Name: Sachin Badgujar, University of North Carolina, Email: Sachin.d.badgujar@gmail.com

Objective: Analyze 20 Tampon products from Amazon.com and plot number of reviews per month for past 1 year.

Implementation: Objective is achieved using python 2.7 and below with packages-

1. amazonproduct
2. selenium
3. datetime
4. csv
5. pandas

The program is using Amazon Product API to fetch all the information for the tampon product. The first 20 search results are taken into consideration. However, the API does not fetch all the reviews of the product. To achieve this, program takes help of Selenium web driver to scrap the Amazon website. This is as per the assumption that the sole purpose of results is to increase the sales of Amazon.com.

Special Conditions: The program needs several changes in order to run in different environment-

1. You need credentials for your Amazon Product API
2. Python Packages mentioned in Implementation
3. As the number of products increases, program becomes unstable in terms of web-scrapping

Input/Output: Program outputs 2 files

1. AmazonReviewsVerifiedPurchaser.csv – Contains Product ID, Name, Best Seller Rank, Date of Review
2. PerMonthReviews.csv – This contains the number of reviews per month per product.

I have used third-party web tools to visualize the data and to interact with different patterns in it. This is just for demo purpose. Number of product reviews are shown for last 12 months.

Results-
https://vizydrop.com/shared/drop/58ec247bd8169a017fcbf842?authkey=900cf05799f2695023fe

The product information on the graph is in format – 'ASIN – BestSellerRank – ProductName(30 Chars)' For example, 'B000GCNBLW - 4740 - Tampax Cardboard Applicator Ta'. The interactive to see the individual patterns by clicking on Category on right-hand side or resetting it.