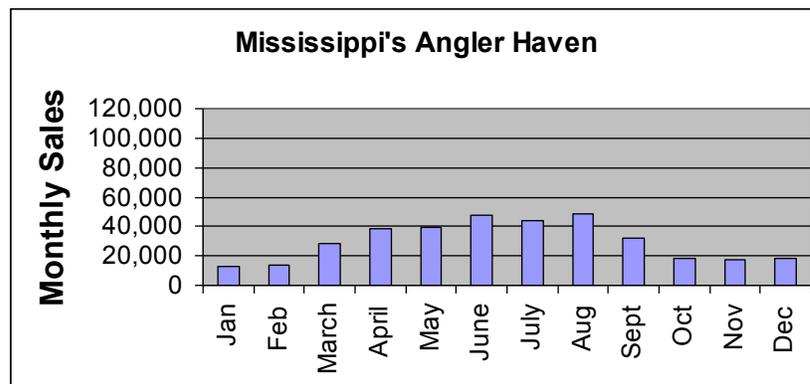
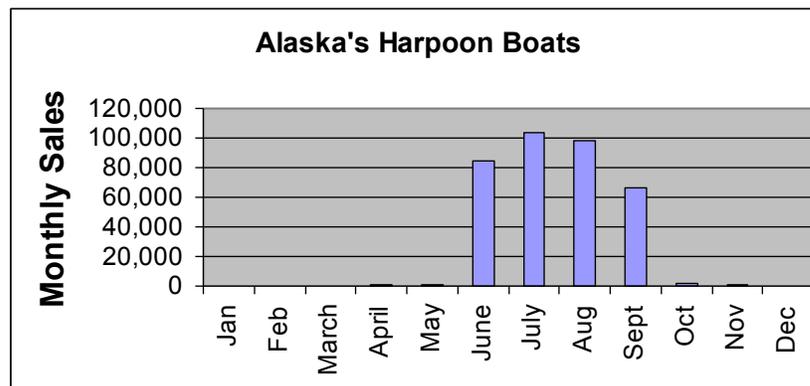
As described above, Angler Haven only outsold Harpoon Boats for the year by \$540. Therefore, increase any monthly sales figure for Harpoon Boats by \$540 (e.g. switch January's sales figure from \$0 to \$540). This strategy would bring both companies to the same end of year sales total of \$358,950 which when divided out across the same 12-month period would produce the same average or mean value. Similarly, deduct \$540 from one of Angler Haven's sales figures to produce a mean equivalent to that of Harpoon Boats.

Based on these statistics, what job do you recommend Rodney accept? Give statistical evidence that supports your recommendation.

Answers will vary.

Data Exhibit C - Shape of Data

Sketch a bar graph for each boat dealership's month sales data.



For each dealership, describe the shape and variability in the data.

Alaska's Harpoon Boats distribution shows extreme highs and extreme lows in the sales data ranging from \$0 to just over \$100,000. There are obvious spikes during the summer months while the rest of the months hover more towards \$0.

Mississippi's Angler Haven's distribution is a bell-shaped distribution reflecting more consistent data values. The graph does show a peak in the summer months but it is not as drastic as compared to Harpoon Boats. There are no gaps in the data. The data ranges from just above \$10,000 up to nearly \$50,000.

Based on this representation, what job do you recommend Rodney accept? Give statistical evidence that supports your recommendation.

Answers will vary.

Data Exhibit D – Box-and-Whisker Plot

Sketch a box-and-whisker plot for each boat dealership's monthly sales data on one scale. Be sure to include the five-number summary of data for each plot (e.g., minimum value, lower quartile value, median value, upper quartile value, maximum value).

Five number summaries	Harpoon Boats	Angler Haven
Minimum value	0	12500
Q1	225	17600
Median	1080	30000
Q3	75250	41500
Maximum value	104000	49000

Compare the length of the "box" for each boat dealership. What does this length represent and what does it mean in Rodney's situation?

The length of the box represents the values that lie between the first quartile and the third quartile. This is known as the Inter Quartile Range, or IQR, and represents 50% of the data distribution. Therefore, the longer the "box"- the greater the spread of the data. The box for Harpoon Boats is significantly larger than the box for Angler Haven. The reason is that the Angler Haven monthly sales were more consistent from month to month compared to Harpoon Boats. This would mean that for six months out of the year, one can expect monthly sales ranging anywhere from \$225 to \$75,250 at Harpoon Boats, compared to Angler Haven which would range from \$17,600 to \$41,500. For Rodney, this would mean he could expect a more consistent monthly salary at Angler Haven, but at Harpoon Boats his potential income would fluctuate greatly from month to month.

Based on these statistics, what job do you recommend Rodney accept? Give statistical evidence that supports your recommendation.

Answers will vary.

Data Exhibit E - Standard Deviation

Calculate the standard deviation for each boat dealership. Describe the process used for calculating this value.

