### Exercises 3.3

1. A group of diners were asked how much they would pay for a meal. Their responses were: $7.50, $25.00, $10.00, $10.00, $7.50, $8.25, $9.00, $5.00, $15.00, $8.00, $7.25, $7.50, $8.00, $7.00, $12.00.
   1. Find the mean, including units.
   2. Find the median, including units.
   3. Based on the mean and the median, would you expect the distribution to be symmetric, skewed left, or skewed right? Explain.
2. The amount of commercials in an hour of television varies by channel. The total length (in minutes) of all commercials from 8 pm to 9 pm in for some selected broadcast and cable channels on a weekday evening were: 10, 12.75, 7, 9, 9.75, 6.5, 12.5, 12.5, 8.75, 17, 10.5, 2.
   1. Find the mean, including units.
   2. Find the median, including units.
   3. Based on the mean and median, would you expect the distribution to be symmetric, skewed left, or skewed right? ADD
3. You recorded the time in seconds it took for 8 participants to solve a puzzle. The times were: 15.2, 18.8, 19.3, 19.7, 20.2, 21.8, 22.1, 29.4.
   1. Find the mean, including units.
   2. Find the median, including units.
   3. Based on the mean and the median, would you expect the distribution to be symmetric, skewed left, or skewed right? Explain.
4. You weigh 9 Oreo cookies, and you find the weights (in grams) are: 3.49, 3.51, 3.51, 3.51, 3.52, 3.54, 3.55, 3.58, 3.61
5. Find the mean, including units.
6. Find the median, including units.
7. Based on the mean and the median, would you expect the distribution to be symmetric, skewed left, or skewed right? Explain. ADD
8. Use the following table is the cost of purchasing a car at a local dealership.  Some of the cars sold were new and some were used.
   1. Calculate find the mean, including units.
   2. Can you figure out how to find the median using the frequency table? See if you can do it without listing out all the data values.
   3. Based on the mean and the median, would you expect the distribution to be symmetric skewed left or skewed right? Explain.

|  |  |
| --- | --- |
| Cost  (Thousands of dollars) | Frequency |
| 15 | 3 |
| 20 | 7 |
| 25 | 10 |
| 30 | 15 |
| 35 | 13 |
| 40 | 11 |
| 45 | 9 |
| 50 | 7 |

1. ADD As part of a study of email, a researcher counted the length of 34 emails. The lengths of the emails are shown below, rounded to the nearest thousand characters (so a length 0 means that the numbers of characters rounded to 0, not that the message was blank).
   1. Calculate and find the mean, including units.
   2. Can you figure out how to find the median using the frequency table? See if you can do it without listing all the data values.
   3. Based on the mean and the median, would you expect the distribution to be symmetric, skewed left, or skewed right?

\* The data is from Advanced High School Statistics, 2nd Ed. (<https://www.openintro.org/stat/textbook.php?stat_book=aps>).

|  |  |
| --- | --- |
| Length of an email (Thousands of characters) | Frequency |
| 0 | 4 |
| 1 | 5 |
| 2 | 2 |
| 3 | 3 |
| 4 | 3 |
| 5 | 1 |
| 6 | 3 |
| 7 | 3 |
| 8 | 0 |
| 9 | 3 |
| 10 | 3 |
| 11 | 2 |
| 12 | 0 |
| 13 | 0 |
| 14 | 2 |

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. Find the mean and median of each group.
   2. Compare and contrast the two groups.

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

1. ADD The US Census Bureau, in addition to counting the population of the US every 10 years, conducts yearly informational surveys, such as the American Community Survey (ACS). For the 2012 ACS, a randomly chosen group of 20 respondents (10 males, 10 females) answered a question about their incomes.

Males: $53,000; $70,000; $12,800; 30,000; $4,500; $42,000; $48,000; $60,000; $108,000; $11,000

Females: $1,600; $1,200; $20,000; $25,000; $670; $29,000; $44,000; $30,000; $5,800; $50,000

1. Find the mean and median of each group.
2. Compare and contrast the two groups.

\*Data from Advanced High School Statistics, 2nd Edition. Section 2.1 # 1 Exercise. (<https://www.openintro.org/stat/textbook.php?stat_book=aps>)

1. An experiment compared the ability of three groups of participants to remember briefly-presented chess positions. The data are shown below. The numbers represent the average number of pieces correctly remembered from three chess positions.
2. Make a histogram for each group.
3. Find the mean of each group.
4. Find the median of each group.
5. Compare the shapes of the distributions as well as the centers of the three groups.

