# Chapter 3: Statistics

## Section 3.2 Describing Data

### Exercises 3.2

1. True or False: The bars of a histogram should always touch. ADD
2. True or False: The bars of a bar graph should always touch. ADD
3. Is the data described categorical or quantitative?
   1. In a study you ask the subjects their age in years.
   2. In a study you ask the subjects their gender.
   3. In a study you ask the subjects their ethnicity.
   4. The daily high temperature of a city over several weeks.
   5. A person’s annual income.
4. Is the data described categorical or quantitative?
   1. In a study you ask the subjects how many siblings they have.
   2. In a study you ask the subjects what their favorite movie genre is.
   3. In a study to measure the subjects’ blood pressure.
   4. The daily rainfall in a city over several weeks.
   5. In a study you ask the subjects the amount they spend on housing each month. ADD
5. What types of graphs are used for categorical data? ADD
6. What types of graphs are used for quantitative data? ADD
7. A group of adults were asked how many children they have in their family. The bar graph to the right shows the number of adults who indicated each number of children.
   1. How many adults had 3 children?
   2. How many adults where questioned?
   3. What percentage of the adults questioned had 0 children?
8. Jasmine was interested in how many days it would take a DVD order from Netflix to arrive at her door. The graph shows the data she collected.
   1. How many movies took 2 days to arrive?
   2. How many movies did she order in total?
   3. What percentage of the movies arrived in one day?
9. This relative frequency bar graph shows the percentage of students who received each letter grade on their last English paper. The class contains 20 students. How many students earned an A on their paper?
10. This relative frequency bar graph shows the percentage of each drink type served over the weekend at a local coffee shop. There were 120 drinks served in total. How many served drinks were lattes? ADD
11. Corey categorized his spending for this month into four categories: Rent, Food, Fun, and Other. The percentages he spent in each category are pictured here. If he spent a total of $2,600 this month, how much did he spend on rent?
12. Habiba categorized the amount of time spent each week into 5 categories: Work, Travel, Housework, Leisure, and Sleep. If there are a total of 168 hours each week, how many hours does Habiba spend travelling each week? ADD
13. In a survey[[1]](#footnote-1), 1012 adults were asked whether they personally worried about a variety of environmental concerns. The number of people who indicated that they worried “a great deal” about some selected concerns is listed below.
    1. Is this categorical or quantitative data?
    2. Make a bar chart for this data.
    3. Can we make a pie chart for this data?

|  |  |
| --- | --- |
| **Environmental Issue** | **Frequency** |
| Pollution of drinking water | 597 |
| Contamination of soil and water by toxic waste | 526 |
| Air pollution | 455 |
| Global warming | 354 |

1. In a survey, 2056 adults were asked about their views on immigration. The percent of people who responded that immigrants to the United States are making each of the following situations in the country better are listed below.
   1. Is this categorical or quantitative data?
   2. Make a relative frequency bar chart for this data.
   3. Can we make a pie chart for this data? ADD

|  |  |
| --- | --- |
| **Situation** | **Relative Frequency (%)** |
| Food, music and the arts | 57 |
| The economy in general | 43 |
| Social and moral values | 31 |
| Job opportunities for you and your family | 19 |
| Taxes | 20 |
| Crime | 7 |

1. The following table is from a sample of five hundred homes in Oregon that were asked the primary source of heating in their home.

|  |  |
| --- | --- |
| **Type of Heat** | **Relative Frequency (%)** |
| Electricity | 33 |
| Heating Oil | 4 |
| Natural Gas | 50 |
| Firewood | 8 |
| Other | 5 |

* 1. How many of the households heat their home with firewood?
  2. What percent of households heat their home with natural gas? ADD

1. The following table is from a sample of 50 undergraduate students at Portland State University.

|  |  |
| --- | --- |
| **Class** | **Relative Frequency (%)** |
| Freshman | 18 |
| Sophomore | 13 |
| Junior | 23 |
| Senior | 46 |

* 1. What percent of the sampled students are below senior class?
  2. How many of the sampled students are freshmen? ADD

1. A group of adults were asked how many cars they had in their household.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 4 | 2 | 2 | 1 | 2 | 3 | 3 | 1 | 4 | 2 | 2 |
| 0 | 2 | 1 | 3 | 2 | 2 | 1 | 2 | 3 | 5 | 1 | 2 |

* 1. Is this categorical or quantitative data?
  2. Make a frequency table for the data using class intervals of size 2 starting from 0.
  3. Construct a histogram of the data.
  4. Describe the shape of the data.

