

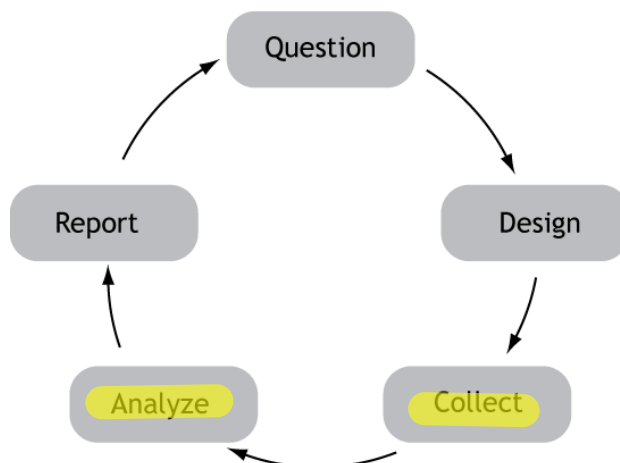
Unit 3B

Analyzing Data

Learning Outcomes

1. Students interpret a variety of graphical displays of statistical information.
2. Students estimate center, shape, spread, and unusual features of graphical displays and use these characteristics to describe distributions.
3. Students analyze the appropriateness and usefulness of statistical graphical displays.
4. Students prepare appropriate statistical reports and communicate in oral and written form.
5. Students collect sets of data and create a variety of displays.
6. Students compare and contrast multiple data sets.

Research Cycle



Explore different ways to sort, analyze, and visualize data.

Why is this important?

- different types of data

U3B-5

Histograms

bin size - in a bar graph, the *size* (not the space) of the interval used. Think of each bar as a *bin* for data



distort - to misrepresent data either through the misuse of statistical analysis or presentation of data. Poorly/complex constructed graphs may also distort data.

frequency - the rate at which something occurs or is repeated over a particular period of time or in a given sample

frequency table - *data*

interval width - in a frequency diagram, the *size* (not the space) of the interval used.

skewness - a measure of the tendency for the distribution of values to be more spread out on one side than the other.

- What do you think categorical data means?

To what word is categorical related?

information (words / non numerical)

- What do you think quantitative data means?

To what word is quantitative related?

#'s

- What do you think univariate means?

What does uni mean?

1 variable (frequency)

- What do you think bivariate means?

What does bi mean? 2 variables

f
x . . . y . . .

Univariate means "one variable" (one type of data)

Example: Travel Time (minutes): 15, 29, 8, 42, 35, 21, 18, 42, 26

The variable is Travel Time 8, 15, 18, 21, 26, 29, 35, 42, 42

We can do lots of things with univariate data:

Find a central value using mean, median and mode
 Find how spread out it is using range, quartiles and standard deviation
 Make plots like Bar Graphs, Pie Charts and Histograms

Bivariate means "two variables", in other words there are two types of data

With bivariate data you have two sets of related data that you want to compare:

Example:

An ice cream shop keeps track of how much ice cream they sell versus the temperature on that day.

The two variables are Ice Cream Sales and Temperature.

Here are their figures for the last 12 days. What would be the best way to display this data?



Ice Cream Sales vs Temperature	
Temperature °C	Ice Cream Sales
14.2°	\$215
16.4°	\$325
11.9°	\$185
15.2°	\$332
18.5°	\$406
22.1°	\$522
19.4°	\$412
25.1°	\$614
23.4°	\$544
18.1°	\$421
22.6°	\$445
17.2°	\$408

- a. What year were you born?
- b. What is your gender?
- c. How many text messages did you send yesterday?
0-30 31-50 51-70 71-90 over 90
- d. How many people under the age of 18 live in your house?
- e. Which is your favorite food from the following choices:
pizza, hamburgers, sushi, salad, chicken, other?

- f. Which is your favorite drink from the following?
Milk, Water, juice, tea, soft drink, lemonade
- g. How many hours do you work in an average week?
0 1-5 6-10 11-15 16-20 Over 20
- h. What is the average number of hours you spend on the Internet per week?
0 1-5 6-10 11-15 16-20 Over 20
- i. How many hours a week do you listen to music?
0 1-5 6-10 11-15 16-20 Over 20
- j. What is your favorite kind of music?
Rock Classical Country Pop Latin

- k. List your gender and your shoe size.
- l. List your gender and number of books read this month.
- m. Of the following, what is your favorite subject?
Math Science English Social Studies
- n. How many hours a week do you do homework?
0-5 6-10 11-15 16-20 over 20
- o. List your gender and how many hours of sleep you got last night .
0 1-3 4-6 7-9 10-12 over 12

In pairs create a bar graph for one of these sets of data.