**Case Study 1: Real-time Analytics for E-Commerce Platform Using Scala and Apache Spark**

**Business Scenario:**

You have been provided some Amazon products with their scoring by different users. Find out the average rating of each product.

**Data Set:**A text file consisting of Amazon data is provided in the following format. Use Scala with Spark/SparkSQL to find out the desired resultant.

**product/productId: B0000DF3IX  
product/title: Paprika Hungarian Sweet  
product/price: unknown  
review/userId: A244MHL2UN2EYL  
review/profileName: P. J. Whiting "book cook"  
review/helpfulness: 0/0  
review/score: 5.0  
review/time: 1127088000  
review/summary: Sweet Paprika: A sweet ingredient!  
review/text: While in Hungary we were given a recipe for Hungarian Goulash. It needs sweet paprika. This was terrific in that dish and others. I will purchase it again when I need more.**

**Case Study 2: Credit Card transaction analysis using Scala and Spark**

**Business Scenario:**

You have been provided with a credit card dataset containing information about credit card transactions. Your task is to analyze the dataset using Scala and Spark/Spark SQL and derive meaningful insights. A couple of lists or reports are to be generated from this data as specified.

**Data Sets and Data Dictionary:**

The data is given in csv files. Below is the data dictionaries.

A screen shot of a computer code

Description automatically generated

**Problem Statement:**

1. List the number of members who are eligible for credit card.
2. List the number of members who are eligible and active in the bank.
3. List the credit card users belonging to Spain
4. Save the above output as comma separated files.