1. The table below shows the scores on a math test.

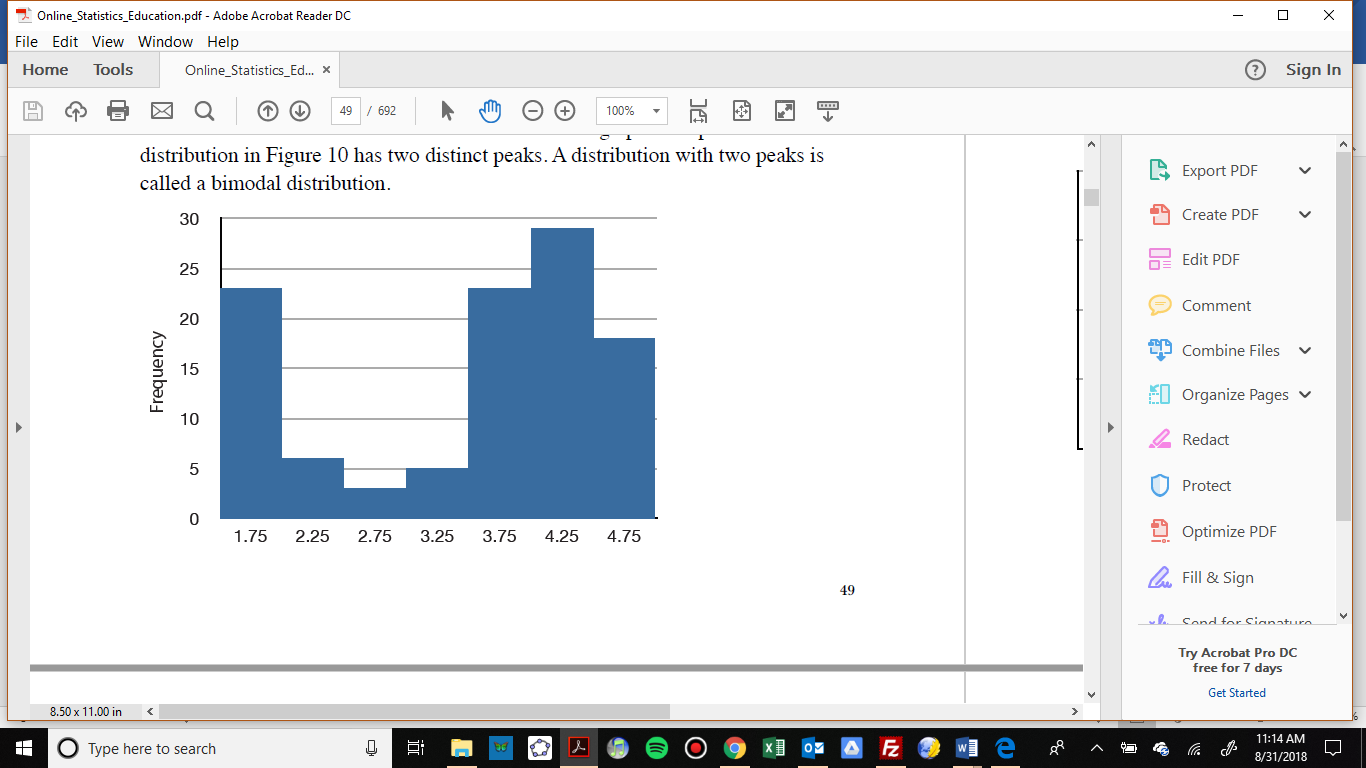
|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 82 | 55 | 51 | 97 | 73 | 79 | 100 | 60 | 71 | 85 | 78 | 59 |
| 90 | 100 | 88 | 72 | 46 | 82 | 89 | 70 | 100 | 68 | 61 | 52 |

* 1. Is this categorical or quantitative data?
  2. Make a frequency table for the data using class intervals of size 10 starting from 40.
  3. Construct a histogram of the data.
  4. Describe the shape of the data.

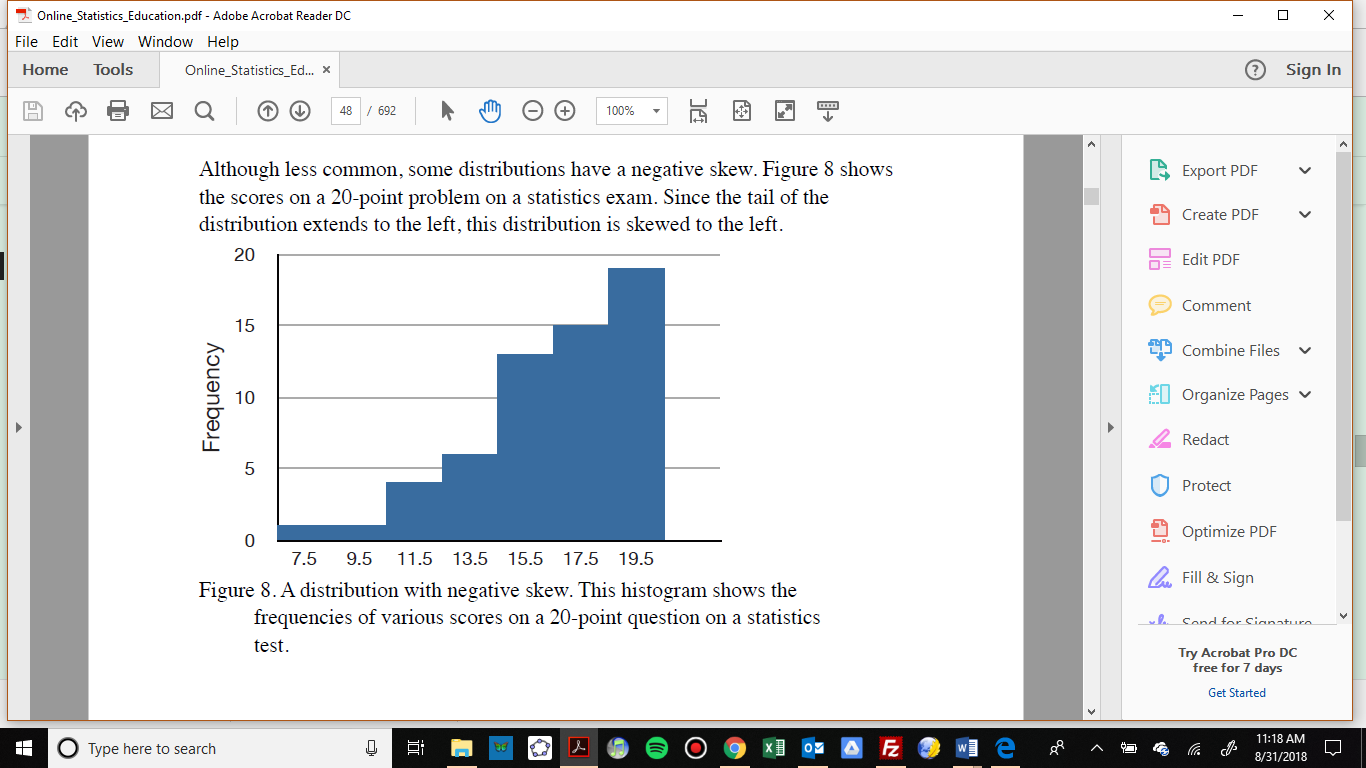
1. This graph shows the number of adults and kids who prefer each type of soda. There were 130 adults and kids surveyed. Discuss some ways in which the graph could be improved.
2. A poll was taken asking people if they agreed with the positions of the 4 candidates for a county office. Does this pie chart present a good representation of this data? Explain.
3. Why is this a misleading or poor graph? ADD This is a bar graph and is titled "Favorite Drinks". The x-axis is labeled Drink Type and includes colas, lemon flavored, root beer, teas, coffee, and other. The y-axis is labeled frequency and has no scale.
4. Why is this a misleading or poor graph? ADD