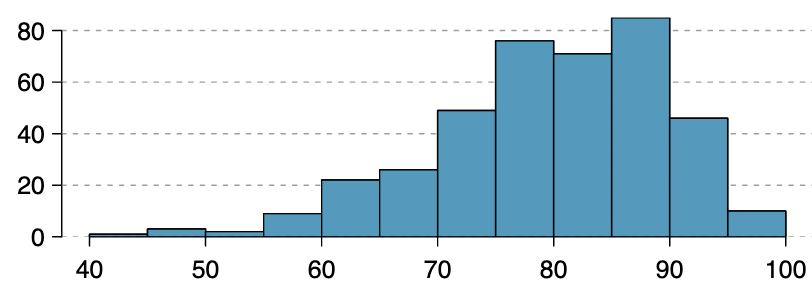
|  |  |  |
| --- | --- | --- |
| Non-players | Beginners | Tournament Players |
| 22.1 | 32.5 | 40.1 |
| 22.3 | 37.1 | 45.6 |
| 26.2 | 39.1 | 51.2 |
| 29.6 | 40.5 | 56.4 |
| 31.7 | 45.5 | 58.1 |
| 33.5 | 51.3 | 71.1 |
| 38.9 | 52.6 | 74.9 |
| 39.7 | 55.7 | 75.9 |
| 43.2 | 55.9 | 80.3 |
| 43.2 | 57.7 | 85.3 |

1. There is evidence that smiling can attenuate judgments of possible wrongdoing. This phenomenon termed the "smile-leniency effect" was the focus of a study by Marianne LaFrance & Marvin Hecht in 1995[[1]](#footnote-1). The following data are measurements of how lenient the sentences were for three different types of smiles and one neutral control. A higher number indicates greater leniency. The same subject was used for all of the conditions so that may affect the results.
2. Make four histograms, one for each smile type and one for the neutral control.
3. Find the mean for each type of smile and the neutral control.
4. Find the median for each type of smile and the neutral control.
5. Compare the shapes of the distributions as well as the centers for each type of smile and control.

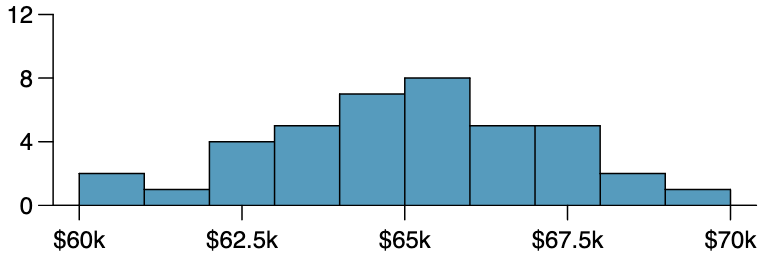
|  |  |  |  |
| --- | --- | --- | --- |
| False  Smile | Felt Smile | Miserable  Smile | Neutral Control |
| 2.5 | 7 | 5.5 | 2 |
| 5.5 | 3 | 4 | 4 |
| 6.5 | 6 | 4 | 4 |
| 3.5 | 4.5 | 5 | 3 |
| 3 | 3.5 | 6 | 6 |
| 3.5 | 4 | 3.5 | 4.5 |
| 6 | 3 | 3.5 | 2 |
| 5 | 3 | 3.5 | 6 |
| 4 | 3.5 | 4 | 3 |
| 4.5 | 4.5 | 5.5 | 3 |
| 5 | 7 | 5.5 | 4.5 |
| 5.5 | 5 | 4.5 | 8 |
| 3.5 | 5 | 2.5 | 4 |
| 6 | 7.5 | 5.5 | 5 |
| 6.5 | 2.5 | 4.5 | 3.5 |
| 3 | 5 | 3 | 4.5 |
| 8 | 5.5 | 3.5 | 6.5 |
| 6.5 | 5.5 | 8 | 3.5 |
| 8 | 5 | 5 | 4.5 |
| 6 | 4 | 7.5 | 4.5 |
| 6 | 5 | 8 | 2.5 |
| 3 | 6.5 | 4 | 2.5 |
| 7 | 6.5 | 5.5 | 4.5 |
| 8 | 7 | 6.5 | 2.5 |
| 4 | 3.5 | 5 | 6 |
| 3 | 5 | 4 | 6 |
| 2.5 | 3.5 | 3 | 2 |
| 8 | 9 | 5 | 4 |
| 4.5 | 2.5 | 4 | 5.5 |
| 5.5 | 8.5 | 4 | 4 |
| 7.5 | 3.5 | 6 | 2.5 |
| 6 | 4.5 | 8 | 2.5 |
| 9 | 3.5 | 4.5 | 3 |
| 6.5 | 4.5 | 5.5 | 6.5 |

