CHAPTER SIX: SPREADSHEETS   Test: Tuesday, February 25

Spreadsheets are an extremely useful and versatile tool. Some basic knowledge allows many basic tasks to be done beautifully and almost effortlessly, such as simple data management and graphing. However, a background in mathematical concepts and logical thinking, such as learned in this course, can make spreadsheets much more powerful, because they can be programmed to make recalculations instantly based on new inputs. In particular, a solid understanding of functions and arguments is essential. There is an online tutorial in lieu of notes for the first section of this chapter.

6-A   Formatting and Formulas   Friday • 2/7

Select the Sheets Tutorial app from the course website.

❶ Format cells, rows, and columns.

❷ Identify a function's arguments.

❸ Copy formulas into additional cells.

❹ Use basic one-argument functions such as `sum`, `average`, and `count`.

❺ Use functions with criterion arguments such as `if`, `countif`, and `sumif`.

❻ Use data-lookup functions such as `vlookup` and `hlookup`.

❼ Use data validation to ensure valid data entry.

❽ Find and use a function appropriate for a given purpose.

❾ Apply conditional formatting based on specified criteria.

❿ Create a pivot table.

6-B   Graphs   Thursday • 2/13

data • level of measurement • nominal • ordinal • interval • ratio • time plot • scatter plot • bar graph • circle graph • histogram

❶ Classify data by level of measurement.

❷ Choose an appropriate type of graph for a data set.

❸ Make a circle graph by hand.

❹ Make a time plot, scatter plot, or bar graph by hand.

❺ Use a spreadsheet to make a graph.

6-B   Graphs

DATA is a plural word that essentially means *observed values*.

Data can be classified into four LEVELS OF MEASUREMENT, listed below from lowest to highest:

NOMINAL Data can be sorted into categories, such as yes and no or red, green, and blue, but they are not necessarily rankable.

ORDINAL Data can be put in order, such as first to last or strongest to weakest, but the difference between positions may not be known. For example, the first place finisher may have beaten the second place finisher by five seconds or by two minutes.

INTERVAL Data can be subtracted to find a meaningful difference, such as 60° being 40° warmer than 20°, but may not be able to be divided to find a meaningful ratio. In this case, it would not make sense to say that 60° is three times as warm as 20°.

RATIO Data can be divided to find a meaningful ratio, such as $60 per hour being three times as much as $20 per hour. For ratio data, zero is the baseline, so a value of zero means there is none of what is being measured.

Any variable can be treated as a lower level of measurement. For example, temperatures are interval data but also can but put in order from coldest to warmest.

❶ Classify data by level of measurement.

     1. All data can be considered nominal.

     2. If the data can be put in order based on what they are measuring, they are ordinal.

     3. If the data have specific numeric values independent of the other data, they are interval. For example, *9 minutes* is a specific amount, but *slower* or *7th place* could represent different values depending on the other data.

     4. If a value of zero would mean there were none of the variable (even if such a value is not possible in reality), then meaningful ratios can be found and the data are ratio.

     ❶ For each of the following variables, give a possible way it could be measured and what level of measurement applies.

a) color	b) temperature	c) iPhone capacity	d) Yelp rating
color name	degrees Celsius	gigabytes	number of stars
nominal	interval	ratio	ordinal

Different types of graphs are used for different types of data.