The standard deviation for Harpoon Boats is \$42,147.00

The standard deviation for Angler Haven is \$13,034.60

To calculate the standard deviation, first find the mean of the data. Then subtract the mean from each data value to determine its distance from the mean. Next, square the differences. Squaring the differences ensures all the values are positive. Now find the total sum of all the squares. Find the average of these sums by dividing the sum by 11 (one less than the total number of data entries); the result is known as the variance. Finally, take the square root of the variance to obtain the standard deviation.

For each dealership, what percent of the monthly sales fell within one standard deviation of the mean?

	$\bar{x} - \sigma$	\bar{x}	$\bar{x} + \sigma$	Monthly sales within $\bar{x} \pm \sigma$	Percent of Monthly sales within $\bar{x} \pm \sigma$
Harpoon Boats	-\$12,280	\$ 29,867	\$ 72,014	Jan, Feb, Mar, Apr, May, Sept, Oct, Nov, Dec	75%
Angler Haven	\$ 16,877.90	\$ 29,912.50	\$ 42,947.10	Mar, Apr, May, Sept, Oct, Nov, Dec	58.3%

Compare the standard deviation for each boat dealership. Explain what these values mean in terms of Rodney's situation.

The standard variation for Harpoon Boats is more than three times greater than that of Angler Haven. This signifies that the monthly sales data at Harpoon Boats is not as closely centered around the mean value as compared to Angler Haven. In other words, the greater the standard deviation, the greater the variability in the data. For Rodney, it means that he will have to choose between a workplace where the monthly sales are significantly more stable or work for a company whose monthly sales history is more volatile (+/- \$13,000 vs. +/- \$42,147 from the average monthly sales figure).

Based on these statistics, which job do you recommend? Give statistical evidence that supports your recommendation.

Answers will vary.

TCCRS Cross-Disciplinary Standards Addressed

Performance Expectation	Getting Started	Investigating	Drawing Conclusions
<i>I. Key Cognitive Skills</i>			
A.1. Engage in scholarly inquiry and dialogue.	✓		
B.2. Construct well-reasoned arguments to explain phenomena, validate conjectures, or support positions.	✓	✓	✓
B.4. Support or modify claims based on the results of an inquiry.		✓	✓
D.1. Self-monitor learning needs and seek assistance when needed.	✓	✓	✓
D.2. Use study habits necessary to manage academic pursuits and requirements.	✓	✓	✓
D.3. Strive for accuracy and precision.	✓	✓	✓
D.4. Persevere to complete and master tasks.	✓	✓	✓
E.1. Work independently.	✓	✓	✓
E.2. Work collaboratively.	✓		✓
<i>II. Foundational Skills</i>			
A.4. Identify the key information and supporting details.	✓		
B.1. Write clearly and coherently using standard writing conventions.			✓
C.6. Design and present an effective product.			✓
D.1. Identify patterns or departures from patterns among data.	✓	✓	✓
D.2. Use statistical and probabilistic skills necessary for planning an investigation, and collecting, analyzing, and interpreting data.	✓	✓	✓

TCCRS Mathematics Standards Addressed

Performance Expectation	Getting Started	Investigating	Drawing Conclusions
<i>IV. Measurement Reasoning</i>			
D.1. Compute and use measures of center and spread to describe data.	✓	✓	✓
D.2. Apply probabilistic measures to practical situations to make an informed decision.	✓	✓	✓

TEKS Standards Addressed

<i>Data Driven - Texas Essential Knowledge and Skills (TEKS): Math</i>
<p>111.36.c.2. The student uses graphical and numerical techniques to study patterns and analyze data. The student is expected to:</p> <p>111.36.c.2.A. interpret information from various graphs, including line graphs, bar graphs, circle graphs, histograms, scatterplots, line plots, stem and leaf plots, and box-and-whisker plots to draw conclusions from the data.</p> <p>111.36.c.2.B. analyze numerical data using measures of central tendency, variability, and correlation in order to make inferences.</p>
<p>111.36.c.4. The student uses probability models to describe everyday situations involving chance. The student is expected to:</p> <p>111.36.c.4.A. compare theoretical and empirical probability.</p>

Data Driven

Introduction

Rodney comes from a long line of boat salesmen. His father, two uncles, a brother, and his grandfather have sold recreational boats for a living. Now Rodney is planning to enter the family business. For his first full time job, Rodney can either work for his brother in Mississippi or move to Alaska to work for his uncle. Rodney wants to choose the job that will yield the highest income.

As a salesperson, Rodney's salary will be based entirely on commission. Rodney has arranged to get the same commission percentage at either job.

In this assignment, you will help Rodney decide whether to work at his uncle's boat dealership near his own hometown of Cedar Lake, Mississippi, or to go to Alaska and work for his brother. Remember, the main factor in Rodney's decision will be job-earning potential. Your task is to analyze the sales figures of the dealerships and the trends of monthly boat sales and then recommend the job Rodney should accept. Your final work product will be a written report that uses statistical evidence to justify your recommendation.

