PI2C_Lucas_Citolin

February 9, 2020

[1]:
```
import pyspark
from pyspark.sql import functions as F
from pyspark.sql.functions import desc
from pyspark.sql.window import Window
import datetime
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
from pyspark import SparkConf, SparkContext
conf = SparkConf().setAppName("Alti bd2") \
                    .setMaster('local') \
                    .set('spark.executor.memory','5G') \
                    .set('spark.cores.max', '24') \
                    .set('spark.sql.tungsten.enabled', 'true')

sc = SparkContext(conf=conf)

sc._jsc.hadoopConfiguration().set('textinputformat.record.delimiter',␣
 ↪'\r\n\r\n')
```

[2]:
```
spark = pyspark.sql.SparkSession.builder \
            .master('local') \
            .appName('Altigran BD2') \
            .getOrCreate()

file = './data/amazon-meta.txt'
data = sc.textFile(file)
```

[3]:
```
spark
```

[3]:
```
<pyspark.sql.session.SparkSession at 0x7f2fcbf0edd0>
```

0.1 Utils

[4]:
```
def parse(text):
    lines = text.split('\r\n')
    obj = {}
    ID = -1
    reading_categories = False
    reading_reviews = False
    for i in lines:

        if reading_categories:
            categories = i.split('|')
            if(len(categories) == 1):
                reading_categories = False
            else:
                obj['CATEGORIES']['categories'].
 ↪append(categories[len(categories)-1])

        if reading_reviews:
            reviewsObj = {}
            reviewsObj['date'] = i.split(' cutomer')[0].replace(' ', '')
            reviews = i.split(': ')
            reviewsObj['customer'] = reviews[1].split(' rating')[0].replace('␣
 ↪', '')
            reviewsObj['rating'] = reviews[2].split(' ')[0].replace(' ', '')
            reviewsObj['votes'] = reviews[3].split(
                ' helpful')[0].replace(' ', '')
            reviewsObj['helpful'] = reviews[4].replace(' ', '')
            obj['REVIEWS']['reviews'].append(reviewsObj)

        elif 'Id:   ' in i:
            obj['ID'] = i.split('Id:')[1].replace(' ', '')
            ID = obj['ID']
        elif 'ASIN:' in i:
            obj['ASIN'] = i.split('ASIN:')[1].replace(' ', '')
        elif 'title:' in i:
            title = i.split('title: ')[1]
            obj['TITLE'] = title
        elif 'group:' in i:
            obj['GROUP'] = i.split('group: ')[1]
        elif 'salesrank:' in i:
            obj['SALESRANK'] = i.split('salesrank: ')[1]
        elif 'similar:' in i:
            similars = i.split(': ')[1].split('  ')
            number_of_similars = similars[0]
            similars.pop(0)
            obj['SIMILARS'] = {'size': number_of_similars, 'similars': similars}
```

```
        elif 'categories:' in i:
            number_of_categories = i.split(': ')[1]
            obj['CATEGORIES'] = {
                'size': number_of_categories, 'categories': []}
            reading_categories = True
        if 'reviews:' in i:
            reviews_split = i.split(': ')
            total = reviews_split[2].split('  ')[0]
            downloaded = reviews_split[3].split('  ')[0]
            avg_rating = reviews_split[4]
            obj['REVIEWS'] = {'total': total, 'downloaded': downloaded,
                              'avg_rating': avg_rating, 'reviews': []}
            reading_reviews = True
    return obj


# Doc = Text Document
def filterById(doc, selected_id):
    lines = doc.split('\n')
    for line in lines:
        if 'Id:   ' in line:
            ID = int(line.split('Id:')[1].replace(' ', ''))
            if( ID == selected_id ):
                return True
            else:
                return None

# Just to print
class bcolors:
    HEADER = '\033[95m'
    OKBLUE = '\033[94m'
    OKGREEN = '\033[92m'
    WARNING = '\033[93m'
    FAIL = '\033[91m'
    ENDC = '\033[0m'
    BOLD = '\033[1m'
    UNDERLINE = '\033[4m'
```

0.2 1. Create the dataframes using RDDs

Professor, eu tentei criar os DataFrames em uma única passada emitindo tuplas com IDs referenciando tabelas (como podes ver no código comentado), funcionou porém acabou com a minha memória.

Para que isso não ocorra no seu pc, fiz da forma B: cada passada nos arquivos constrói uma relação.

[5]:
```
# TUPLE_PRODUCTS_ID = 0
# TUPLE_SIMILARS_ID = 1
# TUPLE_CATEGORIES_ID = 2
# TUPLE_REVIEWS_ID = 3

# KEY = 0
# VALUE = 1

# def emitProductsRelation(obj):
#     try:
#         return [ (TUPLE_PRODUCTS_ID, ( int(obj['ID']), int(obj['ASIN']),␣
 ↪obj['TITLE'], obj['GROUP'], int(obj['SALESRANK']), \
#                 int(obj['REVIEWS']['total']),␣
 ↪int(obj['REVIEWS']['downloaded']), int(obj['REVIEWS']['avg_rating']) ) )]
#     except Exception as e:
#         return []

# def emitSimilarsRelation(obj):
#     try:
#         ID = int(obj['ID'])
#         array = [ ( TUPLE_SIMILARS_ID, ( ID, int(similar) ) ) for similar in␣
 ↪obj['SIMILARS']['similars'] ]
#         return array
#     except Exception as e:
#         return []

# def emitCategoriesRelation(obj):
#     try:
#         ID = int(obj['ID'])
#         array = [ ( TUPLE_CATEGORIES_ID, ( ID, categories ) ) for categories␣
 ↪in obj['CATEGORIES']['categories'] ]
#         return array
#     except Exception as e:
#         return []

# def emitReviewsRelation(obj):
#     try:
#         ID = int(obj['ID'])
#         array = [ (TUPLE_REVIEWS_ID, ( ID, r['date'], r['customer'],␣
 ↪int(r['rating']), int(r['votes']), int(r['helpful']) ) ) \
#                   for r in obj['REVIEWS']['reviews'] ]
#         return array
#     except Exception as e:
#         return []

# # Get each data from the textual document in the following format:
# # ( TUPLE_ID, ARRAY )
```

```
# # With TUPLE_ID: PRODUCT_ID  ||  CATEGORY_ID  || REVIEW_ID
# def emitRelations(doc):
#     obj = parse(doc)
#     products = emitProductsRelation(obj)
#     similars = emitSimilarsRelation(obj)
#     categories = emitCategoriesRelation(obj)
#     reviews = emitReviewsRelation(obj)
#     return products + similars + categories + reviews

# rdd = data.flatMap( emitRelations ) \
#             .groupByKey() \
#             .map( lambda x: ( x[KEY], list(x[VALUEmini])))

# products_relation = rdd.filter( lambda x: x[0] == TUPLE_PRODUCTS_ID ).
↪collect()[0][VALUE]
# similars_relation = rdd.filter( lambda x: x[0] == TUPLE_SIMILARS_ID ).
↪collect()[0][VALUE]
# categories_relation = rdd.filter( lambda x: x[0] == TUPLE_CATEGORIES_ID ).
↪collect()[0][VALUE]
# reviews_relation = rdd.filter( lambda x: x[0] == TUPLE_REVIEWS_ID ).
↪collect()[0][VALUE]

# abc = rdd.collect()
```

[6]:
```
def emitProductsRelation(doc):
    obj = parse(doc)
    try:
        return( int(obj['ID']), obj['ASIN'], obj['TITLE'], obj['GROUP'],␣
↪int(obj['SALESRANK']), \
                int(obj['REVIEWS']['total']), int(obj['REVIEWS']['downloaded']),␣
↪float(obj['REVIEWS']['avg_rating']) )
    except Exception as e:
        pass

rdd = data.map( emitProductsRelation ) \
          .filter(lambda x: x!=None)

productsDF = rdd.toDF(['id', 'asin', 'title', 'group', 'salesrank', 'r_total',␣
↪'r_downloaded', 'r_avg_rating' ])
productsDF.createOrReplaceTempView('products')
```

[7]:
```
def emitSimilarsRelation(doc):
    obj = parse(doc)
    array = []
    try:
        ID = int(obj['ID'])
```

```
        array = [ ( ID, similar ) for similar in obj['SIMILARS']['similars'] ]
        return array
    except Exception as e:
        return array

rdd = data.flatMap( emitSimilarsRelation )

similarsDF = rdd.toDF(['productID', 'asin'])
similarsDF.createOrReplaceTempView('similars')
```

[8]:
```
def emitCategoriesRelation(doc):
    obj = parse(doc)
    array = []
    try:
        ID = int(obj['ID'])
        array = [ ( ID, categories ) for categories in␣
 ↪obj['CATEGORIES']['categories'] ]
        return array
    except Exception as e:
        return array

rdd = data.flatMap( emitCategoriesRelation )

categoriesDF = rdd.toDF(['productID', 'category'])
categoriesDF.createOrReplaceTempView('categories')
```

[9]:
```
def emitReviewsRelation(doc):
    obj = parse(doc)
    array = []
    try:
        ID = int(obj['ID'])
        array = [ ( ID, datetime.datetime.strptime( r['date'], '%Y-%m-%d').
 ↪date(), \
                    r['customer'], int(r['rating']), int(r['votes']),␣
 ↪int(r['helpful']) ) \
                  for r in obj['REVIEWS']['reviews'] ]
        return array
    except Exception as e:
        return array

rdd = data.flatMap( emitReviewsRelation )

reviewsDF = rdd.toDF(['productID', 'date', 'customer', 'rating', 'votes',␣
 ↪'helpful'])
reviewsDF.createOrReplaceTempView('reviews')
```

0.3 Questão A

0.3.1 Dado um produto, listar:

0.3.2 — Os 5 comentários mais úteis e com a maior avaliação.

