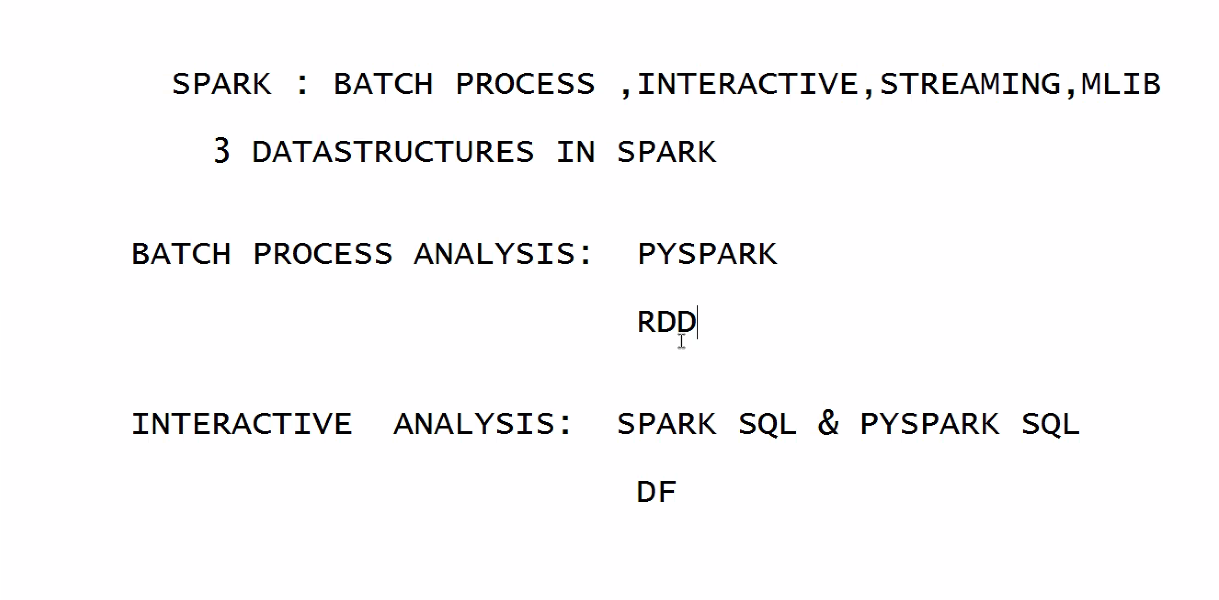
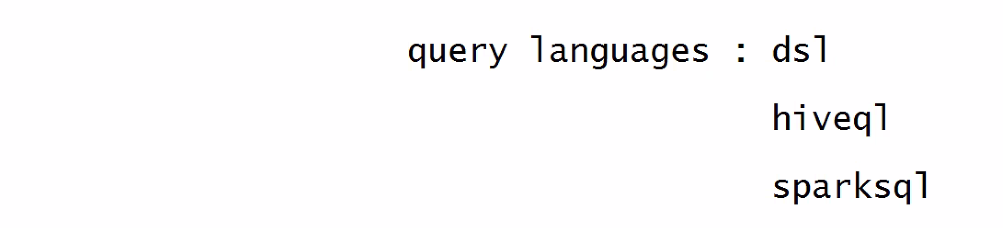
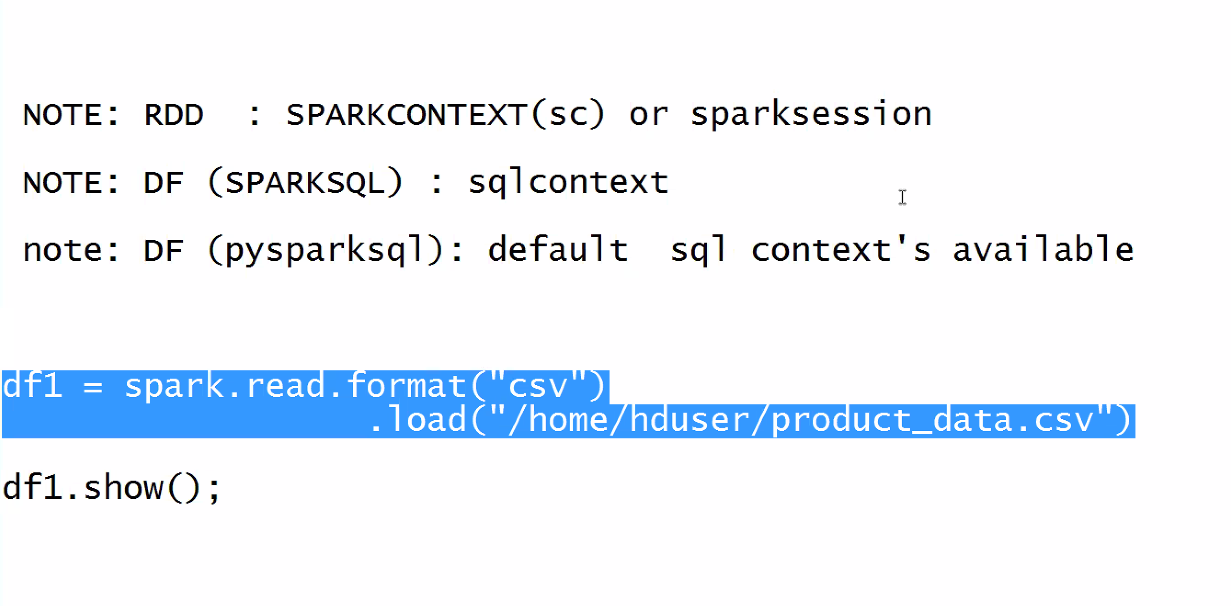
Type of data structures