Directions

Getting Started

1. Read the Introduction above that describes Rodney's situation.
2. Brainstorm with a partner what additional information would be useful in making a recommendation. Be prepared to share your ideas with the class.
3. With your partner, review the data provided in the table below. Be prepared to discuss the following:
 - a. Based on this data, which job would you recommend for Rodney? Be prepared to provide evidence that supports your conclusion.

Total Annual Sales	Alaska	Mississippi
2009	630,000	648,000
2010	775,000	708,000
2011	624,000	732,000
2012	846,000	768,000

- Rodney's father told him that he shouldn't base his decision solely on total annual sales figures. Rodney dug a little deeper and found the following monthly sales figures for the specific cities where his brother and uncle own their businesses.

Monthly Sales (3 rd Quarter)	<i>Harpoon Boats</i> Whalebone, Alaska	<i>Angler Haven</i> Cedar Lake, Mississippi
July	61,000	54,000
August	68,500	59,000
September	46,000	52,000

Analyze with your partner the sales data given above. How does this data influence the recommendation you will make? Be prepared to provide statistical evidence that supports your recommendation.

Investigating

- With your partner, examine the monthly sales data given on page SH-1.
- With your partner, complete the *Data Exhibit Cards A-E* on page SH-2 through SH-6. Calculate the statistical measures or representations as indicated on each card and think deeply about what information can be deduced from the data.

Drawing Conclusions

- With your partner, analyze the data from the *Data Exhibit Cards*. Based on this data, decide which job to recommend.
- With your partner, determine which data representation and/or calculation was most helpful in your decision.
- With your partner, consider other factors that might influence your decision. What other kinds of data might you need to research before finalizing your decision?
- On your own, prepare a report that details your recommendation for Rodney. Be sure to include any statistical evidence that strongly supports your recommendation. Additionally, if warranted, include any other possible data that Rodney should gather before making his final decision. Be prepared to present your final recommendation and supporting evidence to the class.
- Listen carefully to the oral recommendations given by your peers. If you were Rodney, whose report would you find most convincing? Why?

Monthly Sales Data

Rodney’s father explained to him that the two cities are very different from each other. Whalebone, Alaska is a popular summer resort town. During the summer months, the population of Whalebone, including tourists, swells to 35,000. As the weather turns cold, the population of Whalebone drops to about 1800. From October to April, there are almost no tourists in Whalebone. In contrast, Cedar Lake, Mississippi has warm weather at least 10 months out of the year. Even in the winter, anglers still like to fish in the warm Gulf of Mexico waters. The population of Cedar Lake is about 20,000 with a steady tourist industry year-round. (Pretend you are working on this problem before the 2010 gulf oil disaster. Work through the problem. When you finish, you’ll come back and reflect on how the gulf oil disaster changes the situation.)

Monthly Sales	<i>Harpoon Boats</i> Whalebone, Alaska	<i>Angler Haven</i> Cedar Lake, Mississippi
January	0	12,500
February	0	14,000
March	450	28,000
April	800	38,500
May	1200	39,000
June	84,500	48,000
July	104,000	44,000
August	98,500	49,000
September	66,000	32,000
October	2,000	18,750
November	960	17,200
December	0	18,000
Yearly Total		

Data Exhibit A - Organizing the Data

Draw a graphical representation of the monthly sales data for both boat dealerships.

What information does the representation provide? What information does the representation not provide?

How do you think monthly sales data from previous years would compare to the data shown above? Explain.

Data Exhibit B - Measures of Central Tendency

Calculate the mean and median monthly sales for each boat dealership.

Compare each measure above to the mean and median sales figures for the 3rd quarter months only. In your opinion, explain which statistic best represents a typical monthly sales figure for each dealership?

Compare the mean values for both dealerships. What do you observe? Explain why this occurs?

Describe how to change one value so the means for the two boat dealerships are the same.

Based on these statistics, what job do you recommend Rodney accept? Give statistical evidence that supports your recommendation.

Data Exhibit C - Shape of Data

Sketch a bar graph for each boat dealership's month sales data.

For each dealership, describe the shape and variability in the data.

Based on this representation, what job do you recommend Rodney accept? Give statistical evidence that supports your recommendation.

Data Exhibit D – Box-and-Whisker Plot

Sketch a box-and-whisker plot for each boat dealership's monthly sales data on one scale. Be sure to include the five-number summary of data for each plot (e.g. minimum value, lower quartile value, median value, upper quartile value, maximum value).

Compare the length of the "box" for each boat dealership. What does this length represent and what does it mean in Rodney's situation?

Based on these statistics, which job do you recommend Rodney accept? Give statistical evidence that supports your recommendation.

Data Exhibit E - Standard Deviation

Calculate the standard deviation for each boat dealership. Describe the process used for calculating this value.

What does the standard deviation tell you about a set of data?

For each dealership, what percent of the monthly sales fell within one standard deviation of the mean?

Compare the standard deviation for each boat dealership. Explain what these values mean in terms of Rodney's situation.

Based on these statistics, what job do you recommend Rodney accept? Give statistical evidence that supports your recommendation.