Module 1 - Python Core Strategic JnJ

Johnson&Johnson

databricks

Python Academy Courses

- **Scalability**: Prepares learners to understand Python but in PySpark's logic and distributed systems

- **Standardization**: Ensures consistent skill baselines across global vendors

- **Maintainability**: Designed for review, integration, and LMS compatibility

- **Learning ROI**: Aligns with experiential learning, includes hands-on exercises

Added Content and Matters

- **Exception Handling** – Builds robust, fault-tolerant PySpark pipelines

- **Context Managers and File I/O** – Enables working with distributed files and managing resources

- **Iterators and Generators** – Bridges Python logic to PySpark's lazy evaluation model

- **String Processing** – Essential for text cleaning in real-world data workflows

- **APIs and JSON** – Prepares learners for common data ingestion and integration tasks

- **Advanced Data Structures (Expanded)** – Supports handling of complex DataFrame schemas in PySpark

Sessions 1.1. Basics

Objectives

- Familiarize with Databricks notebook environment
- Understand Python variables and data types
- Practice basic operations and expressions
- Learn notebook-specific features
- Access to Databricks workspace
- Basic understanding of programming concepts

?

Python Academy Courses

Module 1 - Core Python/

```
0-Welcome-Introduction.ipynb
session1.1_Basics.ipynb
session1.2_Modules_and_Packages.ipynb
session1.3_Data_Structures.ipynb
session1.4_Advanced_Data_Structures.ipynb
session1.5_Conditions_and_Loops.ipynb
session1.6_Functions.ipynb
session1.7_Dates_and_Times.ipynb
session1.8_Regular_Expressions.ipynb
session1.9_Classes.ipynb
session1.10_Decorators.ipynb
session1.11_Virtual_Environments.ipynb
session1.12_Exception_Handling.ipynb
session1.13_Context_Managers.ipynb
session1.14_Iterators_and_Generators.ipynb
session1.15_String_Processing.ipynb
session1.16_APIs_and_JSON.ipynb
```

🔗 1-Welcome-Introduction

Course Structure Notes

- **Modules**
  Represent the major **phases of upskilling**:
  - Python Core
  - PySpark Fundamentals
  - Databricks for Production Pipelines

- **Sessions (within each Module)**
  - Mirror content delivery either in **Databricks notebooks** or within the **Moodle LMS**
  - Each session focuses on a focused topic or competency

- **Activities (per Session)**
  Include a variety of learning formats to maximize engagement and retention:
  - **Quizzes** – Concept checks and progress tracking
  - **Notebooks** – Hands-on coding exercises (Databricks or downloadable)
  - **Assessments** – Applied challenges or final tasks
  - **Generative AI Assistant** – Embedded support for Q&A, reflection, and code guidance
  - **Videos & Audios** – Concept walkthroughs, expert commentary
  - **Certifications** – Final recognition upon completion of each module or track

> Sessions

> New section