Project: **Market Analysis in Banking Domain**

**Background:**

Your client, a Portuguese banking institution, ran a marketing campaign to convince potential customers to invest in a bank term deposit scheme.

**Objective:**

The marketing campaigns were based on phone calls. Often, the same customer was contacted more than once through phone, in order to assess if they would want to subscribe to the bank term deposit or not. You have to perform the marketing analysis of the data generated by this campaign.

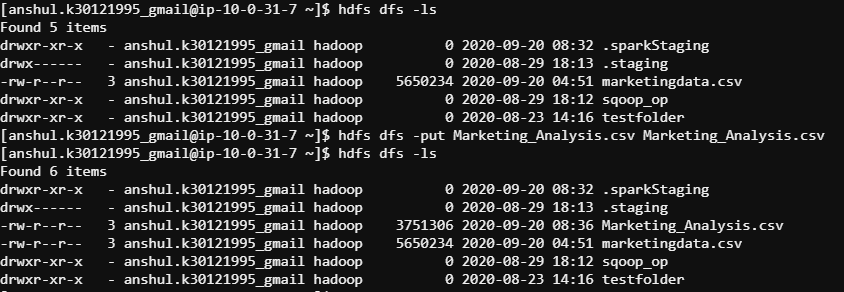
**Question-wise Solution:**

The data size is huge and the marketing team has asked you to perform the below analysis-

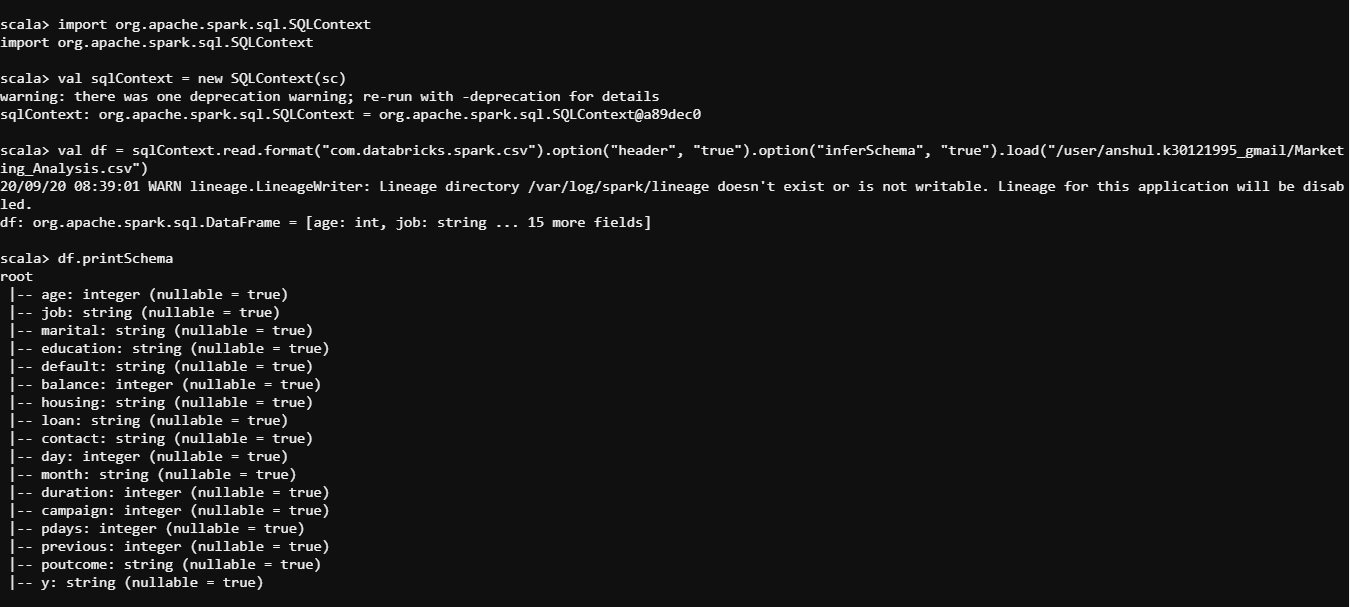
1. **Load data and create a Spark data frame:**

