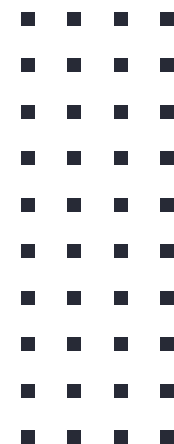


Cogs in the Amazon e-Customer Relations



Samy Palaniappan

Metis Project 5

March, 2020



Minimal Churn

How ? What are they doing right ?
Insights into the customer's head
- from their reviews.



Objective

1. **Identify** the needs and wants of the online customer
2. **Fill** customer needs



Specific Business Questions

1. Identify the **likes / dislikes** of customer about product.



Specific Business Questions

1. Identify the **likes / dislikes** of customer about product.
2. **Vine reviewers** and their role in customer relations ? – Are they able to predict what the customer feelings ?

Specific Business Questions

1. Identify the **likes / dislikes** of customer about product.
2. **Vine reviewers** and their role in customer relations ? – Are they able to predict what the customer feelings ?
3. Build a **recommender** system

Methodology - Tools

Amazon Reviews Dataset



11 Mil Reviews - 8 GB



PostgreSQL



Coding Platform



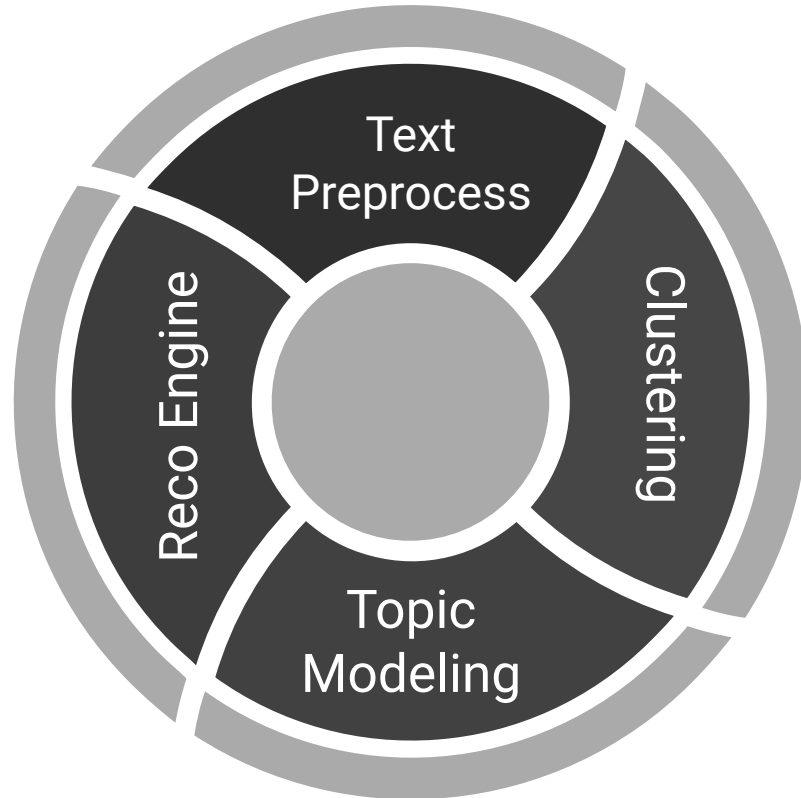
ML Libraries

spaCy

gensim



Methodology - Workflow



Methodology - Text Preprocessing / Clustering

Cleaning

Regex to
remove
symbols

Methodology - Text Preprocessing / Clustering

Cleaning

Regex to
remove
symbols

Stop Words

Spacy to tag
POS (ADV, ADJ),
bigrams, &
extend SW list

Doc - Word

DTM with
Count-Vectorizer
& TFIDF (min_df,
max_df)

Methodology - Text Preprocessing / Clustering

Cleaning

Regex to
remove
symbols

Stop Words

Spacy to tag
POS (ADV, ADJ),
bigrams, &
extend SW list

Doc - Word

DTM with
Count-Vectorizer
& TFIDF (min_df,
max_df)

Clustering

Kmeans - to
identify ideal
number of
clusters

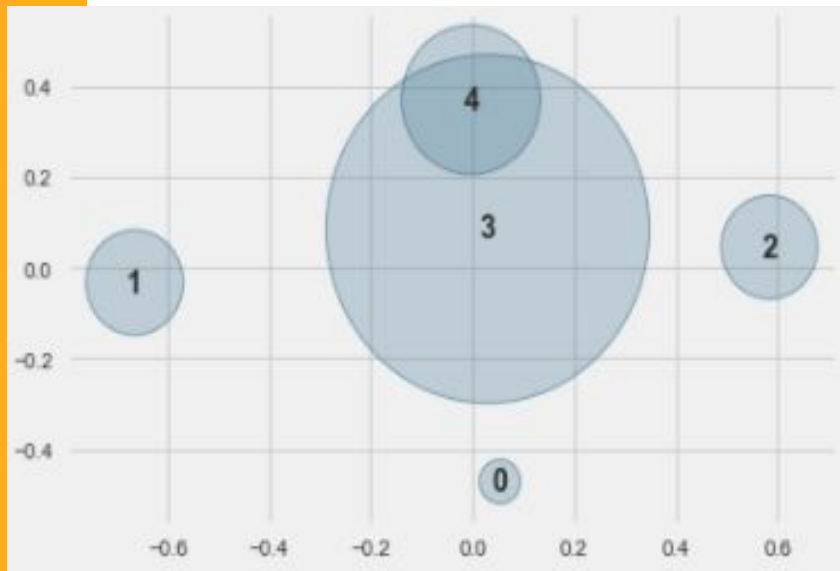
The slide features a dark blue background with a large orange horizontal band. On the left, there are four thick horizontal bars. On the right, there are four thick horizontal bars and a grid of small squares. The main title is in bold dark blue text, and the subtitle is in white text.

