General Assembly

GETTING STARTED WITH SQL

##Spreadsheets vs. Databases

- Spreadsheets (like those used in Excel or Google Sheets) can't accommodate the volume, complexity, and variety of data most companies need.
- Most businesses store their information in databases, or a centralized collection of multiple spreadsheets.
- **Relational databases** contain data sets (referred to as tables), organized in a specific way.
- A schema is like a roadmap that represents a database's structure, as well as the relationship between its tables.

##SQL

- **Structured Query Language (SQL)** is a type of computer language that allows you to query a database and retrieve information.
- Some different SQL database options include PostgreSQL, MySQL, Oracle, SQLite

##Data Analysis Workflow:

- **Communicate**: Before accessing a database, analysts should generally have an idea of what they need in order to solve their problem.
- **Wrangle**: As an analyst, obtaining data often means accessing a SQL database and using SQL to gather the data you need.
- **Explore**: After obtaining the data, SQL is an efficient way to summarize large data sets and understand if more data is needed.
- **Analyze**: Using SQL, analysts can organize and aggregate data in a format that makes it easy to clearly and accurately analyze.

General Assembly