



# Implementing GAISE recommendations

Robert Carver, Stonehill College

John Gabrosek, Grand Valley State University

Megan Mocko, University of Florida

Jeff Witmer, Oberlin College

Beverly Wood, Ambry-Riddle Aeronautical University

# GAISE 2016 Recommendations

- ▶ Teach statistical thinking.
  - ▶ Teach statistics as an investigative process of problem-solving and decision-making.
  - ▶ Give students experience with multivariable thinking.
- ▶ Focus on conceptual understanding.
- ▶ Integrate real data with a context and a purpose.
- ▶ Foster active learning.
- ▶ Use technology to explore concepts and analyze data.
- ▶ Use assessments to improve and evaluate student learning.


# Your Investigative Task

1. First 5 minutes – Get an Understanding of the Subject Matter / Data
2. Next 5 – 10 minutes **Determine a question** – Make sure that you are looking at 3 variables
3. Explore **Data**. Use visualizations to help understand relationships.
4. Find a visualization to share.
  - a) Prepare a 3 minute presentation to share what you found.
  - b) Prepare 1 or 2 slides (saved as a PDF). The slide should include a visualization of the data.
  - c) Explain how it can be used in your classroom.
  - d) Reflect on how this meets GAISE 2016 College Report.
    - ▶ How could you incorporate active learning, the investigative process, etc?
5. By \_\_\_\_, slide should be either dropped in the dropbox (ask one of the presenters) or use a USB.



# Pick a Data Set

1. Youth Health
2. Highway Fatalities
3. Lacrosse
4. Airline Data (On Time)
5. Missing Children



The following slides will give more detail on each data set and in some cases give an example of a visualization.

# Youth Health

- ▶ 2 Files

- ▶ Data File: YRBS2015.csv (A filtered version of 2008 – 2013 could be used.)

- ▶ Data Dictionary: YRBS\_2015\_sadc\_documentation.pdf (Can also be found online at [https://www.cdc.gov/healthyyouth/data/yrbs/pdf/2015/2015\\_yrbs\\_sadc\\_documentation.pdf](https://www.cdc.gov/healthyyouth/data/yrbs/pdf/2015/2015_yrbs_sadc_documentation.pdf))

# Highway Fatalities

- ▶ 8 Files
- ▶ Files good for an introductory course (where the goal is to get students to think about the investigative process, a little bit of data cleaning because some values (such as age 999) are obviously wrong, the use of multivariable thinking, and summarizing a relatively simple data set.)
  - ▶ FARS2011.csv – Data set used in the session.
  - ▶ FARS2011.xlsx – Data set used in the session.
  - ▶ FARS2011 Doc – This is the data dictionary for the FARS2011 file used in the session.
  - ▶ FARS\_Visualiations – Example visualizations from the FARS2011.csv data set

# Highway Fatalities

- ▶ Files good for an advanced course such as R programming, that requires merging data sets from different files, use the files below:
  - ▶ FARS2011\_2.csv – Same data set as used in session but without the state names (only has codes) and without the number of drivers for each state.
  - ▶ FARS2011\_2.xlsx – Same data set as used in session but without the state names (only has codes) and without the number of drivers for each state.
  - ▶ FARS\_StateCodes.xlsx – Excel file that connects state codes to state names.
  - ▶ FARS\_Licensed\_Drivers\_2011.xlsx – File that has the number of licensed drivers in each state broken down by gender and by age group.



# Lacrosse

- ▶ Four Files
  - ▶ Data File: Lacrosse\_women.csv
  - ▶ Data Dictionary: Lacrosse\_women\_Data dictionary
  - ▶ Visualization Examples: Lacrosse\_USCOTS2017.docx
  - ▶ Lacrosse\_USCOTS2017.Rmd (File from which the above was knitted)

# Airline Data (on time)

One File, multiple sheets called Airlines\_ONTIME.xlsx

- ▶ First sheet: Data File: Transtats\_ONTIME
- ▶ Second sheet: Data Dictionary called Variables includes the variable description as well as a description of where the data came from
- ▶ Third sheet: Example Visualization, includes the visualizations found by the two groups during the breakout session in USCOTS along with ideas about what might be done to explore multivariate thinking in class

# Missing Children

## ▶ 5 Files

- ▶ Full Original Data set: Missing\_Children\_ActiveCases.csv
- ▶ Dataset with born between 1/1/1999 or later (Additional variable age when reported missing added to data set) Missing-Children\_ActiveCases\_bornafter1\_1\_9.xlsx
- ▶ Census data by state: MissingChildren\_State\_PopulationSize.xlsx
- ▶ Data Dictionary : Missing\_Children\_Data\_dictionary
- ▶ Example Visualizations from the smaller data set: MissingChildren\_Visualizations