



**Grade 6 Math Circles**  
February 4/5, 2014  
*Introduction to Statistics*

Statistics (or Stats) is a branch of math that deals with collecting and analysing data. Statistics is used almost everywhere. It can be used to solve crimes, predict the outcome of an event, and is even used in most major sports leagues. One very important application of statistics is the Census.

## **Introduction to the Census**

A Census is a survey that collects data on a particular population. For example, the Canadian Census gathers information about people who live in Canada. A Census usually asks questions about the size of your family, the languages you speak, where you live, and more. Statisticians analyse this data to find out how many people live in a particular area, how many people speak each language, and other important information.

In Canada, there is a Census every five years. The last one was May 2011, which means the next one will be May 2016. There are two different types: the Census questionnaire (previously known as the short-form Census) and the National Household Survey (previously known as the long-form Census).

## **Presenting Data**

There are many different ways to present data. One way is a table. However, depending on the type of data it may be more useful to display it in a graph. Graphs allow us to visualize what the data represents, the relation between different categories, and any overall trends.

## Tables

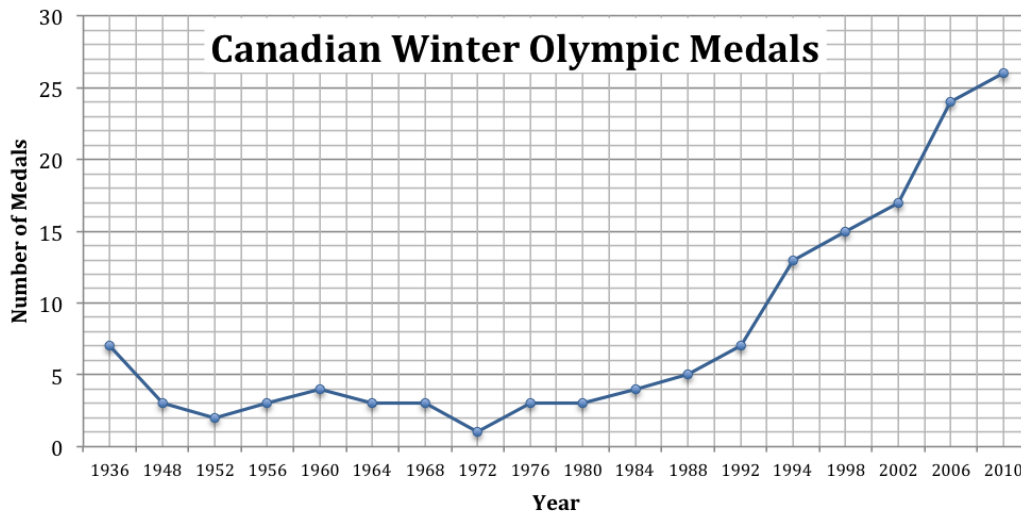
Tables are useful when you want to show exact values in a neat and organized way. Tables are generally used as an initial method to present data. They make data easy to find. Tables also give exact values.

### Canadian Winter Olympic Medals (since 1936)

| Year | Medals |  | Year | Medals |
|------|--------|--|------|--------|
| 1936 | 7      |  | 1980 | 3      |
| 1948 | 3      |  | 1984 | 4      |
| 1952 | 2      |  | 1988 | 5      |
| 1956 | 3      |  | 1992 | 7      |
| 1960 | 4      |  | 1994 | 13     |
| 1964 | 3      |  | 1998 | 15     |
| 1968 | 3      |  | 2002 | 17     |
| 1972 | 1      |  | 2006 | 24     |
| 1976 | 3      |  | 2010 | 26     |

## Line Graph

Line graphs are usually used to show trends. They are usually only used to compare similar data.



| Year | Medals | Year | Medals |
|------|--------|------|--------|
| 1936 | 7      | 1980 | 3      |
| 1948 | 3      | 1984 | 4      |
| 1952 | 2      | 1988 | 5      |
| 1956 | 3      | 1992 | 7      |
| 1960 | 4      | 1994 | 13     |
| 1964 | 3      | 1998 | 15     |
| 1968 | 3      | 2002 | 17     |
| 1972 | 1      | 2006 | 24     |
| 1976 | 3      | 2010 | 26     |