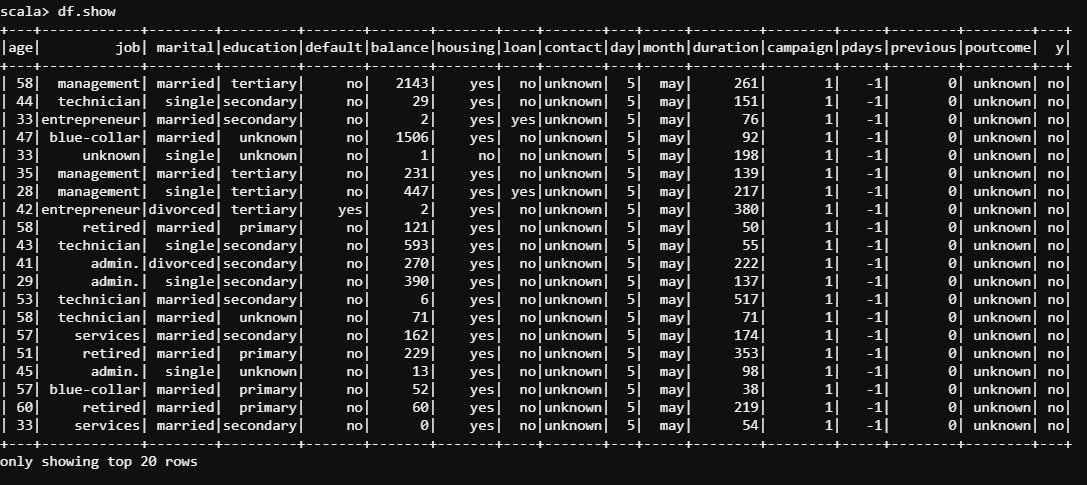
Code in the screenshot below –

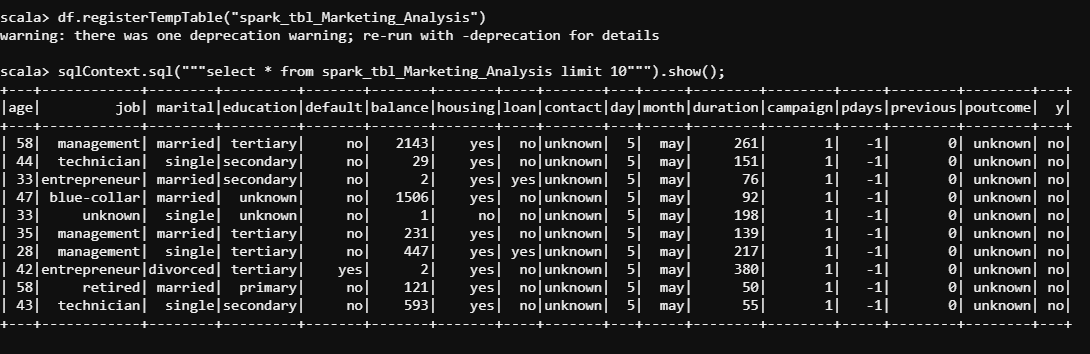
1. Adding the data file to Hadoop Distributed File System



1. Loading Data and Creating Spark Data Frame

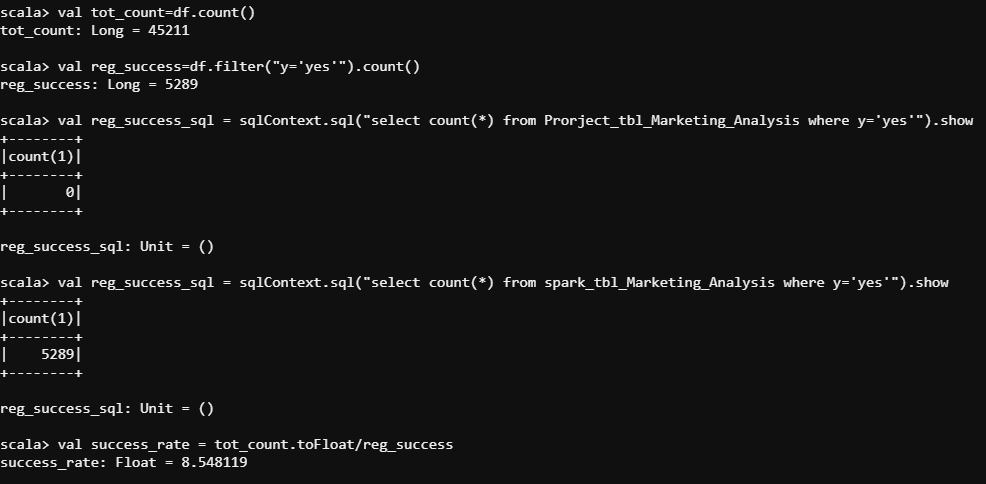






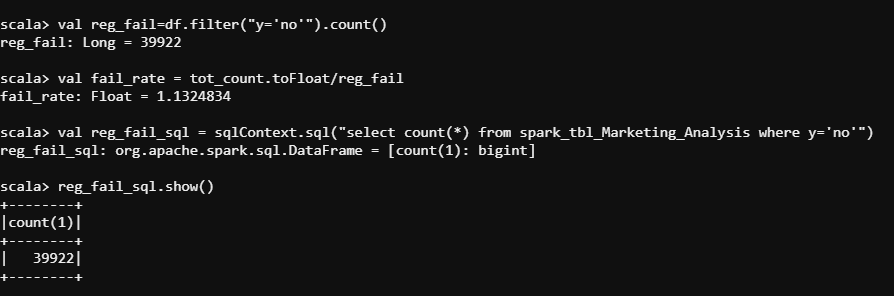
1. **Give marketing success rate (No. of people subscribed / total no. of entries):**