0.3.3 — Os 5 comentários mais úteis e com a menor avaliação.

0.4 Usando SQL

Estratégia: Criar uma **VIEW** com os 5 comentários mais úteis. - Ordenar view por rating de forma ascendente - Ordenar view por rating de forma decrescente

[10]:
```
SELECTED_ID = 2

most_useful_comments_view = spark.sql("""
    SELECT *
        FROM reviews r
        WHERE r.productID = {}
        SORT BY r.helpful DESC
        LIMIT 5
""".format(SELECTED_ID))
most_useful_comments_view.createOrReplaceTempView('most_useful_comments')
```

0.4.1 [SQL] - Os 5 comentários mais úteis e com a MAIOR avaliação.

[11]:
```
spark.sql('SELECT * FROM most_useful_comments ORDER BY rating DESC').show(1000)
```

```
+---------+----------+--------------+------+-----+-------+
|productID|      date|      customer|rating|votes|helpful|
+---------+----------+--------------+------+-----+-------+
|        2|2002-01-24|A13SG9ACZ9O5IM|     5|    8|      8|
|        2|2002-05-23|A1GIL64QK68WKL|     5|    8|      8|
|        2|2002-02-06|A2P6KAWXJ16234|     4|   16|     16|
|        2|2002-03-23|A3GO7UV9XX14D8|     4|    6|      6|
|        2|2004-02-11|A1CP26N8RHYVVO|     1|   13|      9|
+---------+----------+--------------+------+-----+-------+
```

0.4.2 [SQL] - Os 5 comentários mais úteis e com a MENOR avaliação.

[12]:
```
spark.sql('SELECT * FROM most_useful_comments ORDER BY rating ASC').show(1000)
```

```
+---------+----------+--------------+------+-----+-------+
|productID|      date|      customer|rating|votes|helpful|
```

```
+---------+----------+--------------+------+-----+-------+
|        2|2004-02-11|A1CP26N8RHYVVO|     1|   13|      9|
|        2|2002-03-23|A3GO7UV9XX14D8|     4|    6|      6|
|        2|2002-02-06|A2P6KAWXJ16234|     4|   16|     16|
|        2|2002-01-24|A13SG9ACZ9O5IM|     5|    8|      8|
|        2|2002-05-23|A1GIL64QK68WKL|     5|    8|      8|
+---------+----------+--------------+------+-----+-------+
```

0.5 Usando DataFrame DSL

[13]:
```
most_useful_comments_df = reviewsDF.where( reviewsDF.productID == SELECTED_ID )␣
↪\
                                   .orderBy('helpful', ascending=False) \
                                   .limit(5) \
                                   .cache()
```

[14]:
```
answer_a_a = most_useful_comments_df.orderBy('rating', ascending=False)
answer_a_b = most_useful_comments_df.orderBy('rating', ascending=True)
```

0.5.1 [DSL] - Os 5 comentários mais úteis e com a MAIOR avaliação

[15]:
```
answer_a_a.show(1000)
```

```
+---------+----------+--------------+------+-----+-------+
|productID|      date|      customer|rating|votes|helpful|
+---------+----------+--------------+------+-----+-------+
|        2|2002-01-24|A13SG9ACZ9O5IM|     5|    8|      8|
|        2|2002-05-23|A1GIL64QK68WKL|     5|    8|      8|
|        2|2002-02-06|A2P6KAWXJ16234|     4|   16|     16|
|        2|2002-03-23|A3GO7UV9XX14D8|     4|    6|      6|
|        2|2004-02-11|A1CP26N8RHYVVO|     1|   13|      9|
+---------+----------+--------------+------+-----+-------+
```

0.5.2 [DSL] - Os 5 comentários mais úteis e com a MENOR avaliação

[16]:
```
answer_a_b.show(1000)
```

```
+---------+----------+--------------+------+-----+-------+
|productID|      date|      customer|rating|votes|helpful|
+---------+----------+--------------+------+-----+-------+
|        2|2004-02-11|A1CP26N8RHYVVO|     1|   13|      9|
|        2|2002-03-23|A3GO7UV9XX14D8|     4|    6|      6|
```

```
|        2|2002-02-06|A2P6KAWXJ16234|     4|   16|     16|
|        2|2002-01-24|A13SG9ACZ9O5IM|     5|    8|      8|
|        2|2002-05-23|A1GIL64QK68WKL|     5|    8|      8|
+---------+----------+--------------+------+-----+-------+
```

0.6 Questão B

0.6.1 Dado um produto, listar os produtos similares com maiores vendas do que ele

0.6.2 Usando SQL

[17]:
```
SELECTED_ID = 2

spark.sql("""
    SELECT *
        FROM similars s
        INNER JOIN products p ON p.asin = s.asin
        WHERE p.salesrank < (
            SELECT products.salesrank FROM products WHERE products.id = {0}
        )
            AND s.productID = {0}
        ORDER BY p.salesrank

""".format(SELECTED_ID)) \
   .show(1000)
```

```
+---------+----------+------+----------+--------------------+-----+---------+---
----+------------+------------+
|productID|      asin|    id|      asin|
title|group|salesrank|r_total|r_downloaded|r_avg_rating|
+---------+----------+------+----------+--------------------+-----+---------+---
----+------------+------------+
|        2|0738700940|299250|0738700940|              Lammas| Book|    58836|
10|          10|         4.5|
|        2|1567184960| 62291|1567184960|Yule: A Celebrati…| Book|   103012|
35|          35|         3.5|
|        2|0738700525|170507|0738700525|Midsummer: Magica…| Book|   159277|
6|           6|         4.5|
+---------+----------+------+----------+--------------------+-----+---------+---
----+------------+------------+
```

0.6.3 Usando DataFrame DSL

[18]:
```
# Projetar o salesrank do produto selecionado
minimal_salesrank = productsDF.select(productsDF.salesrank) \
                              .where(productsDF.id == SELECTED_ID) \
                              .collect()[0][0]

# Retornar produtos similares ao produto selecionado com um salesrank menor
result = similarsDF.where(similarsDF.productID == SELECTED_ID) \
                   .join(productsDF, productsDF.asin == similarsDF.asin) \
                   .where(productsDF.salesrank < minimal_salesrank) \
                   .orderBy( productsDF.salesrank, ascending=True ) \
                   .show()
```

```
+---------+----------+------+----------+--------------------+-----+---------+---
----+------------+------------+
|productID|      asin|    id|      asin|
title|group|salesrank|r_total|r_downloaded|r_avg_rating|
+---------+----------+------+----------+--------------------+-----+---------+---
----+------------+------------+
|        2|0738700940|299250|0738700940|              Lammas| Book|    58836|
10|          10|         4.5|
|        2|1567184960| 62291|1567184960|Yule: A Celebrati…| Book|   103012|
35|          35|         3.5|
|        2|0738700525|170507|0738700525|Midsummer: Magica…| Book|   159277|
6|           6|         4.5|
+---------+----------+------+----------+--------------------+-----+---------+---
----+------------+------------+
```

0.7 Questão C

0.7.1 Dado um produto, mostrar a evolução diária das médias de avaliação ao longo do intervalo de tempo coberto no arquivo de entrada