---

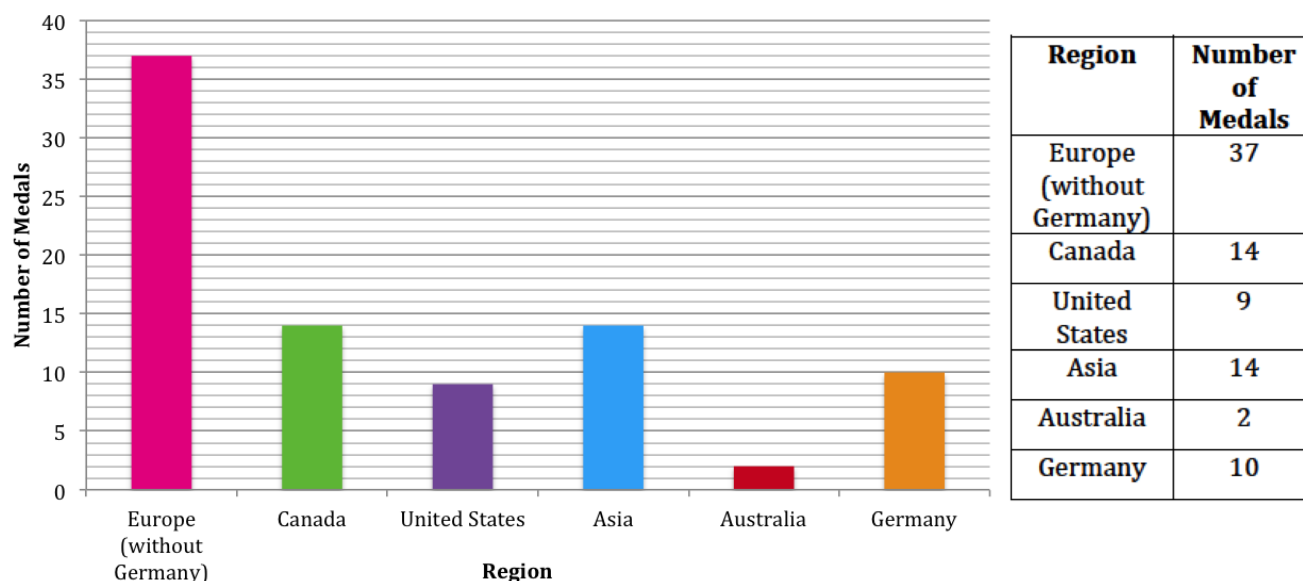
To make a Line Graph:

1. Start with a grid.
  2. Title each axis and give it an appropriate scale.
  3. Draw a dot above each category. The height of the dot is based on its value.
  4. Connect the dots moving left to right.
  5. Title the Graph.
-

## Bar Graphs

Bar graphs are useful to show numerical relations visually. They allow for easy comparisons between different categories. Bar graphs work best with a small number of categories

### 2010 Winter Olympic Gold Medals



Data from: <http://olympic.ca/games/2010-vancouver/>

---

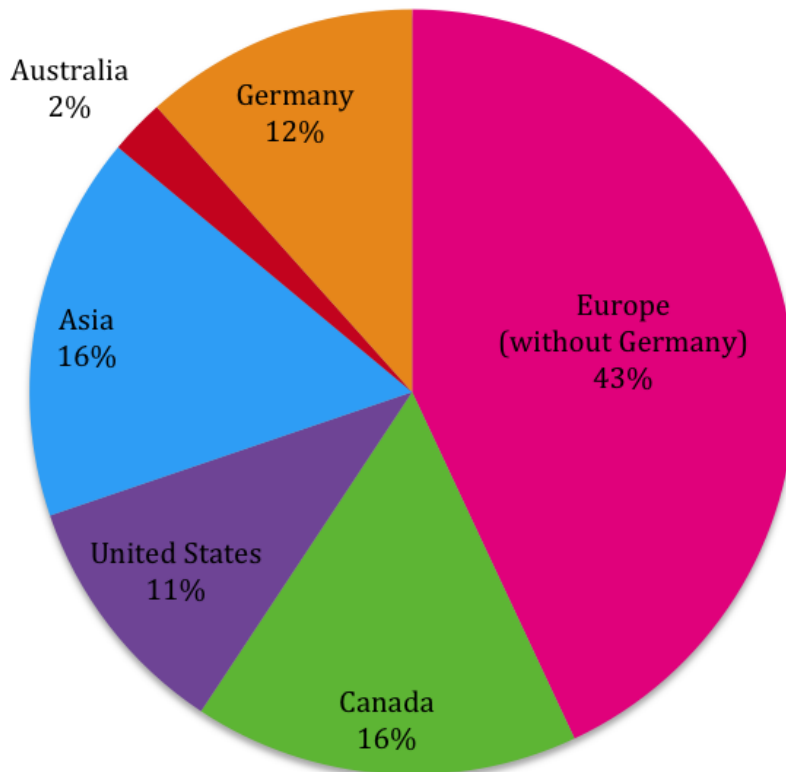
To make a bar graph:

