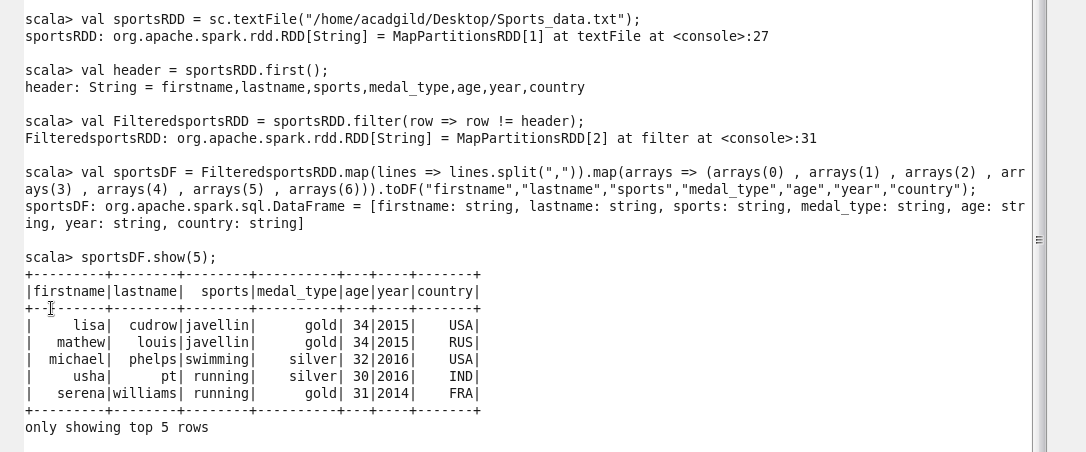
1. Change firstname , lastname columns into

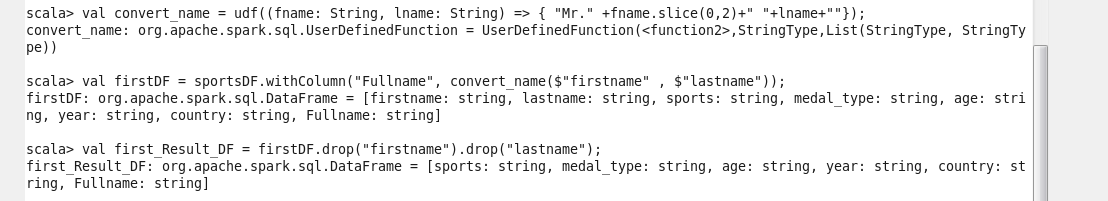
Mr.first\_two\_letters\_of\_firstname<space>lastname

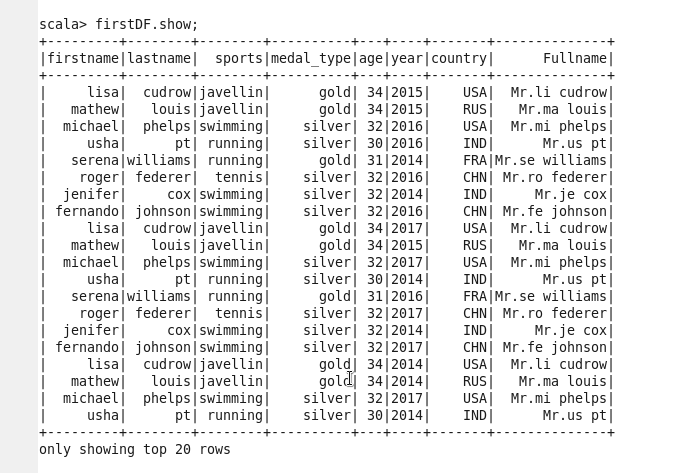
For example: Michael , phelps becomes Mr.mi phelps

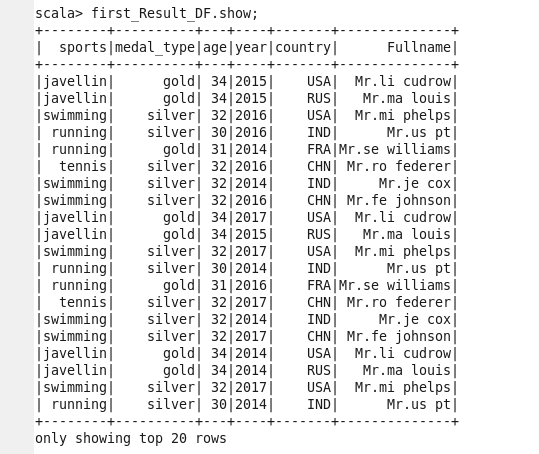
**Loading RDD and creating the dataframe for RDD:**



Creating the UDF which convert the first name with Mr.prefix and also drop the First and last name columns:







1. Add a new column called ranking using udfs on dataframe, where:

Gold medalist, with age >=32 are ranked as pro

Gold medalist, with age <=31 are ranked amateur

Silver medalist, with age >=32 are ranked as expert

Silver medalist, with age <=31 are ranked rookie