1. Make up three data sets with 5 values each that have:
   1. The same mean but different medians.
   2. The same median but different means.
2. The frequency table below shows the number of women’s shoes that were sold in an hour at a local shoe store.
   1. Would you treat this data as categorical or quantitative?
   2. How would the bar graph be different from the histogram?
   3. Treat the data as quantitative and find the mean and the median. Are these useful statistics?

|  |  |
| --- | --- |
| Shoe Size | Frequency |
| 5 | 4 |
| 6 | 4 |
| 7 | 6 |
| 8 | 6 |
| 9 | 5 |

1. ADD At the end of the term, 400 students take a final exam, and their scores (as percentages) are plotted in a histogram.  
   
   1. Is the graph shown above skewed right, skewed right, or symmetric? Explain why you chose your answer.
   2. Based on your answer to part (a), which would you expect: that the mean is less than the median, the mean is greater than the median, or that the mean and median are equal?

\*Graph and description from Advanced High School Statistics, Section 2.2 Exercises

1. ADD The following graph shows the distribution of yearly incomes of 40 patrons at a college coffee shop.  
   
2. Is the graph shown above skewed right, skewed right, or symmetric? Explain why you chose your answer.
3. Based on your answer to part (a), which would you expect: that the mean is less than the median, the mean is greater than the median, or that the mean and median are equal?

\*Graph and description from Advanced High School Statistics, Section 2.2 Exercises

1. ADD For each of the following distributions, would you expect that the mean is less than the median, less than the median, or equal to the median? Explain your reasoning.
2. Household incomes in the US
3. Weights of newborn babies
4. The number of children in a household in the US
5. Medical costs for all adults in the US
6. Medical costs for adults in the US older than 65

### Solutions for 3.3 Exercises

**1a.**

In Excel:

There are 15 times shown, so .

The mean is:

**1b.**

In Excel:

There are 15 times shown, so .

We start by listing the data in order:

$5.00, $7.00, $7.25, $7.50, $7.50, $7.50, $8.00, $8.00, $8.25, $9.00, $10.00, $10.00, $12.00, $15.00, $25.00

**1c.**

Since the mean is greater than the median, we would expect the distribution will be skewed right.

**2a.**

In Excel:

minutes

There are 12 times shown, so .

The mean is:

**2b.**

In Excel:

minutes

There are 12 times shown, so .

We start by listing the data in order:

**2c.**

Because the mean and median are approximately equal, we would expect that the distribution is symmetric.

**3a.**

In Excel:

seconds

There are 8 times shown, so .

**3b.**

In Excel:

There are 8 times shown, so .

The times are given already in order:

15.2, 18.8, 19.3, 19.7, 20.2, 21.8, 22.1, 29.4

**3c.**

Since the mean and median are approximately equal, we would expect that the distribution is symmetric.

**4a.**

In Excel:

grams

There are 9 weights shown, so .

**4b.**

In Excel:

grams

There are 9 weights shown, so .

The weights are given already in order:

3.49, 3.51, 3.51, 3.51, 3.52, 3.54, 3.55, 3.58, 3.61

**4c.**

Since the mean and median are approximately equal, we would expect the distribution to be symmetric.

**5a.**

In GeoGebra Classic, enter the costs into the column A and frequencies into column B of the spreadsheet and use the “One Variable Analysis” function. Then use the “Show Statistics” option.

The sum of the frequencies is 75, so .

**5b.**

In GeoGebra Classic, enter the costs into the column A and frequencies into column B of the spreadsheet and use the “One Variable Analysis” function. Then use the “Show Statistics” option.

Since there are 75 values (an odd number), we know that the median will be the single middle data value. Because , we know it will be the 38th value in the list. The 38th value is 35, so the median is 35 thousand dollars.

