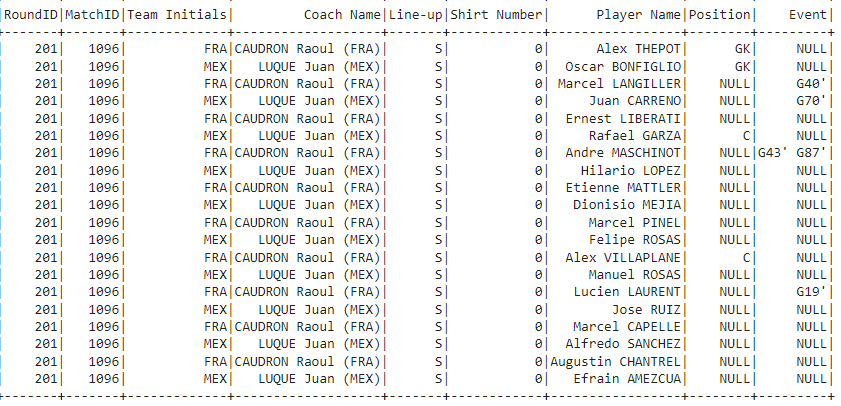
For World cup players dataset:

1. Reading Data from file

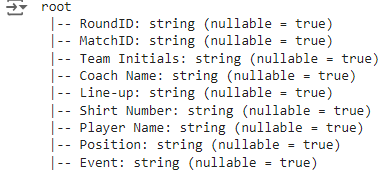
df=spark.read.csv("WorldCupPlayers.csv",header=True)

df.show()



2. Schema of Dataframe

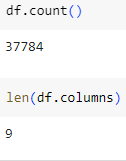
df.printSchema()



3.Count the number of Rows and Columns of Dataframe.

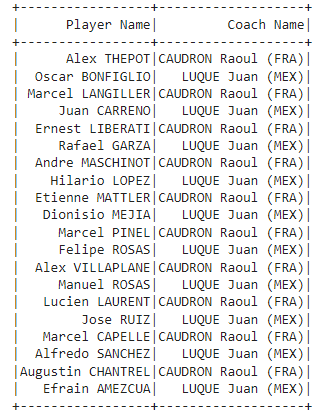
df.count()

len(df.columns)



4.Show the player name and coach name

df.select("Player Name","Coach Name").show()



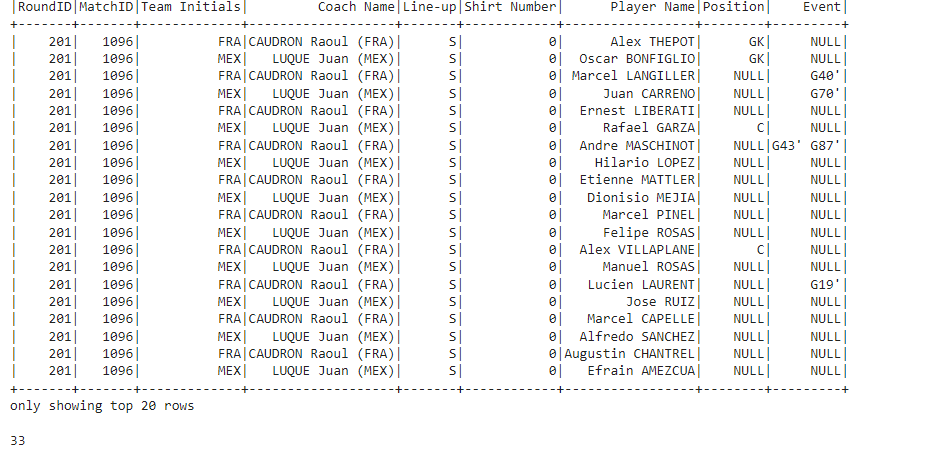
5.Show the records where Match ID must be equal to 1096 and calculate

how many records are there in the filtered output.

df1=df.filter(df.MatchID=="1096")

