Exercise 4.1

Answers 4.1

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

Python is well-liked among data analysts because to its ease of learning, vast user base, and abundance of data analysis libraries.

Simple to understand
Python is easy to understand and write because of its simple, natural language-like syntax.
Python is easy to learn, even for those who have little to no programming expertise.

Big community
Python's libraries and frameworks are contributed to by a sizable and vibrant community of developers and data scientists.
Stack Overflow, mailing groups, and user-contributed code and documentation are all resources available to users for assistance.

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).

- **Netflix:** Python is utilized for platform auto-remediation and test automation, allowing for the quick and effective rollout of new features.
- **Google:** Python supports the majority of Google's web services. For instance, analytical algorithms are utilized in search engines and on YouTube for a variety of purposes, including data analysis and video processing.
- **Spotify**: Spotify uses robust Python to improve its Radio and Discover features for users. Even though computer experts acknowledge Python's superior productivity in development, about 80% of Spotify's backend services are created fully in this language:
- **Amazon:** Python is used by the business in its recommendation system, which uses AI and ML to examine consumer purchasing patterns and make pertinent product recommendations.

- **Facebook**: Facebook's primary language is Python, which accounts for 21% of the entire codebase. It is an essential part of Facebook's backend infrastructure that keeps the site operational.

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.**

  To do this, I would utilize Excel. Excel has many capabilities that make it a useful tool for handling small datasets. Data visualization using charts and data filtering are both simple processes.

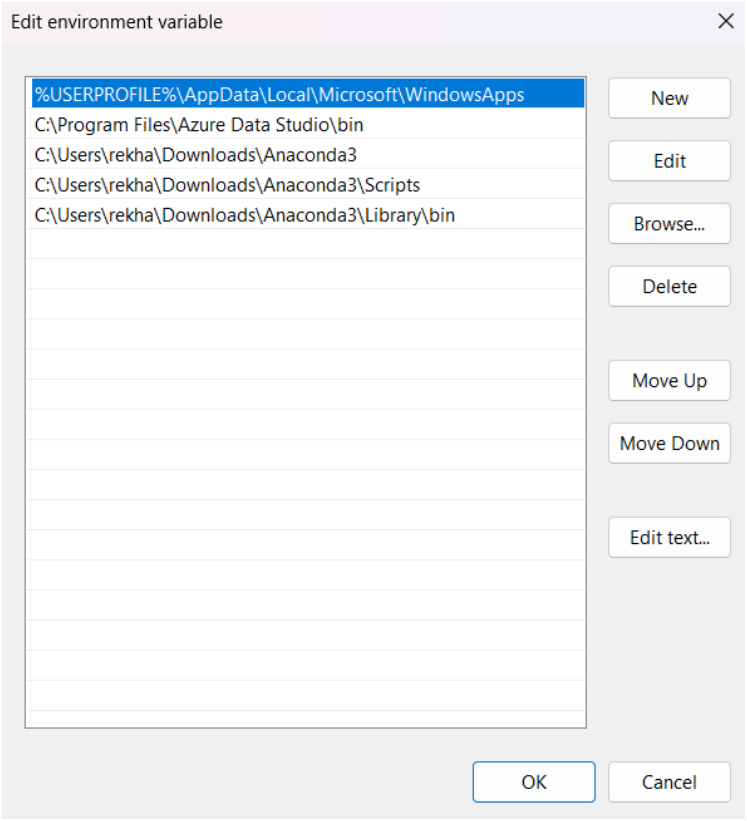
- **You need to retrieve some portion of data from a very large database.**

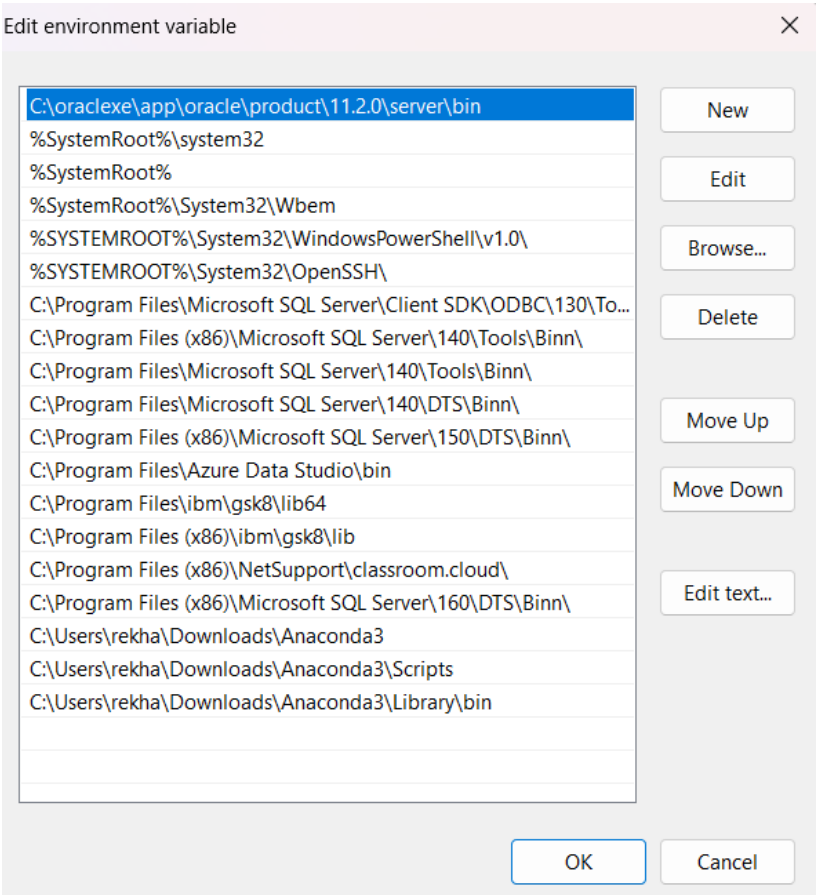
  I'll be using SQL for this. Large-scale datasets benefit greatly from SQL's efficiency. Data can be retrieved with ease using straightforward queries. In order to improve functionality, we can also use JOINS and VIEWS, as well as do data cleansing, filtering, and manipulation.

