

MATH 8: UNIT 1: BIVARIATE DATA

Because statistics hold such a prominent place in modern society, they are a natural starting point for 8th grade math. Students will create and analyze tables, scatter plots, lines of best fit, and other visual representations of bivariate data. *Statistics* means counting what's important to understand something, like various numbers in baseball (batting average, fielding average, earned run average, etc.) *Bivariate data* means looking at two areas of information, often in relationship, such as:

- Hoover MS students playing more video games in summer than during the school year.
- San Jose residents buying more clothing of local sports teams after successful seasons.
- Hoover MS students having higher grades if they study before tests.

Useful text: pages 224 - 232

Section 1: Creating a graph for bivariate data (like video game playing hours, and month of year to show more playing during vacation).

Section 2: Bivariate data relationships and terms:

- a. *Positive, negative, or no relationship* between the two areas measured.
- b. *Outlier* data (unusual relationship with some data points).
- c. *Clusters* of data.
- d. *Line of best fit* to create a straight line on the graph showing the trend of the data.
- e. *Slope* showing how steep a line of best fit is measured.
- f. Completing a two-way table of bivariate data (like 10+ hours of video games Mon-Fri in July versus 10+ hours in September).

Section 1: Creating a graph for bivariate data

- a. A table of data is one way to show information. We'll create several in class for your notes.
- b. We then use these tables to plot the data on a graph. This can be called a *scatterplot* because, well, dots are scattered on the graph. We'll practice on the notes with our examples.
- c. In creating a graph, use numbers that create the graph/picture so the graph clearly shows any trends. Again, we'll practice so you'll have a feel of selecting number ranges that best show the information.

Section 2: Bivariate data relationships and terms

- a. *Positive, negative, or no relationship* means a positive slope (as x increases, y increases), negative slope (as x increases, y decreases), and then no apparent relationship that what happens to the x measurement has any effect on what y measures.
- b. *Outliers* are data points that are different from a general trend. For example, in general the more education a person has, the higher income they earn. However, there are a few outliers that show the usual pattern isn't always what happens; that is, there are unusual facts almost everywhere in life where we look.
- c. *Clusters* are groupings where several data points are close together; that is, the opposite of outliers where we find what's usual for what data we're measuring.
- d. A *line of best fit* divides the data points into two equal groups as best possible to show the trend of the information. We'll practice to create these.

- e. *Slope* measures how steep a straight line is, and is applied to a line of best fit to use a number that shows steepness. Numbers do that: little numbers show a little steepness (in this case), big numbers show big steepness, and negative numbers show going downhill! Slope means the change of the y-axis over change of the x-axis, or rise/run. We'll practice so you can see these in action, and tell what a slope number means for how steep a graphed line is, and which direction it goes :)
- f. We can use slope in a line of best fit to predict the future in some cases, and general trends outside what we measured. We'll practice this prediction with examples.
- g. Two-way tables of data are another way to show data. We'll practice.