

Math 160, Finite Mathematics for Business

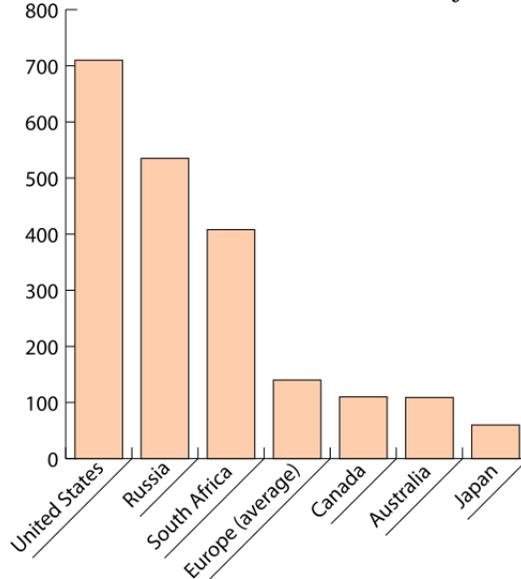
Section 7.1: Visual Representation of Data - Discussion Notes Brian Powers - TA - Fall 2011

Bar Charts

A bar chart or bar graph is a chart with rectangular bars with lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally.

Bar charts are used for plotting discrete (or 'discontinuous') data which has discrete values. Some examples of discontinuous data include 'shoe size' or 'eye color', for which you would use a bar chart. In contrast, some examples of continuous data would be 'height' or 'weight'. A bar chart is very useful if you are trying to record certain information whether it is continuous or not continuous data. Bar charts also look a lot like a histogram. They are often mistaken for each other.

Bar Chart for Incarceration Rate by Country

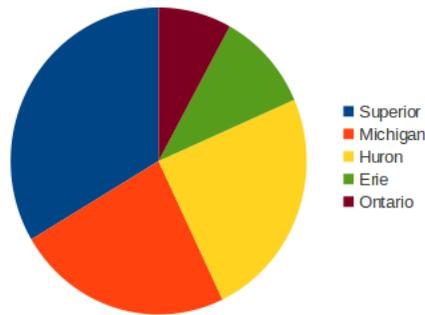


Pie Charts A pie chart (or a circle graph) is a circular chart divided into sectors, illustrating proportion. In a pie chart, the arc length of each sector (and consequently its central angle and area), is proportional to the quantity it represents.

To make the pie chart, make a table of the data, with total given below. Add a column for Percentage, and Degrees. Degrees are calculated as : $\text{Degrees} = \text{Percentage} \times 360$

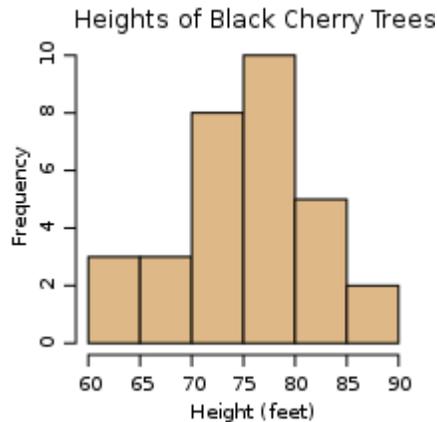
Lake	Area (mi ²)	Percentage	Degrees
Superior	31,700	33.5	120.60
Michigan	22,300	23.6	84.96
Huron	23,100	24.4	87.84
Erie	9,910	10.5	37.8
Ontario	7,550	8.0	28.7
Total	94,560	100.0	360.00

Great Lakes Areas



Histograms

A histogram is a graphical representation showing a visual impression of the distribution of data. A histogram consists of tabular frequencies, shown as adjacent rectangles, erected over discrete intervals (bins), with an area equal to the frequency of the observations in the interval. The height of a rectangle is also equal to the frequency density of the interval, i.e., the frequency divided by the width of the interval. The total area of the histogram is equal to the number of data. A histogram may also be normalized displaying relative frequencies. It then shows the proportion of cases that fall into each of several categories, with the total area equaling 1. The categories are usually specified as consecutive, non-overlapping intervals of a variable. The categories (intervals) must be adjacent, and often are chosen to be of the same size.



Median, Quartiles, Box Plots

The **median** is the numerical value separating the higher half of a sample, a population, or a probability distribution, from the lower half. Exactly half of the data will be greater than the median, and half will be less than the median. To find the median, sort the data from lowest to highest. If you have an odd number of data points, the median is the middle data point. (For example, if you have 25 data points, your median is the 13th ordered data point.) If you have an even number of data points, the median is the average of the two middle data points.

Quartiles are a way of breaking the data into four quarters. The first quartile is the median of the lower half of the data, and the third quartile is the median of the upper half of the data.

The minimum is the first data point when the data is sorted from lowest to highest, and the maximum is the last.

Together the minimum, First Quartile, Median (i.e. Second Quartile), Third Quartile and maximum make up the **5 number summary** of the data.

The **interquartile range** is the difference between the first and third quartile. It is calculated by $Q3 - Q1$.

Give the 5 number summary and interquartile range of the following data set: 34, 11, 32, 22, 45, 75, 34, 22, 45, 83, 31, 34, 55

First we want to order the data: 11, 22, 22, 31, 32, 34, 34, 34, 45, 45, 55, 75, 83

Minimum: 11

Median: because there are 13 data points, the median will be the 7th one (so there are $6 \leq$ the median and $6 \geq$ the median). The median is 34

First Quartile: This is the median of the lower half of the data (11, 22, 22, 31, 32, 34) - because there are 6 data points, the median is the average of 22 and 31 = $\frac{31+22}{2} = 26.2$

Third Quartile: The median of the upper half of the data (34, 45, 45, 55, 75, 83) - the median of this data is $\frac{45+55}{2} = 50$

Maximum: 83

So the 5 number summary is 11, 26.5, 34, 50, 83

The interquartile range is $50 - 26.5 = 23.5$

Box Plots

A box plot is a visual representation for the 5-number summary. These can be made vertically or horizontally. Draw a rectangle from Q1 to Q3 with a vertical line at the median. Then add a horizontal line extending to the right and left to the maximum and minimum with a small vertical line at each.

ex: Draw a box plot for the data set: 34, 11, 32, 22, 45, 75, 34, 22, 45, 83, 31, 34, 55

The 5 number summary is 11, 26.5, 34, 50, 83, the box plot is:

