

# Exploratory Data Analysis – Expectations

---

## Goals:

- Explore the variables in your data
  - If you were presenting your EDA to an audience, what would the audience want to know about the variables in your data?
    - Distributions, relationships, outliers, descriptive statistics, etc.
- Uncover insights about the data
  - What stories does your data tell?
- Answer (or partially answer) your research questions about your data

## Expectations:

- Perform descriptive statistics for variables
  - Frequency counts, contingency table, five number summary, mean, etc.
  - Use appropriate techniques depending on variable type (categorical, quantitative)
- Report on missing values in data
  - How many missing values, in which variables, and how/if this affects your analysis
- Create compelling visualizations that show:
  - Distributions, relationships, outliers, descriptive statistics, etc.
  - Data stories
  - Answers or partial answers to your research questions
    - Histogram, box plot, scatter plot, pairs plot, bar graph, map, etc.
    - Use appropriate techniques depending on variable type (categorical or quantitative)
    - Use appropriate techniques depending on variable (for example, lat and long variables would most likely lead to a map visualization)
    - Visuals should be labeled, titled, with legends (if applicable), etc.
- Add any **necessary** text to explain visuals, trends, data insights, research question (partial) answers, etc.
  - Text should be limited and in sentence form. No long paragraphs!

## Deliverables:

- Code in Python, R, etc.
  - Code should be commented appropriately, readable, and in a format that incorporates code, text, and visualizations (R Markdown or Jupyter Notebook)
- If you use a different tool to complete EDA:

- Tool is capable of creating dashboards, for example Tableau or Exploratory.io: you may create dashboards that incorporate text and visualizations.
  - Make sure that your dashboards are public so that viewers without an account for that tool can see your dashboards
- Tool is not capable of creating dashboards: you may create a document that incorporates code (if used), text, and visualizations.
  - This may involve copying and pasting code and/or visualizations into a document.
- Deliverables should be well-organized and visual appealing
- Post your deliverables on your GitHub page in your capstone project repository.
  - For dashboards or other items you cannot post, post a link to them instead.