

PS2 McGuire

kmmcguire

January 2020

- Statistical programming languages: Includes both scripted and compiled languages, though scripted languages such as R, Python, and Julia are better suited to many data science specific applications.
- Data collection: Can include both collection through traditional means or methods such as web scraping.
- Data management: Includes tools for handling and accessing large data sets such as SQL databases or Hadoop.
- Visualization: Visualization can often be done directly through packages integrated with scripting languages.
- Modeling: Models can be built by combining statistical methods with either programming languages or additional software packages.