**Business Requirements Specification (BRS)**

1. **BACKGROUND OF THE PROJECT:**

In our project, we analyzed data from the UCI Machine Learning Repository called BankMarketing Data Set. The data is related to direct marketing campaigns of a Portuguese Banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to assess if the product (bank term deposit) would be (’yes’) or not (’no’) subscribed. There were available four datasets,we chose bank-full.csv data set, which contains all examples for the older version of this dataset.

Banking - Analysis of Bank Marketing. The bank direct marketing data analysis is important work that helps the banks predict whether customers will sign long term deposits with the banks. The method that can predict such customers' needs can be profitable to the banks for improving their marketing campaign strategies.

1. **APPROACH:**

In order to optimize marketing campaigns with the help of the dataset, we will have to take the following steps:

* The Expertise to evaluate Banking dataset→

The following key capabilities are required to acquire significant insight from the data:

1. Data can be viewed from a central site (PYSPARK)
2. Updates on banking statistics in real time
3. Tools for data visualization (GRAFANA)
4. Data can be accessed at any time and from any location.
5. Data analysis based on time.

* The flows goes like→

1. **Import data from dataset and perform initial high-level analysis:**Analyze the set of rows, see if any values are missing, and examine the dataset columns and their values in relation to the campaign outcome.
2. **Drop unwanted data:** Remove extraneous columns, deal with missing and inaccurate values, and convert category columns into dummy variables to clean up the data.
3. **Machine Learning Techniques:** Use machine learning techniques to forecast the outcome of a marketing effort and to identify elements that influence its performance.

**3. GOALS OF PROJECT**

In general, datasets containing marketing data can be utilized for one of two purposes:

1. Prediction of the results of the marketing campaign for each customer and clarification of factors which affect the campaign results. This helps to find out ways to make marketing campaigns more efficient.
2. Finding out customer segments, using data for customers, who subscribed to term deposit. This helps to identify the profile of a customer, who is more likely to acquire the product and develop more targeted marketing campaigns.

**4. PROBLEM STATEMENT**

Analyze the business perspectives and provide the hidden inference from the data.Also deploy the machine learning model for predicting the subscription of the client.

**5. SOLUTION/ PROPOSED SYSTEM**

* The main motive in the banking data is to analyze the hidden interference and targeting results.
* To find the term deposit and include business Queries.The project's major goal is to attract clients by offering bank services such as loans, credit and debit cards, digital financial services, and even personal services. Most commercial banks, on the other hand, provide some basic modern services.

**6. USERS OF PROJECT**

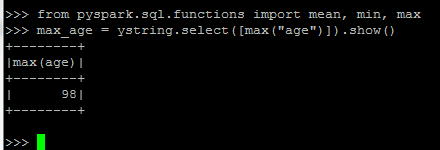
* Banks provide bank services to attract customers, from giving loans, credit and debit cards, digital financial services, and even personal services. However, some essential modern services are offered by most commercial banks.
* Client - a possible definition Customers are those who turn to bank services, but first you need to be known more closely. A customer is a person or entity receiving or benefiting from the bank, and who uses a company or banking institution to change currency.
* The customer is one who has a bank account. "Customers are bank depositors, who entrust their money, banking facilities, for storage and fruiting, for which they receive interest and applicants or those who need temporary, some additional money for paying interest or commission"

