Reading 3

Skills:

- Details

- Make associations

- Establish comparisons

Getting started: Have you used Python? What are some characteristics of this programming language?

THE DATA SCIENCE DEBATE BETWEEN R AND PYTHON



There is a very hot debate on the topic of which language is more suitable for data science: R or Python. Both R and Python have their own specific features suitable for data science and analytics applications. There may be some situations where one language is more preferred than the other, but it does not mean that the other language is useless.

What Are R and Python?

R is an open-source language which was developed during mid-1990s as a variation of the S language. It was developed by two statisticians, Canadian Robert Gentleman and New Zealand Ross Ihaka. It was

LIVEWORKSHEETS

designed in order to update the programming experience. Nowadays, it is being used extensively for research, enterprise and academics. Due to its usage in many fields, it is one of the most popular statistical programming languages. It is quite simple to use, but it can be a little difficult for those completely new to programming. However, they can learn more from the different resources available on the internet.

Python was created during the early 1990s by Dutch programmer Guido Van Rossum. It focuses on ease of coding and more adaptability. Python is being widely used by those programmers who want to have greater control over the codes they make for faster and more efficient data analysis. The programming language is very easy to use and learn. It is also very flexible and can be used to generate exactly what the user wants to create.

How Are They Different from Other Languages?

The work of data analysis is very important, and the process must be flexible. For this reason, the process must be interactive so that it remains efficient. However, the language must also be very flexible, interactive and easy to use. R is a very flexible language. While other languages are used for some exact purpose and cannot work for anything else, R can actually work for a wide range of purposes, especially in the scientific fields. Another thing that differentiates R from other statistical programming languages is its interactivity. R has a very powerful mechanism that can be used for quickly creating data structures. R is also a very powerful graphics medium, unlike textual programming languages, which is an important feature since graphics are very useful, especially in the field of statistics and data analysis. Thus, R can be used to produce many different types of graphs easily.

Python is also an excellent choice for data analysis. It is very adjustable when compared to languages like Perl or Ruby, as it can be customized by the use of modules. It has many features too. It is also a graphical language, which allows it to have visual libraries, and that aspect helps to visualize graphs and statistical data easily. Another thing that differentiates it from other languages is its easy-to-use syntax.

Use Cases for R and Python

There are many use cases of both R and Python for data analysis. For example, ForecastWatch.com collects data from different weather forecast sites and rates the sites according to their accuracy. This allows better weather forecasting and allows weather forecasters to compare their accuracy with

LIVEWORKSHEETS

others. Python was used for every component of this service due to its flexibility, which comes from its ability to use many standard libraries. Python has also been used to power social networks for EZTrip.com and Gusto.com. They required a system to help their customers report on their journeys while improving their online booking system. While their existing booking system already worked quite well, it couldn't handle multiple requests efficiently. However, since Python was used, it became much faster due to better data analysis and management facilities. This helped them create a better user interface based on the user's queries.

R is also being used in many places like social networking sites and crowdfunding sites. R's visualization ability is making it the favorite of many data analytics organizations too. R is currently being used in the ANZ Bank for analyzing the risks of crediting, and Facebook is also using R to analyze large numbers of status updates.

Both of these languages have an ever-active community which is always happy to help, and both of them are being constantly updated in order to adjust new features and technologies.

**Adapted from https://www.techopedia.com/2/32005/trends/big-data/the-debate-between-r-and-python*

Read the following descriptions. Write R if they apply to R, P if they apply to Python, or B if they apply to both.

1. It is not a complex language
2. It provides interactive functions.
3. It is perfect for graphic inputs.
4. It was created as the alternative to another language.
5. It is especially used for statistics and data analysis.
6. It was developed first.
7. It allows programmers to have more command of the code.
8. It is very adaptable.
9. Its linguistic structure is easy to use.
10. It was created by someone from the Netherlands.

LIVEWORKSHEETS

What do you think?

From your own experience, what are some important aspects that must be taken into account when programming?

LIVEWORKSHEETS