

Choosing the Right Tool For the Job

As a data analyst, you will usually have to decide which program or solution is right for the particular project you are working on.

Depending on which phase of the data analysis process you're in, you will need to use different tools. For example, if you are focusing on creating complex and eye-catching visualization, then those visualization tools are the best choice. If, for instance, you are focusing on *organizing, cleaning, and analyzing data* then you will probably be choosing between *spreadsheets and databases using Queries*

Spreadsheets and databases both offer way to **store, manage, and use data**.

Some Differences

Spreadsheets

- Software Applications
- Structure data in "Row and Column" format
- Organize information in cells
- Provides access to a limited amount of data
- Manual Data Entry
- Generally only one active user at a time
- Controlled by the User

Databases

- Data stores - accessed using "Query Language" (eg.SQL)
- Structure data using rules and relationships
- organize information in complex collection
- Provide access to huge amounts of data
- Strict and consistent data entry
- Multiple Users
- Controlled by a database management system

You don't have to choose one or the other because each serves its only purpose, Generally, data analysts work with a combination of the two, as both tools are very useful in data analytics. For example, you can store data in a database, then export it to a spreadsheet for analysis. Or, if you are collecting information in a spreadsheet, and it becomes too much for that particular platform, you can import it into a database.

There are also **Programming Languages** such as **R** and **Python** that give you even greater control of your data, it's analysis, and the visualizations you create with it.