**7. REQUIREMENTS**

**FUNCTIONAL REQUIREMENTS**

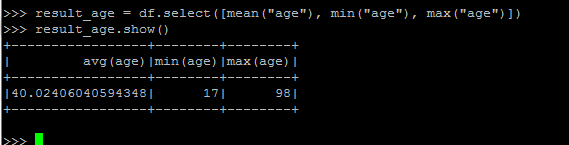
**BUSINESS QUERIES:**

1. Maximum, Mean, and Minimum age of average targeted customer

****

**result\_age = df.select([mean("age"), min("age"), max("age")])**

**result\_age.show()**

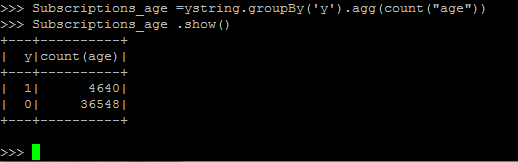
****

**The average age of the target customer**

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

**Subscriptions\_age =ystring.groupBy('y').agg(count("age"))**

**Subscriptions\_age.show()**

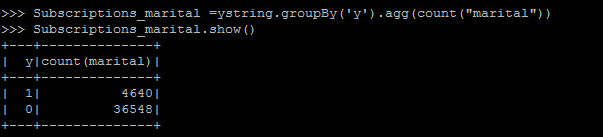
****

**Age matters in the term deposit**

**5. Check if marital status mattered for subscription to deposit**

**Subscriptions\_marital =ystring.groupBy('y').agg(count("marital"))**

**Subscriptions\_marital.show()**

****

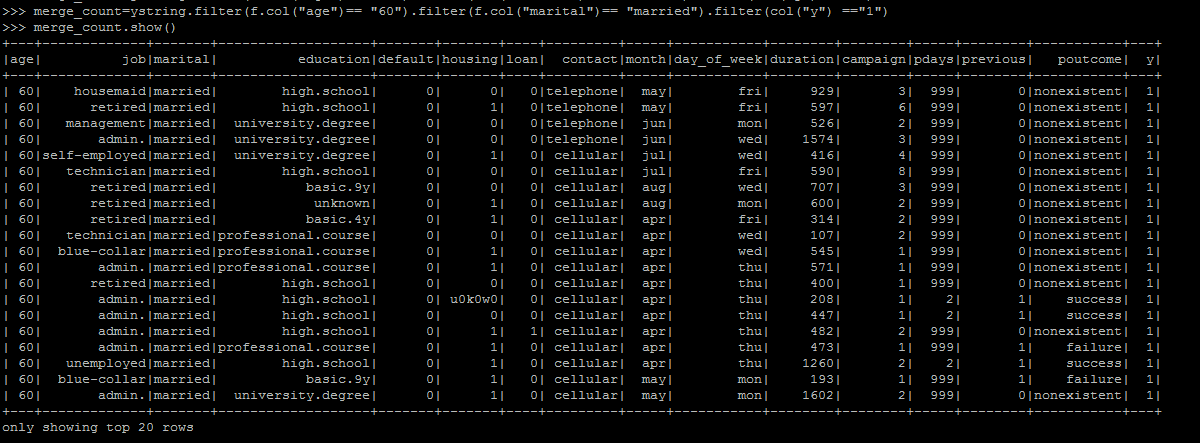
**marital status mattered for subscription to deposit**

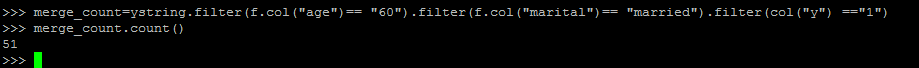
**6. Check if age and marital status together mattered for subscription to deposit scheme**

**merge\_count=ystring.filter(f.col("age")== "60").filter(f.col("marital")== "married").filter(col("y") =="1")**

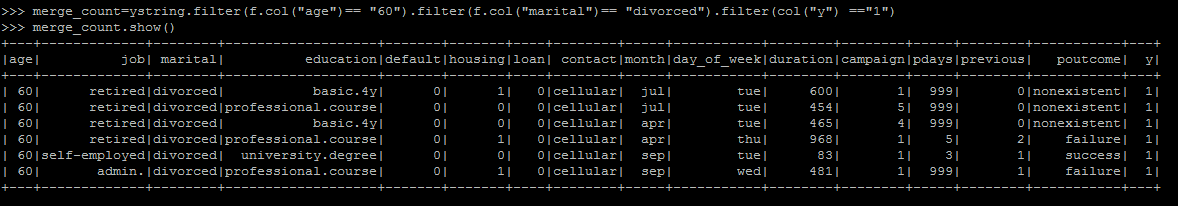
**merge\_count.show()**

* **Married users→ checking the probability of the married client in terms of subscriptions.**

****

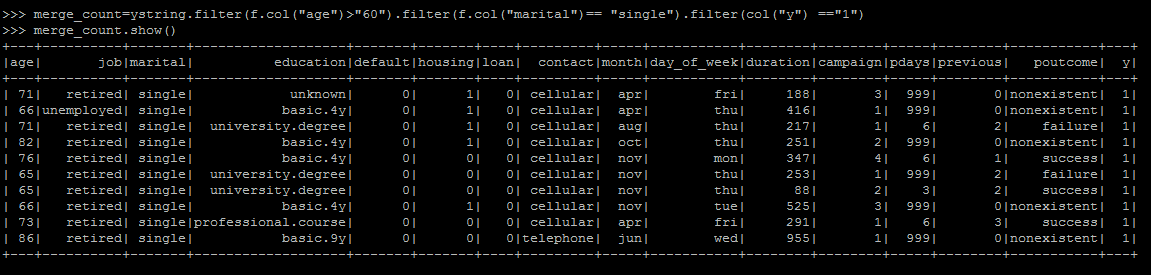
****

* **Divorced user→ Checking the probability of the divorced client in terms of subscriptions.**

