Unit 7 Notes – Data Visualization and Data Storytelling

First Video Slides: <https://github.com/cs109/2015/blob/master/Lectures/03-EDA.pdf>

Downloaded on Hard Drive

**Exploratory Data Analysis and Visualization**

* Always have an initial question before going into the data
  + This can change, but ask a scientist you need to go in with a hypothesis
* The initial example was the antibiotic table
  + Which antibiotics are most effective against different bacteria? (gram negative vs. gram positive)
  + Which antibiotics should be combined?
* Visualize that data and come up with more questions before apply any statistics
  + Explain and explore
* Can use seaborn for visualization in Python
  + Go to Seaborn website for tutorial
* Only hard rule: bar charts start at zero
  + If you change axis from zero-base
* Chart Suggestions – A Thought Starter
* Histograms: Always play with the bin size
* Create a lot of visualizations quickly for one thing without attachment, so you can figure out what works and what doesn’t
* Don’t use color intensity or colors to show quantitative differences, could show cardinal differences or categorical
  + Position and length are better for quantitative differences
* Color brewer in seaborn to get a color scale (easy to read, not offensive, no color blindness issues)

Second Video Slides: <https://github.com/cs109/2015/blob/master/Lectures/06-StoryTelling.pdf>

Downloaded on Hard Drive

**Storytelling and Effective Communication**

* Essential Question
  + What is the goal (of communication in DS)?
  + Who cares (and why should they care)?
* IMAC
  + Inferential goal
  + Model
  + Algorithms
  + Conclusions and checking
* Know your audience, tell a story, be careful with notation, be a great writer, have a sense of direction (sign-posting, following a path)
* It is good if your story (about data) brings about new questions
* Use visuals
* The presentation should be an advertisement for your paper
* What do people know beforehand? Did they have biases?
* Framing – Why should I care?
  + Everyone has an opinion and a point of view (share your own)
    - Still, keep your audience in mind
  + Engage people’s emotions (but keep your integrity)