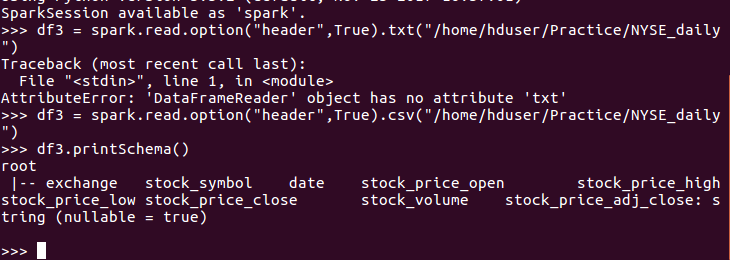








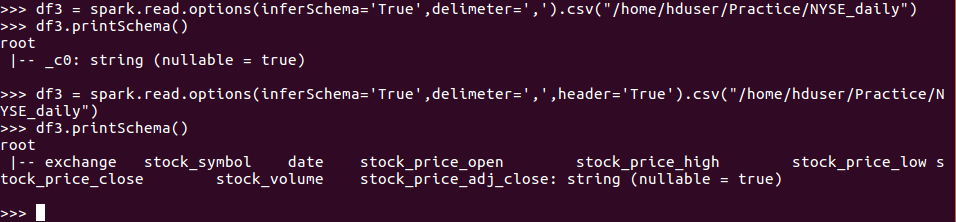
Load file with headers



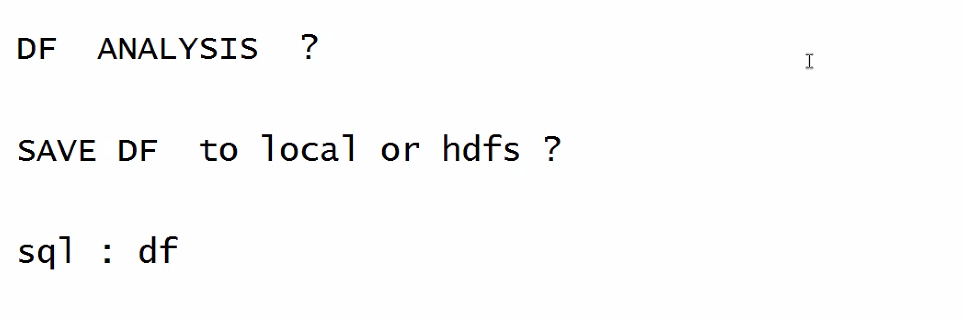
df3 = spark.read.option("header",True).csv("/home/hduser/Practice/NYSE\_daily")

df3 = spark.read.options(inferSchema='True',delimeter=',').csv("/home/hduser/Practice/NYSE\_daily")