****

****

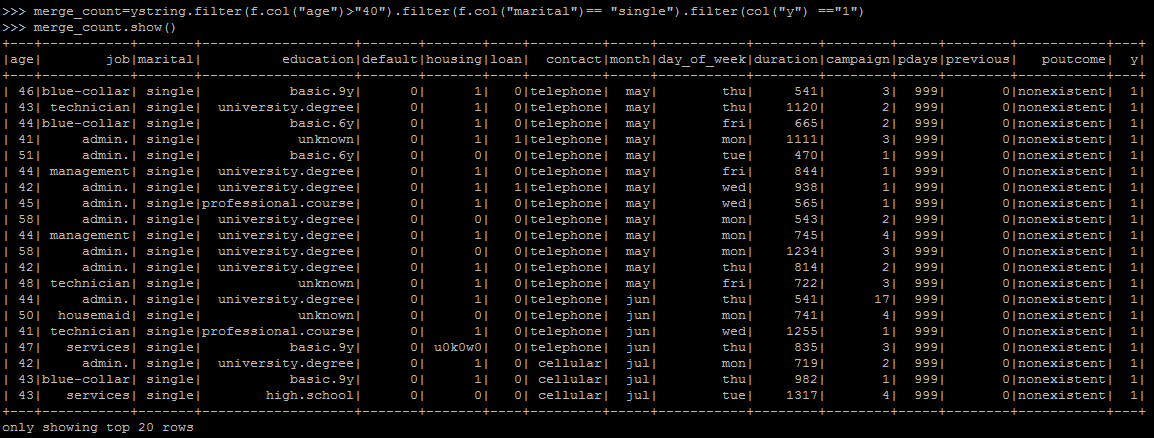
* **Singles users → Checking the probability of the single client in terms of subscriptions.**

****

****

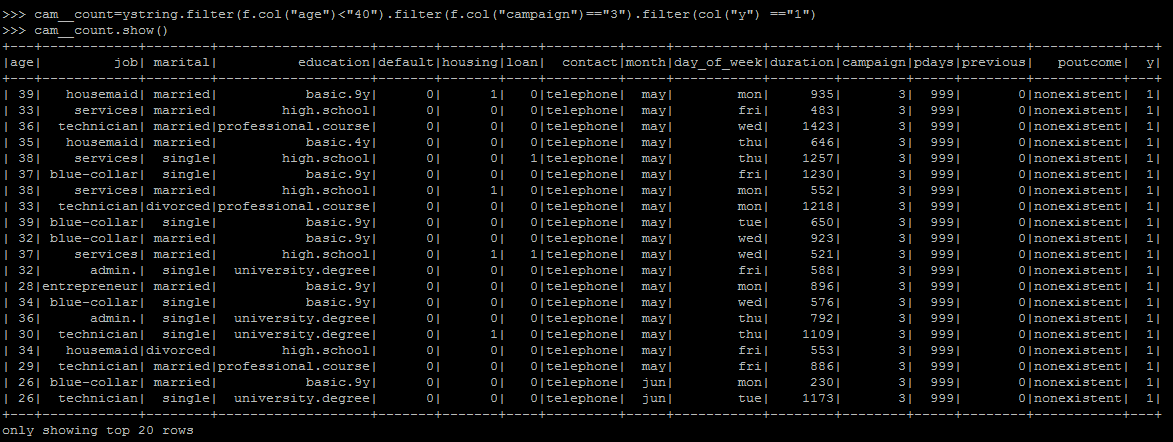
**7. Does age and find the right age effect on campaigns.**

