Lab 19 Databricks Spark

1. Provision Azure Databricks Workspace

A screenshot of a computer

AI-generated content may be incorrect.

1. Create a cluster

A computer screen with a message

AI-generated content may be incorrect.

1. Import notebook

A screenshot of a computer

AI-generated content may be incorrect.

1. Ingest data

A screenshot of a computer

AI-generated content may be incorrect.

1. Query data in files

A screenshot of a computer

AI-generated content may be incorrect.

1. Include headers

A computer screen with a white screen

AI-generated content may be incorrect.

1. Verify datatypes

A computer screen with a computer screen

AI-generated content may be incorrect.

1. Filter a dataframe

A computer screen with a white screen

AI-generated content may be incorrect.

1. Aggregate and group data in a dataframe

A computer screen with a white box

AI-generated content may be incorrect.

1. Query data using Spark SQL

A computer screen with a computer screen

AI-generated content may be incorrect.

1. Visualize data using matplotlib

A computer screen with a bar graph

AI-generated content may be incorrect.

1. Visualize data using Seaborn

A computer screen shot of a graph

AI-generated content may be incorrect.

1. Delete Azure Databricks resources

A computer screen with a message

AI-generated content may be incorrect.

Summary:

Azure Databricks is a unified, open analytics platform for building, deploying, sharing, and maintaining enterprise-grade data, analytics, and AI solutions at scale. Azure Databricks is built on Apache Spark offering a highly scalable solution. Spark has multi language support as it supports Java, Scala, Python and SQL. Spark is highly scalable and can help with data cleansing, manipulation, statistical analysis, machine learning , data analytics and visualization. We begin the lab by provisioning Databricks resources. We create a Spark cluster and import a notebook. We connect the notebook to our cluster and run all the cells. The lab ends with deleting the Azure Databricks resources.