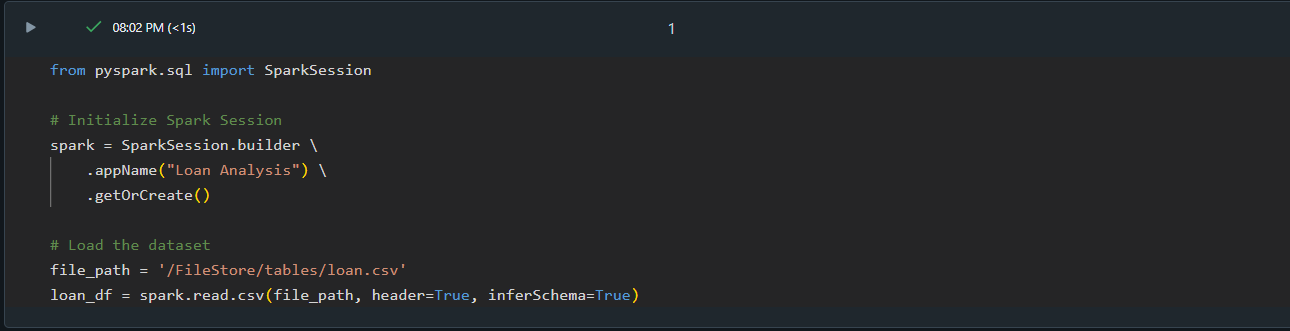
**Case Study Pyspark**

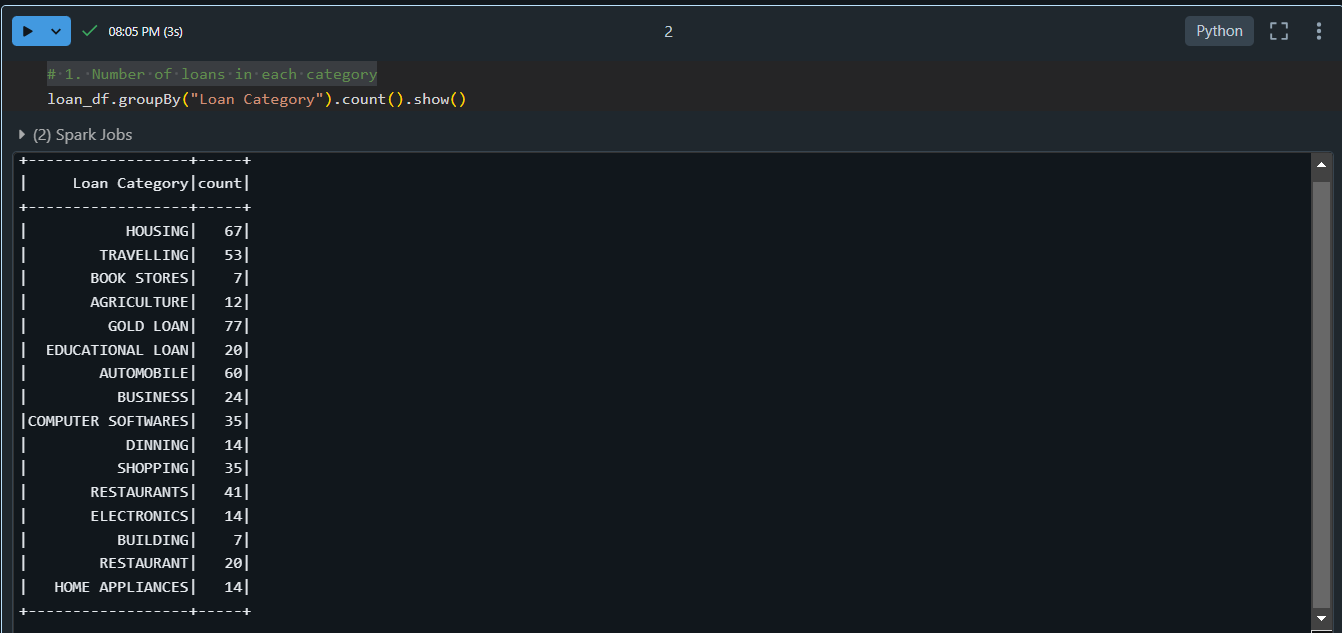
**Divya Sree Murali**

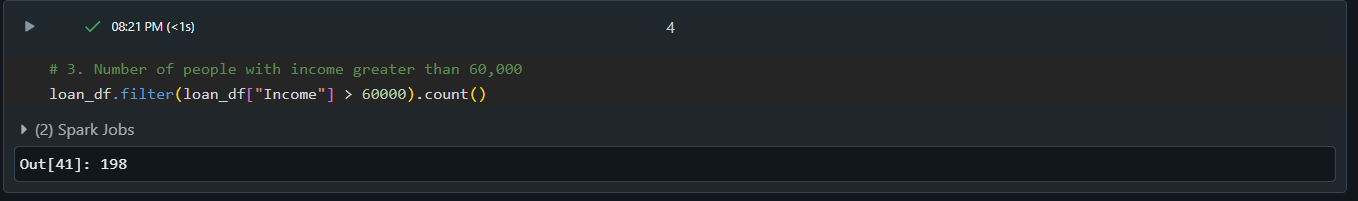
**25-11-2024**

**Initializing Spark Session and loading dataset**

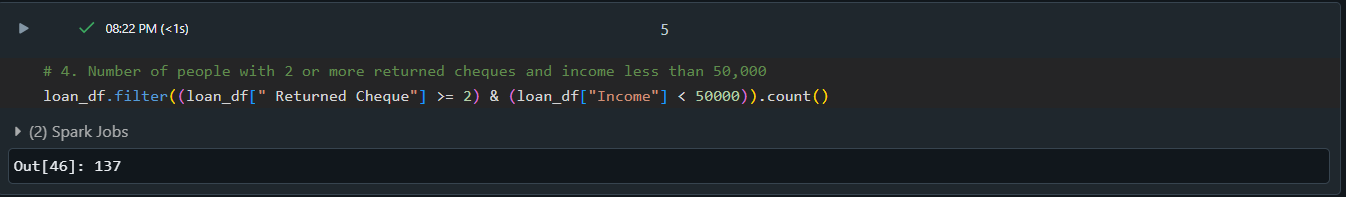