**Age of less than 40**

****

****

**Greater than 40 age**

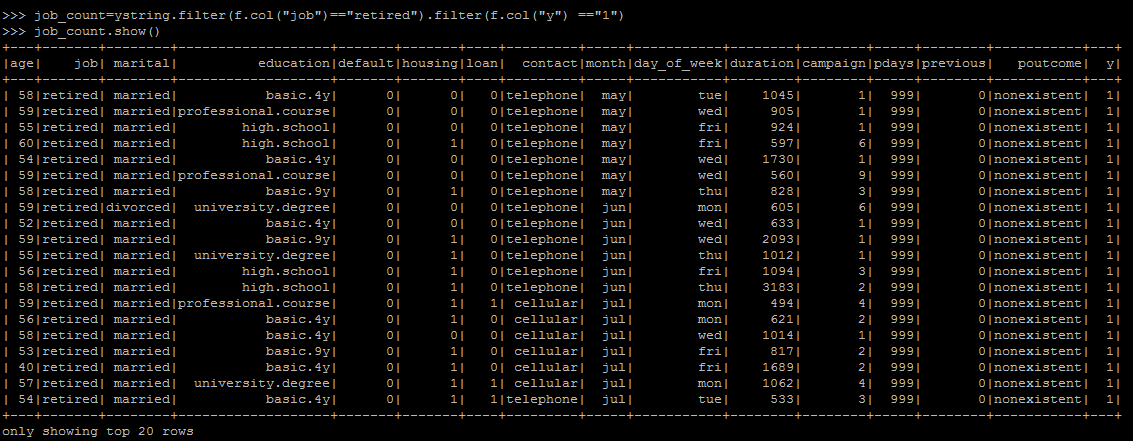
****

**SOLUTIONS→**

1. **Are retired people more likely to subscribe for term deposits? ( maybe they prefer safe investments)**

**obs->** No. The number of Subscriptions by retired people are moderate. But we can see a positive result of the campaign on senior people. That is the proportion of yes/no is more for retired people.

**Retired**

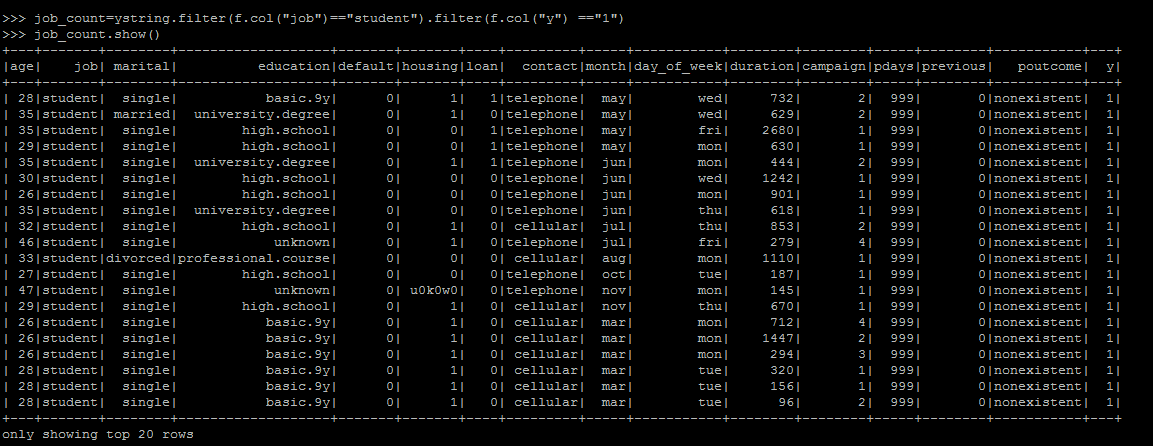


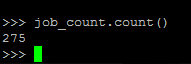
**Count**



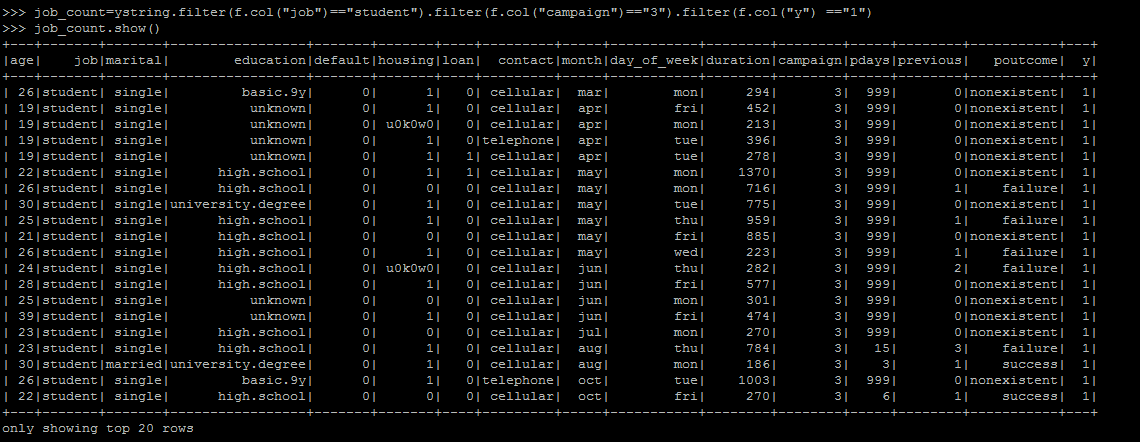
**Similarly, it is also interesting to see a very positive effect of the marketing campaign on students.**