0.7.2 Usando SQL

[19]:
```
SELECTED_ID = 19367

answer_c_sql = spark.sql("""
    SELECT
        r.date as date,
        r.rating as rating,
        ( SELECT p.r_avg_rating FROM products p WHERE p.id = {0} ) as avg_rating
    FROM reviews r
```

```
    WHERE r.productID = {0}
    ORDER BY r.date
""".format(SELECTED_ID)) \
   .show(1000)

# answer_c_sql.show()
# answer_c = answer_c_sql.rdd.map(tuple).collect()
```

```
+----------+------+----------+
|      date|rating|avg_rating|
+----------+------+----------+
|1998-09-21|     5|       4.0|
|1998-10-15|     5|       4.0|
|1998-10-21|     5|       4.0|
|1998-10-23|     2|       4.0|
|1998-11-01|     5|       4.0|
|1998-11-02|     5|       4.0|
|1998-11-02|     5|       4.0|
|1998-11-02|     5|       4.0|
|1998-11-04|     5|       4.0|
|1998-11-10|     2|       4.0|
|1998-11-11|     5|       4.0|
|1998-11-24|     5|       4.0|
|1998-11-27|     2|       4.0|
|1998-12-01|     1|       4.0|
|1998-12-06|     4|       4.0|
|1998-12-11|     5|       4.0|
|1998-12-11|     5|       4.0|
|1998-12-25|     5|       4.0|
|1998-12-26|     5|       4.0|
|1998-12-27|     5|       4.0|
|1998-12-31|     5|       4.0|
|1999-01-10|     5|       4.0|
|1999-01-10|     5|       4.0|
|1999-01-12|     1|       4.0|
|1999-01-17|     5|       4.0|
|1999-01-17|     5|       4.0|
|1999-01-21|     2|       4.0|
|1999-01-23|     3|       4.0|
|1999-01-26|     5|       4.0|
|1999-01-28|     5|       4.0|
|1999-02-04|     1|       4.0|
|1999-02-07|     5|       4.0|
|1999-02-12|     5|       4.0|
|1999-02-15|     5|       4.0|
|1999-02-24|     5|       4.0|
|1999-03-24|     4|       4.0|
```

```
|1999-03-24|     5|       4.0|
|1999-03-26|     5|       4.0|
|1999-04-07|     5|       4.0|
|1999-04-10|     2|       4.0|
|1999-04-16|     5|       4.0|
|1999-04-22|     3|       4.0|
|1999-04-23|     5|       4.0|
|1999-04-24|     5|       4.0|
|1999-04-29|     5|       4.0|
|1999-05-10|     5|       4.0|
|1999-06-17|     5|       4.0|
|1999-07-05|     5|       4.0|
|1999-07-07|     4|       4.0|
|1999-07-21|     4|       4.0|
|1999-07-26|     2|       4.0|
|1999-07-30|     5|       4.0|
|1999-08-11|     3|       4.0|
|1999-08-13|     5|       4.0|
|1999-08-16|     3|       4.0|
|1999-08-17|     5|       4.0|
|1999-08-19|     5|       4.0|
|1999-08-19|     5|       4.0|
|1999-09-02|     4|       4.0|
|1999-09-10|     5|       4.0|
|1999-09-10|     1|       4.0|
|1999-09-13|     5|       4.0|
|1999-09-14|     5|       4.0|
|1999-09-20|     4|       4.0|
|1999-09-21|     5|       4.0|
|1999-09-26|     2|       4.0|
|1999-09-26|     5|       4.0|
|1999-09-26|     4|       4.0|
|1999-10-01|     5|       4.0|
|1999-10-12|     5|       4.0|
|1999-10-12|     5|       4.0|
|1999-10-14|     5|       4.0|
|1999-10-16|     4|       4.0|
|1999-10-21|     4|       4.0|
|1999-10-23|     2|       4.0|
|1999-10-28|     5|       4.0|
|1999-11-05|     3|       4.0|
|1999-11-07|     4|       4.0|
|1999-11-07|     4|       4.0|
|1999-11-08|     5|       4.0|
|1999-11-08|     5|       4.0|
|1999-11-08|     4|       4.0|
|1999-11-09|     3|       4.0|
|1999-11-09|     5|       4.0|
```

```
|1999-11-15|     5|       4.0|
|1999-11-15|     5|       4.0|
|1999-11-23|     1|       4.0|
|1999-11-26|     5|       4.0|
|1999-11-28|     3|       4.0|
|1999-12-02|     5|       4.0|
|1999-12-02|     5|       4.0|
|1999-12-12|     4|       4.0|
|1999-12-13|     5|       4.0|
|1999-12-13|     5|       4.0|
|1999-12-15|     5|       4.0|
|1999-12-16|     5|       4.0|
|1999-12-21|     5|       4.0|
|1999-12-21|     1|       4.0|
|1999-12-21|     5|       4.0|
|2000-01-21|     5|       4.0|
|2000-01-23|     1|       4.0|
|2000-01-25|     5|       4.0|
|2000-01-26|     5|       4.0|
|2000-01-30|     5|       4.0|
|2000-01-30|     4|       4.0|
|2000-01-31|     4|       4.0|
|2000-02-07|     5|       4.0|
|2000-02-10|     1|       4.0|
|2000-02-24|     2|       4.0|
|2000-02-25|     5|       4.0|
|2000-02-29|     5|       4.0|
|2000-03-05|     5|       4.0|
|2000-03-11|     5|       4.0|
|2000-03-13|     5|       4.0|
|2000-03-14|     5|       4.0|
|2000-03-18|     5|       4.0|
|2000-03-21|     5|       4.0|
|2000-03-27|     4|       4.0|
|2000-03-29|     5|       4.0|
|2000-04-02|     5|       4.0|
|2000-04-18|     4|       4.0|
|2000-04-20|     5|       4.0|
|2000-04-20|     2|       4.0|
|2000-04-20|     3|       4.0|
|2000-04-28|     5|       4.0|
|2000-05-10|     1|       4.0|
|2000-06-26|     5|       4.0|
|2000-06-27|     5|       4.0|
|2000-07-02|     1|       4.0|
|2000-07-19|     5|       4.0|
|2000-07-20|     3|       4.0|
|2000-07-22|     5|       4.0|
```

```
|2000-07-22|     3|       4.0|
|2000-08-01|     2|       4.0|
|2000-08-04|     5|       4.0|
|2000-08-04|     4|       4.0|
|2000-08-12|     5|       4.0|
|2000-08-20|     2|       4.0|
|2000-08-24|     5|       4.0|
|2000-08-30|     5|       4.0|
|2000-09-04|     3|       4.0|
|2000-09-07|     5|       4.0|
|2000-09-24|     3|       4.0|
|2000-09-27|     2|       4.0|
|2000-10-04|     4|       4.0|
|2000-10-05|     4|       4.0|
|2000-10-06|     2|       4.0|
|2000-10-07|     3|       4.0|
|2000-10-23|     5|       4.0|
|2000-10-26|     5|       4.0|
|2000-10-26|     5|       4.0|
|2000-11-01|     3|       4.0|
|2000-11-11|     1|       4.0|
|2000-11-17|     5|       4.0|
|2000-11-20|     2|       4.0|
|2000-11-29|     5|       4.0|
|2000-12-14|     1|       4.0|
|2000-12-17|     2|       4.0|
|2000-12-17|     2|       4.0|
|2000-12-18|     3|       4.0|
|2000-12-29|     3|       4.0|
|2001-01-01|     5|       4.0|
|2001-01-02|     5|       4.0|
|2001-01-04|     1|       4.0|
|2001-01-14|     2|       4.0|
|2001-01-16|     3|       4.0|
|2001-01-23|     4|       4.0|
|2001-01-24|     5|       4.0|
|2001-02-02|     4|       4.0|
|2001-02-09|     2|       4.0|
|2001-03-04|     1|       4.0|
|2001-03-05|     4|       4.0|
|2001-03-08|     5|       4.0|
|2001-03-09|     5|       4.0|
|2001-03-14|     5|       4.0|
|2001-03-16|     5|       4.0|
|2001-03-23|     2|       4.0|
|2001-03-31|     5|       4.0|
|2001-04-05|     2|       4.0|
|2001-04-11|     4|       4.0|
```

```
|2001-04-17|     5|       4.0|
|2001-04-23|     4|       4.0|
|2001-04-30|     5|       4.0|
|2001-05-01|     5|       4.0|
|2001-05-03|     3|       4.0|
|2001-05-16|     4|       4.0|
|2001-05-19|     5|       4.0|
|2001-05-20|     5|       4.0|
|2001-05-26|     5|       4.0|
|2001-05-29|     5|       4.0|
|2001-05-31|     5|       4.0|
|2001-05-31|     5|       4.0|
|2001-06-01|     4|       4.0|
|2001-06-02|     1|       4.0|
|2001-06-03|     5|       4.0|
|2001-06-09|     5|       4.0|
|2001-06-13|     5|       4.0|
|2001-06-13|     5|       4.0|
|2001-06-13|     1|       4.0|
|2001-06-16|     4|       4.0|
|2001-06-17|     5|       4.0|
|2001-06-22|     3|       4.0|
|2001-07-08|     1|       4.0|
|2001-07-09|     4|       4.0|
|2001-07-11|     5|       4.0|
|2001-07-11|     5|       4.0|
|2001-07-15|     5|       4.0|
|2001-07-17|     4|       4.0|
|2001-07-17|     5|       4.0|
|2001-07-24|     5|       4.0|
|2001-08-01|     5|       4.0|
|2001-08-06|     1|       4.0|
|2001-08-07|     5|       4.0|
|2001-08-09|     4|       4.0|
|2001-08-12|     5|       4.0|
|2001-08-15|     5|       4.0|
|2001-08-23|     1|       4.0|