This is a bar graph and is titled "Profit During First Half of Year". The x-axis is labeled Year and includes January, February, March, April, May, and June. The y-axis is labeled Profits in $. The scale goes from 0 to 4500 by 500. The bar for each month is labeled with the following values: January 4230, February 3760, March 2670, April -1320, May 750, and June 1560. The bar for April is colored red while the remaining bars are colored blue. 

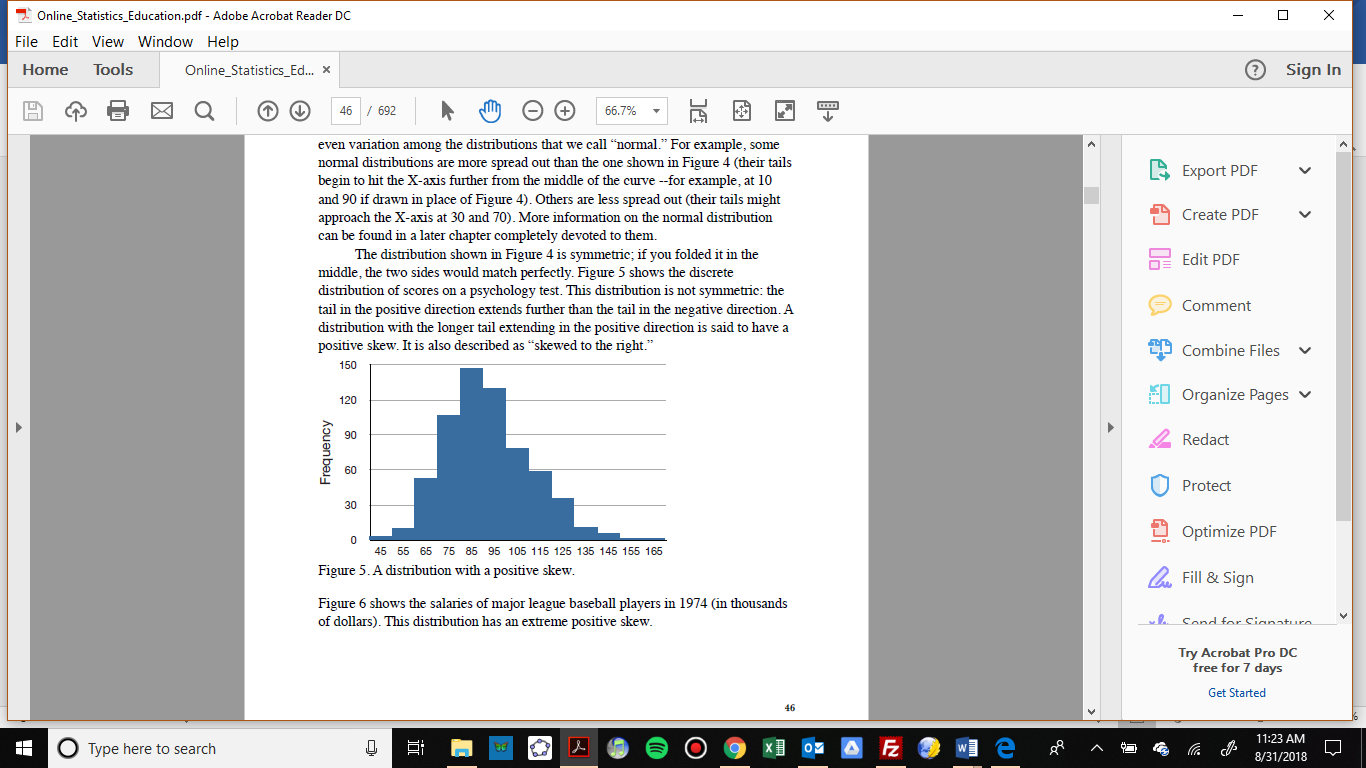
1. Match each description to one of the graphs.
   1. Normal distribution
   2. Positive or right skewed
   3. Negative or left skewed
   4. Bimodal