**Student**

****

****

**With campaign**

****

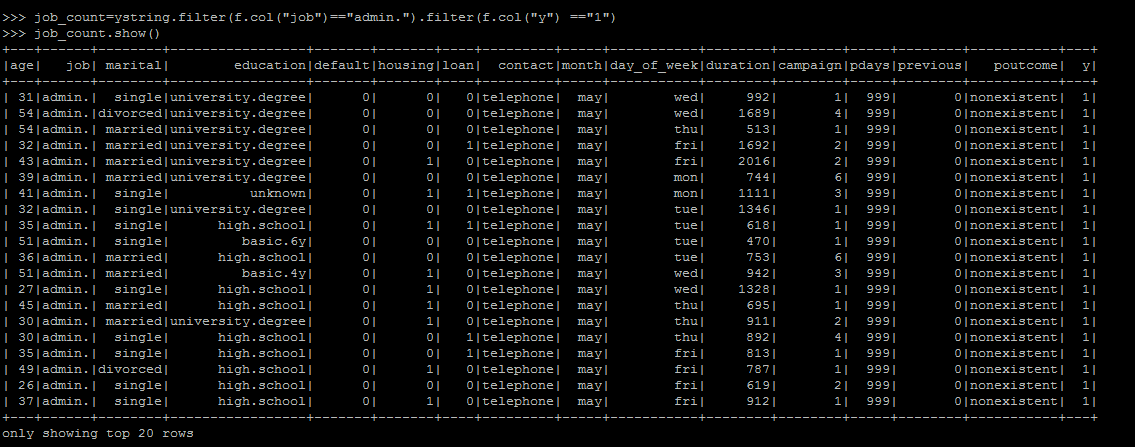
****

1. **Do salaried people prefer it more than business owners who would invest money into their business rather than putting it in the bank?**

Obs → Yes, Salaried People(admin, service, technician, blue collar jobs etc.) are much more interested than entrepreneurs and self employed people.

**job\_count = ystring.filter(f.col(“job”)==”admin.”).filter(f.col(“y”) == “1”)**

**job\_count .show()**



1. **Are married people more likely to subscribe for term deposits? (They may prefer having savings for their children?)**

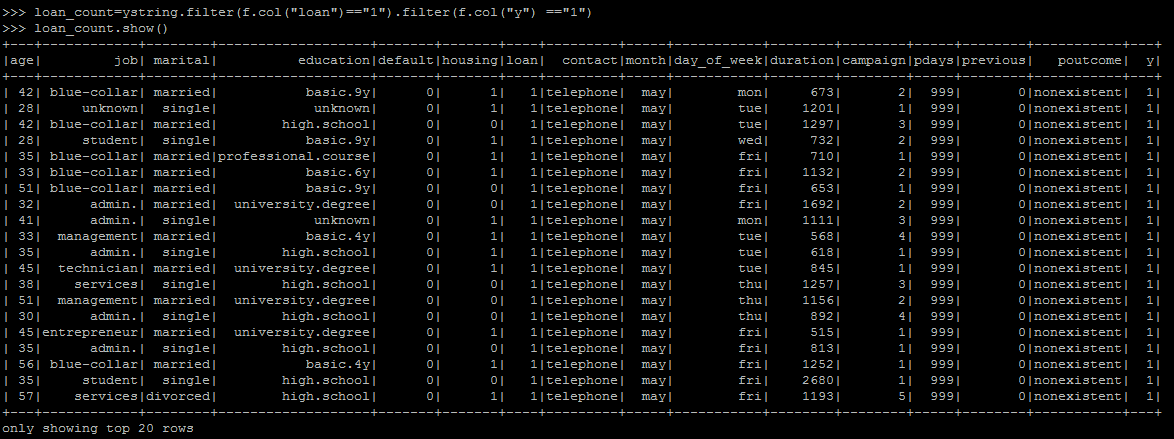
Obs→ No, Proportion of single subscribers is more than 3%.

