IBM Course Seven
Data Analysis Using Python

Week One

Objectives

01
Describe Python data acquisition and analysis techniques

02
Analyze Python data using a dataset

03
Identify three Python libraries and describe their uses.

Read data using Python's Pandas package.

Read and Watch!

◈ Read the article and make notes/add to your existing notes. Compare your notes with your peers.

◈ Watch the short video and answer the following questions:

  ⬨ Summarize data science in two sentences. Use layman's terms.

  ⬨ Why is exploratory data analysis the most important step?

  ⬨ Explain the connection between visualization and communication.

  ⬨ Brainstorm another authentic scenario where the use of data science would result in enhanced business practice.

How much do you remember?

Test your knowledge of definitions!

Article Summary

- Python

- Lists

- Functions

- Methods

What is a Python Library?

◈ A Python Library is a collection of functions and methods that can be called upon without needing to write code.

◈ Each library houses built-in modules that provide different functionalities; all functionalities can be directly used.

◈ Python Data Analysis Libraries have Three Groups:

  ▫ Scientific Computing Libraries

  ▫ Data Visualization

  ▫ Algorithmic Libraries

Scientific Computing Libraries

- Pandas

- NumPy

- SciPy

Data Visualization

- Matplotlib
- Seaborn

Algorithmic Libraries

- Scikit-learn
- Statsmodels