

Problem Set 2 Buchman

blakebuchman03

February 2021

1 Tools of a Data Scientist

Collecting Data - Being able to collect data is obviously a necessary tool because you can't manipulate and use the data until you have it. This is where everything starts and gets your project off of the ground.

Web Scraping - A form of collecting data, being able to use this is very useful because it makes the process so much faster. If you can do web scraping then you can pull more data in a much quicker way so you can start working with it.

Cleaning Data - Cleaning data is essential if you want to be efficient. Being able to clean data makes it to where you can actually work with your data and have true conclusions from your data, avoiding repeated entries or missing variables.

SQL Ability - Very important skill to have because a lot of jobs want you to have knowledge of SQL. It is also very important to have the ability to work with because it is everywhere, no matter what you're doing as a data analyst / scientist you will probably end up working with SQL at some point.

Visualization - Very important tool in conveying your results to decision makers. Visualization makes it easier to understand for decision makers that might now know exactly what they are looking at. No one, especially decision makers, wants to look through a ton of code or data, so being able to visualize it makes it easier to understand and takes less time to convey.

Statistical Modeling - This is essential because this is the goal of your project when using data. Being able to create a model and then understand the model is something you must have. If you don't have this tool, then you will not be able to be successful and or effective as a data analyst.