1. Draw two lines (each called an **axis** ) that connect at a point (it should look like half a rectangle).
  2. Label each axis with a title.
  3. Label the **vertical axis** or **y-axis** with numbers. Be sure to write the units if they are not already in the title.
  4. Label the **horizontal axis** or **x-axis** with categories.
  5. Draw a bar above each category. The height of the bar should be the number that corresponds with that category. The width of the bar isn't important, but should be the same for all categories.
  6. Give the graph a title.
-

## Pie Chart

Pie charts are useful when you want to show the relation between the parts of something. Pie charts allow us to see how the parts come together to make a whole.

### 2010 Winter Olympic Gold Medals



| Region                   | Number of Medals |
|--------------------------|------------------|
| Europe (without Germany) | 37               |
| Canada                   | 14               |
| United States            | 9                |
| Asia                     | 14               |
| Australia                | 2                |
| Germany                  | 10               |

Data from: <http://olympic.ca/games/2010-vancouver/>

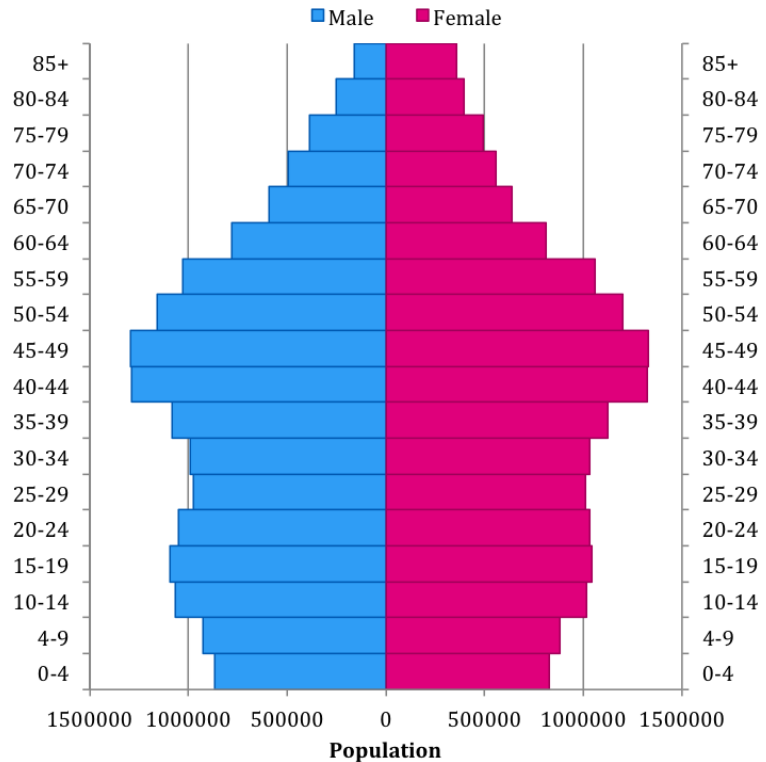
To draw a pie chart:

1. Start with a circle.
2. Calculate the percent for each category (if not already given).  
Here is how:
  - (a) Calculate the total of all categories (add up all the values).
  - (b) Divide the value of a category by the total and multiply by 100. This is the percent.
3. Divide the circle according to the calculated percents.
4. Label each section with the category and percent.
5. Give the graph a title.

