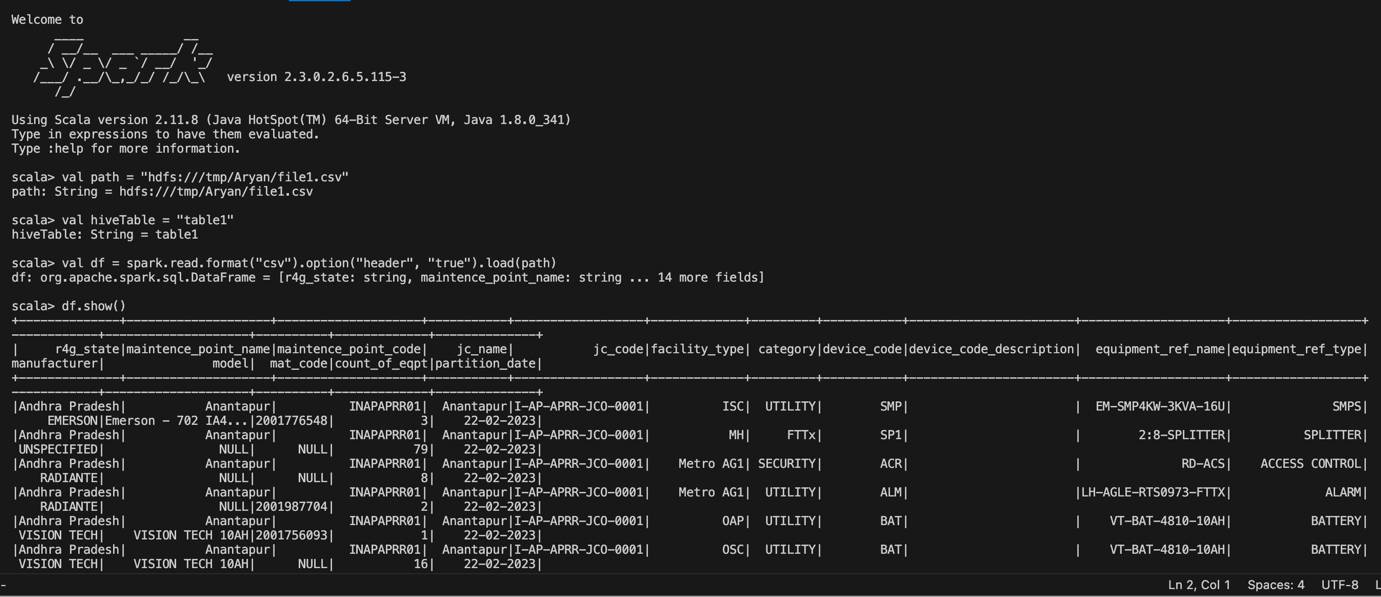
Spark Assignment 2

**Difference b/w SparkContext and SparkSession:**

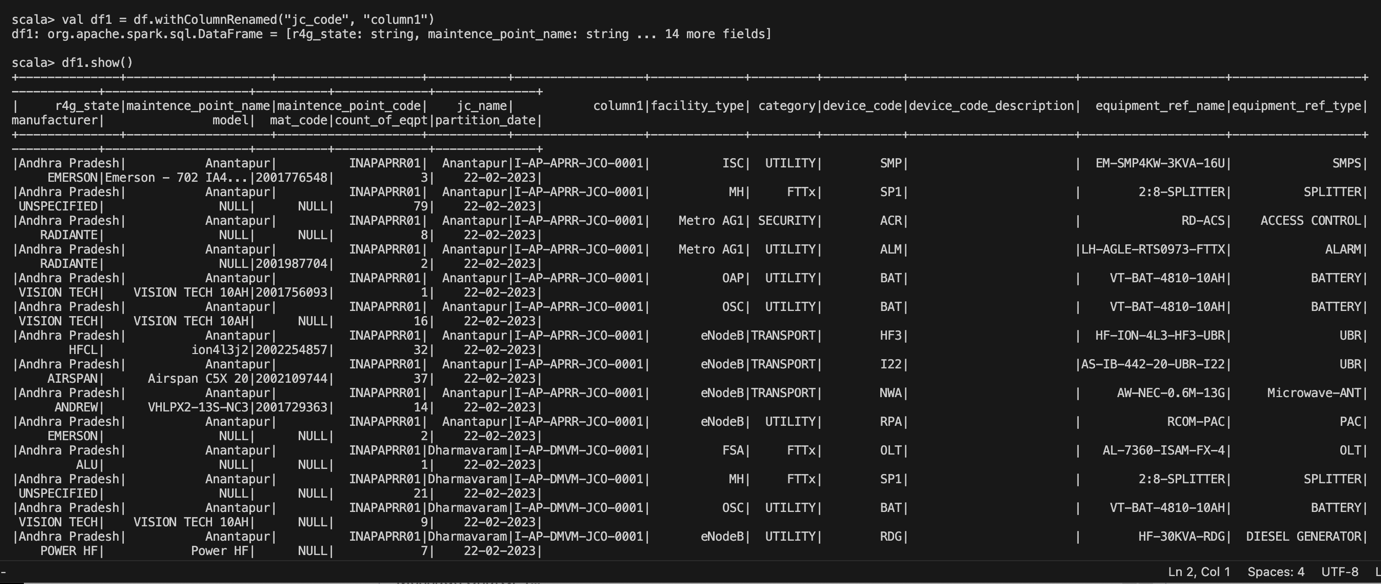
SparkContext is more focused on the low-level aspects of Spark and RDDs, SparkSession provides a higher-level API for working with structured data using DataFrames and Datasets.