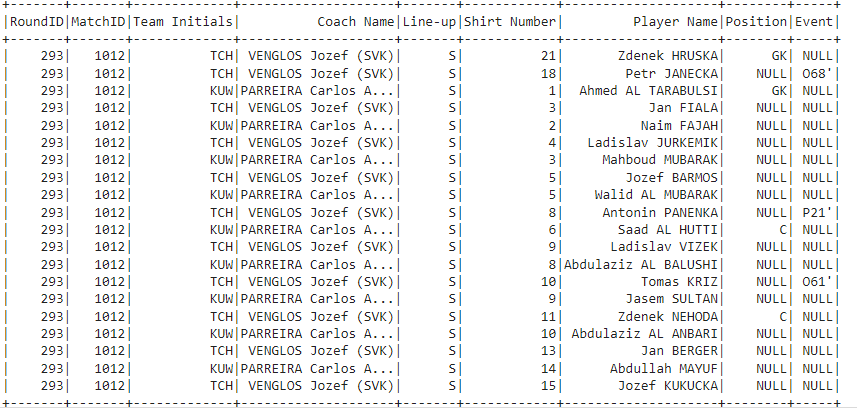
df1.show()

df1.count()

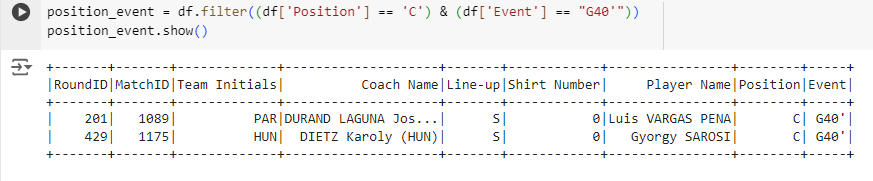


6. order the data by match id

df.orderBy("MatchID").show()

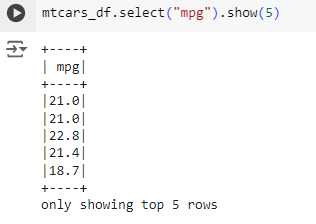


7. Show the records of position C and event is G40'

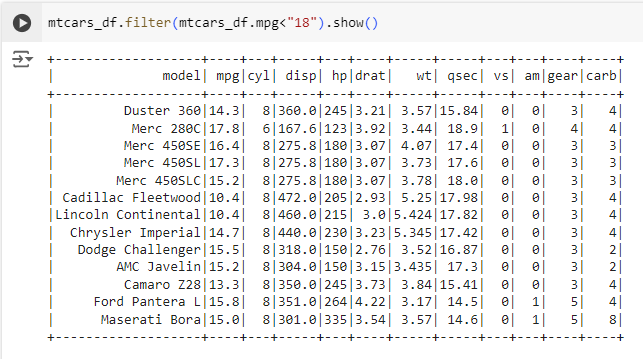


for mt cars file:

1.Find the top five records of mpg attribute

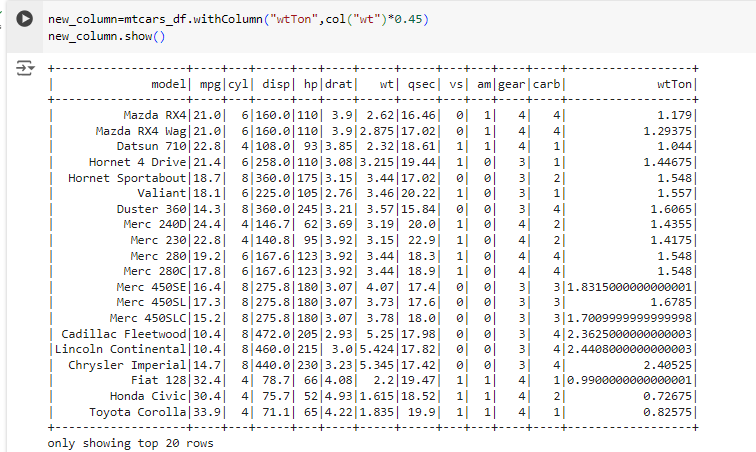


2.Find out the records where the value of mpg is less than 18

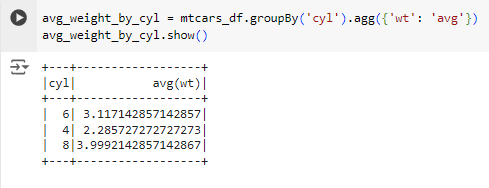


3.Add a new column named as "wtTon" to dataframe with values of weight

column multiplied by 0.45.



4. Compute the average weight of cars by their cylinders



5. Write a sql query to find gears where the value of cylinder lies

between 4 and 9.

