Detached | File ▾ | Edit ▾ | View: Standard ▾ | Permissions | Run All | Clear ▾ | Publish | Comments | Experiment | Revision history

Cmd 1

databricks
Academy

Cmd 2

Hands-on with Databricks

Objective: *Familiarize yourself with the Databricks platform, the use of notebooks, and basic SQL operations in Databricks.*

In this lab, you will complete a series of exercises to familiarize yourself with the content covered in Lesson 0.1.

Cmd 3

Exercise 1

In order to execute code with Databricks, you need to have your notebook attached to an active cluster.

Ensure that:
1. You have created a cluster following the walkthrough of the video in this lesson.
2. Your cluster's Databricks Runtime Version is 7.2 ML.
3. Your cluster is active and running.
4. This notebook is attached to your cluster.

Cmd 4

Exercise 2

The fundamental piece of a Databricks notebook is the command cell. We use command cells to write and run our code.

Complete the following:
1. Insert a command cell beneath this one.
2. Write `1 + 1` in the command cell.
3. Run the command cell.
4. Verify that the output of the executed code is `2`.

Cmd 5

```
1   # ANSWER
2   1 + 1
```

Cmd 6

Exercise 3

Command cells can also be used to add comments using a lightweight markup language named *markdown*. (That's how these command cells are written).

Complete the following:
1. Double-click on this command cell.
2. Notice the *magic command* at the top of the command cell that enables the use of markdown.
3. Insert a command cell beneath this one and add the magic command to the first line.
4. Write `THE MAGIC COMMAND FOR MARKDOWN IS _____` with the magic command filling the blank.

Cmd 7

ANSWER

THE MAGIC COMMAND FOR MARKDOWN IS %md.

Cmd 8

Exercise 4

Throughout this course, we will be using a setup file in each of our notebooks that connects Databricks to our data.

Complete the following:
1. Run the below command cell to execute the setup file.
2. Insert a SQL command cell beneath the command cell containing the setup file.
3. Query all of the data in the table **dsfda.ht_daily_metrics** using the query `SELECT * FROM dsfda.ht_daily_metrics`.
4. Examine the displayed table to learn about its columns and rows.

Cmd 9

```
1   %run "../../Includes/Classroom-Setup"
```

Cmd 10

```
1   %sql
2   -- ANSWER
3   SELECT * FROM dsfda.ht_daily_metrics
```

Cmd 11

Exercise 5

Throughout this course, we will need to manipulate data and save it as new tables using Delta, just as we did in the video during the lesson.

Complete the following:

1. Insert a new SQL command cell beneath this one.
2. Write a SQL query to return rows from the **dsfda.ht_users** table where the individual's lifestyle is `"Sedentary"`.
3. Use the SQL query to create a new Delta table named **dsfda.ht_users_sedentary** and store the data in the following location: `"/dsfda/ht-users-sedentary"`.

Cmd 12

```sql
%sql
-- ANSWER
CREATE TABLE dsfda.ht_users_sedentary
USING DELTA LOCATION "/dsfda/ht-users-sedentary"
AS (
  SELECT *
  FROM dsfda.ht_users
  WHERE lifestyle = 'Sedentary'
)
```

Cmd 13

Great job! You've completed the first lesson of the Data Science Fundamentals with Databricks course.

Please proceed to the next lesson to begin Module 2: An Introduction to Data Science.

Cmd 14

Shift+Enter to run