

Spark and SQL Spark test for Equifax Collections & Recoveries

Requirements:

- The test must:
 - Be coded in JAVA 8.
 - Use spark-core and spark-sql as part of the solution for the required reports.

Deliverables:

- Java executable in GitHub
- Source java code in GitHub

Problem

The Numeror bank provides a csv named: accounts.csv where each row has the next values: group, account, client number and amount, as well the bank delivers a json named: customer.json where each record represents client number, customer name and age.

Report 1:

Bank requires a report where for each client id in accounts must be linked with their respective customer for example:

6, 1991000580488888,24040005804,560.00

Must be linked with:

```
{"client":"24040005804","name":"Gollum","age":60}
```

So at the end the report should look like:

Gollum, total accounts: 8, account1: 1991000580488888, account2: 9991000589488888, account3: 1014000581415759, ...

Report 2:

Bank requires a report where for each client id summarize the amount in each product, where product is based on group, follow the next map

Group → Product

1 → Mortgage

6 → Credit Card

8 → Personal Loan

So at the end the report should look like:

Gollum, Mortgage: 48000935.00, Credit Card: 1120.00, Personal Loan: 96135.00

Report 3:

Bank requires a report with seniors' customers (for the bank seniors customers are any client with more than 100 years old). And that report should offer a discount based on customer age as follow:

Max age of all seniors' customer provide 40% discount

Min age of all seniors' customer provide 10% discount

The discount will be a linear curve function, therefore $\text{discount} = m \cdot \text{age_of_senior} + b$ (so is part of the task to calculate m and b).

At the end the report should look like

Sauron, Account (discount): 1014002279606632 (216.80), 10140022796066329 (21680.00)