

# Introductory Statistics – Day 2

## Descriptive Statistics

# Class Check-In

- Did everyone find the Moodle preview?
- Did everyone find the first reading?
- Did everyone read the first reading?
- Did everyone get logged into WeBWorK for the online homework?

## What type of study is this?

A group of students wants to determine whether talking on a cell phone while driving causes people to get into accidents. They request records from the Helena Police Department for all motor vehicle accidents over the past year and calculate the percentage of accidents in which cell phones were a factor.

- A. Observational Study
- B. Controlled Experiment
- C. Survey
- D. None of these

If this is a controlled experiment: What is the treatment?  
Can you conclude a cause/effect relationship?

## **Activity 1:**

*Descriptive statistics and basic graphs in Excel and TinkerPlots*

Download the `VideoGames.xls` data set from the class website. The columns in the video game correspond with the attributes on the next slide.

column name	description and units
name	game title
platform	type of gaming system
year of release	when the game first appeared for this platform
genre	game type: e.g. sports, shooter, etc.
NA sales	sales in North America, in millions of dollars
global sales	total sales, in “millions” of dollars
critic score	average critic score, maximum of 100
critic count	number of critics
user score	average user score, maximum of 10
user count	number of users who provided scores
rating	E = everyone E10+ = everyone ten and older T = teen M= mature

The above table is sometimes referred to as a

*code book or data dictionary.*

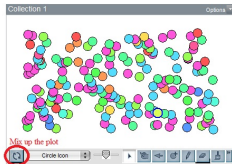
When you generate a data set, it is polite to include a code book so that others can identify what you measured and how.

## Now let's tinker with the data

- Open the data set in Excel. Look through the rows and columns to see what sort of data is available.
- Use `ctrl+A` to select all of the data, and use `ctrl+C` to copy the data to the clipboard in Excel.
- Open TinkerPlots and drag an empty table into the workspace. Use `ctrl+V` to paste the data into the table in TinkerPlots.
- *Tinker with your data.* Within TinkerPlots, drag a plot into the workspace. Spend 5 minutes exploring your data. Click on different columns in your table. Click and drag on points in your plot. Check in with a neighbor and see how they are organizing their data.

**Explore with purpose:** Download the Practice Page document from Moodle.

- Explore with TinkerPlots and generate a graph for each question.



- You can reset your plot to an unsorted cloud of data using the *Mix up plot* command in the lower left corner of the plot.
- If your cloud of dots is all grey, you have not selected an attribute to explore. Select a column of your choice in your table to see the data colored by that category.
- To copy an image from TinkerPlots: Edit → Copy As Picture. Then copy (ctrl+C) into Word document.
- Feeling stuck? Ask a neighbor or ask the teacher

## **Activity 2: Descriptive statistics in Excel**

Open the `VideoGames.xls` data set in Excel. Now we will use features of Excel to find common descriptive statistics.

### *Part I: Measuring Center*

- A. Find the mean, median, and mode for the user ratings and for the critic ratings for all data in the sample.
- B. Among the three measures of center (mean, median, and mode) which would be the most useful for summarizing this data and why?



## **Activity 2, continued.**

### *Part II: Measuring Spread*

- A. Find the standard deviation and variance for both the user ratings and critic ratings.
- B. The five number summary of a single variable data set consists of the minimum, lower quartile, median, upper quartile, and maximum. Find the five number summary for both the user ratings and the critic ratings.

### *Part III: Graphing in Excel*

- A. Explore the basic graphing options in Excel, under the Insert tab.
- B. Are there any graphs you can be make easily to summarize the data?

## Activity 2, continued.

### *Part IV: More Excel Tools*

- A. Use the descriptive statistics tool in Excel. What types of information are provided by this tool?
- B. Choose the *Data Analysis* option in the *Data* tab. Select *Descriptive Statistics* from the list of options. Then highlight the range of data you would like summarized, check the boxes *Labels in first row* and *Summary statistics*, and click OK.

*Note: If you do not see Data Analysis in your Data tab, do a quick online search for how to add the Data Analysis toolpack for Excel.*