## Population of Canada (2006)

### Population Pyramid

Population pyramids are useful to show how a population is changing and how it will change in the future. It is easy to see whether the population is growing, shrinking, or ageing. You can also compare the number of males and females. In population pyramids, being able to determine the precise numbers isn't as important as the general shape of the graph.



Data from: <http://www12.statcan.ca/census-recensement/2006/dp-pd/prof/92-591/index.cfm?Lang=E>

To make a population pyramid:

1. Begin with a grid (either a full grid or one similar to the example above). Label the horizontal axis "Population".
2. Draw a vertical line that divides the graph in half. Label this point "0" on the horizontal axis. Label one half "Male" and the other half "Female".
3. Continue labelling the right half of the horizontal axis as you would label the vertical axis of a bar graph. Label the left half using the same numbers but starting with the largest number on the left, and getting smaller as you move toward 0.  
Note: the horizontal axis can be labelled with the number of people or the percent of people.
4. Label the vertical axes with age categories on both sides.
5. Turn the page so the horizontal axis becomes the vertical axis. Follow the remaining instructions for a bar graph.
6. Turn the page the other way and do the same for the other gender.
7. Give the graph a title.

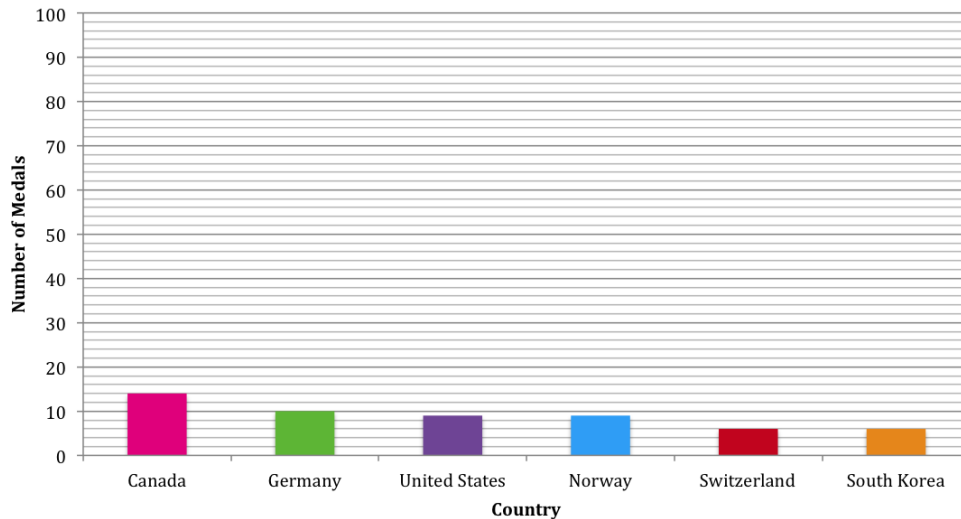
# Manipulating Graphs

Who got the most gold medals in each graph below? \_\_\_\_\_

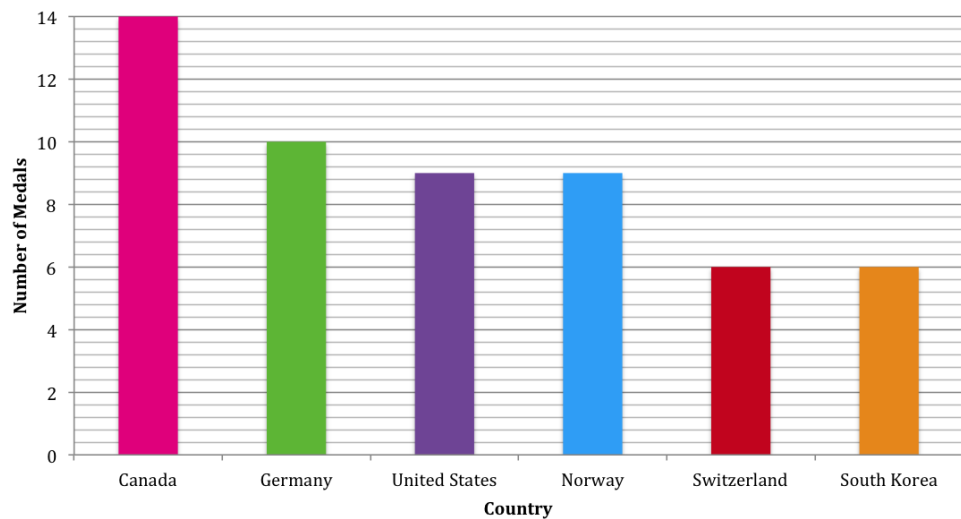
Was it close or did they win by a lot? \_\_\_\_\_