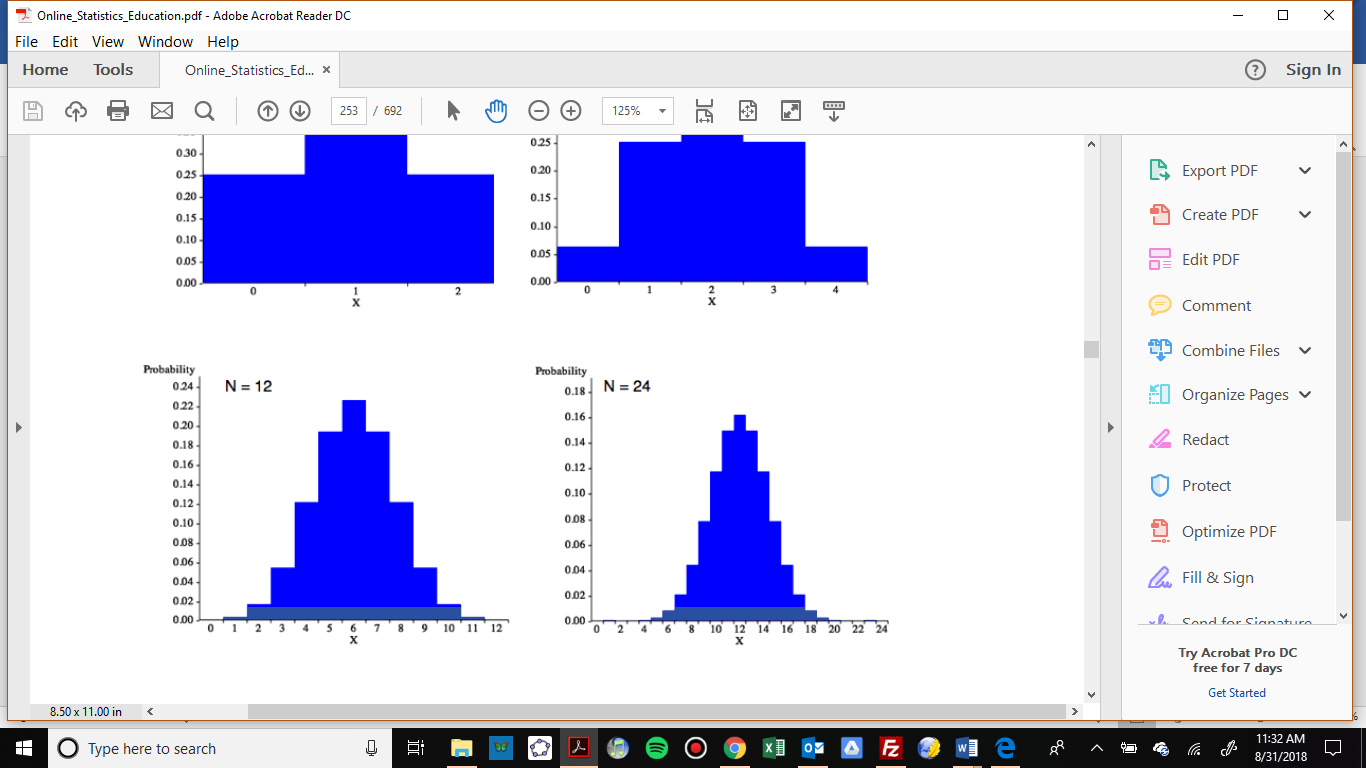
The frequency of times between eruptions of the Old Faithful geyser.



Scores on a 20-point statistics quiz.

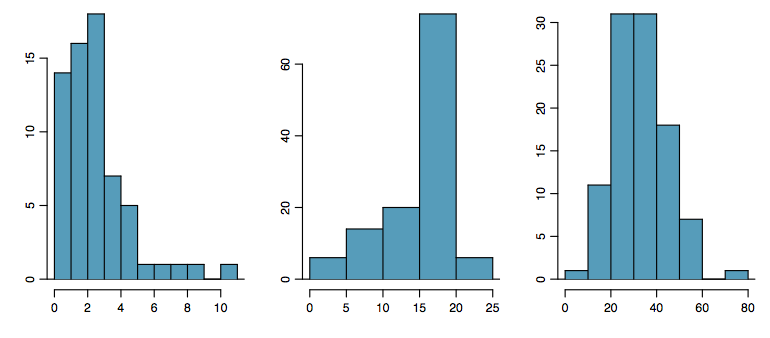


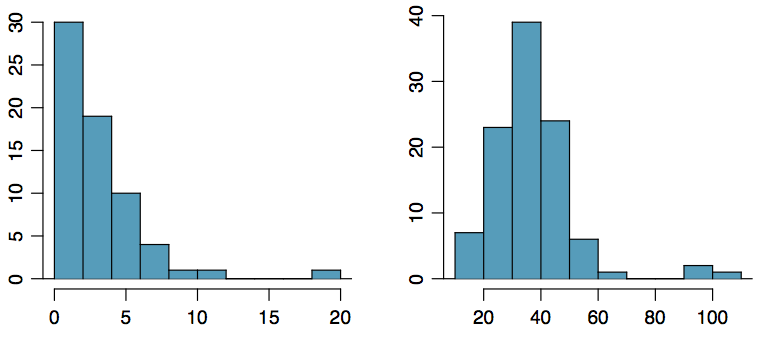
Distribution of scores on a psychology test.