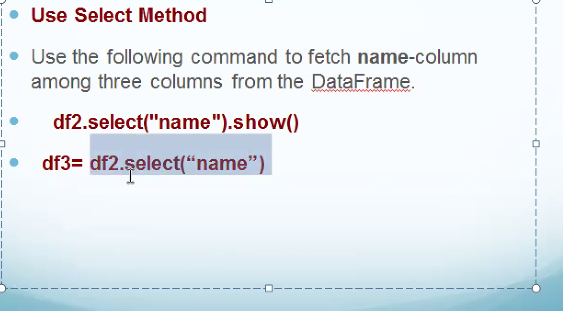
df3 = spark.read.options(inferSchema='True',delimeter=',',header='True').csv("/home/hduser/Practice/NYSE\_daily")

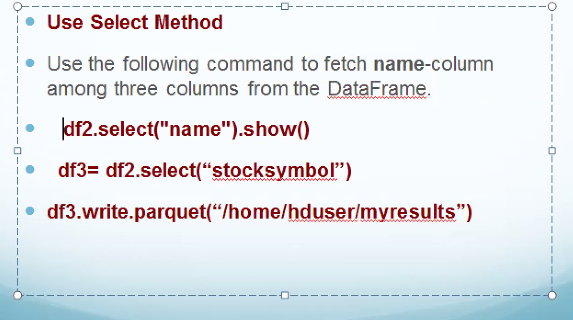


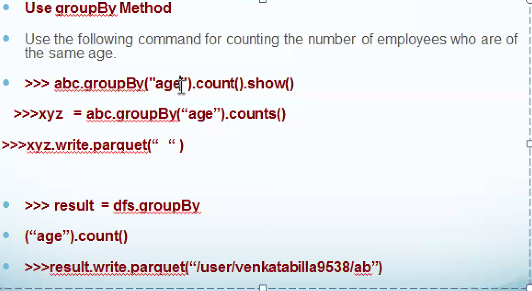