**5c.**

Since the mean is less than the median, we would expect the distribution to be skewed left.

**6a.**

In GeoGebra Classic, enter the costs into the column A and frequencies into column B of the spreadsheet and use the “One Variable Analysis” function. Then use the “Show Statistics” option.

We are told that there are 34 emails in the sample, so .

**6b.**

In GeoGebra Classic, enter the costs into the column A and frequencies into column B of the spreadsheet and use the “One Variable Analysis” function. Then use the “Show Statistics” option.

Since there are 34 values (an even number), we know that the median will be the mean of the two middle values. Because , we know the two middle values are the 17th and 18th values. The 17th value is 4, and the 18th value is 5, so the median is 4.5 thousand characters.

**6c.**

Since the mean is greater than the median, we expect the distribution to be skewed right.

**7a.**

For Researcher 1:

In Excel:

months

months

The mean for Researcher 1’s patients is 23.6 months, and the median for Researcher’s 1 patients is 24 months.

For Researcher 2:

In Excel:

months

months

The mean for Researcher 2’s patients is 22.8 months, and the median is 22 months.

**7b.**

Both the mean and median for Researcher 1’s patients are greater than the mean and median for Researcher 2’s patients. So, on average, Researcher 1’s patients have a longer life time after starting the cancer treatment than Researcher 2’s patients.

**8a.**

For Males:

In Excel:

The mean for males is $43,930, and the median for males is $45,000.

For Females:

The mean for females is $20,727, and the median for females is $22,500.

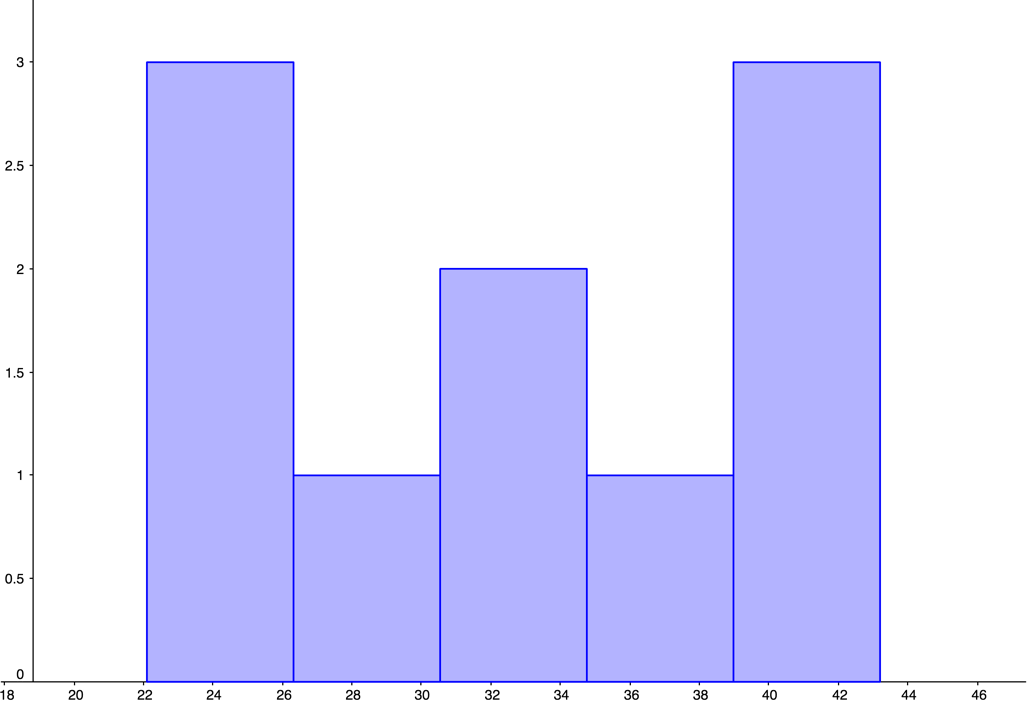
**8b.**

Both the mean income and median income for the males in the sample were twice as large as the mean income and median income for the females in the sample. The difference between the mean and median income for the females was slightly larger than the mean and median income for the males, so the distribution of incomes for the females is possibly more skewed than the distribution for the males.