|2001-08-24|     5|       4.0|
|2001-08-27|     3|       4.0|
|2001-08-27|     5|       4.0|
|2001-09-03|     1|       4.0|
|2001-09-03|     2|       4.0|
|2001-09-10|     3|       4.0|
|2001-09-20|     5|       4.0|
|2001-10-02|     2|       4.0|
|2001-10-03|     4|       4.0|
|2001-10-08|     2|       4.0|
|2001-10-20|     3|       4.0|
```

```
|2001-10-21|     5|       4.0|
|2001-10-22|     5|       4.0|
|2001-10-24|     3|       4.0|
|2001-11-05|     5|       4.0|
|2001-11-08|     5|       4.0|
|2001-11-11|     4|       4.0|
|2001-11-27|     4|       4.0|
|2001-12-10|     5|       4.0|
|2001-12-13|     5|       4.0|
|2001-12-14|     5|       4.0|
|2001-12-21|     4|       4.0|
|2001-12-23|     3|       4.0|
|2001-12-30|     5|       4.0|
|2002-01-06|     1|       4.0|
|2002-01-08|     5|       4.0|
|2002-01-09|     1|       4.0|
|2002-01-16|     5|       4.0|
|2002-01-20|     5|       4.0|
|2002-01-22|     3|       4.0|
|2002-02-01|     1|       4.0|
|2002-02-03|     5|       4.0|
|2002-02-05|     3|       4.0|
|2002-02-05|     1|       4.0|
|2002-02-23|     3|       4.0|
|2002-02-25|     1|       4.0|
|2002-03-02|     5|       4.0|
|2002-03-03|     4|       4.0|
|2002-03-08|     1|       4.0|
|2002-03-11|     5|       4.0|
|2002-03-11|     5|       4.0|
|2002-03-12|     1|       4.0|
|2002-03-26|     1|       4.0|
|2002-03-27|     5|       4.0|
|2002-03-28|     4|       4.0|
|2002-03-31|     2|       4.0|
|2002-04-10|     5|       4.0|
|2002-04-13|     3|       4.0|
|2002-04-16|     4|       4.0|
|2002-04-18|     1|       4.0|
|2002-04-23|     3|       4.0|
|2002-04-24|     5|       4.0|
|2002-04-25|     2|       4.0|
|2002-04-27|     4|       4.0|
|2002-04-28|     1|       4.0|
|2002-05-05|     5|       4.0|
|2002-05-05|     3|       4.0|
|2002-05-11|     5|       4.0|
|2002-05-17|     5|       4.0|
```

```
|2002-05-17|     5|       4.0|
|2002-05-24|     5|       4.0|
|2002-05-25|     5|       4.0|
|2002-05-28|     2|       4.0|
|2002-05-29|     4|       4.0|
|2002-07-07|     5|       4.0|
|2002-07-21|     1|       4.0|
|2002-07-29|     2|       4.0|
|2002-08-02|     5|       4.0|
|2002-08-04|     5|       4.0|
|2002-08-16|     3|       4.0|
|2002-08-17|     5|       4.0|
|2002-08-22|     5|       4.0|
|2002-08-25|     4|       4.0|
|2002-08-27|     1|       4.0|
|2002-09-04|     5|       4.0|
|2002-09-22|     3|       4.0|
|2002-09-29|     4|       4.0|
|2002-10-07|     1|       4.0|
|2002-10-14|     2|       4.0|
|2002-10-21|     5|       4.0|
|2002-10-24|     5|       4.0|
|2002-11-02|     4|       4.0|
|2002-11-06|     5|       4.0|
|2002-11-23|     4|       4.0|
|2002-12-01|     5|       4.0|
|2002-12-12|     4|       4.0|
|2002-12-14|     5|       4.0|
|2002-12-21|     1|       4.0|
|2002-12-22|     4|       4.0|
|2002-12-26|     1|       4.0|
|2002-12-27|     3|       4.0|
|2002-12-31|     5|       4.0|
|2002-12-31|     5|       4.0|
|2003-01-05|     3|       4.0|
|2003-01-08|     5|       4.0|
|2003-01-14|     1|       4.0|
|2003-01-22|     4|       4.0|
|2003-01-27|     4|       4.0|
|2003-02-07|     4|       4.0|
|2003-02-13|     4|       4.0|
|2003-02-24|     2|       4.0|
|2003-03-01|     4|       4.0|
|2003-03-20|     5|       4.0|
|2003-03-20|     5|       4.0|
|2003-04-10|     4|       4.0|
|2003-04-11|     1|       4.0|
|2003-04-14|     4|       4.0|
```

```
|2003-04-14|     4|       4.0|
|2003-04-21|     5|       4.0|
|2003-04-24|     1|       4.0|
|2003-05-05|     2|       4.0|
|2003-05-20|     5|       4.0|
|2003-05-25|     4|       4.0|
|2003-05-28|     3|       4.0|
|2003-06-12|     5|       4.0|
|2003-06-23|     1|       4.0|
|2003-06-28|     2|       4.0|
|2003-06-29|     5|       4.0|
|2003-07-06|     5|       4.0|
|2003-07-11|     5|       4.0|
|2003-07-20|     4|       4.0|
|2003-07-21|     1|       4.0|
|2003-08-07|     2|       4.0|
|2003-08-12|     1|       4.0|
|2003-08-18|     1|       4.0|
|2003-08-22|     5|       4.0|
|2003-08-31|     1|       4.0|
|2003-09-01|     2|       4.0|
|2003-09-05|     3|       4.0|
|2003-09-17|     1|       4.0|
|2003-09-17|     1|       4.0|
|2003-09-18|     2|       4.0|
|2003-09-18|     2|       4.0|
|2003-09-20|     4|       4.0|
|2003-10-09|     1|       4.0|
|2003-10-13|     4|       4.0|
|2003-10-18|     1|       4.0|
|2003-10-24|     5|       4.0|
|2003-11-08|     5|       4.0|
|2003-11-09|     5|       4.0|
|2003-11-13|     1|       4.0|
|2003-11-13|     1|       4.0|
|2003-11-14|     1|       4.0|
|2003-12-03|     2|       4.0|
|2003-12-22|     4|       4.0|
|2004-01-05|     5|       4.0|
|2004-01-06|     3|       4.0|
|2004-01-18|     4|       4.0|
|2004-01-22|     3|       4.0|
|2004-01-26|     5|       4.0|
|2004-01-31|     1|       4.0|
|2004-02-03|     5|       4.0|
|2004-02-12|     2|       4.0|
|2004-02-14|     5|       4.0|
|2004-02-23|     5|       4.0|
```

```
|2004-02-26|     4|       4.0|
|2004-03-02|     4|       4.0|
|2004-03-03|     3|       4.0|
|2004-03-13|     1|       4.0|
|2004-03-27|     5|       4.0|
|2004-04-03|     5|       4.0|
|2004-04-15|     4|       4.0|
|2004-04-28|     1|       4.0|
|2004-04-28|     5|       4.0|
|2004-05-04|     5|       4.0|
|2004-05-09|     3|       4.0|
|2004-05-16|     1|       4.0|
|2004-05-16|     1|       4.0|
|2004-05-24|     2|       4.0|
|2004-05-25|     5|       4.0|
|2004-05-26|     1|       4.0|
|2004-05-31|     3|       4.0|
|2004-06-23|     5|       4.0|
|2004-06-30|     3|       4.0|
|2004-07-02|     5|       4.0|
|2004-07-05|     5|       4.0|
|2004-07-07|     4|       4.0|
|2004-07-14|     5|       4.0|
|2004-07-22|     5|       4.0|
|2004-08-01|     4|       4.0|
|2004-08-03|     1|       4.0|
|2004-08-03|     5|       4.0|
|2004-08-14|     3|       4.0|
|2004-09-01|     2|       4.0|
|2004-09-05|     3|       4.0|
|2004-09-06|     5|       4.0|
|2004-09-14|     5|       4.0|
|2004-09-14|     4|       4.0|
|2004-09-15|     5|       4.0|
|2004-09-17|     3|       4.0|
|2004-09-18|     1|       4.0|
|2004-09-28|     3|       4.0|
|2004-10-05|     5|       4.0|
|2004-10-13|     5|       4.0|
|2004-10-19|     2|       4.0|
|2004-10-19|     5|       4.0|
|2004-11-17|     5|       4.0|
|2004-11-25|     4|       4.0|
|2004-11-30|     5|       4.0|
|2004-12-06|     1|       4.0|
|2004-12-06|     5|       4.0|
|2004-12-08|     5|       4.0|
|2004-12-19|     5|       4.0|
```

```
|2004-12-22|     1|       4.0|
|2004-12-22|     5|       4.0|
|2004-12-28|     5|       4.0|
|2004-12-31|     3|       4.0|
|2005-01-22|     2|       4.0|
|2005-01-29|     1|       4.0|
|2005-02-07|     5|       4.0|
|2005-02-26|     3|       4.0|
|2005-03-01|     5|       4.0|
|2005-03-06|     5|       4.0|
|2005-03-08|     3|       4.0|
|2005-03-12|     4|       4.0|
|2005-03-28|     4|       4.0|
|2005-04-01|     4|       4.0|
|2005-04-09|     5|       4.0|
|2005-04-19|     2|       4.0|
|2005-04-20|     2|       4.0|
|2005-04-21|     5|       4.0|
|2005-04-25|     3|       4.0|
|2005-04-29|     5|       4.0|
|2005-05-07|     5|       4.0|
|2005-05-14|     5|       4.0|
|2005-05-21|     4|       4.0|
|2005-05-29|     5|       4.0|
|2005-05-29|     5|       4.0|
|2005-06-08|     3|       4.0|
|2005-06-11|     5|       4.0|
|2005-06-12|     3|       4.0|
|2005-06-22|     1|       4.0|
+----------+------+----------+
```

