**Week 5 Discussion**

In today’s competitive data science climate, it is important to have as many tools in your toolbelt as possible. Compare and contrast the benefits of using either R or Python for statistical analysis and machine learning. In which cases should you use R? In which cases should you use Python?

Datacamp.com describes “Python [as] a general-purpose, open-source programming language used in various software domains, including data science, web development, and gaming” and “R [as] an open-source programming language specifically created for statistical computing and graphics.”[[1]](#endnote-1) They further state that R was written for statistics and has tools that allow complex statistical analysis in relatively few lines of code and allows for merging dataflows to further allow complex analysis. Alternatively, Python is useful for creating machine learning applications and pipelines that will integrate with the web.

With this is mind I would recommend R for completing complex analysis on a single set of data that I may not need to change repeatedly and rerun. I would also use R to create visual representations of data. And use Python to create programs that can be called up and used repeatedly with new information and for machine learning applications.

1. https://www.datacamp.com/blog/python-vs-r-for-data-science-whats-the-difference [↑](#endnote-ref-1)