Did Canada win a lot of gold medals? \_\_\_\_\_

**Winter Olympic Gold Medals (Top 6) #1**



**Winter Olympic Gold Medals (Top 6) #2**

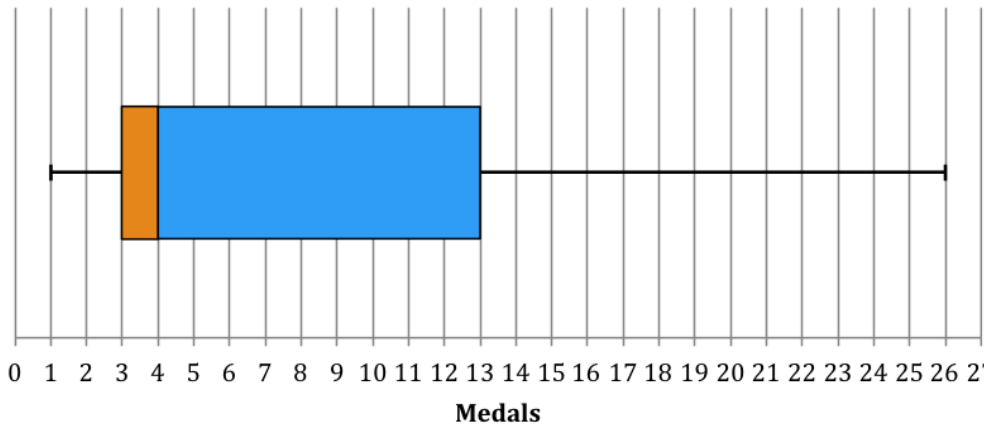


Both graphs are in fact based on the same data, but with different scales. There are times when you expect a set of data to show a specific trend, but what if the graph of your data doesn't emphasize what you want it to? For example, say you wanted to show that Canada had a lot more gold medals. The first graph doesn't show this very well; in fact, it shows that Canada barely had more gold medals. To force the data to show what they want it to, many people, including large companies, will manipulate their graphs. Although the data isn't being changed, Statisticians can change things like the scale of a graph to convey a different message. **It is important to note that differences and trends can only be exaggerated or minimized, they CAN NOT be changed.**

## Box Plot

Box plots present data through their quartiles. This shows how evenly the data is distributed. Box plots can only be used to present similar data.

### Canadian Winter Olympic Medals (1936-2010)



| Year | Medals | Year | Medals |
|------|--------|------|--------|
| 1936 | 7      | 1980 | 3      |
| 1948 | 3      | 1984 | 4      |
| 1952 | 2      | 1988 | 5      |
| 1956 | 3      | 1992 | 7      |
| 1960 | 4      | 1994 | 13     |
| 1964 | 3      | 1998 | 15     |
| 1968 | 3      | 2002 | 17     |
| 1972 | 1      | 2006 | 24     |
| 1976 | 3      | 2010 | 26     |

- **Quartiles** divide data into four parts based on the number of data points.
- The **1st quartile (Q1)** or **lower quartile** separates the first quarter from the second quarter.
- The **median** or **2nd quartile (Q2)** divides data in half. It is the middle number.
- The **3rd quartile (Q3)** or **upper quartile** separates the third quarter from the fourth quarter.

---

You can find each quartile by following these steps:

1. Arrange the data in order.
2. Find the total number of data points and divide by 2.
3. To find the **median**:
  - (a) If the total is even, find that data point and the one after it. Add them together and divide by 2. This is called taking the **average**.
  - (b) If the total is odd, round up and find that point (find next data point that is greater than your result)

4. To find the **1st quartile** repeat steps 1-3 with all data to the left of the median.
5. To find the **3rd quartile** repeat steps 1-3 with all data to the right of the median.