A TIME SERIES Plot is used for *x-y* data pairs when the independent variable *x* is time on a timeline and the graph represents a function.
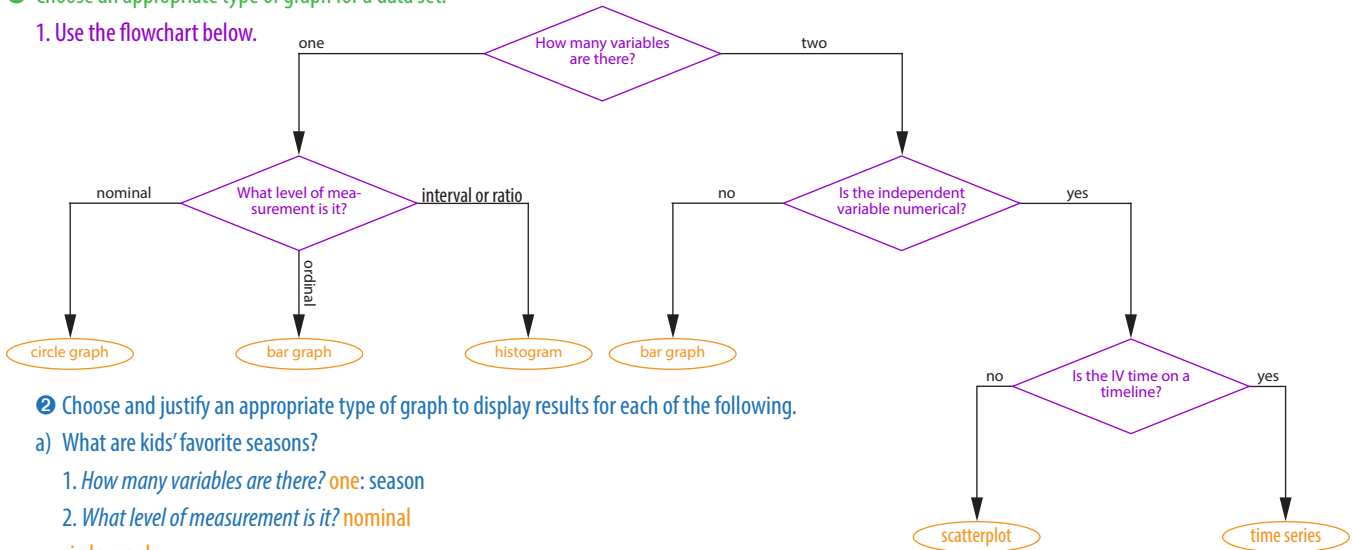
A SCATTER Plot is used for *x-y* data pairs that do not necessarily represent a function (two data points could have the same *x* value).

A BAR Graph is used to compare averages, totals, or frequencies of two or more categories. When there is only one variable, making the total of all categories 100%, a CIRCLE Graph (pie chart) can be used instead.

A HISTOGRAM is a bar graph in which the *x*-axis has equal-width numerical categories that connect (e.g., 0-10, 10-20, 20-30) and the *y*-axis shows the frequency or proportion of data in each range.

❷ Choose an appropriate type of graph for a data set.

1. Use the flowchart below.

How many variables are there?
- one → What level of measurement is it?
  - nominal → circle graph
  - ordinal → bar graph
  - interval or ratio → histogram
- two → Is the independent variable numerical?
  - no → bar graph
  - yes → Is the IV time on a timeline?
    - no → scatterplot
    - yes → time series

❷ Choose and justify an appropriate type of graph to display results for each of the following.

a) What are kids' favorite seasons?
  1. *How many variables are there?* one: season
  2. *What level of measurement is it?* nominal
  circle graph

b) How much do adults pay for food each month?
  1. *How many variables are there?* one: monthly food bill
  2. *What level of measurement is it?* ratio
  histogram

c) How much rainfall has Santa Cruz had in each of the past 30 years?
  1. *How many variables are there?* two: year (independent variable) and amount of rainfall (dependent variable)
  2. *Is the independent variable numerical?* yes
  3. *Is the independent variable time on a timeline?* yes: there is a data point for each of the past 30 years
  time series plot

The most important aspect of a graph is that it makes the results clear at a glance. This can be achieved by using a type appropriate for the data, precisely labeling it, and keeping it neat and not unnecessarily complicated.

❸ Make a circle graph.

1. Multiply each proportion by 360°.
2. Use a protractor to measure out each segment.
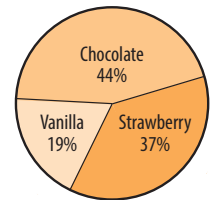3. Title the graph and label the slices.