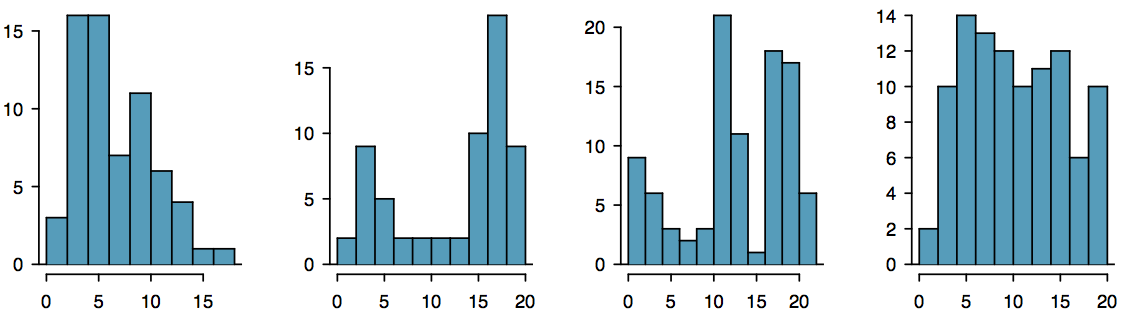


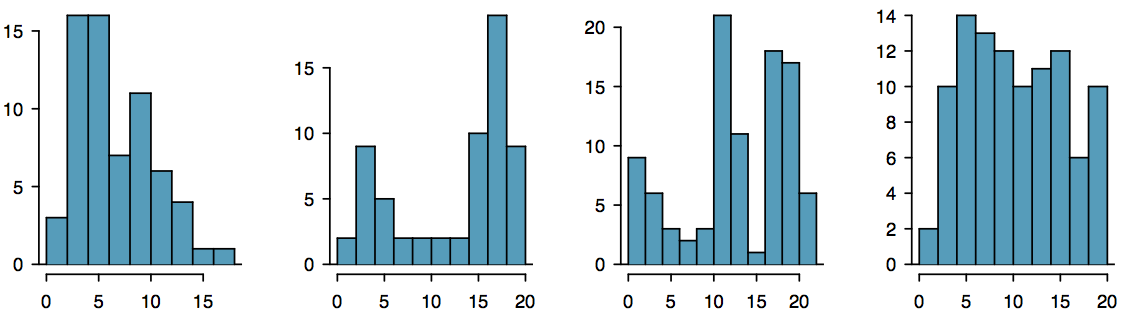
The number of heads in 24 sets of 100 coin flips.

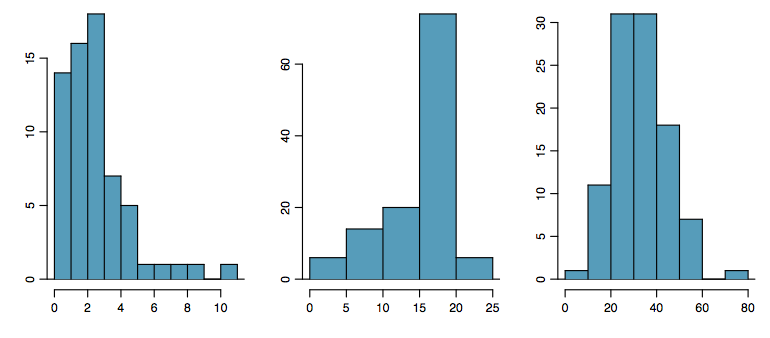
1. Write a sentence or two to describe each distribution in terms of modality, symmetry, and outliers.











1. Studies are often done by pharmaceutical companies to determine the effectiveness of a treatment. Suppose that a new cancer drug is currently under study. Of interest is the average length of time in months patients live once starting the treatment. Two researchers each follow a different set of 40 cancer patients throughout their treatment. The following data (in months) are collected.
   1. Create a histogram for each dataset, using the same class intervals and scales so you can compare them.
   2. Compare and contrast the two distributions.