**RDDs vs Dataframes vs Datasets vs Spark SQL:**

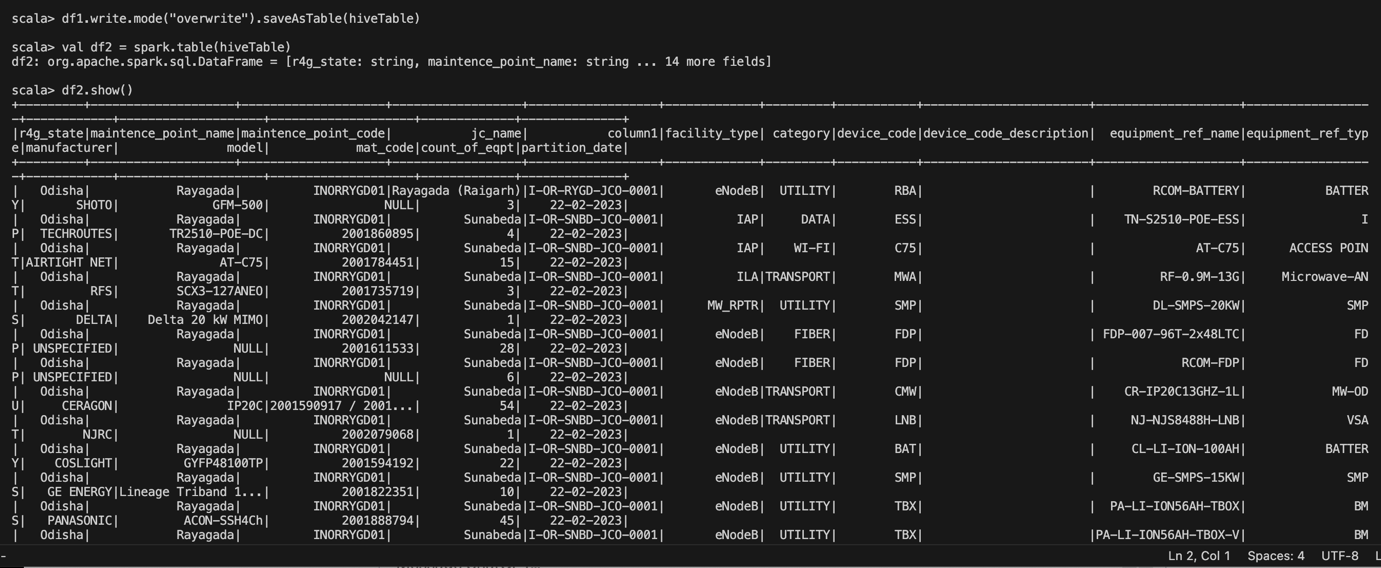
RDDs are the basic abstraction in Spark, DataFrames provide a higher-level and optimized interface for structured data, Datasets combine static typing with DataFrames, and Spark SQL allows you to use SQL syntax to query.

1.

2.



3.



4.

Aryan.Sahoo