# Bài tập APACHESPARK

1. Enable dataproc metastore

A screenshot of a web page

Description automatically generated

1. Create cluster

A screenshot of a computer

Description automatically generated

1. Import file Pyspark

A screenshot of a computer

Description automatically generated

1. Init spark and check version

A screenshot of a computer program

Description automatically generatedA screenshot of a computer

Description automatically generated

1. Read file

A computer screen shot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Interact with columns

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Pyspark & SQL

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. User define function

A screenshot of a computer

Description automatically generated

1. Spark SQL

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Write result to parquet file

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated