Task 4.1 Intro to Programming for Data Analysts

Write 2 to 3 sentences on why Python is so popular among data analysts

Python is popular among data analysts because it is easy to learn and understand. It has a wide range of specialized libraries and tools that make data analysis tasks easier and more efficient. Additionally, Python's versatility allows analysts to integrate data analysis with other tasks like visualization and machine learning.

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).

Here are the some of the top companies in the world known to utilize Python either as a tool for software engineering or for analytics are:

1. Google
2. Netflix
3. Amazon
4. Microsoft
5. Facebook

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.

In order to handle small data set that needs some quick tweaks and minor analysis. I would use spreadsheet software like Microsoft Excel. It is user-friendly and provide a range of features for data manipulation and analysis. Visualisation is also simple and easy in excel.
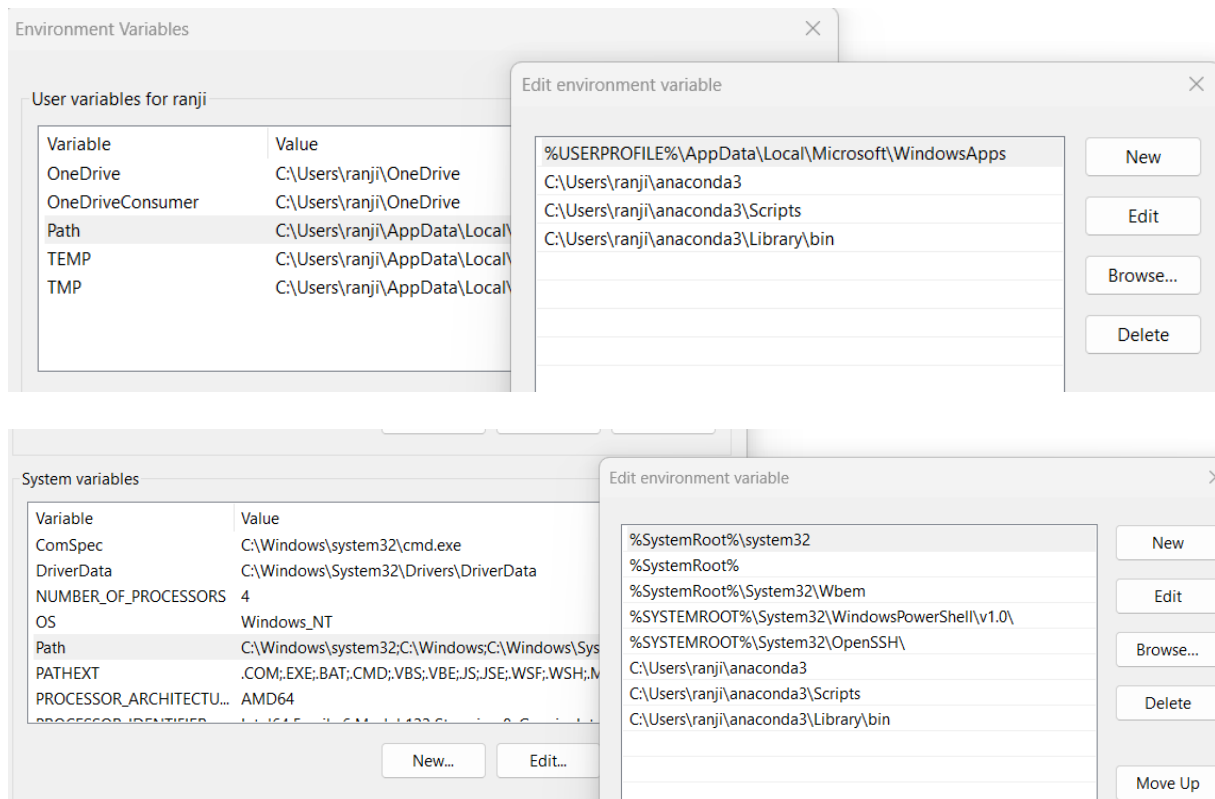
b. You need to retrieve some portion of data from a very large database.

To retrieve small portion of data from large database I would use SQL database.

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.

Python is likely the most suitable option in this scenario, given the large size of the dataset and the need for advanced analysis. However, SQL could also serve as a viable tool for either a portion or the entirety of the analysis.

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.