MA103: Mathematical Modeling & Intro to Calculus
Descriptive Modeling of a Data Point 2

Lesson Objectives: Cadets will

1. calculate z-scores and interpret them as measures of relative standing.
2. standardize datasets and identify usual and unusual data.

Admin Notes / Agenda

- Academic security brief
- Sit in groups - when are you meeting out of class? Agree to a time / location and record here.
- Scavenger Hunt corrections due. Quiz tomorrow - WPR 1 next Friday.
- Data set assigned is team number.
- Turn in (online) PSL 1 today.
- Lesson review - CDT Pickett.

Download the IOCT dataset from my instructor webpage. Open the file in Excel, make sure that you are not working in your internet browser.

Today we will explore the dataset using the math modeling triangle for descriptive statistics. Does the data represent a population or sample? Create meaningful visualizations that help to communicate the story of the dataset. Find all measures of center and spread for each variable.

1. **Transform**
    - Given:
        - What questions should we ask about the given data? Does the dataset represent a population or a sample? What research questions could we answer with this data?
    - Find: Select your own for the find step.
    - Explore:

  - Define variables: make a data dictionary

  - Make assumptions: Clean the data. What assumptions are **necessary** and **reasonable**?

  - Model type: explore visualizations that could communicate the story of the dataset, explore all measures of center and spread for all variables, standardize the data and identify possible outliers. Explore possible visualizations using the standardized data.

- Final Model: select visualizations and descriptive statistics to model the data
  - What questions should we ask about our model?
  - What story are we telling about the data with our selected visualizations and statistics? How would that story change if we selected different visualizations or statistics?
  - How do the data points you identified as outliers affect your chosen statistics and visualizations?

2. **Solve**

- Calculate the selected statistics and finalize visualizations. If you have already done calculations in the transform step, write out the values and possible outliers formally for the solve step.

3. **Interpret**

- What questions should we ask ourselves when we communicate results? Did we answer the research question? Are the visualizations and statistics (models) we selected valid? Is our data valid? Are we clearly communicating the data (ethical framework)?