GeoGebra was used to create the histograms. You should check with your instructor to see if histograms are to be hand-drawn or computer generated. Answers will vary depending on the size of the margins and the programs you are using.

9a)

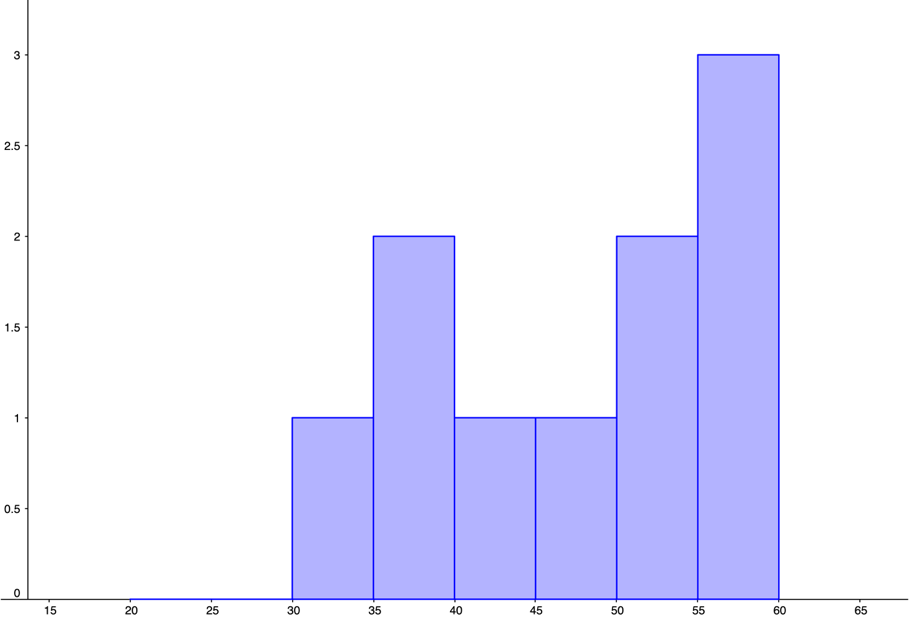
Histogram for Average Number of Pieces Correctly Remembered by Non-players



Frequency

Number of pieces

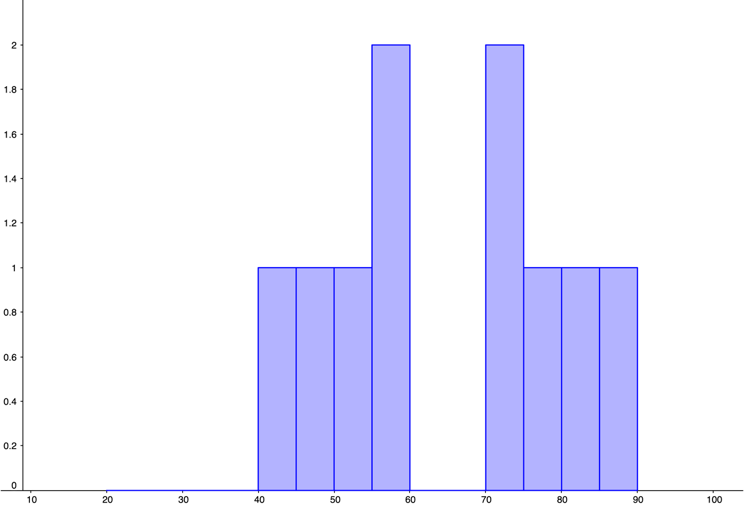
Histogram for Average Number of Pieces Correctly Remembered by Beginners



Frequency

Number of pieces

Histogram for Average Number of Pieces Correctly Remembered by Tournament Players



Frequency

Number of pieces

**9b.**

The mean number of pieces correctly remembered for non-players was 33.0 pieces.

The mean number of pieces correctly remembered for beginners was 46.8 pieces.

The mean number of pieces correctly remembered for tournament players was 63.9 pieces.

**9c.**

The median number of pieces correctly remembered for non-players was 32.6 pieces.

The median number of pieces correctly remembered for beginners was 48.4 pieces.

The median number of pieces correctly remembered for tournament players was 64.6 pieces.

**9d.**

The distribution for non-players appears to be uniform. The distribution for beginners looks unimodal and left-skewed. The distribution for tournament players appears bimodal and symmetric.

The mean and median number of pieces correctly remembered were both greatest for tournament players, with non-players having the smallest mean and median of pieces correctly remembered.