❸ Out of 70 kids, 31 preferred chocolate, 26 preferred strawberry, and 13 preferred vanilla.

chocolate: $\frac{31}{70} \cdot 360° = 159°$ strawberry: $\frac{26}{70} \cdot 360° = 134°$ vanilla: $\frac{13}{70} \cdot 360° = 67°$

Kids' Favorite Ice Cream Flavors

Chocolate 44%
Vanilla 19%
Strawberry 37%

❹ Make a time series plot, scatter plot, or bar graph.

    1. Label and scale the independent variable on the *x*-axis. For a bar graph, label each level of the independent variable, and put the bars in a logical order, such as first to last or highest to lowest.

    2. Label and scale the dependent variable on the *y*-axis. The scale should represent an appropriate range for the variable, starting at zero unless there is a specific reason to do otherwise, and it should be in natural, equal increments with consistent decimal places.

    3. For a bar chart, draw the bars. If proportions are being compared, divide to convert each bar's frequency to a proportion.

       For a time series plot or scatter plot, plot the points.

       For a time series plot, connect the points.

    4. Give the graph a title that clarifies its purpose.

    ❹ Make up realistic data and create each of the following:

      a) a time series plot showing Matt's height through age 15

      b) a scatter plot showing how long 25 kids of different ages could hold their breath

      c) a bar chart comparing middle school boys and girls on whether or not they regularly play video games
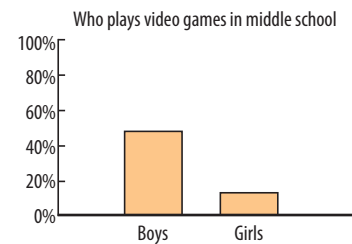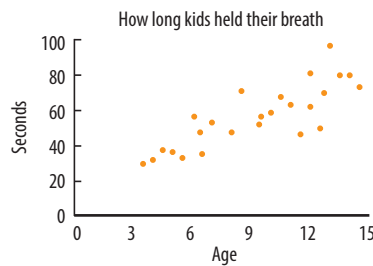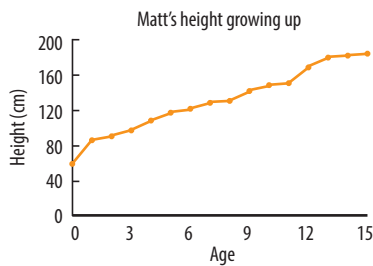
Spreadsheets can be used to quickly make clear, precise graphs. However, care must be taken to fix default settings that are not appropriate for your data. The directions below will vary some between programs and between different versions of the same program.

❺ Use a spreadsheet to make a graph.

    1. Enter your data or calculated statistics. In most cases you can enter labels into cells now or add them to the graph later.

    2. Select the cells containing the information to be put into the graph.

    3. Choose *Chart...* from the *Insert* menu.

    4. Choose the appropriate type of chart (see ❷).

    5. Choose *customization*.

    6. Title the graph.

    7. Legends are needed in specific cases, such as for factorial designs (more than one independent variable). Remove the legend if this does not apply.

    8. Keep the text and color formatting simple. In general, a sans-serif font such as Ariel, and one color per independent variable (except for circle graphs), is appropriate.

    9. For a time series plot, scatter plot, or bar graph, label and scale the independent variable on the *x*-axis and the dependent variable on the *y*-axis (see steps 1 and 2 in ❹).

    10. If something needs to be changed once the graph has been placed, try double-clicking on it or two-finger-clicking on it.

    ❺ In Statistics one semester, the average grade for boys was 83.9% and the average grade for girls was 90.0%.

    1. Enter "Boys" in a cell and their percentage in the next cell. Below this, do the same for girls.

    2. Select these four cells.

    4. Choose a bar graph.

    6. Add a title.

    7. There is only one independent variable, so set the legend to *none*.

    9. Normally the independent variable should be labeled at the bottom, but in this case labeling only "boys" and "girls" with no stated variable is sufficiently clear.

        For a grades graph in a college-prep class such as Statistics, 50% is an appropriate minimum for the *y* axis, and 100% is appropriate for the maximum.

        Scaling in percentages with no decimals is clean and simple.