1. **If you already have loans, would you be less likely to subscribe to term deposits?**

Obs→ People having personal loans subscribe less to the term deposits. The same is not true for people with home loans.

**With Loan**

**loan\_count=ystring.filter(f.col("loan")=="1").filter(f.col("y") =="1")**



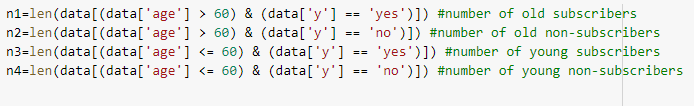


**Without loan**

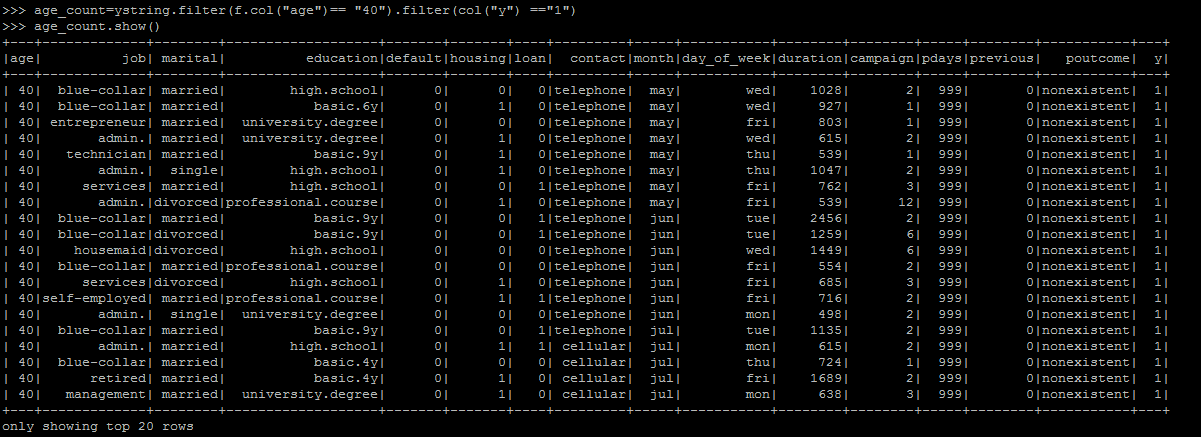
**loan\_count=ystring.filter(f.col("loan")=="0").filter(f.col("y") =="1")**

****

1. **Are younger customers more likely to subscribe to a term deposit compared to old customers (age>60) ?**