0.7.3 Usando DataFrame DSL

[20]:
```
answer_c = reviewsDF.where( reviewsDF.productID == SELECTED_ID ) \
                    .join( productsDF, reviewsDF.productID == productsDF.id ) \
                    .select( reviewsDF.date, reviewsDF.rating, productsDF.
↪r_avg_rating ) \
                    .show(1000)
```

```
+----------+------+------------+
|      date|rating|r_avg_rating|
+----------+------+------------+
|1998-09-21|     5|         4.0|
|1998-10-15|     5|         4.0|
|1998-10-21|     5|         4.0|
|1998-10-23|     2|         4.0|
```

```
|1998-11-01|     5|         4.0|
|1998-11-02|     5|         4.0|
|1998-11-02|     5|         4.0|
|1998-11-02|     5|         4.0|
|1998-11-04|     5|         4.0|
|1998-11-10|     2|         4.0|
|1998-11-11|     5|         4.0|
|1998-11-24|     5|         4.0|
|1998-11-27|     2|         4.0|
|1998-12-01|     1|         4.0|
|1998-12-06|     4|         4.0|
|1998-12-11|     5|         4.0|
|1998-12-11|     5|         4.0|
|1998-12-25|     5|         4.0|
|1998-12-26|     5|         4.0|
|1998-12-27|     5|         4.0|
|1998-12-31|     5|         4.0|
|1999-01-10|     5|         4.0|
|1999-01-10|     5|         4.0|
|1999-01-12|     1|         4.0|
|1999-01-17|     5|         4.0|
|1999-01-17|     5|         4.0|
|1999-01-21|     2|         4.0|
|1999-01-23|     3|         4.0|
|1999-01-26|     5|         4.0|
|1999-01-28|     5|         4.0|
|1999-02-04|     1|         4.0|
|1999-02-07|     5|         4.0|
|1999-02-12|     5|         4.0|
|1999-02-15|     5|         4.0|
|1999-02-24|     5|         4.0|
|1999-03-24|     4|         4.0|
|1999-03-24|     5|         4.0|
|1999-03-26|     5|         4.0|
|1999-04-07|     5|         4.0|
|1999-04-10|     2|         4.0|
|1999-04-16|     5|         4.0|
|1999-04-22|     3|         4.0|
|1999-04-23|     5|         4.0|
|1999-04-24|     5|         4.0|
|1999-04-29|     5|         4.0|
|1999-05-10|     5|         4.0|
|1999-06-17|     5|         4.0|
|1999-07-05|     5|         4.0|
|1999-07-07|     4|         4.0|
|1999-07-21|     4|         4.0|
|1999-07-26|     2|         4.0|
|1999-07-30|     5|         4.0|
```

```
|1999-08-11|     3|         4.0|
|1999-08-13|     5|         4.0|
|1999-08-16|     3|         4.0|
|1999-08-17|     5|         4.0|
|1999-08-19|     5|         4.0|
|1999-08-19|     5|         4.0|
|1999-09-02|     4|         4.0|
|1999-09-10|     5|         4.0|
|1999-09-10|     1|         4.0|
|1999-09-13|     5|         4.0|
|1999-09-14|     5|         4.0|
|1999-09-20|     4|         4.0|
|1999-09-21|     5|         4.0|
|1999-09-26|     2|         4.0|
|1999-09-26|     5|         4.0|
|1999-09-26|     4|         4.0|
|1999-10-01|     5|         4.0|
|1999-10-12|     5|         4.0|
|1999-10-12|     5|         4.0|
|1999-10-14|     5|         4.0|
|1999-10-16|     4|         4.0|
|1999-10-21|     4|         4.0|
|1999-10-23|     2|         4.0|
|1999-10-28|     5|         4.0|
|1999-11-05|     3|         4.0|
|1999-11-07|     4|         4.0|
|1999-11-07|     4|         4.0|
|1999-11-08|     5|         4.0|
|1999-11-08|     4|         4.0|
|1999-11-08|     5|         4.0|
|1999-11-09|     3|         4.0|
|1999-11-09|     5|         4.0|
|1999-11-15|     5|         4.0|
|1999-11-15|     5|         4.0|
|1999-11-23|     1|         4.0|
|1999-11-26|     5|         4.0|
|1999-11-28|     3|         4.0|
|1999-12-02|     5|         4.0|
|1999-12-02|     5|         4.0|
|1999-12-12|     4|         4.0|
|1999-12-13|     5|         4.0|
|1999-12-13|     5|         4.0|
|1999-12-15|     5|         4.0|
|1999-12-16|     5|         4.0|
|1999-12-21|     5|         4.0|
|1999-12-21|     1|         4.0|
|1999-12-21|     5|         4.0|
|2000-01-21|     5|         4.0|
```

```
|2000-01-23|     1|         4.0|
|2000-01-25|     5|         4.0|
|2000-01-26|     5|         4.0|
|2000-01-30|     4|         4.0|
|2000-01-30|     5|         4.0|
|2000-01-31|     4|         4.0|
|2000-02-07|     5|         4.0|
|2000-02-10|     1|         4.0|
|2000-02-24|     2|         4.0|
|2000-02-25|     5|         4.0|
|2000-02-29|     5|         4.0|
|2000-03-05|     5|         4.0|
|2000-03-11|     5|         4.0|
|2000-03-13|     5|         4.0|
|2000-03-14|     5|         4.0|
|2000-03-18|     5|         4.0|
|2000-03-21|     5|         4.0|
|2000-03-27|     4|         4.0|
|2000-03-29|     5|         4.0|
|2000-04-02|     5|         4.0|
|2000-04-18|     4|         4.0|
|2000-04-20|     2|         4.0|
|2000-04-20|     5|         4.0|
|2000-04-20|     3|         4.0|
|2000-04-28|     5|         4.0|
|2000-05-10|     1|         4.0|
|2000-06-26|     5|         4.0|
|2000-06-27|     5|         4.0|
|2000-07-02|     1|         4.0|
|2000-07-19|     5|         4.0|
|2000-07-20|     3|         4.0|
|2000-07-22|     5|         4.0|
|2000-07-22|     3|         4.0|
|2000-08-01|     2|         4.0|
|2000-08-04|     4|         4.0|
|2000-08-04|     5|         4.0|
|2000-08-12|     5|         4.0|
|2000-08-20|     2|         4.0|
|2000-08-24|     5|         4.0|
|2000-08-30|     5|         4.0|
|2000-09-04|     3|         4.0|
|2000-09-07|     5|         4.0|
|2000-09-24|     3|         4.0|
|2000-09-27|     2|         4.0|
|2000-10-04|     4|         4.0|
|2000-10-05|     4|         4.0|
|2000-10-06|     2|         4.0|
|2000-10-07|     3|         4.0|
```

```
|2000-10-23|     5|         4.0|
|2000-10-26|     5|         4.0|
|2000-10-26|     5|         4.0|
|2000-11-01|     3|         4.0|
|2000-11-11|     1|         4.0|
|2000-11-17|     5|         4.0|
|2000-11-20|     2|         4.0|
|2000-11-29|     5|         4.0|
|2000-12-14|     1|         4.0|
|2000-12-17|     2|         4.0|
|2000-12-17|     2|         4.0|
|2000-12-18|     3|         4.0|
|2000-12-29|     3|         4.0|
|2001-01-01|     5|         4.0|
|2001-01-02|     5|         4.0|
|2001-01-04|     1|         4.0|
|2001-01-14|     2|         4.0|
|2001-01-16|     3|         4.0|
|2001-01-23|     4|         4.0|
|2001-01-24|     5|         4.0|
|2001-02-02|     4|         4.0|
|2001-02-09|     2|         4.0|
|2001-03-04|     1|         4.0|
|2001-03-05|     4|         4.0|
|2001-03-08|     5|         4.0|
|2001-03-09|     5|         4.0|
|2001-03-14|     5|         4.0|
|2001-03-16|     5|         4.0|
|2001-03-23|     2|         4.0|
|2001-03-31|     5|         4.0|
|2001-04-05|     2|         4.0|
|2001-04-11|     4|         4.0|
|2001-04-17|     5|         4.0|
|2001-04-23|     4|         4.0|
|2001-04-30|     5|         4.0|
|2001-05-01|     5|         4.0|
|2001-05-03|     3|         4.0|
|2001-05-16|     4|         4.0|
|2001-05-19|     5|         4.0|
|2001-05-20|     5|         4.0|
|2001-05-26|     5|         4.0|
|2001-05-29|     5|         4.0|
|2001-05-31|     5|         4.0|
|2001-05-31|     5|         4.0|
|2001-06-01|     4|         4.0|
|2001-06-02|     1|         4.0|
|2001-06-03|     5|         4.0|
|2001-06-09|     5|         4.0|
```

```
|2001-06-13|     1|         4.0|
|2001-06-13|     5|         4.0|
|2001-06-13|     5|         4.0|
|2001-06-16|     4|         4.0|
|2001-06-17|     5|         4.0|
|2001-06-22|     3|         4.0|
|2001-07-08|     1|         4.0|
|2001-07-09|     4|         4.0|
|2001-07-11|     5|         4.0|
|2001-07-11|     5|         4.0|
|2001-07-15|     5|         4.0|
|2001-07-17|     4|         4.0|
|2001-07-17|     5|         4.0|
|2001-07-24|     5|         4.0|
|2001-08-01|     5|         4.0|
|2001-08-06|     1|         4.0|
|2001-08-07|     5|         4.0|
|2001-08-09|     4|         4.0|
|2001-08-12|     5|         4.0|
|2001-08-15|     5|         4.0|
|2001-08-23|     1|         4.0|
|2001-08-24|     5|         4.0|
|2001-08-27|     3|         4.0|
|2001-08-27|     5|         4.0|
|2001-09-03|     2|         4.0|
|2001-09-03|     1|         4.0|
|2001-09-10|     3|         4.0|
|2001-09-20|     5|         4.0|
|2001-10-02|     2|         4.0|
|2001-10-03|     4|         4.0|
|2001-10-08|     2|         4.0|
|2001-10-20|     3|         4.0|
|2001-10-21|     5|         4.0|
|2001-10-22|     5|         4.0|
|2001-10-24|     3|         4.0|
|2001-11-05|     5|         4.0|
|2001-11-08|     5|         4.0|
|2001-11-11|     4|         4.0|
|2001-11-27|     4|         4.0|
|2001-12-10|     5|         4.0|
|2001-12-13|     5|         4.0|
|2001-12-14|     5|         4.0|
|2001-12-21|     4|         4.0|
|2001-12-23|     3|         4.0|
|2001-12-30|     5|         4.0|
|2002-01-06|     1|         4.0|
|2002-01-08|     5|         4.0|
|2002-01-09|     1|         4.0|
```

```
|2002-01-16|     5|         4.0|
|2002-01-20|     5|         4.0|
|2002-01-22|     3|         4.0|
|2002-02-01|     1|         4.0|
|2002-02-03|     5|         4.0|
|2002-02-05|     3|         4.0|
|2002-02-05|     1|         4.0|
|2002-02-23|     3|         4.0|
|2002-02-25|     1|         4.0|
|2002-03-02|     5|         4.0|
|2002-03-03|     4|         4.0|
|2002-03-08|     1|         4.0|
|2002-03-11|     5|         4.0|
|2002-03-11|     5|         4.0|
|2002-03-12|     1|         4.0|
|2002-03-26|     1|         4.0|
|2002-03-27|     5|         4.0|
|2002-03-28|     4|         4.0|
|2002-03-31|     2|         4.0|
|2002-04-10|     5|         4.0|
|2002-04-13|     3|         4.0|
|2002-04-16|     4|         4.0|
|2002-04-18|     1|         4.0|
|2002-04-23|     3|         4.0|
|2002-04-24|     5|         4.0|
|2002-04-25|     2|         4.0|
|2002-04-27|     4|         4.0|
|2002-04-28|     1|         4.0|
|2002-05-05|     5|         4.0|
|2002-05-05|     3|         4.0|
|2002-05-11|     5|         4.0|
|2002-05-17|     5|         4.0|
|2002-05-17|     5|         4.0|
|2002-05-24|     5|         4.0|
|2002-05-25|     5|         4.0|
|2002-05-28|     2|         4.0|
|2002-05-29|     4|         4.0|
|2002-07-07|     5|         4.0|
|2002-07-21|     1|         4.0|
|2002-07-29|     2|         4.0|
|2002-08-02|     5|         4.0|
|2002-08-04|     5|         4.0|
|2002-08-16|     3|         4.0|
|2002-08-17|     5|         4.0|
|2002-08-22|     5|         4.0|
|2002-08-25|     4|         4.0|
|2002-08-27|     1|         4.0|
|2002-09-04|     5|         4.0|
```

```
|2002-09-22|     3|         4.0|
|2002-09-29|     4|         4.0|
|2002-10-07|     1|         4.0|
|2002-10-14|     2|         4.0|
|2002-10-21|     5|         4.0|
|2002-10-24|     5|         4.0|
|2002-11-02|     4|         4.0|
|2002-11-06|     5|         4.0|
|2002-11-23|     4|         4.0|
|2002-12-01|     5|         4.0|
|2002-12-12|     4|         4.0|
|2002-12-14|     5|         4.0|
|2002-12-21|     1|         4.0|
|2002-12-22|     4|         4.0|
|2002-12-26|     1|         4.0|
|2002-12-27|     3|         4.0|
|2002-12-31|     5|         4.0|
|2002-12-31|     5|         4.0|
|2003-01-05|     3|         4.0|
|2003-01-08|     5|         4.0|
|2003-01-14|     1|         4.0|
|2003-01-22|     4|         4.0|
|2003-01-27|     4|         4.0|
|2003-02-07|     4|         4.0|
|2003-02-13|     4|         4.0|
|2003-02-24|     2|         4.0|
|2003-03-01|     4|         4.0|
|2003-03-20|     5|         4.0|
|2003-03-20|     5|         4.0|
|2003-04-10|     4|         4.0|
|2003-04-11|     1|         4.0|
|2003-04-14|     4|         4.0|
|2003-04-14|     4|         4.0|
|2003-04-21|     5|         4.0|
|2003-04-24|     1|         4.0|
|2003-05-05|     2|         4.0|
|2003-05-20|     5|         4.0|
|2003-05-25|     4|         4.0|
|2003-05-28|     3|         4.0|
|2003-06-12|     5|         4.0|
|2003-06-23|     1|         4.0|
|2003-06-28|     2|         4.0|
|2003-06-29|     5|         4.0|
|2003-07-06|     5|         4.0|
|2003-07-11|     5|         4.0|
|2003-07-20|     4|         4.0|
|2003-07-21|     1|         4.0|
|2003-08-07|     2|         4.0|
```

```
|2003-08-12|     1|         4.0|
|2003-08-18|     1|         4.0|
|2003-08-22|     5|         4.0|
|2003-08-31|     1|         4.0|
|2003-09-01|     2|         4.0|
|2003-09-05|     3|         4.0|
|2003-09-17|     1|         4.0|
|2003-09-17|     1|         4.0|
|2003-09-18|     2|         4.0|
|2003-09-18|     2|         4.0|
|2003-09-20|     4|         4.0|
|2003-10-09|     1|         4.0|
|2003-10-13|     4|         4.0|
|2003-10-18|     1|         4.0|
|2003-10-24|     5|         4.0|
|2003-11-08|     5|         4.0|
|2003-11-09|     5|         4.0|
|2003-11-13|     1|         4.0|
|2003-11-13|     1|         4.0|
|2003-11-14|     1|         4.0|
|2003-12-03|     2|         4.0|
|2003-12-22|     4|         4.0|
|2004-01-05|     5|         4.0|
|2004-01-06|     3|         4.0|
|2004-01-18|     4|         4.0|
|2004-01-22|     3|         4.0|
|2004-01-26|     5|         4.0|
|2004-01-31|     1|         4.0|
|2004-02-03|     5|         4.0|
|2004-02-12|     2|         4.0|
|2004-02-14|     5|         4.0|
|2004-02-23|     5|         4.0|
|2004-02-26|     4|         4.0|
|2004-03-02|     4|         4.0|
|2004-03-03|     3|         4.0|
|2004-03-13|     1|         4.0|
|2004-03-27|     5|         4.0|
|2004-04-03|     5|         4.0|
|2004-04-15|     4|         4.0|
|2004-04-28|     1|         4.0|
|2004-04-28|     5|         4.0|
|2004-05-04|     5|         4.0|
|2004-05-09|     3|         4.0|
|2004-05-16|     1|         4.0|
|2004-05-16|     1|         4.0|
|2004-05-24|     2|         4.0|
|2004-05-25|     5|         4.0|
|2004-05-26|     1|         4.0|
```

```
|2004-05-31|     3|         4.0|
|2004-06-23|     5|         4.0|
|2004-06-30|     3|         4.0|
|2004-07-02|     5|         4.0|
|2004-07-05|     5|         4.0|
|2004-07-07|     4|         4.0|
|2004-07-14|     5|         4.0|
|2004-07-22|     5|         4.0|
|2004-08-01|     4|         4.0|
|2004-08-03|     5|         4.0|
|2004-08-03|     1|         4.0|
|2004-08-14|     3|         4.0|
|2004-09-01|     2|         4.0|
|2004-09-05|     3|         4.0|
|2004-09-06|     5|         4.0|
|2004-09-14|     5|         4.0|
|2004-09-14|     4|         4.0|
|2004-09-15|     5|         4.0|
|2004-09-17|     3|         4.0|
|2004-09-18|     1|         4.0|
|2004-09-28|     3|         4.0|
|2004-10-05|     5|         4.0|
|2004-10-13|     5|         4.0|
|2004-10-19|     5|         4.0|
|2004-10-19|     2|         4.0|
|2004-11-17|     5|         4.0|
|2004-11-25|     4|         4.0|
|2004-11-30|     5|         4.0|
|2004-12-06|     1|         4.0|
|2004-12-06|     5|         4.0|
|2004-12-08|     5|         4.0|
|2004-12-19|     5|         4.0|
|2004-12-22|     5|         4.0|
|2004-12-22|     1|         4.0|
|2004-12-28|     5|         4.0|
|2004-12-31|     3|         4.0|
|2005-01-22|     2|         4.0|
|2005-01-29|     1|         4.0|
|2005-02-07|     5|         4.0|
|2005-02-26|     3|         4.0|
|2005-03-01|     5|         4.0|
|2005-03-06|     5|         4.0|
|2005-03-08|     3|         4.0|
|2005-03-12|     4|         4.0|
|2005-03-28|     4|         4.0|
|2005-04-01|     4|         4.0|
|2005-04-09|     5|         4.0|
|2005-04-19|     2|         4.0|
```

```
|2005-04-20|     2|         4.0|
|2005-04-21|     5|         4.0|
|2005-04-25|     3|         4.0|
|2005-04-29|     5|         4.0|
|2005-05-07|     5|         4.0|
|2005-05-14|     5|         4.0|
|2005-05-21|     4|         4.0|
|2005-05-29|     5|         4.0|
|2005-05-29|     5|         4.0|
|2005-06-08|     3|         4.0|
|2005-06-11|     5|         4.0|
|2005-06-12|     3|         4.0|
|2005-06-22|     1|         4.0|
+----------+------+------------+
```

