CS432/532: NoSQL Project Report

Project Title: Amazon Product Consumer Review Analysis

Team Member(s): Sanjana Gosavi , Jigeesha Jain

I. Problem

a. Analyzing the consumer reviews recorded by Amazon over the period of 2010 to 2018 would help to improve the quality of service provided to the consumers, increase productivity and efficiency, have great agility and speed to market keeping "Customer First" approach.

The 2N user interactive data analysis tasks completed are:

-Quarter wise analysis of products based on reviews.

- Quality Analysis based on user entered product in that specific year.

-Comparative analysis on effect of Black Friday

- Popularity based recommendation for Kindle

b. For this project, we have used non structured query language called MongoDB.

c. We have used public dataset from Kaggle.

II. Software Design and Implementation

A. Software Design and NoSQL-Databse and Tools Used

Software Design:

Software used for programming was Jupyter Notebook which is a web based interactive computational environment for creating Notebook documents and the language used for query handling was Python.

NoSQL Database:
The project was implemented using MongoDB, a document-oriented database which gives the flexibility and scalability classified as NoSQL database program. The database query language we have used is non-structured query language that is MongoDB. For analyzing, we linked MongoDB through Python using the package PyMongo.

Tools Used :

- MongoDB Compass: Used to visualize data and how the reviews of various consumers for various products can be analyzed.
- Jupyter Notebook
- MongoDB server

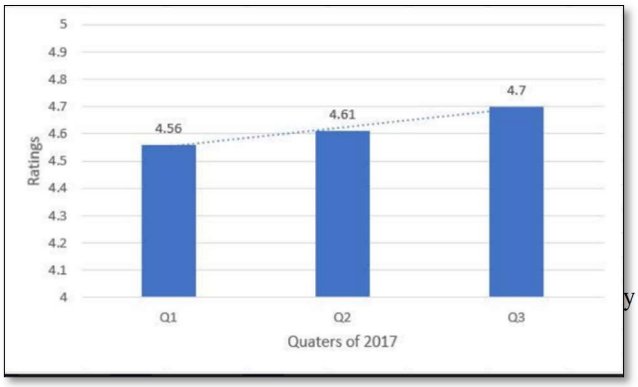
B. Parts that you have implemented

In the project the dataset has been analyzed using various techniques for getting intuitive results.

a. Initially, we ask the user to provide the year as an input and based on that, we can get the quarter wise trend of all the products in that particular year.
b. In the second task, we are deep diving not only in the year but also for a specific product the user can get average rating based on the reviews given so far which enables the consumers as well as organization to better choose and improve the products respectively.
c. In the third task, we are analyzing the effect of Black Friday based on consumer reviews by relatively comparing two different years of user's choice.
d. In our fourth task, we are analyzing the popularity gained by Amazon Kindle keeping the rating threshold as our base.

III. Project Outcome

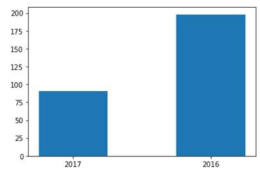
a. The below graph represents quarterly comparison of ratings for the user entered year "2017" for all the products.



y

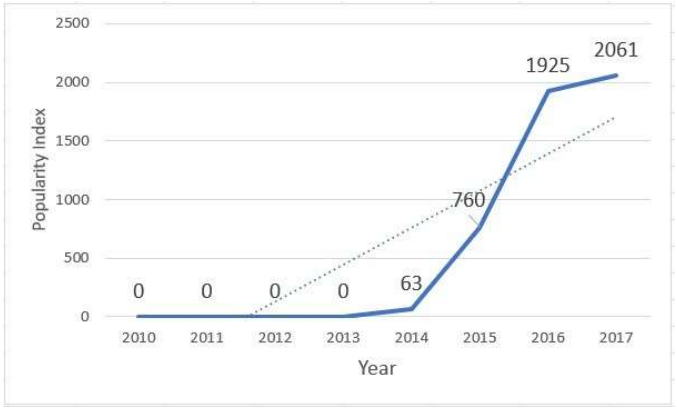
user i.e., 2016 and 2017 and we have tried to plot the same graph in Jupyter Notebook using the matplotlib package.

```
Enter the number between 1 and 5 as a choice and choice 6 for exit:1
Enter the year:2017
The average of Quarter 1 for the given year is:  4.564058887942572
The average of Quarter 2 for the given year is:  4.616766467065868
The average of Quarter 3 for the given year is:  4.709523809523381
Enter the number between 1 and 5 as a choice and choice 6 for exit:2
Enter the product:All-New Fire HD 8 Tablet, 8 HD Display, Wi-Fi, 16 GB - Includes Special Offers, Magenta
Enter the year:2017
4.584158415841584
Enter the number between 1 and 5 as a choice and choice 6 for exit:3
Enter the year:2017
Enter the year:2016
91
198
```




```
Enter the number between 1 and 5 as a choice and choice 6 for exit:4
{'2010': 0, '2011': 0, '2012': 0, '2013': 0, '2014': 63, '2015': 760, '2016': 1925, '2017': 2061}
Enter the number between 1 and 5 as a choice and choice 6 for exit:6
```

c. The graph below represents the popularity index of Amazon Kindle over the period of 2010 to 2017. This graph shows that the amazon product kindle has been gaining popularity greatly since 2015 onwards.



REFERENCES

[1] https://jupyter.brynmawr.edu/services/public/dblank/Jupyter%20Notebook%20Users%20Manual.ipynb

[2] https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products

[3] https://www.mongodb.com/languages/python

[4] https://docs.mongodb.com/manual/tutorial/

[5] https://www.tutorialspoint.com/mongodb/index.htm