**1.Number of loans in each category**

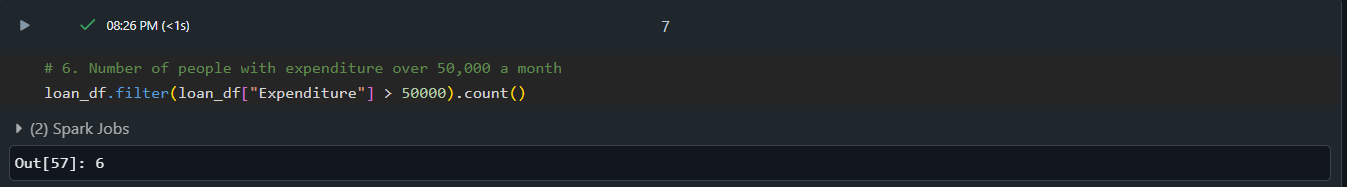


**3. Number of people with income greater than 60,000** 

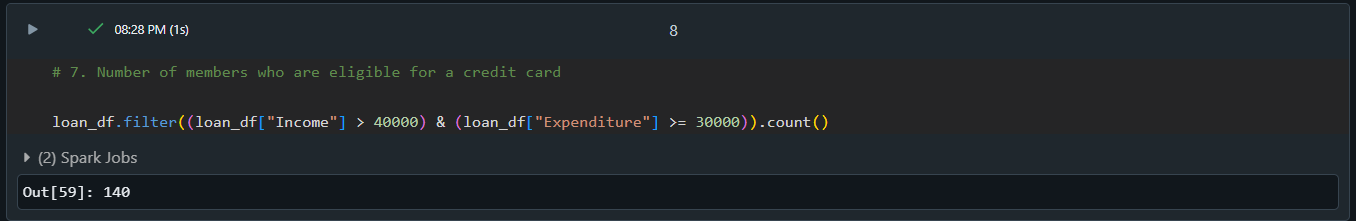
**4.Number of people with 2 or more returned cheques and income less than 50000**



