Chapter 2: Graphical Summaries of Data
2.1 Graphical Summaries for Qualitative Data

Frequency:

Frequency distribution:

Example 2.1
The following are survey results from Fall 2014 Statistics class regarding eye color. Construct a frequency distribution.

<table>
<thead>
<tr>
<th>Blue</th>
<th>Brown</th>
<th>Brown</th>
<th>Brown</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown</td>
<td>Hazel</td>
<td>Hazel</td>
<td>Brown</td>
<td></td>
</tr>
<tr>
<td>Blue</td>
<td>Brown</td>
<td>Blue</td>
<td>Hazel</td>
<td>Brown</td>
</tr>
<tr>
<td>Hazel</td>
<td>Brown</td>
<td>Blue</td>
<td>Brown</td>
<td></td>
</tr>
<tr>
<td>Hazel</td>
<td>Hazel</td>
<td>Blue</td>
<td>Brown</td>
<td></td>
</tr>
</tbody>
</table>

Solution

Relative frequency:

Relative frequency distribution:
Relative Frequency =

Example
Construct a relative frequency distribution for the eye color data.

<table>
<thead>
<tr>
<th>Eye Color</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hazel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Solution

Bar graph:

Example
Construct a frequency bar graph and the relative frequency bar graph for the eye color data.

<table>
<thead>
<tr>
<th>Eye Color</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hazel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Solution

Pareto Charts:

Pie Charts:

Example
Construct a pie chart for the eye color data.

<table>
<thead>
<tr>
<th>Eye Color</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown</td>
<td></td>
</tr>
<tr>
<td>Blue</td>
<td></td>
</tr>
<tr>
<td>Hazel</td>
<td></td>
</tr>
</tbody>
</table>
Solution

Do You Know
- How to construct a frequency and relative frequency distribution for qualitative data?
- How to construct the various kinds of bar graphs?

2.2 Frequency Distributions and Their Graphs
To summarize quantitative data, we use a frequency distribution just like those for qualitative data. Frequency distributions have the following:

Classes:

Lower class limit:

Upper class limit:

Class width:
Here’s a random example:

### Choosing Classes

- 
- 
- 
- 
- 

### Constructing a Frequency Distribution

#### Step 1:

#### Step 2:

#### Step 3:

#### Step 4:

#### Step 5:
Note: If you are not told the class width, you can figure it out yourself by doing (largest data value - smallest data value) divided by the number of classes. Round this (always up) to a convenient number. You will always either be given the class width or the number of classes.

**Example**
Following are the number of years lived by selected US presidents after their first inauguration. Construct a frequency distribution starting with a lower class limit of 0 and a class width of 5.

10, 29, 26, 28, 15, 23, 17, 25, 0, 4, 1, 16, 12, 4, 17, 16, 0, 24, 12, 4, 18, 21, 11, 2, 9, 36, 12, 28, 16, 3, 9, 25, 23

**Solution**

Given a frequency distribution, a relative frequency distribution can be constructed by computing the relative frequency for each class. Now create a relative frequency of the previous data. (in your notes, put it above here).

Recall: Relative Frequency =

**Histograms**
Once we have a frequency distribution or a relative frequency distribution, we can put the information in graphical form by constructing a…

**Histogram:**
Example
Construct a frequency histogram and relative frequency histogram for the president data.

Solution

Choosing the Number of Classes
There are no hard and fast rules for choosing the number of classes. There are two principles that can guide the choice:

•

•

Note: there are ways to do a histogram on the calculator, but honestly, it is quicker and easier (in my opinion) to just do it by hand, assuming you have already created a frequency distribution.

Shape of a Distribution from the Histogram
Symmetric:

Skewed right (positively skewed):

Skewed left (negatively skewed):
**Example:** which histogram below is symmetric, skewed right, skewed left?

![Histograms](image)

**Mode:**

**Unimodal:**

**Bimodal:**

![Histograms](image)

**Class Midpoints:**

Class Midpoint =

**Example:** Find all class midpoints for presidential data.

**Frequency polygon:**

![Histograms](image)
Example
Construct a frequency polygon for the presidential data.
Solution

Relative frequency polygon:

Ogives:

Example
Construct an ogive for the presidential data.
Solution
Do You Know
- How to construct frequency and relative frequency distributions for quantitative data?
- How to construct frequency and relative frequency histograms
- How to determine the shape of a distribution from a histogram?
- How to construct frequency and relative frequency polygons?
- How to construct an ogive?

2.3 More Graphs for Quantitative Data
**Stem-and-Leaf Plots**
Stem-and-leaf plots and dotplots illustrate the shape of the data set, while allowing every value in the data set to be seen.

In a stem-and-leaf plot, the rightmost digit is the leaf, and the remaining digits form the stem. Consider the values 14.8 and 2,739:

![Stem-and-Leaf Plot Example](image)

**Procedure to Construct a Stem-and-Leaf Plot**

Step 1:

Step 2:

Step 3:

**Example**
The following are quiz scores from a Calculus quiz I once gave. Create a stem and leaf plot.
10, 12, 16, 25, 19, 21, 22, 27, 29, 25, 29, 17, 5, 12, 13, 29, 23, 12, 15, 7, 8, 23, 25, 27, 12, 17, 18, 12, 19, 18, 23

**Solution**
Sometimes you can “split the stem” and have two or more lines for each stem. Do this with above quiz score data.

**Dotplots:**

**Example – Dotplot**
Construct a dotplot of the Calculus quiz data.

**Solution**

**Do You Know**
- How to construct various types of stem-and-leaf plots?
- How to construct dotplots?
2.4 Misleading Graphs

Introduction
Statistical graphs, when properly used, are powerful forms of communication. Unfortunately, when graphs are improperly used, they can misrepresent the data, and lead people to draw incorrect conclusions.

Misrepresentation
There are three common forms of misrepresentation. They are

- Incorrect Position of the Vertical Scale
  The baseline of a chart or plot is the value at which the horizontal axis intersects with the vertical axis.

  ![Baseline example](image)

  With charts or plots that represent how much or how many of something, _______________ ____________________________.

- The Area Principle (I stole the rest of this section straight from the book…sorry!)
  When amounts are compared by constructing an image for each amount, the areas of the images must be proportional to the amounts. For example, if one amount is twice as much as another, its image should have twice as much area as the other image. When the Area Principle is violated, the images give a misleading impression of the data.

  Example – The Area Principle
Note that the price in 2008 is about 3.5 times the price in 2000. In a bar chart, the area of the bar for 2008 is also about 3.5 times that of the area of the bar for 2000.

This bar chart follows the Area Principle.

<table>
<thead>
<tr>
<th>Year</th>
<th>Cost of Jet Fuel per Gallon</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>$0.90</td>
</tr>
<tr>
<td>2008</td>
<td>$3.16</td>
</tr>
</tbody>
</table>

A common mistake is to vary both dimensions (height and width) of the image. Following is a comparison of the cost of jet fuel that uses a picture of an airplane.

Three-dimensional Graphs
Newspapers and magazines often present three-dimensional bar charts because they are visually impressive. Unfortunately, in order to make the tops of the bars visible, these charts are often drawn as though the reader is looking down on them. This makes the bars look shorter than they really are. *(Note: I think this is sort of ridiculous. Just me)*

The following figures present the same cost of jet fuel data. However, in the three-dimensional bar chart, the bars appear shorter than they really are.
Do You Know
- The three most common forms of misrepresentation?