Why Data Science?

Programming with Python

3 min

After interacting with the database, it is time to analyze the data further and eventually visualize the data. And SQL cannot get us there.

Python is a general-purpose programming language. It can do almost all of what other languages can do with comparable, or faster, speed. It is often chosen by Data Analysts and Data Scientists for prototyping, visualization, and execution of data analyses on datasets.

There's an important question here. Plenty of other programming languages, like R, can be useful in the field of data science. Why are so many people choosing Python?

One major factor is Python's versatility. [There are over 125,000 third-party Python libraries.](#) These libraries make Python more useful for specific purposes, from the traditional (e.g. web development, text processing) to the cutting edge (e.g. AI and machine learning). For example, a biologist might use the [Biopython library](#) to aid their work in genetic sequencing.

Additionally, Python has become a go-to language for data analysis. With data-focused libraries like pandas, NumPy, and Matplotlib, anyone familiar with Python's syntax and rules can use it as a powerful tool to process, manipulate, and visualize data.

Instructions

1. Checkpoint 1 Passed

1.

Catherine just downloaded Python 3 onto her office laptop.

Let's test out a simple piece of code. Run the **script.py** program.

script.py

```
libraries = ["NumPy", "SciPy", "Pandas", "Matplotlib", "Seaborn"]

completion = [100, 100, 96, 0, 0]


libraries.append("scikit-learn")

completion.append(0)


gradebook = list(zip(libraries, completion))


print("Lesson Completion Rates:")

print(gradebook)

print("\n")
```

```python
# What's next?


# gradebook.append(("BeautifulSoup", 0))

# gradebook.append(("Tensorflow", 0))
```

```
Lesson Completion Rates:
[('NumPy', 100), ('SciPy', 100), ('Pandas', 96),
('Matplotlib', 0), ('Seaborn', 0), ('scikit-learn', 0)]
```