0.8 Questão D

0.8.1 Listar os 10 produtos lideres de venda em cada grupo de produto

0.8.2 Usando SQL

[30]:
```
answer_d = spark.sql("""
    SELECT rs.title, rs.group, rs.salesrank
        FROM (
            SELECT title, group, salesrank, ROW_NUMBER() over( Partition BY p.
 ↪group ORDER BY p.salesrank ASC ) as Rank
                FROM products p
                WHERE salesrank != -1
        ) rs WHERE Rank <= 10
""").show(1000)
```

```
+--------------------+------------+---------+
|               title|       group|salesrank|
+--------------------+------------+---------+
|   From Soup to Nuts|       Video|        0|
|The War of the Wo…|       Video|        1|
|   Shirley Valentine|       Video|        2|
|Leslie Sansone - …|       Video|        6|
|Robin Hood - Men …|       Video|        7|
|Richard Simmons -…|       Video|        8|
|     Howard the Duck|       Video|       12|
|     Charlotte's Web|       Video|       14|
|A Tree Grows in B…|       Video|       16|
|  My Neighbor Totoro|       Video|       17|
|  IlluStory Book Kit|         Toy|       59|
|Wizard Card Game …|         Toy|     1890|
```

|Photostory Junior…| Toy| 2288|
|Party Tyme Karaok…| Toy| 4053|
|Party Tyme Karaok…| Toy| 7812|
|Party Tyme Karaok…| Toy| 10732|
|The Songs of Brit…| Toy| 31296|
|R- Photostory Senior| Toy| 45241|
| The Drifter| DVD| 0|
|The House Of More…| DVD| 0|
|1, 2, 3 Soleils: …| DVD| 0|
|Star Wars - Episo…| DVD| 28|
| Band of Brothers| DVD| 47|
|The Little Mermai…| DVD| 49|
| The Wizard of Oz| DVD| 55|
|Fawlty Towers - T…| DVD| 85|
|Star Wars - Episo…| DVD| 85|
|Jerry Seinfeld Li…| DVD| 88|
|Yoga Kit Living Arts| Sports| 4684|
|Baby'S Record Kee…|Baby Product| 1017|
|PRIMA PUBLISHING …| Video Games| 339|
|How to Start and …| Book| 0|
|Common Sense Abou…| Book| 0|
|How To Get The Be…| Book| 0|
|Help Me Talk Righ…| Book| 0|
|Gods on Earth (Th…| Book| 0|
| El arte de amar| Book| 0|
|The Manager's Bib…| Book| 0|
| Dona Barbara| Book| 0|
|The Irish America…| Book| 0|
|The Diligent: A V…| Book| 0|
|Improvisations - …| Music| 0|
| Lucky Man| Music| 0|
| I Need Your Loving| Music| 0|
|That Travelin' Tw…| Music| 0|
| Buzz Buzz| Music| 27|
|A Rush of Blood t…| Music| 33|
Michael Bublé	Music	42
Come Away with Me	Music	46
Songs About Jane	Music	53
Facing Future	Music	55
ClickArt Christia…	Software	200
RINGDISC Wagner: …	Software	327
Zondervan Bible S…	Software	1955
Just Enough Vocal…	Software	3771
WINDOWS NT SERVER…	Software	3828
SPELLING CORRECTOR	CE	39367
Hal Leonard Begin…	CE	69089
Hal Leonard Begin…	CE	71678
FRANKLIN COMP. KJ…	CE	84976

```
+--------------------+-----------+---------+
```