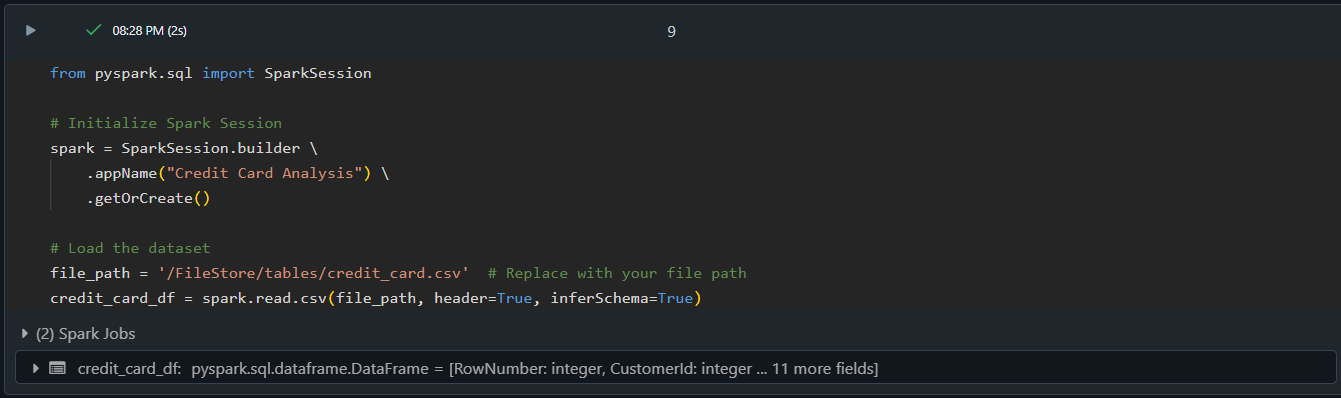
**6.Number of people with expenditure over 50000 a month**

