Databricks 200 Workshop

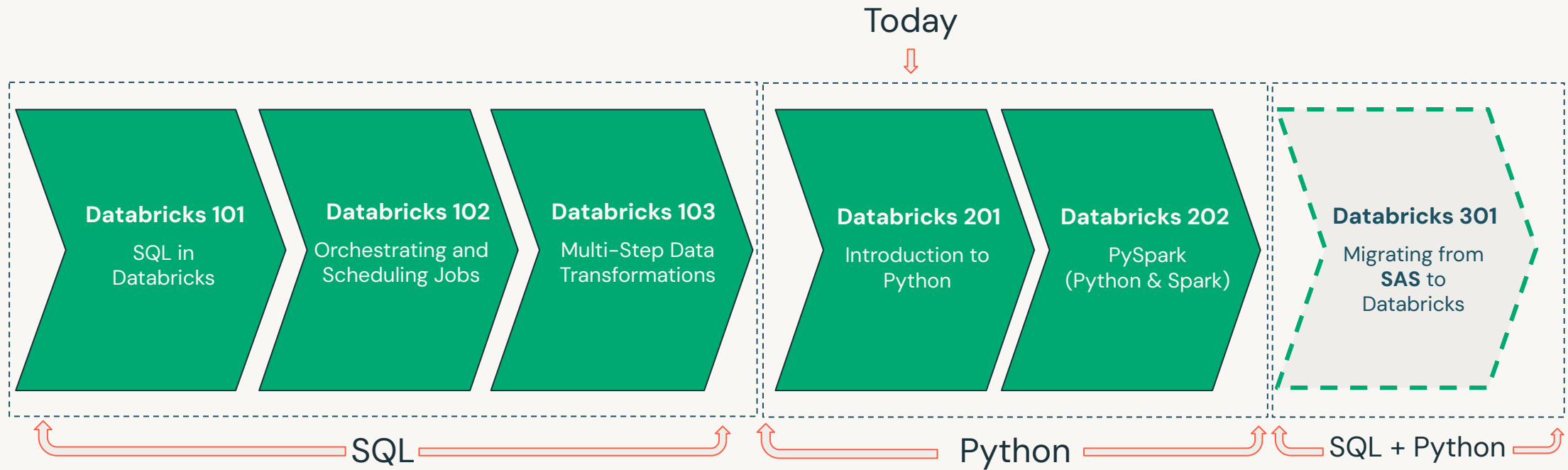## Intro to Python

Dan Davis – Databricks

Learning Objectives

Things you'll be able to do after this course

- Use Python basics
- Use Python built-in data types and functions
- Manipulate various data structures
- Utilize Databricks Workspace as a programming environment

SAS Enablement Pathway

Today

Databricks 101
SQL in Databricks

Databricks 102
Orchestrating and Scheduling Jobs

Databricks 103
Multi-Step Data Transformations

Databricks 201
Introduction to Python

Databricks 202
PySpark (Python & Spark)

Databricks 301
Migrating from **SAS** to Databricks

SQL

Python

SQL + Python

Databricks Resources

Start here

Start here today: [Analyst Enablement Workspace](#)

Start here for resources: [D&A Reference Hub](#)