Results - Customer Topics

Finding what the customers care about ?

Customer Influencers - Topics (NMF-TFIDF)

Inter-Cluster Distance (K-means)



Topic 0: Life, Work, Recommended, New

['time' 'life' 'year' 'recommend' 'new']

Topic 1: Children Family

['loved' 'daughter' 'son' 'kid' 'absolutely']

Topic 2: Condition, Delivery, Value

['condition' 'arrived' 'price' 'described' 'new']

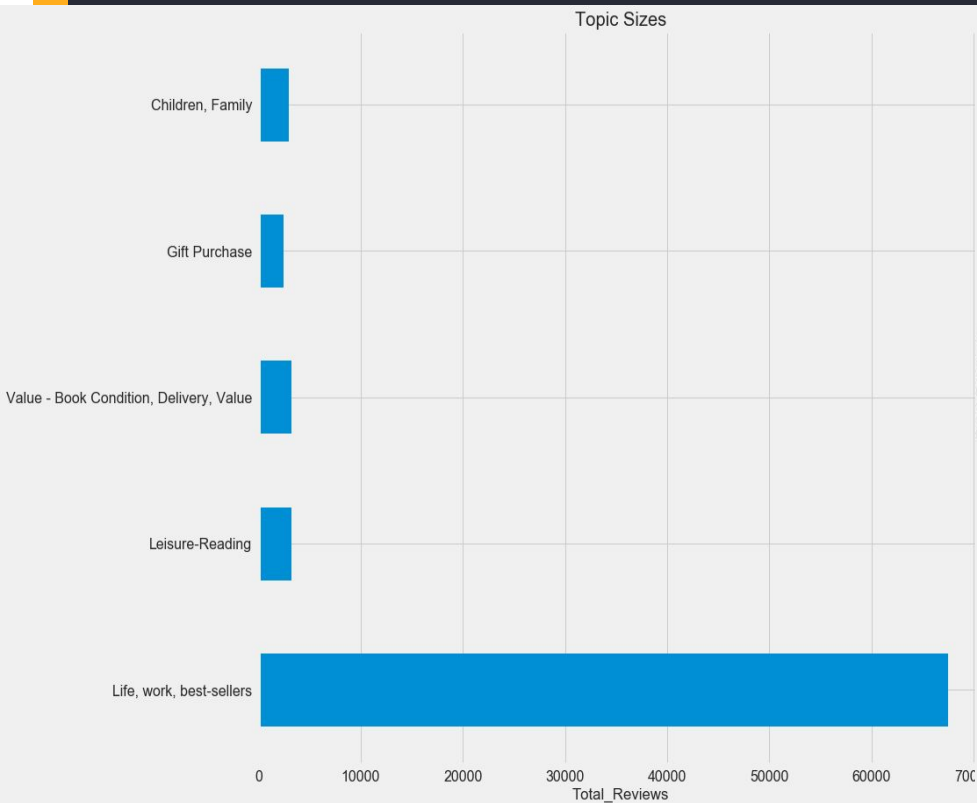
Topic 3: Fiction - Feeling, Series, Character

['series' 'enjoyed' 'character' 'next' 'favorite']

