**Assignment 19.3 Spark**

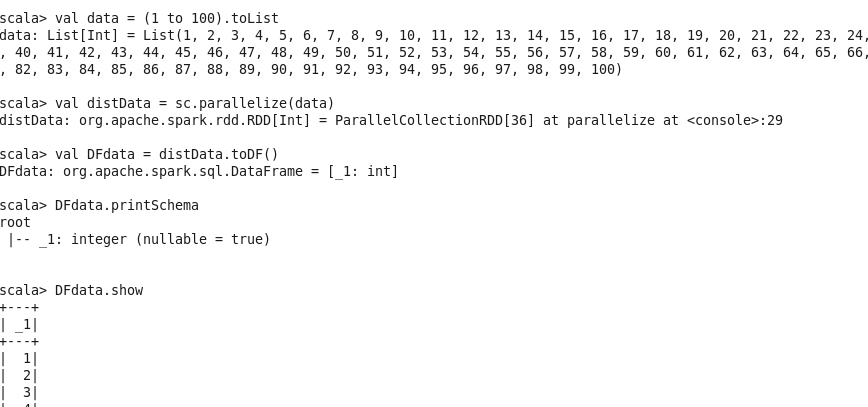
2. Problem Statement

Create a dataframe with 1 to 100 and save as parquet file.

val data = (1 to 100).toList

val distData = sc.parallelize(data)

val DFdata = distData.toDF()



**writing data to parquet file**

DFdata.write.parquet("datarange.parquet")



**Reading data from parquet file**

val parquetDF = sqlContext.read.parquet("datarange.parquet")

parquetDF.printSchema

parquetDF.show

