PS2$_W$hite

paulwhite

February 2022

1 Tools of the Data Scientist

Measurement:
: Data collection and everything to do with it. Most importantly, understanding what data actually measures.

Statistical Programming Languages:
: Languages built to enable people to work with data. Efficient at running statistical analyses.

Visualization tools:
: Tools used to create visual representations of data and their relationships, like ggplot. Makes it easy to present data in an interpretable, digestible format.

Web scraping:
: A method of data collection, this relies on using APIs to pull data from websites, since most data is now stored in the internet.

Big Data Management Software:
: Tools used to be able to actually work with massive amounts of data, like SQL.