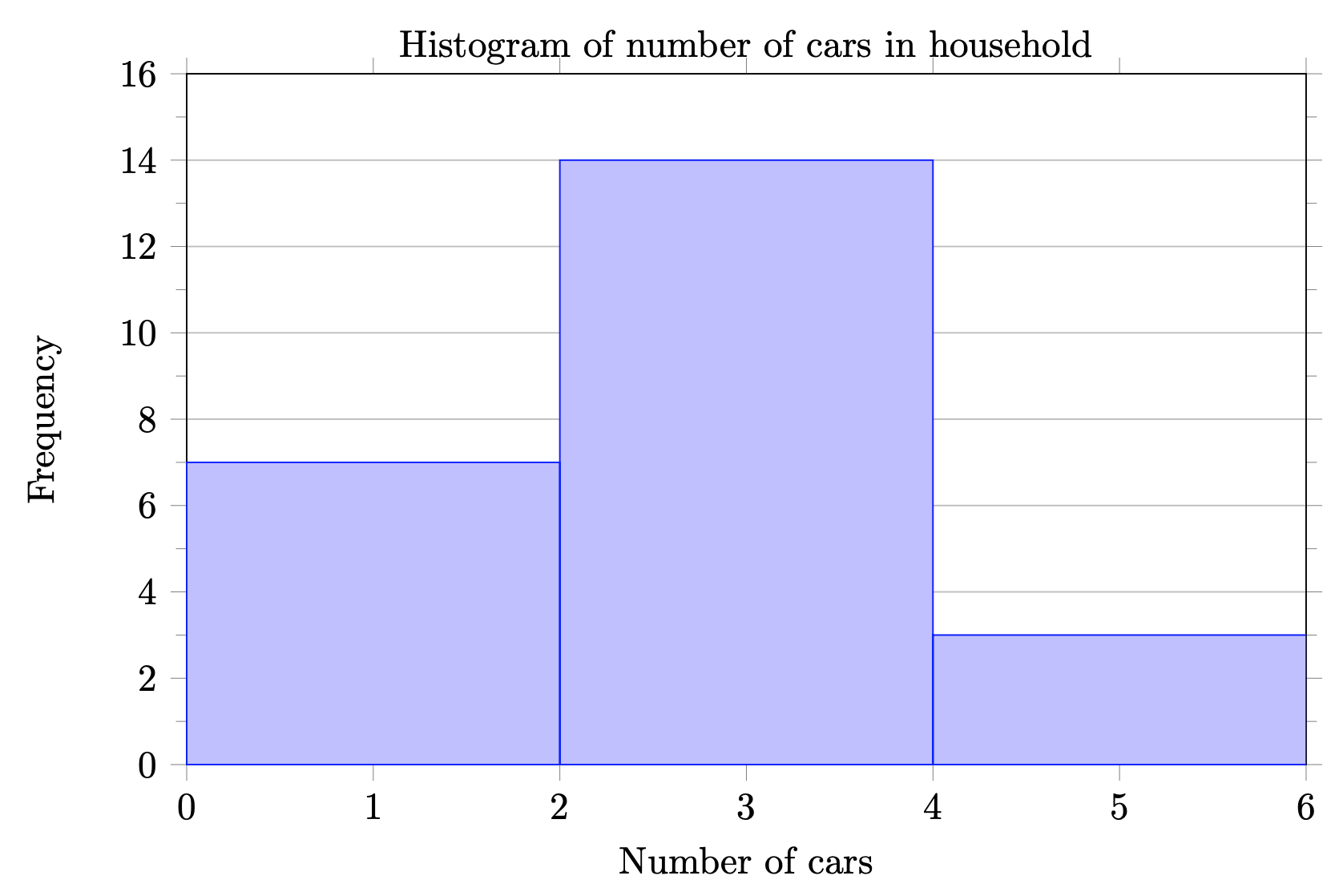
Researcher 1: 3, 4, 11, 15, 16, 17, 22, 44, 37, 16, 14, 24, 25, 15, 26, 27, 33, 29, 35, 44, 13, 21, 22, 10, 12, 8, 40, 32, 26, 27, 31, 34, 29, 17, 8, 24, 18, 47, 33, 34

Researcher 2: 3, 14, 11, 5, 16, 17, 28, 41, 31, 18, 14, 14, 26, 25, 21, 22, 31, 2, 35, 44, 23, 21, 21, 16, 12, 18, 41, 22, 16, 25, 33, 34, 29, 13, 18, 24, 23, 42, 33, 29