Upcoming

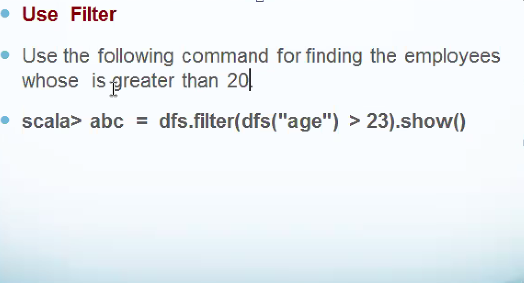


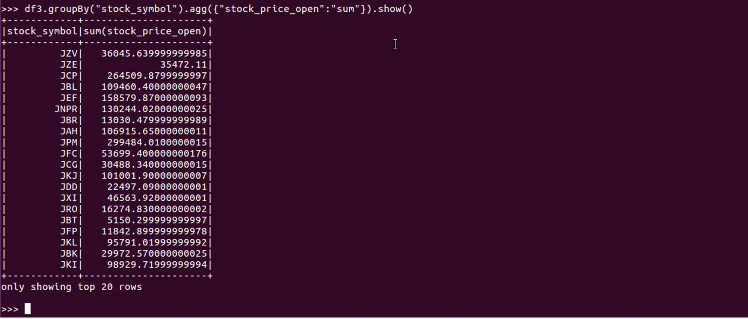
Dsl instead sql

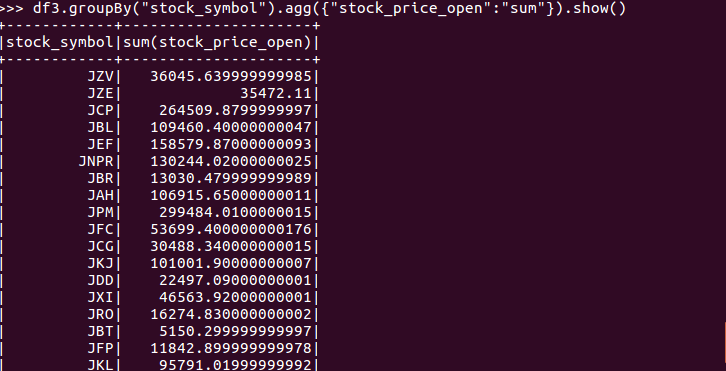




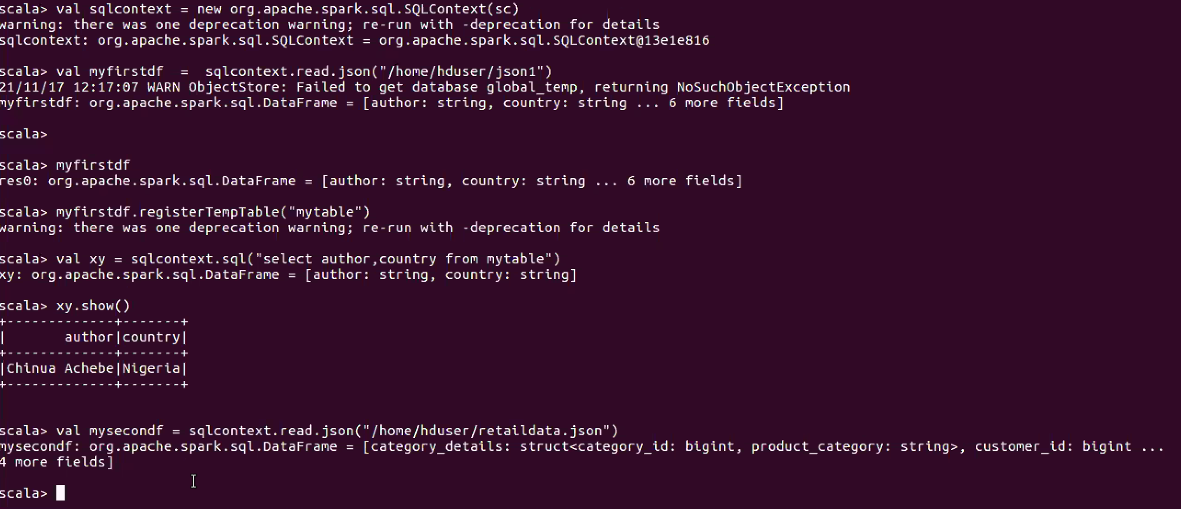








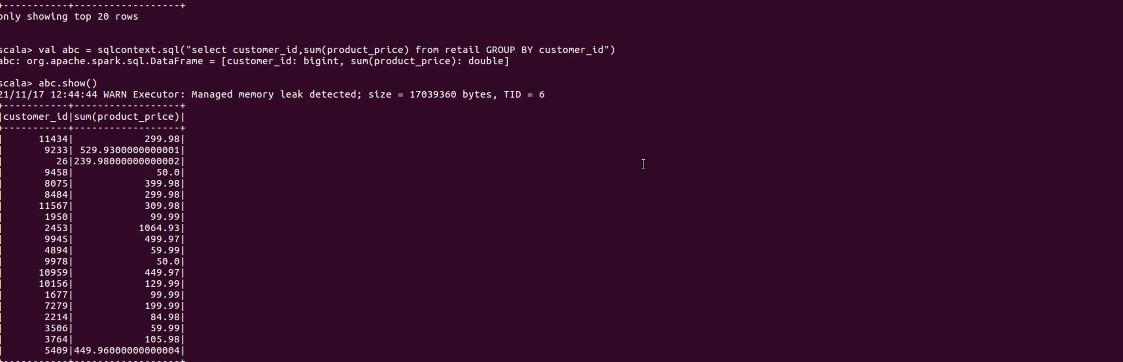
