Databricks Certified Data Engineer Associate

Databricks Notebooks Introduction

Databricks Notebooks Introduction

- Notebooks are a common tool in data science and machine learning for developing code and presenting results
- In Databricks, notebooks are the primary tool for creating data science and machine learning workflows and collaborating with colleagues
- Databricks notebooks provide real-time co-authoring in multiple languages, automatic versioning, and built-in data visualizations

With Databricks notebooks, you can:

- Develop code using Python, SQL, Scala, & R
- Customize environment with libraries of your choice
- Create regularly scheduled jobs to automatically run tasks, including multi-notebook workflows
- Export results and notebooks in .html or .ipynb format
- Use a Git-based repository to store notebooks with associated files & dependencies
- Build and share dashboards
- Open or run a Delta Live Tables pipeline
- Use advanced editing capabilities
- Notebooks are also useful for exploratory data analysis (EDA)

Magic Commands

Magic Commands

- Magic commands in Databricks let you execute the code snippets other than the default language of the notebook
- Magic command start with **%<Choice of Your code snippet language>,** Databricks supports four languages Python, SQL, Scala, & R
- **Magic commands in Databricks:** %python, %sql, %scala, %sh, %fs (or dbutils.fs), %md

THANK YOU