---

**Example** Find the median of the data set: 1, 4, 2, 5, 3, 4, 2

1. Arrange the data in order: 1, 2, 2, 3, 4, 4, 5
2. Find the total number of data points and divide by 2. There are 7 data points.  
 $7 \div 2 = 3.5$  or 3.5
3. Since the total is odd, take the next highest point, which is the 4th point. The fourth number in the list is 3. This is the middle number or median.
  - If we remove the 3 from the list we have an even number of points (1, 2, 2, 4, 4, 5). Since there are 6 points and  $6 \div 2 = 3$ , we take the middle numbers which are the 3rd and 4th points (2 and 4). To find the average, we add the points and divide by 2:  $2 + 4 = 6$  and  $6 \div 2 = 3$ . So the median of this list is 3.

---

To construct a box plot:

1. Start with a grid (or just vertical lines).
  2. Title the horizontal axis and number it with an appropriate scale.
  3. Arrange the data in order and calculate the first quartile, median, and third quartile.
  4. Draw a small vertical line to mark the lowest value and the highest value.
  5. Draw a longer vertical line to mark the first and third quartiles.
  6. Connect the tops and bottoms of the first and third quartile marks to form a box.
  7. Draw a vertical line to mark the median. This line should connect with the box to divide it into two parts (the parts might not be even as in the graph above).
  8. Draw a horizontal line to connect the lowest value mark to the box. Do the same for the highest value mark.
  9. Give your graph a title.
-

# Problem Set

“\*” indicates challenge question

1. The following table shows temperatures (in °C) for Prince George, BC. Use the table to make a line graph and answer the following questions.

|        |         |         |         |         |         |         |         |
|--------|---------|---------|---------|---------|---------|---------|---------|
| Jan. 1 | Jan. 2  | Jan. 3  | Jan. 4  | Jan. 5  | Jan. 6  | Jan. 7  | Jan. 8  |
| 1      | 1       | 1       | -7      | -9      | -5      | -5      | 0       |
| Jan. 9 | Jan. 10 | Jan. 11 | Jan. 12 | Jan. 13 | Jan. 14 | Jan. 15 | Jan. 16 |
| 1      | 0       | 0       | 1       | 5       | 5       | 6       | 4       |

Source: [http://climate.weather.gc.ca/climateData/dailydata\\_e.html?StationID=48370&Month=1&Day=21&Year=2014&timeframe=2&type=line&MeasTypeID=maxtemp](http://climate.weather.gc.ca/climateData/dailydata_e.html?StationID=48370&Month=1&Day=21&Year=2014&timeframe=2&type=line&MeasTypeID=maxtemp)

- (a) Is there a general trend in the temperature? If so, what is it?
  - (b) What does this line graph tell you about the weather?
  - (c) Was it useful to draw a line graph as opposed to just having the table?
2. Lauren’s class of 28 students just had a big test. Her teacher wants to analyze how the class did. He chooses to draw a bar graph. 15 people in the class got an A, 7 people got B’s, and 3 people got C’s. How did the class do? Was a bar graph the best choice to analyze this data? Suggest a better graph or explain why a bar graph is best. (Hint: don’t forget to include the people who failed)
  3. Pick an appropriate topic and draw a pie chart. Use your pie chart to draw three conclusions about your topic.

4. Use the following data to draw a population pyramid and answer the questions.

- (a) How is the population changing? Explain.
- (b) Why are population pyramids usually close to being symmetrical? What are some reasons they wouldn't be symmetrical?
- (c) Why aren't they perfectly symmetrical?