Code in the screenshot below –



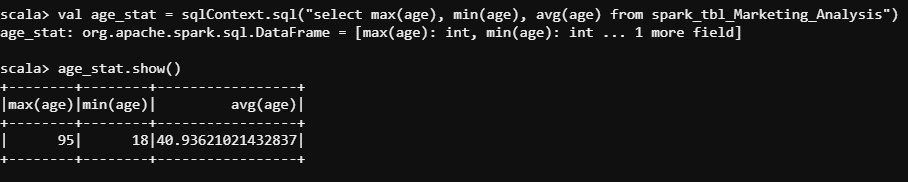
1. **Give marketing failure rate:**

Code in the screenshot below –



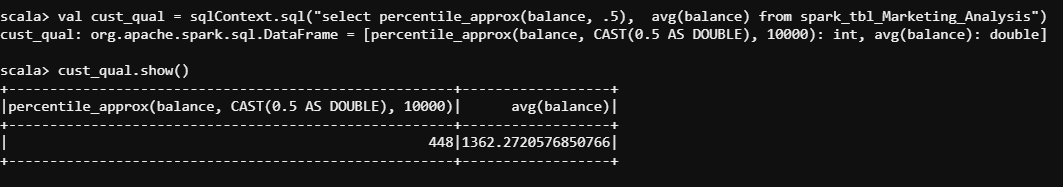
1. **Give the maximum, mean, and minimum age of the average targeted customer:**

Code in the screenshot below –



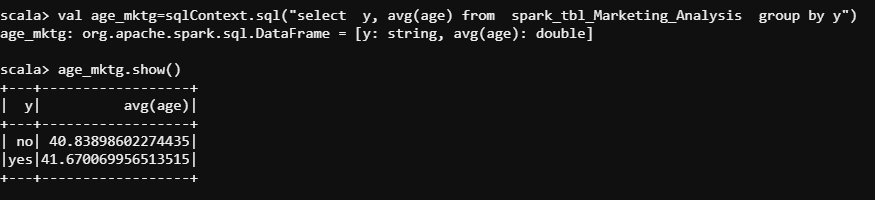
1. **Check the quality of customers by checking average balance, median balance of customers:**

Code in the screenshot below –



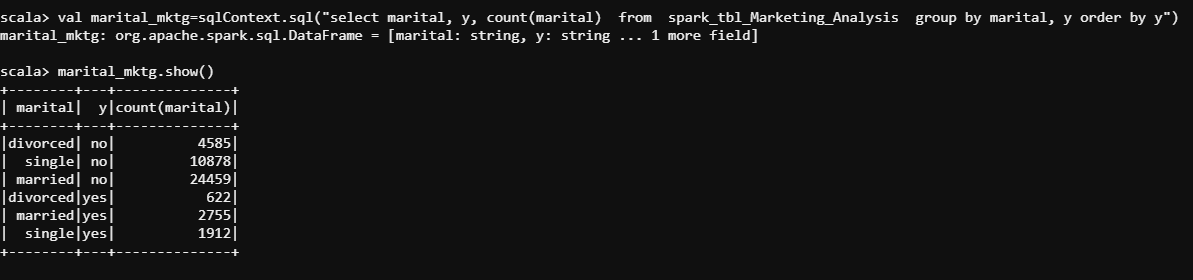
1. **Check if age matters in marketing subscription for deposit:**

Code in the screenshot below –



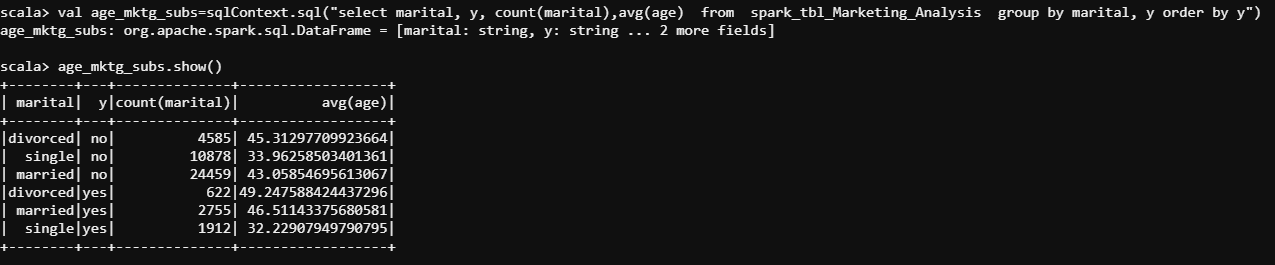
1. **Check if marital status mattered for a subscription to deposit:**

Code in the screenshot below –



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

Code in the screenshot below –



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

Code in the screenshot below –

