

Jeremy Obach
CareerFoundry DA Immersion
Task 4.1

2. Python is popular among data analysts for a couple of major reasons: for one, it's a very flexible, general purpose language that can be used for other types of programming – because of this, it can be more easily integrated into more complex pieces of software. The other major advantage of Python is the growing community and support for Data Analyst functions – this manifests as improvements to the language itself, and in the form of libraries, or general sets of prewritten functions for various purposes, like data cleaning or visualization. Finally, the community of Python user data analysts means that there is a ton of online support in the form of blogs, videos, articles, and forums, in addition to the libraries.

3. Five top companies that utilize Python:

1. Google
2. Netflix
3. Spotify
4. Instagram
5. Shopify

I interact with each of these companies' products literally every day; I think I used all of them but Netflix just this morning – Cyber Week deals have me weak...

4a. For a small data set needing only quick tweaks and minor analysis plus some filtering and a quick chart, Excel is the tool for the job. Though it has its limitations, especially regarding the amount rows of data that it can process, all the tasks mentioned above are comfortably within its wheelhouse.

4b. For pulling data from large databases, SQL, or structured query language, is likely how you would interact with it. SQL is for structured data though; if it was a large database with unstructured data, you would probably need a different tool/language.

4c. Python to the rescue! This dataset is too large for Excel and needs to be prepared for analysis so not just pulled from a database, ruling out SQL. R, a programming language designed specifically for data analysis, might also be an appropriate choice, but given the reasons listed in Question #2, I would choose to learn Python for this task.

5. Included the terminal (?) that popped up in additions to the browser launched page.

CF


Go to current Exercise

My dashboard > Data Analytics Immersion > Exercise 4.1

Next >

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

Jeremyobach - jupyter_mac.command - python - jupyter_mac.command - 132x40

Last login: Fri Dec 1 09:51:02 on ttys000
/Users/jeremyobach/anaconda3/bin/jupyter_mac.command : exit!
(base) Jeremyobach@Jeremy-Mac: ~ % /Users/jeremyobach/anaconda3/bin/jupyter_mac.command : exit!
[10:13:18.921 NotebookApp] Writing notebook server cookie secret to /Users/jeremyobach/Library/Jupyter/runtime/notebook_cookie_secret

Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions.
https://jupyter-notebook.readthedocs.io/en/latest/migrate_to_notebook7.html
Please note that updating to Notebook 7 might break some of your extensions.
[W 10:13:20.174 NotebookApp] Loading JupyterLab as a classic notebook (v6) extension.
[I 2023-12-01 10:13:20.176 LabApp] JupyterLab extension loaded from /Users/jeremyobach/anaconda3/lib/python3.11/site-packages/jupyterlab
[I 2023-12-01 10:13:20.176 LabApp] JupyterLab application directory is /Users/jeremyobach/anaconda3/share/jupyter/lab
[10:13:24.443 NotebookApp] Serving notebooks from local directory: /Users/jeremyobach
[10:13:24.443 NotebookApp] Jupyter Notebook 6.5.4 is running at:
[10:13:24.443 NotebookApp] http://localhost:8888/?token=0849538e274ed8340c81d3d88f5fa338dc3138e722ea819
[10:13:24.443 NotebookApp] or http://127.0.0.1:8888/?token=0849538e274ed8340c81d3d88f5fa338dc3138e722ea819
[10:13:24.443 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 10:13:24.444 NotebookApp]

To access the notebook, open this file in a browser:
file:///Users/jeremyobach/Library/Jupyter/runtime/observer-23446-open.html
Or copy and paste one of these URLs:
http://localhost:8888/?token=0849538e274ed8340c81d3d88f5fa338dc3138e722ea819
or http://127.0.0.1:8888/?token=0849538e274ed8340c81d3d88f5fa338dc3138e722ea819
0.00s - Debugger warning: It seems that frozen modules are being used, which may
0.00s - make the debugger miss breakpoints. Please pass -Xfrozen_modules=off
0.00s - to python to disable frozen modules.
0.00s - Note: Debugging will proceed. Set PYDEVD_DISABLE_FILE_VALIDATION=1 to disable this validation.
[]

Rubric

Refer to the categories below to see how to meet the requirements of the approved stage

Not Yet

Anaconda Environment isn't set up, and questions posed in the task remain unanswered;

A Little More

Word document submitted containing 1) a description of why Python is popular with data analysts; 2)

Almost There

Word document submitted containing 1) a description of why

Approved

Word document submitted containing 1) a description of why Python is popular with data analysts; 2)

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

Don't show anymore

jupyter

Quit Logout

Files Running Clusters

Select items to perform actions on them.

Upload New

	Name	Last Modified	File size
<input type="checkbox"/>	/		
<input type="checkbox"/>	anaconda3	5 hours ago	
<input type="checkbox"/>	Applications	a day ago	
<input type="checkbox"/>	Creative Cloud Files	a day ago	
<input type="checkbox"/>	Desktop	4 months ago	
<input type="checkbox"/>	Documents	a day ago	
<input type="checkbox"/>	Downloads	5 hours ago	
<input type="checkbox"/>	Mirror	a year ago	
<input type="checkbox"/>	Movies	a month ago	
<input type="checkbox"/>	Music	10 months ago	
<input type="checkbox"/>	Pictures	2 months ago	
<input type="checkbox"/>	Public	a year ago	
<input type="checkbox"/>	thinkerswim	10 months ago	
<input type="checkbox"/>	edb_languagepack_4.app.zip	a month ago	93.6 MB
<input type="checkbox"/>	ILLP EP1 SEG2 (LOVE LANGUAGES+ STUTZ) - 11:23:22, 10:39 AM.m4a	a year ago	16.6 MB
<input type="checkbox"/>	ILLP EP1 SEG2 (LOVE LANGUAGES+ STUTZ) - 11:23:22, 10:48 AM.m4a	a year ago	16.6 MB
<input type="checkbox"/>	ILLP EP1 SEG3 (JONAH MOVIES, ZESTY DRAKE) - 11:23:22, 10:59 AM.m4a	a year ago	23.7 MB
<input type="checkbox"/>	ILLP_EP1_SEG1 - 11:23:22, 10:38 AM.m4a	a year ago	11.5 MB