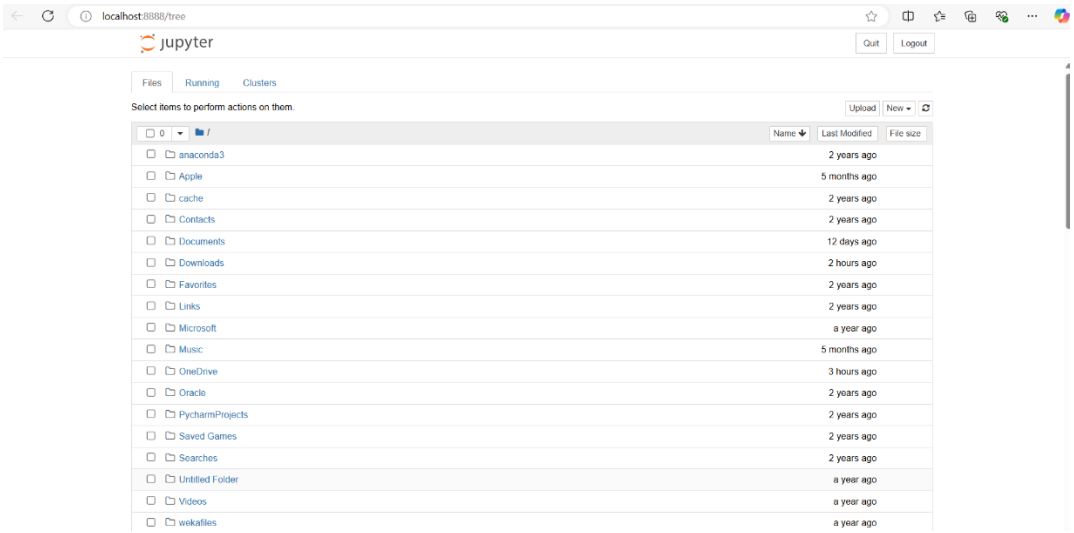
- **You have a data set with 15,000,000 rows and 350 columns that needs to be sorted and prepared for a more advanced analysis.**

  I'm going to use Python for this work. Python is a programming language with several built-in tools and packages that facilitate easy and seamless analysis of large datasets.

5. Set up the environment variables on your computer and copy them into your document together with your answers to steps 2 through 4.


Edit environment variable
%USERPROFILE%\AppData\Local\Microsoft\WindowsApps
C:\Program Files\Azure Data Studio\bin
C:\Users\rekha\Downloads\Anaconda3
C:\Users\rekha\Downloads\Anaconda3\Scripts
C:\Users\rekha\Downloads\Anaconda3\Library\bin
New
Edit
Browse...
Delete
Move Up
Move Down
Edit text...
OK
Cancel


Edit environment variable
C:\oraclexe\app\oracle\product\11.2.0\server\bin
%SystemRoot%\system32
%SystemRoot%
%SystemRoot%\System32\Wbem
%SYSTEMROOT%\System32\WindowsPowerShell\v1.0\
%SYSTEMROOT%\System32\OpenSSH\
C:\Program Files\Microsoft SQL Server\Client SDK\ODBC\130\To...
C:\Program Files (x86)\Microsoft SQL Server\140\Tools\Binn\
C:\Program Files\Microsoft SQL Server\140\Tools\Binn\
C:\Program Files\Microsoft SQL Server\140\DTS\Binn\
C:\Program Files (x86)\Microsoft SQL Server\150\DTS\Binn\
C:\Program Files\Azure Data Studio\bin
C:\Program Files\ibm\gsk8\lib64
C:\Program Files (x86)\ibm\gsk8\lib
C:\Program Files (x86)\NetSupport\classroom.cloud\
C:\Program Files (x86)\Microsoft SQL Server\160\DTS\Binn\
C:\Users\rekha\Downloads\Anaconda3
C:\Users\rekha\Downloads\Anaconda3\Scripts
C:\Users\rekha\Downloads\Anaconda3\Library\bin
New
Edit
Browse...
Delete
Move Up
Move Down
Edit text...
OK
Cancel


localhost:8888/tree
jupyter
Quit
Logout
Files
Running
Clusters
Select items to perform actions on them.
Upload
New
0
/
Name
Last Modified
File size
anaconda3 2 years ago
Apple 5 months ago
cache 2 years ago
Contacts 2 years ago
Documents 12 days ago
Downloads 2 hours ago
Favorites 2 years ago
Links 2 years ago
Microsoft a year ago
Music 5 months ago
OneDrive 3 hours ago
Oracle 2 years ago
PycharmProjects 2 years ago
Saved Games 2 years ago
Searches 2 years ago
Untitled Folder a year ago
Videos a year ago
wekafiles a year ago