David Kinney DSC 550

Professor Deitenbeck Winter 2019

3.4 Reflection: Week 3 Reflection

I can state without exaggeration that visualization is the part of Data Science that drew me in. Oh, all those alluring graphs! Ironically, the graphs that had the most aesthetic impact on me are the same graphs I have never used. In truth, I have yet to wade into the deep end of the pool. My graph/plotting world so far has been limited to scatterplots, bar charts (vertical, horizontal, stacked, 2d-binned, and histograms), line charts and box and whiskers plots. In the world of finance however, I have found I don’t need much beyond that set of tools.

My first exposure to visualization in Data Science was with Hadley Wickham’s **ggplot2** package for the R programming language. Modeled on the “Grammar of Graphics” [1] I found plotting to be so intuitive: data, geometric objects (geoms) and aesthetics.

As fate would happen, academics and career conspired to steer me back to Python. Therefore, you can imagine how happy I was to stumble upon **plotnine**, a veritable clone of **ggplot**.

All of this is to say, *all* of this week was a struggle to me. I do not find matplotlib user-friendly. I understand that matplotlib grew from the MATLIB ecosystem, which I imagine is great if you’re a “real” scientist. Simple, basic plots are straightforward enough, but once past that basic level, the simplest embellishments require too much digging.

In closing, of course this is every much my opinions and my experience. At my real job, I will continue to use **plotnine**. Fortunately, at the Bank, the emphasis is not on the tools you use, it’s on the output you produce. For the class, I of course will continue to put forth the effort in regard to matplotlib; as painful as it may be.

**References**

[1] <http://vita.had.co.nz/papers/layered-grammar.pdf>