Sql instead of dsl



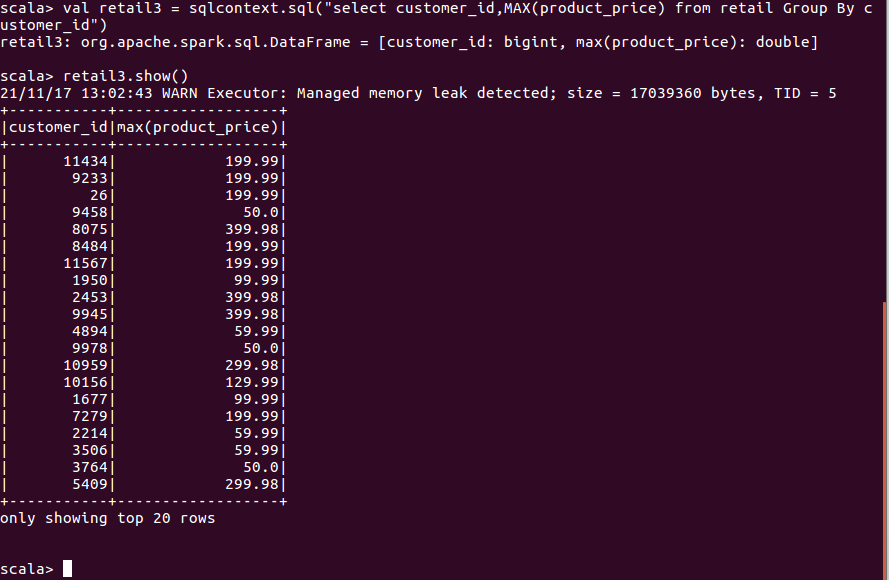
Before sql ops store as temp table



Total product price by customer

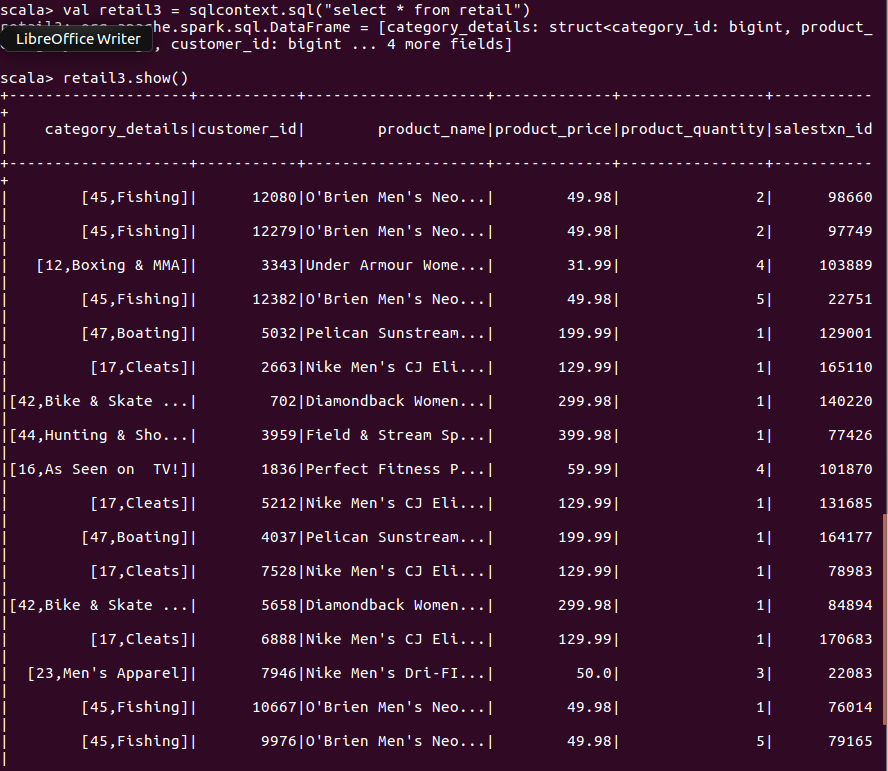


Max price

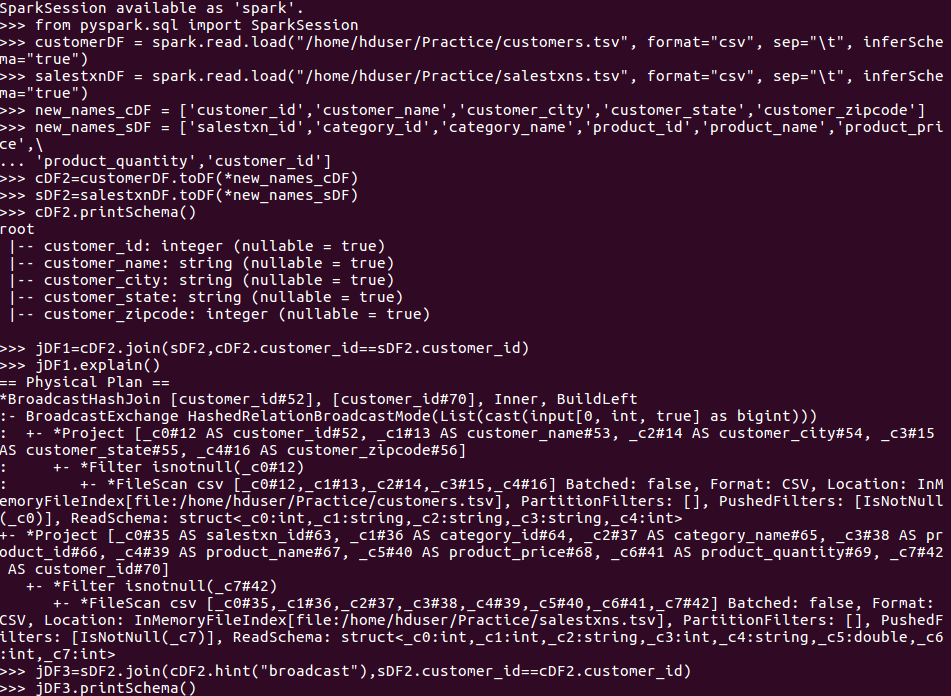


Show temp table