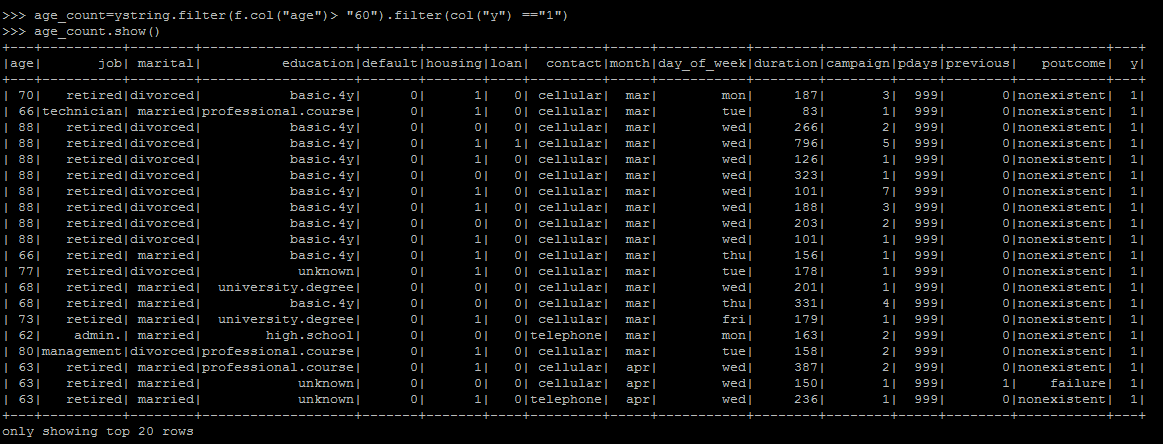
****

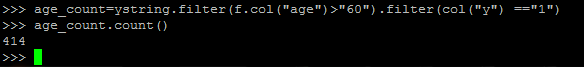
****

****

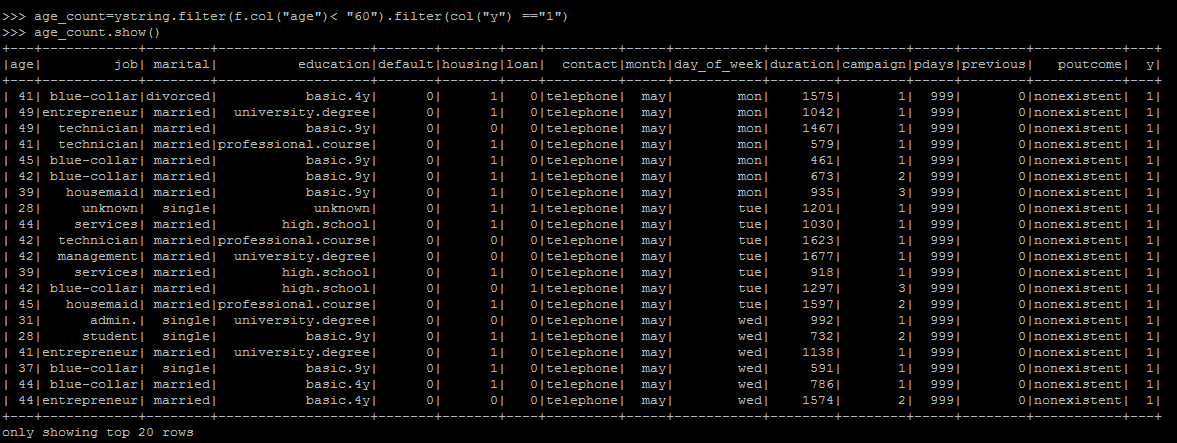
**It looks like the effect of marketing on old people is much more positive than the corresponding effect on young people! So old people can be a major group we would like to target!**