Other Resources

- Databricks Free [Self-Paced Learning](#)
- Databricks Paid [Instructor-Led Learning](#)
- KP Insights: [Coursera](#)
- D&A Reference Hub: [Python Training Guide](#) (List of online Python courses)
- **D&A Reference Hub: [Copilot Chat Instructions](#) (AI Assisted Code Generation)**
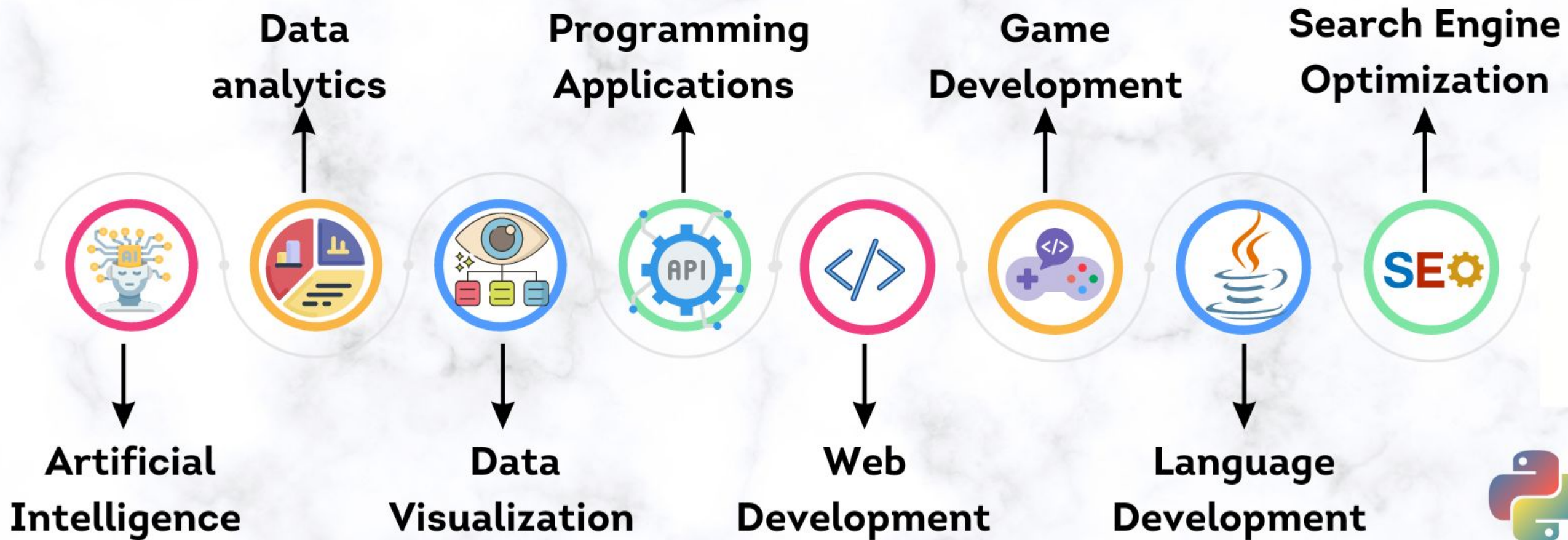
What is Python?

What is Python?

- A **multi-purpose language** with a **simple**, and **beginner-friendly** syntax
- Object-Oriented
- **Fastest** growing language in the world
- Contains other widely used **libraries:**
  - **Library:** Pre-written code and functions that extend basic functions of Python
    - NumPy: Mathematics library
    - Pandas: Data analysis, manipulation, and processing.
    - Matplotlib: Visualizations library
    - Scikit-learn: ML library
    - TensorFlow: Deep learning library

Why Python?

Why Python?

- **Broad Adoption & Community Support**
  - Huge community, low barrier to entry
  - Fastest growing language in the world
  - Simple to read, understand and use
- **Flexibility & Integration**
  - Open Source & Free
  - Cross-Platform, portability
    - Databricks, Spark and other platforms
  - APIs & Automation
- **Rich Ecosystem of Libraries**
  - So many libraries!
- **Career Upside**
  - Python skills are in high demand!

What is Python used for?

98th PERCENTILE

Data analytics

Programming Applications

Game Development

Search Engine Optimization

Artificial Intelligence

Data Visualization

Web Development

Language Development

Python in Databricks

Recap

What did we cover today...

- Data Types (e.g. strings, integers)
- Variables
  - `a = 1`
- Print Statements
  - `print("hello world")`
- F-string formatting
  - `print("hello world")`
- Conditions: if statements
  - `If else`
- Functions
  - `Def my_function()`
- Dataframes

Python Resources at KP

- Databricks Free [Self-Paced Learning](#)
- Databricks Paid [Instructor-Led Learning](#)
- KP Insights: [Coursera](#)
- D&A Reference Hub: [Python Training Guide](#) (List of online Python courses)
- **D&A Reference Hub: [Copilot Chat Instructions](#) (AI Assisted Code Generation)**