| Age   | Male      | Female    | Age   | Male      | Female    |
|-------|-----------|-----------|-------|-----------|-----------|
| 0-4   | 903,447   | 859,660   | 50-54 | 1,366,005 | 1,350,538 |
| 5-9   | 908,338   | 866,122   | 55-59 | 1,232,232 | 1,249,219 |
| 10-14 | 941,478   | 891,818   | 60-64 | 1,052,840 | 1,080,541 |
| 15-19 | 1,080,088 | 1,020,640 | 65-69 | 873,309   | 911,774   |
| 20-24 | 1,205,180 | 1,139,365 | 70-74 | 622,577   | 705,736   |
| 25-29 | 1,184,001 | 1,126,839 | 75-79 | 454,792   | 569,009   |
| 30-34 | 1,175,643 | 1,127,243 | 80-84 | 332,154   | 474,015   |
| 35-39 | 1,128,232 | 1,087,462 | 85-89 | 187,113   | 330,726   |
| 40-44 | 1,141,548 | 1,122,110 | 90-94 | 77,090    | 169,892   |
| 45-49 | 1,258,158 | 1,253,805 | 95-99 | 22,878    | 63,126    |
|       |           |           | 100+  | 4,303     | 17,165    |

5. For each of the following, list the minimum data requirements needed to effectively present the data.

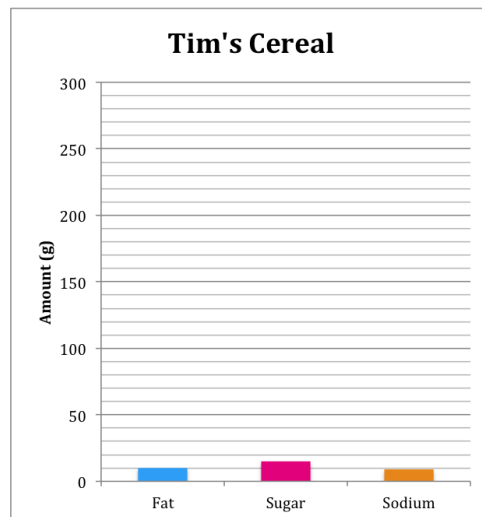
- (a) Table
- (b) Bar Graph
- (c) Pie Chart
- (d) Population Pyramid
- (e) Line Graph
- (f) Box Plot

6. Which of the ways to present data from question 5 is

- (a) the least specific (has the least requirements). Why?
- (b) the most specific (has the most requirements). Why?

7. Given the following data, which types of graphs could you use?
  - (a) Total population of Canada over the past 10 years.
  - (b) Breakdown of a population in 2013.
8. What was done to the graph "Winter Olympic Gold Medals (Top 6) #1" (in the Manipulating Data section) to make it look like all the countries had a similar amount of gold medals?
9. What could have been done better in the graph "2010 Winter Olympic Gold Medals" (in the Bar Graphs section)? (Hint: look at the categories)
10. In each case, choose which graph would be better. Why?
  - (a) Billy wants to show how his company has grown over the past 20 years. Should he use a bar graph or a line graph?
  - (b) Emily wants to compare how many boys and girls are in her class. Should she use a pie chart or a population pyramid?
  - (c) \*Nadine has to graph the amount of papers in 20 different boxes. Should she use a box plot, pie chart or a line graph?
11. Which graph(s) can't be manipulated easily? Why?
12. \*How can you use grouping to manipulate data?
13. Describe how you can change each of the following graphs to show the desired result.
  - (a) Bar graph
    - i. Make all the categories look similar
    - ii. Show a big change between categories
  - (b) Population pyramid
    - i. Make all the categories look similar
    - ii. Show a big change between categories
  - (c) Line graph
    - i. Make it look like the graph is steady (little or no change)
    - ii. Show that there is an increase