### Solutions for Exercises 3.2

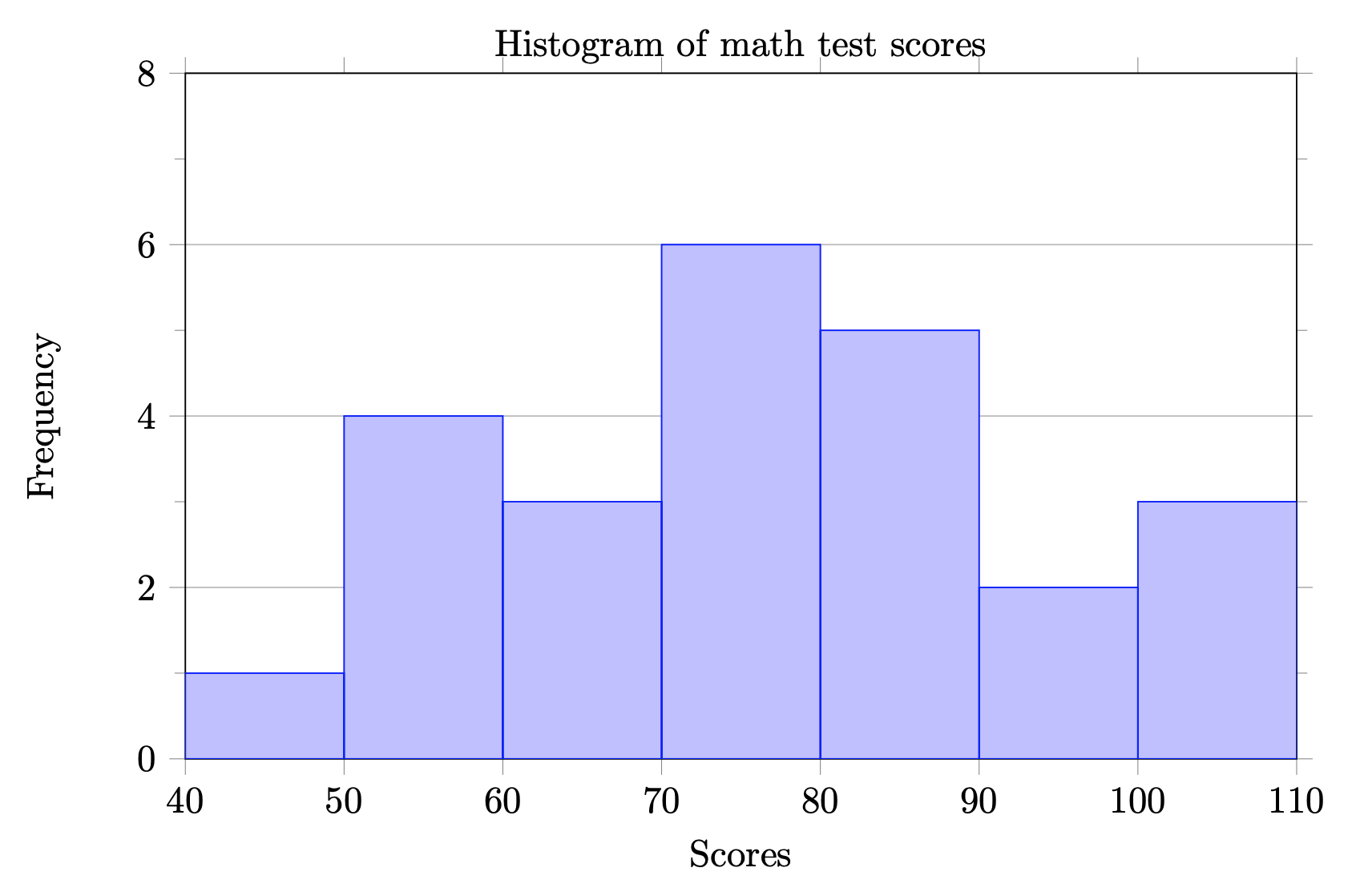
1. True
2. False
   1. Quantitative
   2. Categorical
   3. Categorical
   4. Quantitative
   5. Quantitative
   6. Quantitative
   7. Categorical
   8. Quantitative
   9. Quantitative
   10. Quantitative
3. Bar graphs and pie charts are used for categorical data.
4. Histograms are used for quantitative data.
   1. 2 adults had 3 children.
   2. 15 adults were questioned.
   3. 33.33% of the adults questioned had 0 children.
   4. 8 movies took 2 days to arrive.
   5. She ordered 19 movies in total.
   6. 21.05% of the movies arrived in one day.
5. 5 students earned an A on their paper.
6. 24 served drinks were lattes.
7. Corey spent $676 on rent this month.
8. Habiba spends 11.8 hours travelling each week.
   1. This data is categorical.
   2. No, we cannot make a pie chart out of this data. The total of the relative frequencies is 1932, but only 1012 adults were asked. So some adults selected multiple options.
   3. This data is categorical.
   4. No, we cannot make a pie chart out of this data. The relative frequencies add to 177%, which is more than 100%.
   5. 40 households heat their home with firewood.
   6. 50% of the households heat their home with natural gas.
   7. 54% of the students are below senior class.
   8. 9 of the sampled students are freshmen.
   9. This data is quantitative.

|  |  |
| --- | --- |
| **Number of cars in household** | **Frequency** |
| 0 – 1 | 7 |
| 2 – 3 | 14 |
| 4 - 5 | 3 |



* 1. This data is unimodal and skewed right, with no outliers.
  2. This is quantitative data.

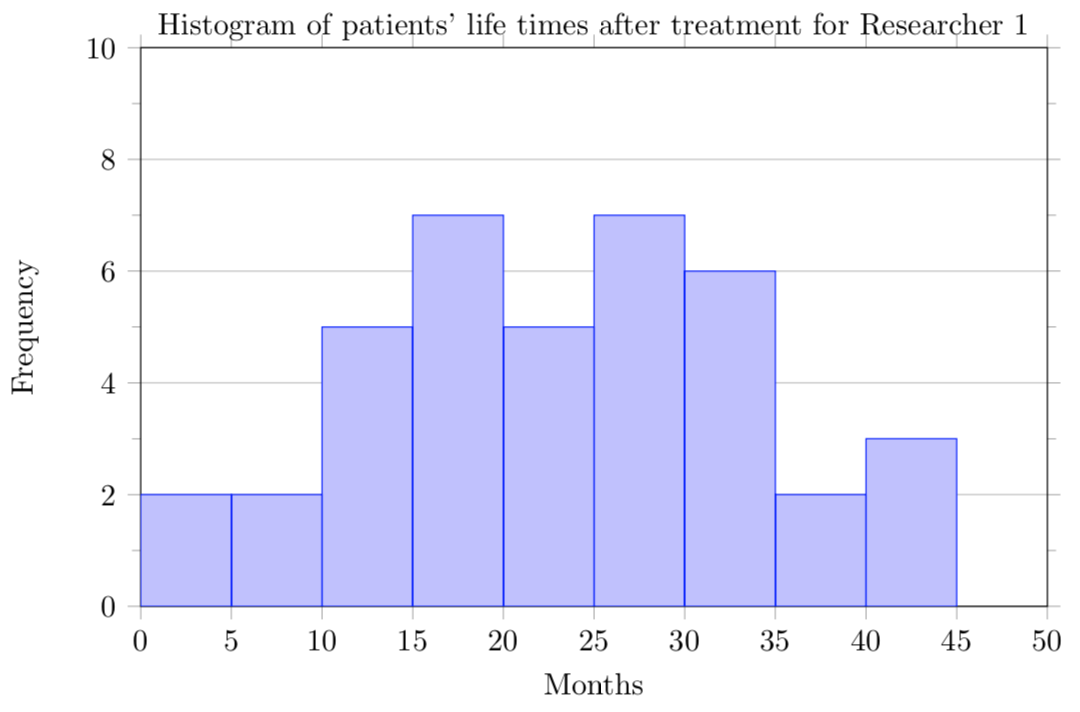
|  |  |
| --- | --- |
| **Math test score** | **Frequency** |
| 40 – 49 | 1 |
| 50 – 59 | 4 |
| 60 – 69 | 3 |
| 70 – 79 | 6 |
| 80 – 89 | 5 |
| 90 – 99 | 2 |
| 100 - 109 | 3 |