**10a.**

Histogram for Smile-Leniency of False Smiles

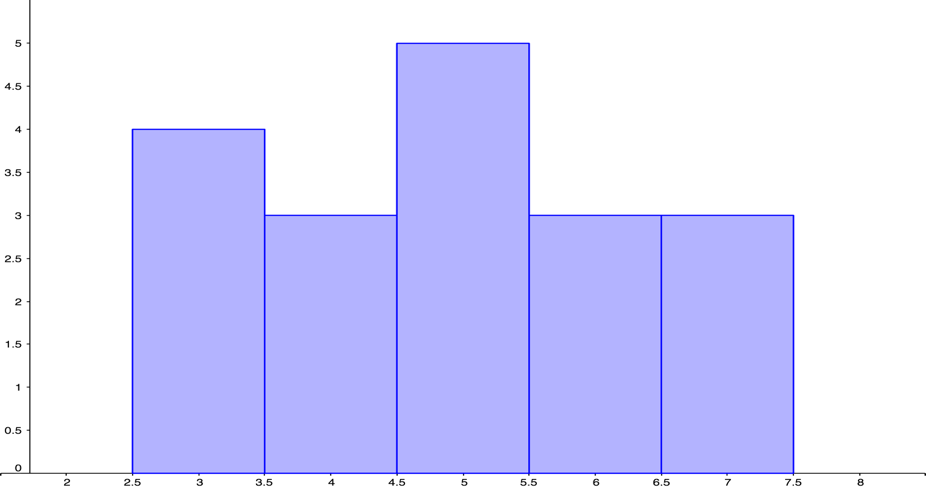
This is a histogram. The y-axis is labeled frequency and goes from 0 to 6. The x-axis is labeled Leniency and goes from 2.5 to 8. There are 5 bars, with the following heights: 6, 2, 4, 5, 1. 


Frequency

Leniency

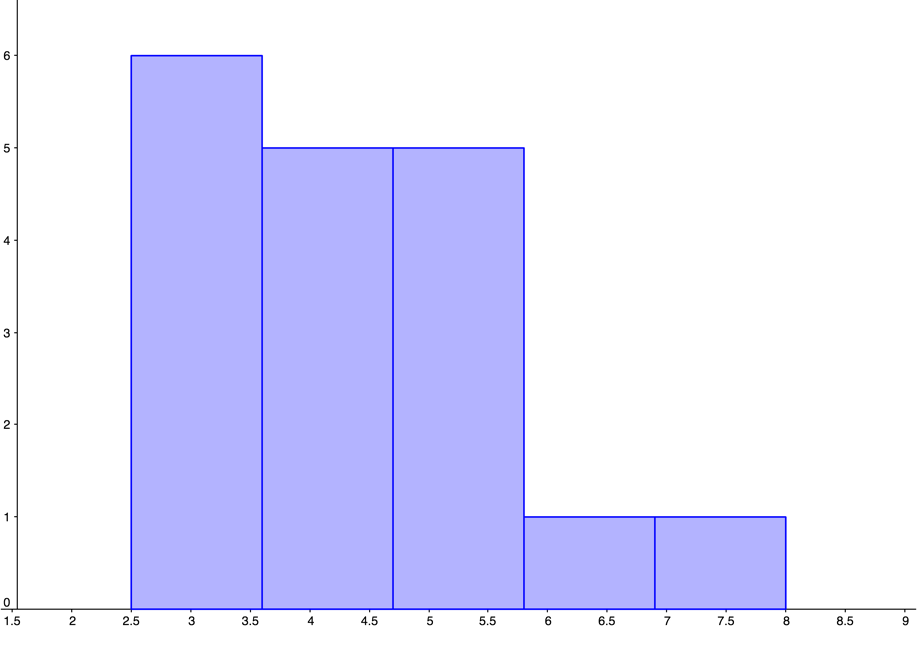
Histogram for Smile-Leniency of Felt Smiles

