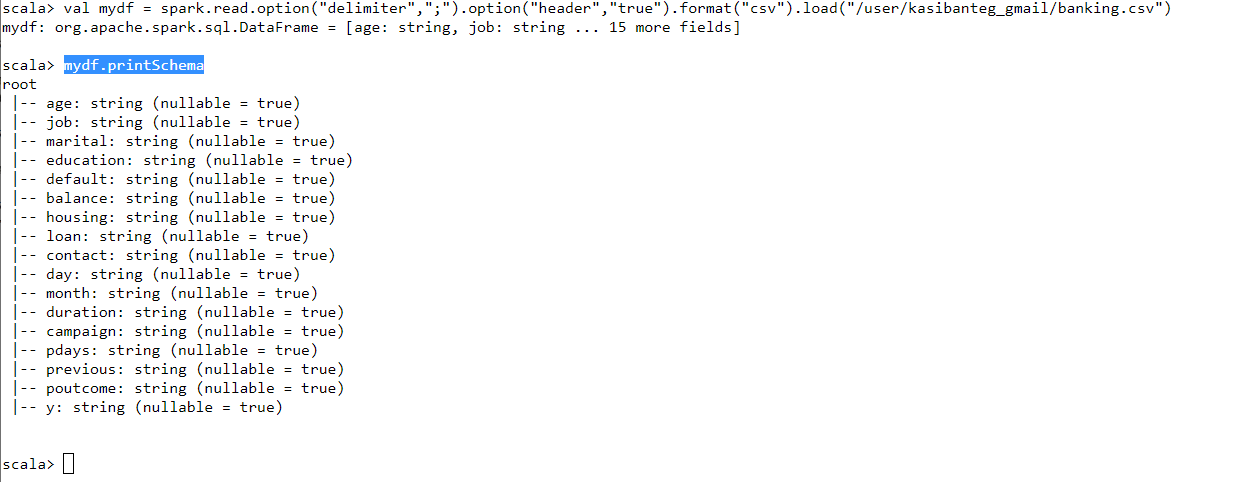
1. **Load data and create a Spark data frame**

**spark2-shell**

****

**// Register the DataFrame as a SQL temporary view**

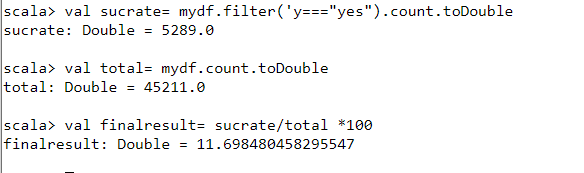
**mydf.createOrReplaceTempView("bankDomain")**

1. **Give marketing success rate (No. of people subscribed / total no. of entries)**

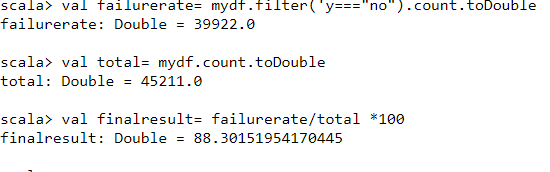
**// Register the DataFrame as a SQL temporary view**

scala> mydf.createOrReplaceTempView("bankDomain")

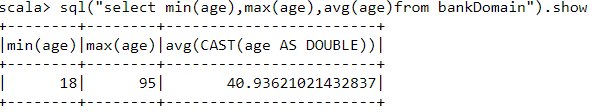
Success Rate



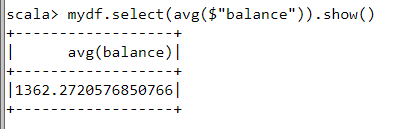
Failure Rate



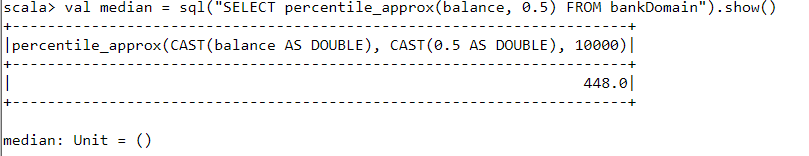
**Give the maximum, mean, and minimum age of the average targeted customer**

****

**Check the quality of customers by checking average balance, median balance of customers**



---median

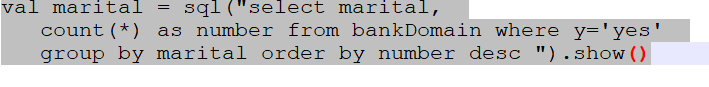


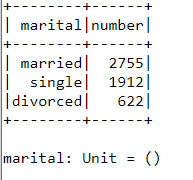
**Check if age matters in marketing subscription for deposit**



YES, Age Matters Between 30 to 36 , this is a promising age Range

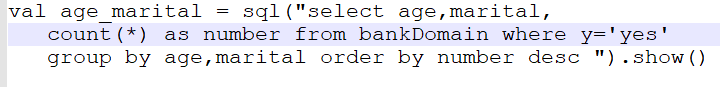
**Check if marital status mattered for a subscription to deposit**

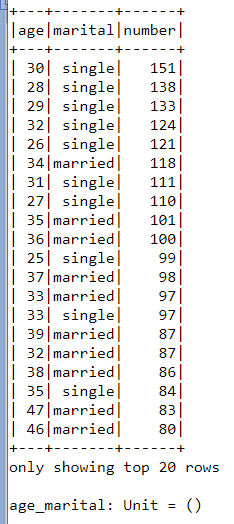
****

****

YES, Marital status Matters Married Subcription was high followed by singles and lastly divorced

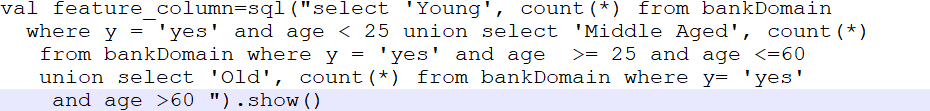
**Check if age and marital status together mattered for a subscription to deposit scheme**

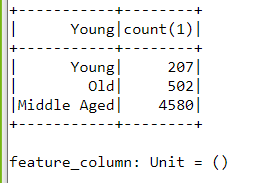




YES, Marital status and Age Mattered, singles from the age of 25 to 35 had more Subcription followed by married from age 30 to 47 but the subscription rate goes on reducing as one advances in the age and Marital status

**Do feature engineering for the bank and find the right age effect on the campaign**

****

****

So we can conclude from the Feature Engineering that It is the ‘Middle Aged’ people between age 25 and 60 who should be the targeted customers as they subscribe the most