4.1: Intro to Programming for Data Analysts

1. **Step 1 Set up a document where you can write your answers to the following questions. You'll want to conduct some research online to help you develop your answers.**
2. **Write 2 to 3 sentences on why Python is so popular among data analysts.**
    - Multi-purpose programming language with extensive collection of libraries and huge community
    - Open Source
    - Easy to learn as it focuses on simplicity and readability. In other words, simpler syntax.
    - It can be directly integrated into business intelligence tools via API.
3. **After doing some research, name the 5 top companies in the world that use Python (either as a tool for software engineering or for analytics).**
    - Instagram
    - Instacart
    - Google
    - Netflix
    - Spotify
4. **For each of the following scenarios, explain what tool you would use and why**
    - You have a small data set that needs some quick tweaks and minor analysis. You'll need to filter some columns and make a quick chart: Excel because it is easy and fast to handle small dataset.
    - You need to retrieve some portion of data from a very large database: SQL because database contains many interconnected tables. Using SQL queries, we can quickly retrieve information we need only by knowing the ERD.
    - You have a data set with 15,000,000 rows and 350 columns that needs to be sorted and prepared for a more advanced analysis: Python because python is designed for advanced data analysis using big dataset.
5. **Download Anaconda.**
6. **Set up the environment variables on your computer and copy them into your document together with your answers to steps 2 through 4.**
7. **Launch Jupyter.**
8. **Take a screenshot of the page that opens in your browser upon launching Jupyter.**