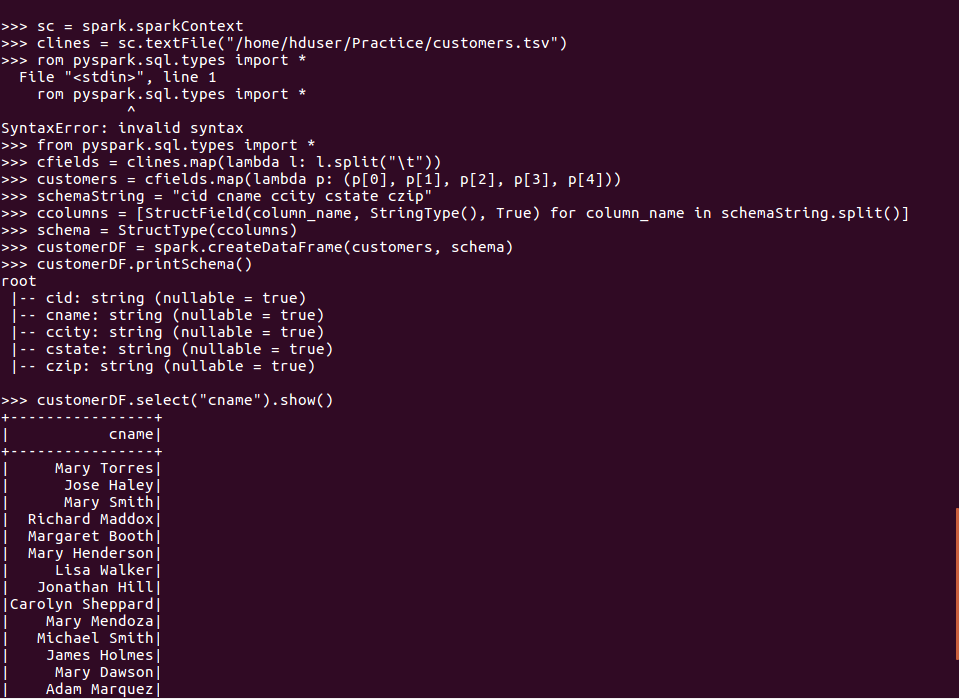
val retail3 = sqlcontext.sql("select \* from retail")



Pyspark join exercise



Load text file and create a dataframe from that example



Further, follow ipynb examples downloaded from jigsaw