Frequency



Leniency

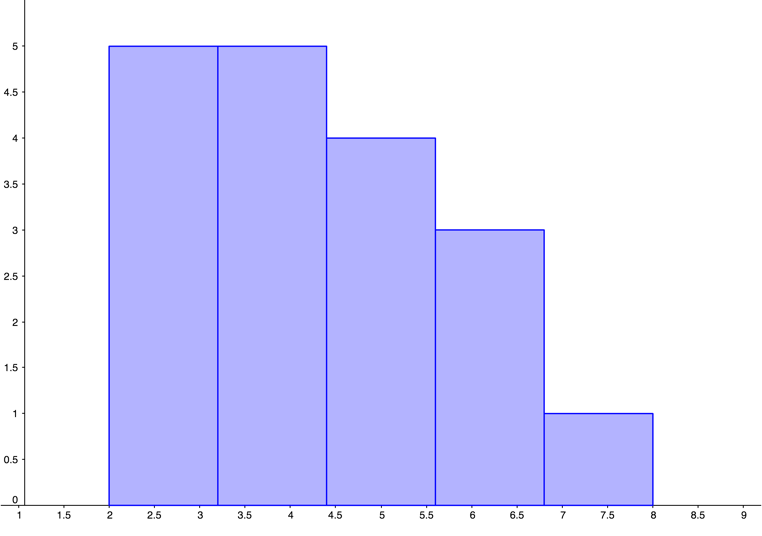
Histogram for Smile-Leniency of Miserable Smiles



Frequency

Leniency

Histogram for Smile-Leniency of Neutral Control Group



Frequency

Leniency

**10b.**

The mean leniency for the false smile group was 5.4.

The mean leniency for the felt smile group was 4.9.

The mean leniency for the miserable smile group was 4.9.

The mean leniency for the neutral control group was 4.1.

**10c.**

The median leniency for the false smile group was 5.5.

The median leniency for the felt smile group was 4.8.

The median leniency for the miserable smile group was 4.8.

The median leniency for the neutral control group was 4.0.

**10d.**

Answers will vary depending on the graphs created. The shape of the false smile is bimodal, whereas the miserable smile and neutral control groups are both unimodal. The miserable smile and neutral control both appear to be skewed to the right, however, comparing the mean to the median we see that there is not much of a difference. Therefore, none of the graphs are skewed. The felt smile group is the most visually uniform. The measures of center (mean and median) have the center of the false smile near 5.5, the felt smile and miserable smile both near 4.8, and the control near 4.

**11a.**

There are many possible answers for this problem. Three data sets with 5 values each that have the same mean but different medians are:

0, 0, 0, 0, 10

0, 0, 2, 4, 4

0, 1, 1, 1, 7

**11b.**

There are many possible answers for this problem. Three data sets with 5 values that have the same median but different means are:

10, 10, 10, 10, 10

0, 0, 10, 15, 20

1, 5, 10, 10, 10

**12a.**

**Argument for categorical: Because it is not clear that the shoe sizes represent a measure, the data can be considered categorical.**

**Argument for quantitative: Because shoe size is a measurement that corresponds to the length of someone’s foot, it can be treated as quantitative data.**

**12b. Each graph would have frequency along the y-axis. In a bar graph the bars would have spaces between them and each bar would be labeled with the shoe size. In a histogram there would not be any spaces between the bars and the shoe sizes could be the scale on the x-axis.**

**12c.**

The mean shoe size to be 7.2 and the median shoe size to be 7.

The mean shoe size is:

Since there are 25 values (an odd number), we know that the median will be the single middle data value. Because , we know it will be the 13th value in the list. The 13th value is 7, so the median shoe size is 7.

**The mean and median shoe size might be useful statistics to the store. If shoe size is positively correlated to height, then a shoe store with a comparatively larger mean or median shoe size could determine that their clients are, on average, comparatively taller. (Other answers are possible.)**

**13a.**

**This graph is skewed left.**

**13b.**

**I expect that the mean is less than the median because the graph is skewed left.**

**14a.**

**This graph is symmetric. The graph has a single peak between $65,000 an $66,000, and there are approximately an equal number of data values on either side of this peak.**

**14b.**

**I expect that the mean and median are equal because the graph is symmetric.**

**15.**

**Answers will vary.**

1. LaFrance, M., & Hecht, M. A. (1995) Why smiles generate leniency. Personality and Social Psychology Bulletin, 21, 207-214. Adapted from [www.onlinestatbook.com](http://www.onlinestatbook.com), by David M. Lane, et al, used under [CC-BY-SA 3.0](https://creativecommons.org/licenses/by-sa/3.0/us/). [↑](#footnote-ref-1)