0.8.3 Usando DataFrame DSL

[32]:
```
productsDF.where( productsDF.salesrank != -1 ) \
          .withColumn('row_number', F.row_number().over(Window.
 ↪partitionBy('group').orderBy('salesrank'))) \
          .select('title', 'group', 'salesrank') \
          .where('row_number <= 10') \
          .show(1000)
```

```
+--------------------+-----------+---------+
|               title|      group|salesrank|
+--------------------+-----------+---------+
|   From Soup to Nuts|      Video|        0|
|The War of the Wo…|      Video|        1|
|   Shirley Valentine|      Video|        2|
|Leslie Sansone - …|      Video|        6|
|Robin Hood - Men …|      Video|        7|
|Richard Simmons -…|      Video|        8|
|     Howard the Duck|      Video|       12|
|     Charlotte's Web|      Video|       14|
|A Tree Grows in B…|      Video|       16|
|  My Neighbor Totoro|      Video|       17|
|  IlluStory Book Kit|        Toy|       59|
|Wizard Card Game …|        Toy|     1890|
|Photostory Junior…|        Toy|     2288|
|Party Tyme Karaok…|        Toy|     4053|
|Party Tyme Karaok…|        Toy|     7812|
|Party Tyme Karaok…|        Toy|    10732|
|The Songs of Brit…|        Toy|    31296|
|R- Photostory Senior|        Toy|    45241|
|         The Drifter|        DVD|        0|
|The House Of More…|        DVD|        0|
|1, 2, 3 Soleils: …|        DVD|        0|
|Star Wars - Episo…|        DVD|       28|
|    Band of Brothers|        DVD|       47|
|The Little Mermai…|        DVD|       49|
|    The Wizard of Oz|        DVD|       55|
|Fawlty Towers - T…|        DVD|       85|
|Star Wars - Episo…|        DVD|       85|
|Jerry Seinfeld Li…|        DVD|       88|
|Yoga Kit Living Arts|     Sports|     4684|
|Baby'S Record Kee…|Baby Product|     1017|
|PRIMA PUBLISHING …| Video Games|      339|
|How to Start and …|       Book|        0|
```

```
|Common Sense Abou…|        Book|        0|
|How To Get The Be…|        Book|        0|
|Help Me Talk Righ…|        Book|        0|
|Gods on Earth (Th…|        Book|        0|
|     El arte de amar|        Book|        0|
|The Manager's Bib…|        Book|        0|
|        Dona Barbara|        Book|        0|
|The Irish America…|        Book|        0|
|The Diligent: A V…|        Book|        0|
|Improvisations - …|       Music|        0|
|           Lucky Man|       Music|        0|
|  I Need Your Loving|       Music|        0|
|That Travelin' Tw…|       Music|        0|
|           Buzz Buzz|       Music|       27|
|A Rush of Blood t…|       Music|       33|
|       Michael Bublé|       Music|       42|
|   Come Away with Me|       Music|       46|
|    Songs About Jane|       Music|       53|
|       Facing Future|       Music|       55|
|ClickArt Christia…|    Software|      200|
|RINGDISC Wagner: …|    Software|      327|
|Zondervan Bible S…|    Software|     1955|
|Just Enough Vocal…|    Software|     3771|
|WINDOWS NT SERVER…|    Software|     3828|
|  SPELLING CORRECTOR|          CE|    39367|
|Hal Leonard Begin…|          CE|    69089|
|Hal Leonard Begin…|          CE|    71678|
|FRANKLIN COMP. KJ…|          CE|    84976|
+--------------------+------------+---------+
```

0.9 Questão E

0.9.1 Listar os 10 produtos com a maior média de avaliações úteis positivas

0.9.2 Usando SQL

[24]:
```
spark.sql("""
    SELECT p.title, r.productID, AVG(r.helpful) as helpful_mean
        FROM reviews r
        INNER JOIN products p ON p.id = r.productID
        WHERE r.rating >= 5
        GROUP BY productID, p.title
        ORDER BY helpful_mean DESC
        LIMIT 10
""").show(1000)
```

```
+--------------------+---------+------------------+
|               title|productID|      helpful_mean|
+--------------------+---------+------------------+
|Easy Adult Piano …|   320534|             320.0|
|Small Engine Repa…|    81098|             247.0|
|T'ai Chi for Olde…|   110544|             233.0|
|The Story About Ping|   312621|231.85294117647058|
|The Story About P…|   411581| 231.7058823529412|
|The Story about P…|   403662| 231.6764705882353|
|The Story about Ping|   357193|231.64705882352942|
|The Glucose Revol…|   537547|             206.0|
|More Than Just Ho…|   523037|             203.0|
|Abbott & Costello…|   243584|             197.0|
+--------------------+---------+------------------+
```

0.9.3 Usando DataFrame DSL

[25]:
```
reviewsDF.where( reviewsDF.rating >= 5 ) \
         .join( productsDF, productsDF.id == reviewsDF.productID ) \
         .select( reviewsDF.productID, productsDF.title, reviewsDF.helpful ) \
         .groupBy( reviewsDF.productID, productsDF.title ) \
         .agg({'helpful':'avg'}) \
         .orderBy('avg(helpful)', ascending=False) \
         .limit(10) \
         .show(1000)
```

```
+---------+--------------------+------------------+
|productID|               title|      avg(helpful)|
+---------+--------------------+------------------+
|   320534|Easy Adult Piano …|             320.0|
|    81098|Small Engine Repa…|             247.0|
|   110544|T'ai Chi for Olde…|             233.0|
|   312621|The Story About Ping|231.85294117647058|
|   411581|The Story About P…| 231.7058823529412|
|   403662|The Story about P…| 231.6764705882353|
|   357193|The Story about Ping|231.64705882352942|
|   537547|The Glucose Revol…|             206.0|
|   523037|More Than Just Ho…|             203.0|
|   243584|Abbott & Costello…|             197.0|
+---------+--------------------+------------------+
```

0.10 Questão F

0.10.1 Listar as 5 categorias de produto com a maior média de avaliações úteis positivas por produto

0.10.2 Usando SQL

[33]:
```
spark.sql("""
    SELECT c.category, AVG(r.helpful) as helpful_mean
        FROM reviews r
        INNER JOIN products p ON p.id = r.productID
        INNER JOIN categories c ON p.id = c.productID
        WHERE r.rating >= 5
        GROUP BY c.category
        ORDER BY helpful_mean DESC
        LIMIT 5
""").show(1000)
```

```
+--------------------+----------------+
|            category|    helpful_mean|
+--------------------+----------------+
|Dominican Republi…|97.7872340425532|
|Casual Users[502030]|            84.5|
|Casual Users[727732]|            84.5|
|Jack Benny Progra…|            80.0|
|Adventures of Ozz…|            80.0|
+--------------------+----------------+
```

0.10.3 Usando DataFrame DSL

[34]:
```
reviewsDF.where( reviewsDF.rating >= 5 ) \
         .join( productsDF, productsDF.id == reviewsDF.productID ) \
         .join( categoriesDF, productsDF.id == categoriesDF.productID ) \
         .select( categoriesDF.category, reviewsDF.helpful ) \
         .groupBy( categoriesDF.category ) \
         .agg({'helpful':'avg'}) \
         .orderBy('avg(helpful)', ascending=False) \
         .limit(5) \
         .show(1000)
```

```
+--------------------+----------------+
|            category|    avg(helpful)|
+--------------------+----------------+
|Dominican Republi…|97.7872340425532|
|Casual Users[502030]|            84.5|
```

```
|Casual Users[727732]|            84.5|
|Jack Benny Progra…|            80.0|
|Adventures of Ozz…|            80.0|
+--------------------+----------------+
```