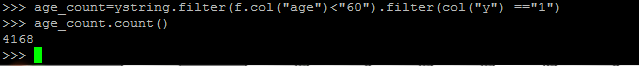
**Greater than 60**

****

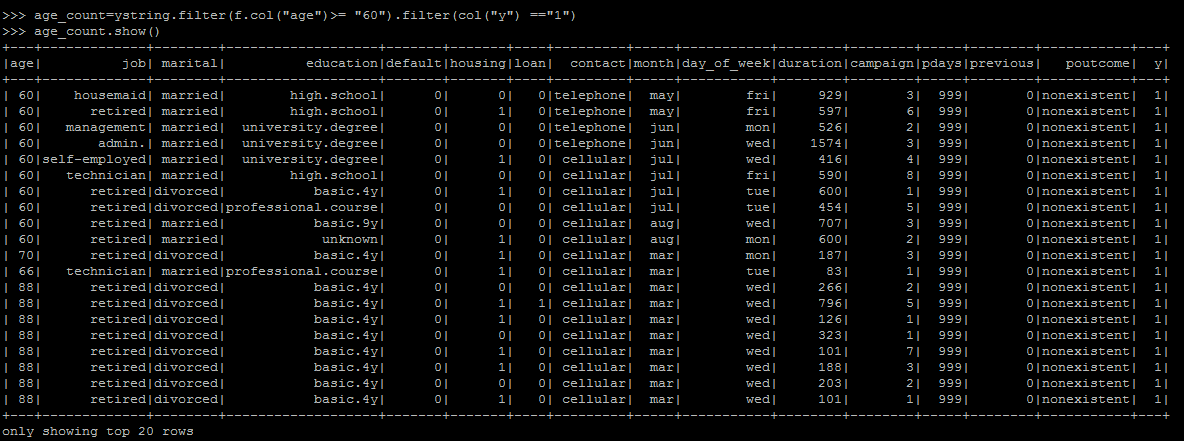
****

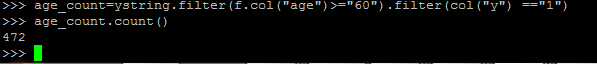
**Less than 60**

****

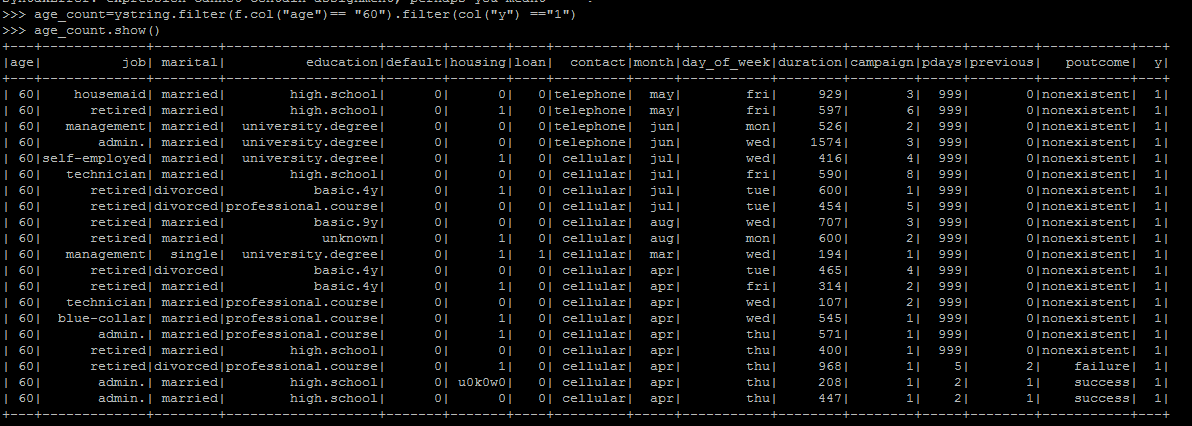
****

**Greater than equal to**

****

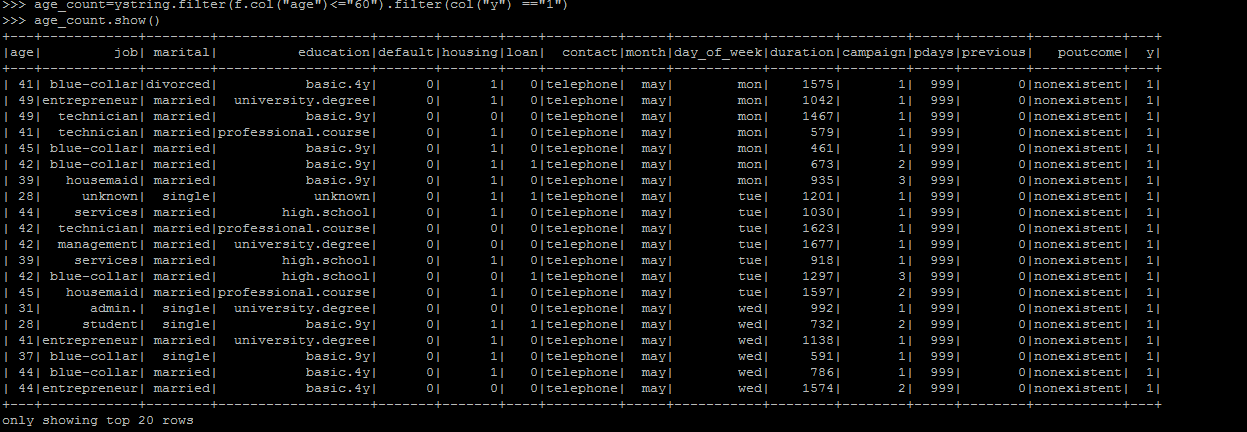
****

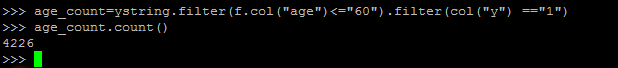
**Equal to**

****

****

**Less than equal to**

****

****

**Exploratory Data Analysis:**

1. Analysis the dataset
2. The number of missing/null values
3. The number of unique values in each variable

**NON- FUNCTIONAL REQUIREMENTS**

**Performance Requirements**

* Performance Requirements: Increase Customer Satisfaction Internet banking system must allow customers to access banking services 24 hours a day, 365 days a year with minimum downtime period for backup and maintenance.
* Expand Product Offerings: The new services allow banks to capture a larger percentage of their customers’ asset base.
* The internet banking system will provide facilities for banks to offer new services and products onto its homepage.
* Reduce Overall Costs: It will help to reduce a bank’s costs in two fundamental ways: it minimizes the cost of processing transactions and reduces the number of branches required to service an equivalent number of customers.

**8. CONCLUSIONS:**

As a result, we can conclude that people with jobs as students and retired seem to have deposited more. On analysis of marital status, it means to have been more counted of singles in term deposit followed by married. Additionally, loans played a vital role in the analysis. It looks like the effect of marketing on old people is much more positive than the corresponding effect on young people. So old people can be a major group we would like to target.

**TEAM D**

**Lasya Tlv**

**Sangeetha D**

**Santhosh. J**

**Shifana Tasneem M**