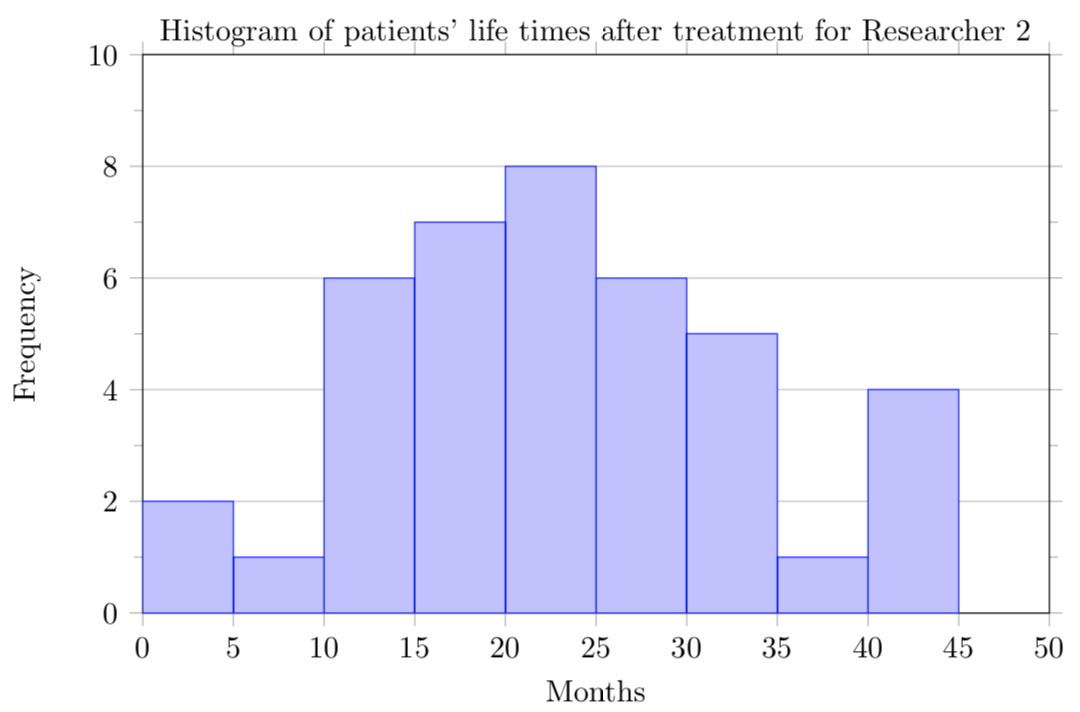


* 1. The data is unimodal. There are no obvious outliers. The data is either symmetric or skewed left.

1. The graph would be more effective at displaying the true differences between the categories if the vertical scale started at 0. The vertical axis is missing a label and units, so we can’t tell if those are frequencies or relative frequencies. A flat bar graph (instead of the 3d graph) would be easier to read.
2. No, this chart does not present a good representation of this data. The percentages in a pie chart must add to 100%, but these add to 193%. A bar chart would be appropriate for this data.
3. This is a poor graph because the vertical axis does not have a numerical scale, so we cannot know how many have each drink as a favorite drink. We also don’t know if the bottom vertical line represents 0, which is potentially misleading.
4. This is a misleading graph because the $-1320 in April represents a loss (instead of a profit), but the height of the bar for April looks like a profit. Also, the horizontal axis label is “Year”, but the bar labels are months.
   1. Normal distribution – The number of heads in 24 sets of 100 coin flips.
   2. Positive or right skewed – Distribution of scores on a psychology test.
   3. Negative or left skewed – Scores on a 20-point statistics quiz.
   4. Bimodal – The frequency of times between eruptions of the Old Faithful geyser.
   5. This distribution is unimodal and right skewed. There is a possible outlier between 10 and 11.
   6. This distribution is unimodal and symmetric. There are no outliers.
   7. This distribution is unimodal and right skewed. There are no obvious outliers.
   8. This distribution is multimodal and left skewed. There are no outliers.
   9. This distribution is unimodal and left skewed. There are no obvious outliers.

Histogram of survival time for patients for Researcher 1





* 1. The data for patients of both researchers is symmetric. Researcher’s 1 patients’ data appears to be unimodal, but Researcher 2’s patients’ data may be bimodal or multimodal. The data for Researcher 1’s patients does not have any outliers, but the data for Researcher’s 2 may have outliers between 0 and 5 months or 40 and 45 months.

1. Gallup Poll. March 5-8, 2009. <http://www.pollingreport.com/enviro.htm> [↑](#footnote-ref-1)