0.11 Questão G

0.11.1 Listar os 10 clientes que mais fizeram comentários por grupo de produtos

0.11.2 Usando SQL

[28]:
```
answer_d = spark.sql("""
    SELECT rs.group, rs.customer, rs.comments
        FROM (
            SELECT
                abc.group,
                abc.customer,
                abc.comments,
                ROW_NUMBER() over( Partition BY abc.group ORDER BY abc.comments␣
 ↪DESC ) as Rank
            FROM (
                SELECT
                    p.group,
                    r.customer,
                    COUNT(*) as comments
                FROM products p
                INNER JOIN reviews r ON p.id = r.productID
                GROUP BY p.group, r.customer
            ) abc
        ) rs WHERE Rank <= 10
""").show(1000)
```

```
+-----------+--------------+--------+
|      group|      customer|comments|
+-----------+--------------+--------+
|      Video| ATVPDKIKX0DER|   72581|
|      Video|A3UN6WX5RRO2AG|   15814|
|      Video|A2NJO6YE954DBH|    1775|
|      Video| AU8552YCOO5QX|    1205|
|      Video|A3P1A63Q8L32C5|     737|
|      Video|A20EEWWSFMZ1PN|     720|
|      Video|A16CZRQL23NOIW|     668|
|      Video|A3LZGLA88K0LA0|     614|
|      Video|A2QRB6L1MCJ53G|     606|
|      Video|A152C8GYY25HAH|     583|
|        Toy| AH4M07U4YC695|       2|
```

```
|          Toy| ATVPDKIKX0DER|       2|
|          Toy|A1SB7SB31ETYZH|       2|
|          Toy|A1A2RJXP6T26PD|       1|
|          Toy|A2YO9AKVAHDR9I|       1|
|          Toy|A2PJ7WLZ38F47S|       1|
|          Toy|A20AL96IIDAEBU|       1|
|          Toy|A20R67CABJH79P|       1|
|          Toy| AU8PR9XJ17CCB|       1|
|          Toy| AQJ2XVMHXGN9A|       1|
|          DVD| ATVPDKIKX0DER|   63148|
|          DVD|A3UN6WX5RRO2AG|   15549|
|          DVD|A2NJO6YE954DBH|    1366|
|          DVD| AU8552YCOO5QX|    1213|
|          DVD|A3P1A63Q8L32C5|     859|
|          DVD|A3LZGLA88K0LA0|     856|
|          DVD| A82LIVYSX6WZ9|     683|
|          DVD|A152C8GYY25HAH|     675|
|          DVD|A16CZRQL23NOIW|     651|
|          DVD|A1CZICCYP2M5PX|     650|
|       Sports|A2RHSQZ7MAKKCO|       1|
|       Sports|A1W180Y9O1FALI|       1|
|       Sports|A3O8EZOX2P399L|       1|
|       Sports|A18ZVYTEDAOF9A|       1|
|       Sports| AL62LOJKDES3M|       1|
|Baby Product|A37TFIP0OMKGMW|       1|
|Baby Product|A2LAH8VX720175|       1|
|Baby Product| AI9SB5VKUFXDC|       1|
| Video Games|A3C811U31YG6FS|       1|
| Video Games|A226EDS7WDF7S1|       1|
| Video Games|A1M4NJYP0WNL8Q|       1|
|         Book| ATVPDKIKX0DER|  643185|
|         Book|A3UN6WX5RRO2AG|  154531|
|         Book|A14OJS0VWMOSWO|    9589|
|         Book|   AFVQZQ8PW0L|    5441|
|         Book|A1K1JW1C5CUSUZ|    3562|
|         Book|A2NJO6YE954DBH|    2055|
|         Book|A3QVAKVRAH657N|    1651|
|         Book|A1NATT3PN24QWY|    1535|
|         Book|A1D2C0WDCSHUWZ|    1508|
|         Book|A2ODBHT4URXVXQ|    1469|
|        Music| ATVPDKIKX0DER|  166149|
|        Music|A3UN6WX5RRO2AG|   15875|
|        Music| A9Q28YTLYREO7|    2760|
|        Music|A2U49LUUY4IKQQ|    1258|
|        Music|A1GN8UJIZLCA59|    1154|
|        Music|A2NJO6YE954DBH|    1128|
|        Music|A1J5KCZC8CMW9I|    1031|
|        Music|A3MOF5KF93Q6WE|     989|
```

```
|       Music| AXFI7TAWD6H6X|     814|
|       Music|A38U2M9OAEJAXJ|     780|
|    Software| AK9MWTH6LJF64|       1|
|    Software|A39UZ9VVRJW4P8|       1|
|    Software|A1T6PXM2M3N84A|       1|
|    Software|A1F8RIBFWRYM3Y|       1|
|    Software|A23DFB8IUTIZM0|       1|
|    Software|A3D5ICIQ8STPCH|       1|
|    Software|A36T3O4TIC1YDQ|       1|
|    Software|A1XQB8IU7S8WEU|       1|
|    Software|A2IOZWBVRO575O|       1|
|    Software| A37UFPGDSSMEV|       1|
|          CE|A1SFX3CR838F36|       1|
|          CE|A2IX9TMXDBUCYV|       1|
|          CE|A1J62O1S6QTHZJ|       1|
|          CE|A13JU90C7AU3RT|       1|
|          CE|A1328SYT22GA4U|       1|
+------------+--------------+--------+
```

0.11.3 Usando DataFrame DSL

[29]:
```
productsDF.join( reviewsDF, productsDF.id == reviewsDF.productID ) \
          .select( productsDF.group, reviewsDF.customer ) \
          .groupBy( reviewsDF.customer, productsDF.group ) \
          .agg({'customer':'count'}) \
          .withColumn('row_number', F.row_number().over(Window.
 ↪partitionBy('group').orderBy(desc('count(customer)')))) \
          .where('row_number <= 10') \
          .select( productsDF.group, reviewsDF.customer, 'count(customer)' ) \
          .show(1000)
```

```
+------------+--------------+---------------+
|       group|      customer|count(customer)|
+------------+--------------+---------------+
|       Video| ATVPDKIKX0DER|          72581|
|       Video|A3UN6WX5RRO2AG|          15814|
|       Video|A2NJO6YE954DBH|           1775|
|       Video| AU8552YCOO5QX|           1205|
|       Video|A3P1A63Q8L32C5|            737|
|       Video|A20EEWWSFMZ1PN|            720|
|       Video|A16CZRQL23NOIW|            668|
|       Video|A3LZGLA88K0LA0|            614|
|       Video|A2QRB6L1MCJ53G|            606|
|       Video|A152C8GYY25HAH|            583|
|         Toy| AH4M07U4YC695|              2|
|         Toy|A1SB7SB31ETYZH|              2|
```

```
|          Toy| ATVPDKIKX0DER|             2|
|          Toy| AQJ2XVMHXGN9A|             1|
|          Toy| AH16IHWEMA61J|             1|
|          Toy|A1LEF9EM2DFDP2|             1|
|          Toy|A1ABBKXKUZF85X|             1|
|          Toy|A1FPVUL053AKXO|             1|
|          Toy|A3O19HBWE10FFJ|             1|
|          Toy|A2OAL96IIDAEBU|             1|
|          DVD| ATVPDKIKX0DER|         63148|
|          DVD|A3UN6WX5RRO2AG|         15549|
|          DVD|A2NJO6YE954DBH|          1366|
|          DVD| AU8552YCOO5QX|          1213|
|          DVD|A3P1A63Q8L32C5|           859|
|          DVD|A3LZGLA88K0LAO|           856|
|          DVD| A82LIVYSX6WZ9|           683|
|          DVD|A152C8GYY25HAH|           675|
|          DVD|A16CZRQL23NOIW|           651|
|          DVD|A1CZICCYP2M5PX|           650|
|       Sports|A18ZVYTEDAOF9A|             1|
|       Sports|A1W180Y9O1FALI|             1|
|       Sports|A3O8EZOX2P399L|             1|
|       Sports|A2RHSQZ7MAKKCO|             1|
|       Sports| AL62LOJKDES3M|             1|
|Baby Product| AI9SB5VKUFXDC|             1|
|Baby Product|A2LAH8VX720175|             1|
|Baby Product|A37TFIP0OMKGMW|             1|
| Video Games|A3C811U31YG6FS|             1|
| Video Games|A1M4NJYPOWNL8Q|             1|
| Video Games|A226EDS7WDF7S1|             1|
|         Book| ATVPDKIKX0DER|        643185|
|         Book|A3UN6WX5RRO2AG|        154531|
|         Book|A14OJSOVWMOSWO|          9589|
|         Book|   AFVQZQ8PW0L|          5441|
|         Book|A1K1JW1C5CUSUZ|          3562|
|         Book|A2NJO6YE954DBH|          2055|
|         Book|A3QVAKVRAH657N|          1651|
|         Book|A1NATT3PN24QWY|          1535|
|         Book|A1D2C0WDCSHUWZ|          1508|
|         Book|A2ODBHT4URXVXQ|          1469|
|        Music| ATVPDKIKX0DER|        166149|
|        Music|A3UN6WX5RRO2AG|         15875|
|        Music| A9Q28YTLYREO7|          2760|
|        Music|A2U49LUUY4IKQQ|          1258|
|        Music|A1GN8UJIZLCA59|          1154|
|        Music|A2NJO6YE954DBH|          1128|
|        Music|A1J5KCZC8CMW9I|          1031|
|        Music|A3MOF5KF93Q6WE|           989|
|        Music| AXFI7TAWD6H6X|           814|
```

```
|       Music|A38U2M9OAEJAXJ|            780|
|    Software|A183K8JAQJW8LZ|              1|
|    Software|A1F8RIBFWRYM3Y|              1|
|    Software|A1XQB8IU7S8WEU|              1|
|    Software|A2I0ZWBVR0575O|              1|
|    Software|A3D5ICIQ8STPCH|              1|
|    Software|A36T3O4TIC1YDQ|              1|
|    Software| A37UFPGDSSMEV|              1|
|    Software|A1EIVBXG3RD15O|              1|
|    Software|A1T6PXM2M3N84A|              1|
|    Software|A36NHJPD24UMGJ|              1|
|          CE|A1SFX3CR838F36|              1|
|          CE|A2IX9TMXDBUCYV|              1|
|          CE|A1328SYT22GA4U|              1|
|          CE|A1J62O1S6QTHZJ|              1|
|          CE|A13JU9OC7AU3RT|              1|
+------------+--------------+---------------+
```