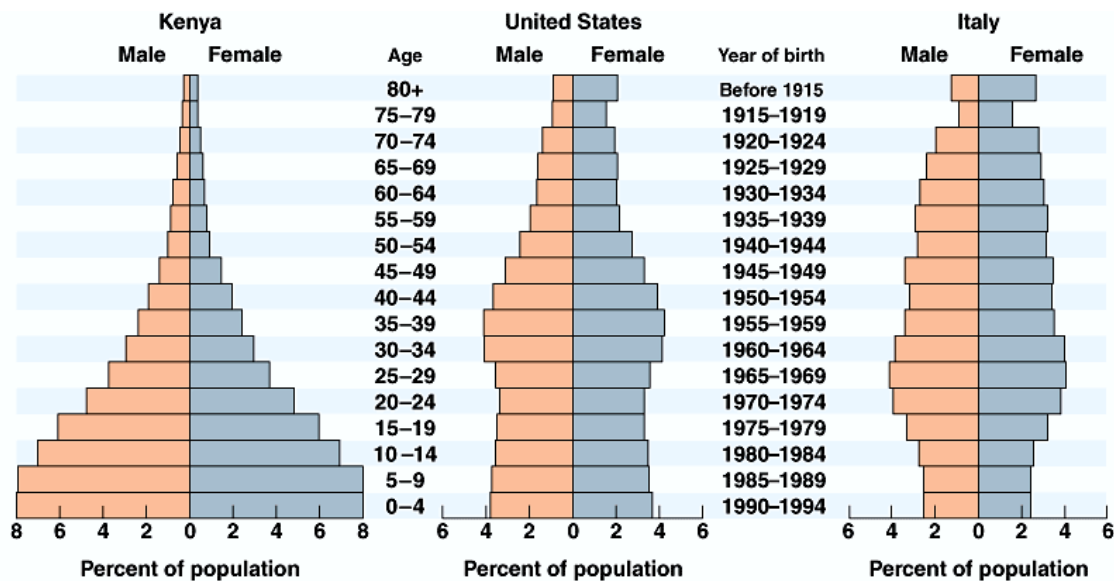
- (d) \*Pie chart (Hint: think about manipulating the categories)
- Show that a category makes up less of the whole (has a smaller percentage)
  - Show that a category makes up less of the whole (has a smaller percentage)
- (e) \*Box plot
- A small range of data
  - A large range of data
14. When drawing a bar graph or pie chart it is important to pick good categories. For example, instead of drawing a bar graph with 20 bars, you might group the data into 5 different categories. Find a different way of grouping the following.
- Letters of the alphabet
  - Soccer, football, hockey, golf, and dance teams.
  - Schools
  - TV shows
  - \*Can you find another way to group each of the above categories?
15. Tim works for a cereal company. He created this graph to show how healthy his cereal is. Tina is a doctor and wants to show that Tim's cereal isn't really that healthy. Help Tina by redrawing the graph.



16. \*Can you think of a place where statistics is not used? This could be an industry, or a sport, or a company, or anything else.

17. **Demography** is a branch of statistics that deals with human populations. Demographers analyze population pyramids to determine how a population is changing. Using the population pyramids below, match each country with a description of how their population is changing.

- |                   |                    |
|-------------------|--------------------|
| (a) Kenya         | (i) Slow growth    |
| (b) United States | (ii) Decreasing    |
| (c) Italy         | (iii) Rapid growth |



Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.  
<http://blue.utb.edu/paullgj/geog3320/lectures/AgeStructures.gif>

18. \*In ice hockey players are given a plus/minus rating after each game. Players get a point added to their rating if they do something good, such as scoring a goal. A point is subtracted if they do something bad, such as letting the other team score a goal. Plus/minus ratings from each game are added together to get a season total. Below is the plus/minus rating for the Toronto Maple Leafs this season. Use this data to draw a box plot.

|    |    |    |     |     |   |     |    |    |     |    |    |    |    |    |
|----|----|----|-----|-----|---|-----|----|----|-----|----|----|----|----|----|
| -2 | -4 | 2  | -12 | -12 | 5 | -12 | 10 | -2 | -12 | -4 | 4  | 0  | -8 | -2 |
| -5 | 9  | -3 | 0   | -1  | 0 | -4  | -8 | -1 | 1   | -5 | -6 | -3 | -1 | -2 |

Source: <http://mapleleafs.nhl.com/club/stats.htm>

- (a) Using your box plot, determine where most of the data is (ie. is it grouped toward the front, back, evenly spaced, etc.)

- (b) How would the median change if the following were applied? How would the box plot change?
- i. 2 points were added to each player's plus/minus rating
  - ii. Each player's rating was doubled.
  - iii. All negative ratings became positive (-2 becomes 2, -5 becomes 5, 1 stays the same, etc.)
  - iv. All numbers switched signs (-2 becomes 2, 1 becomes -1, etc.)
- (c) What conclusions can you draw about the Maple Leafs' season so far?