**Discussion 1: Introduction**

1. Why do Data Scientists use Python?

* simple, easy to learn
* highly readable
* good libraries/packages
* large userbase to answer questions
* while Python has slower execution time than lower-level languages, total time for a project is coding time + execution time – most data science projects are only run a few times so saving time on coding has better gains than execution time gains

1. What is Tableau used for?

* initial data exploration
* data visualization
* summaries

1. What is Github used for?

* allows multiple people to work on different offline copies of the same project and merge their results together
* public hosting of work for free (sort of like a portfolio/blog for coding projects)

1. Why use Jupyter notebooks?

* easy to view workflow of project
* allows visual text representations and code runs
* errors don’t necessarily require re-run of whole program
* easy to do quick checks and scratch work

1. What is a markdown cell?

Markdown cells exist in Jupyter notebooks so that you can incorporate text, links, and images between code cells

1. What is a code cell?

Code cells exist in Jupyter notebooks in order to run the set of instructions written in a computer programming language that tells your computer what to do

1. What are Python packages?

Python packages are a group of pre-made, reusable code that often revolve around a single topic that you can import into your programs

1. Why use functions?

When it makes sense to:

* avoid repeated code
* hide complex code
* make the script more readable
* name a chunk of code
* test out portions of code

1. What is the difference between return and print when using functions?

return statements: allow the programmer to save that value to a variable or do whatever you want with it – it’s how a function can give a value back after being called

print statements: prints out a value to the screen and returns None… after this value is printed, it’s not stored anywhere and cannot be used again