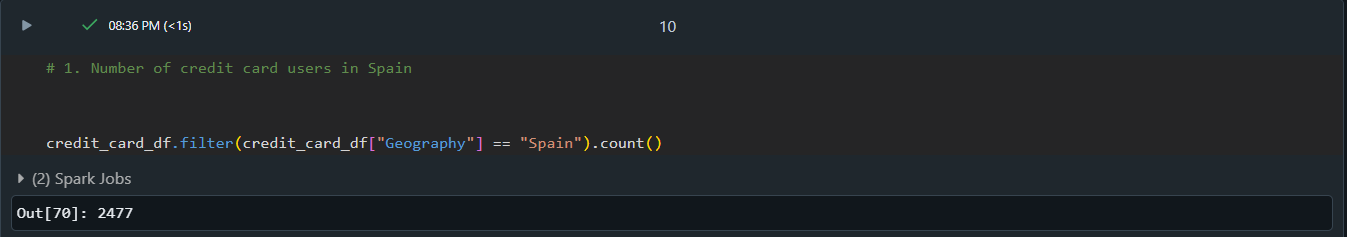


**7.Number of members who are eligible for credit card**

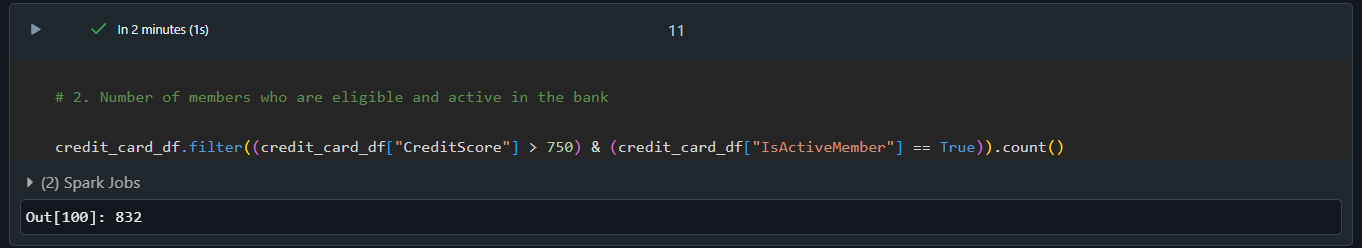


**Credit card Dataset**

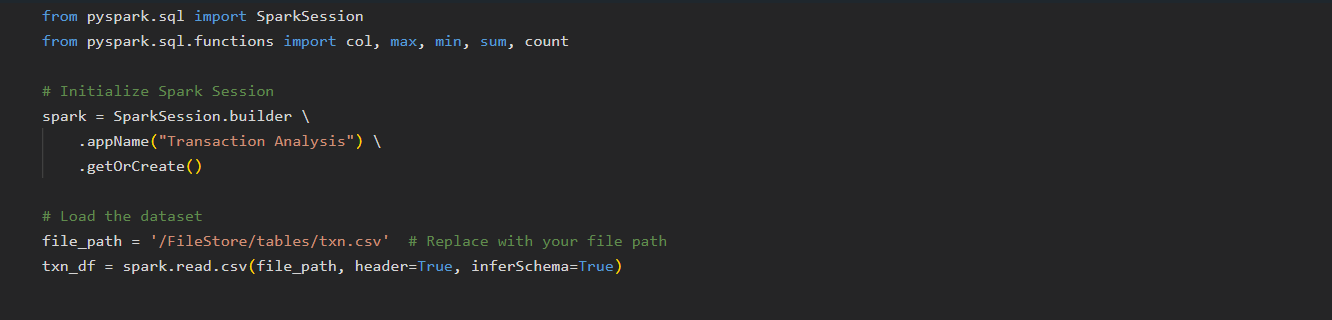


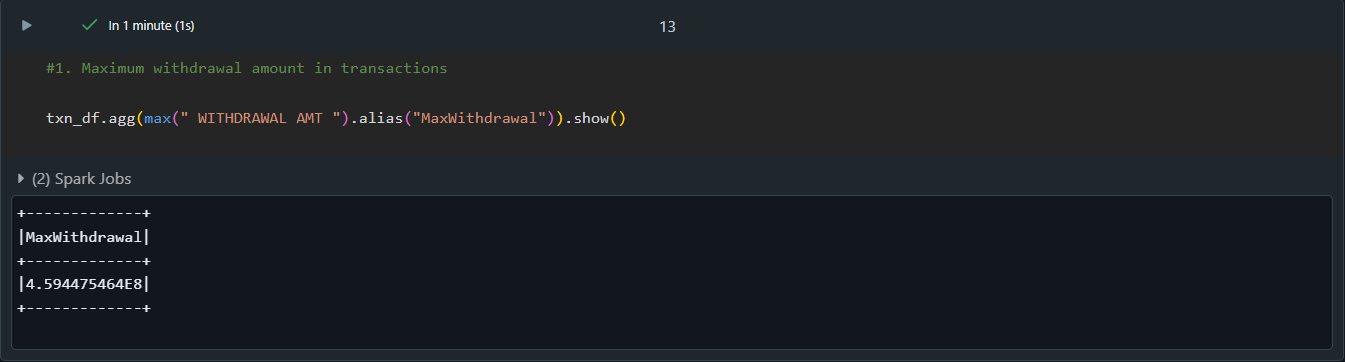
**1. Number of credit card users in Spain**

**2.Number of members who are eligible and active in the bnak**



**Transaction dataset**



**1. Maximum withdrawal amount in transactions**

**2.Minimum withdrawal amount of account**

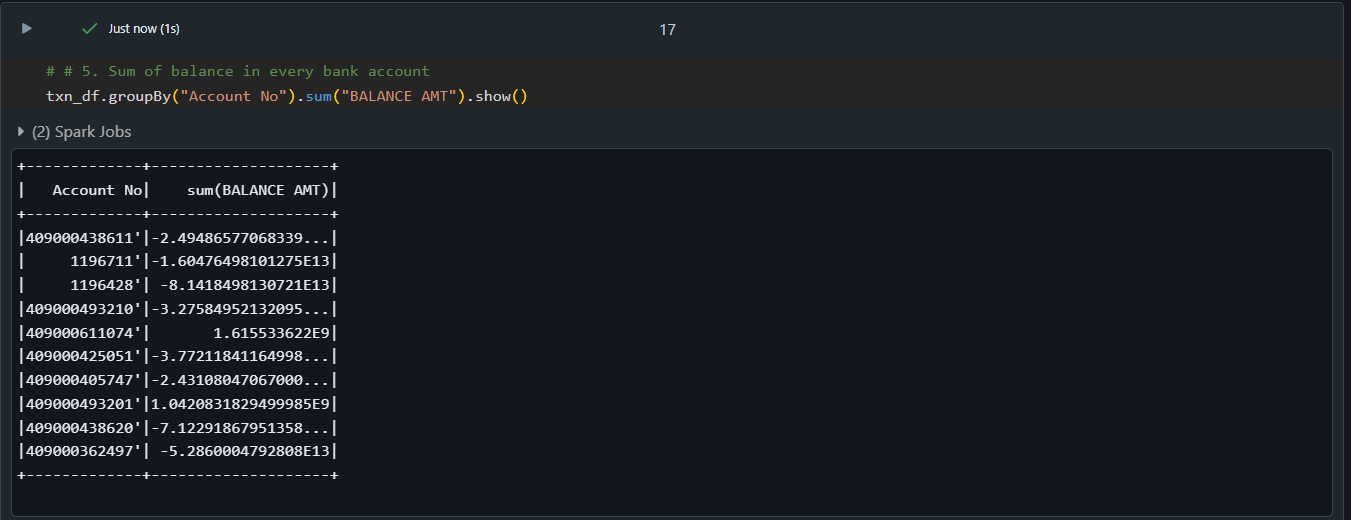


**3. Maximum deposit amount of an account**

**4.Minimum deposit amount of an account**



**5.Sum of balance in every bank account**



**6.Number of Transactions on each date**

