4.1: Intro to Programming for Data Analysts

Task 1: Write 2 to 3 sentences on why Python is so popular among data analysts.

Explanations:

The followings are the reasons why Python is so popular among data analysts.

➢ Python is a multi-functional, maximally interpreted programming language with several advantages that are often used to streamline massive, and complex data sets.
➢ It has many completely free libraries that are open to the public
➢ Python's extreme versatility is another powerful attribute that makes it popular among data scientists and analysts.

Task 2: 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).

Explanations:

➢ Instagram
➢ Sportify
➢ Amazon
➢ SurveyMonkey
➢ Facebook

Link : https://www.cleveroad.com/blog/discover-5-leading-companies-that-use-python-and-learn-does-it-fit-your-project/

Task 3: . For each of the following scenarios, explain what tool you would use and why

A. 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.

Explanation: Excel is the best to use, because it is well suited for small data and filtering, though SQL could also be used.

B. You need to retrieve some portion of data from a very large database.

Explanation: Having learnt Excel and SQL in this program, SQL is the best to be use in retrieving portion of data due to its querying ability.

C. 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.

Explanation: Python is the suitable for this task because it is design for large dataset.

Task 4: Download Anaconda.

➢ Set up the environment variables on your computer and copy them into your document together with your answers to steps 2 through 4
➢ Launch Jupyter.
➢ Take a screenshot of the page that opens in your browser upon launching Jupyter.