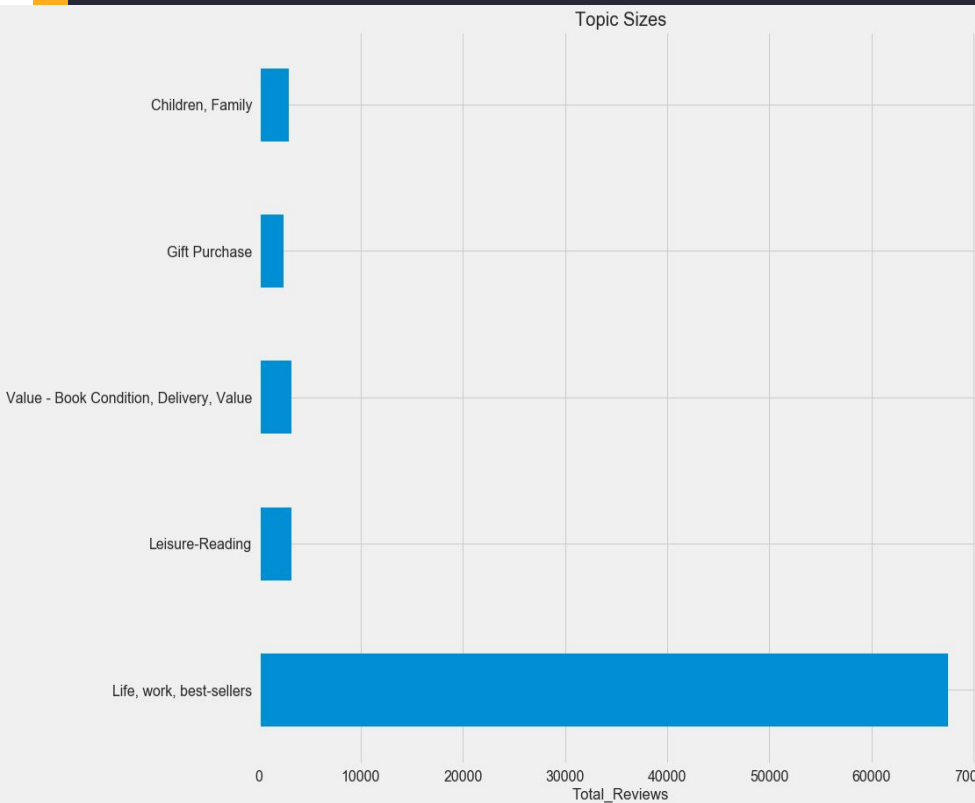
Topic 4: Gift Purchase

['gift' 'friend' 'gave' 'purchased' 'enjoyed']

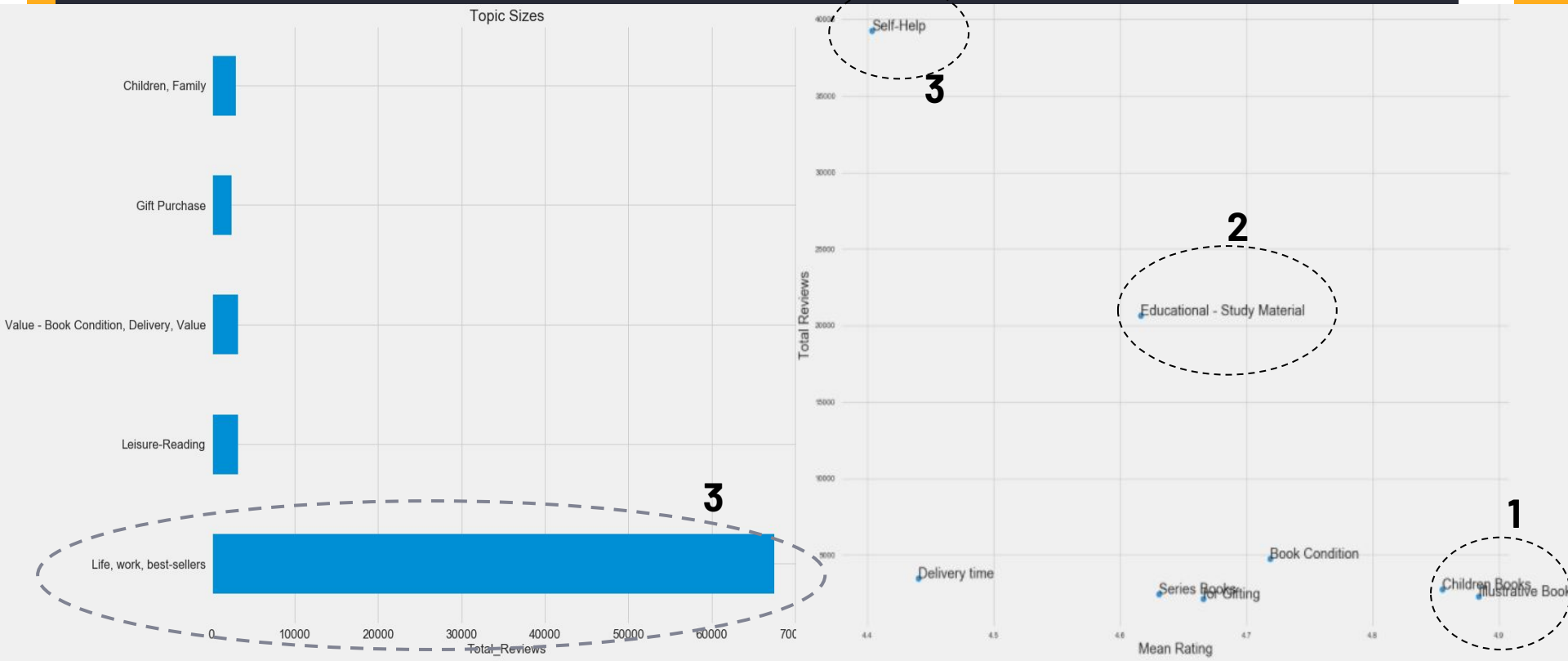
Advice for Authors / Sellers



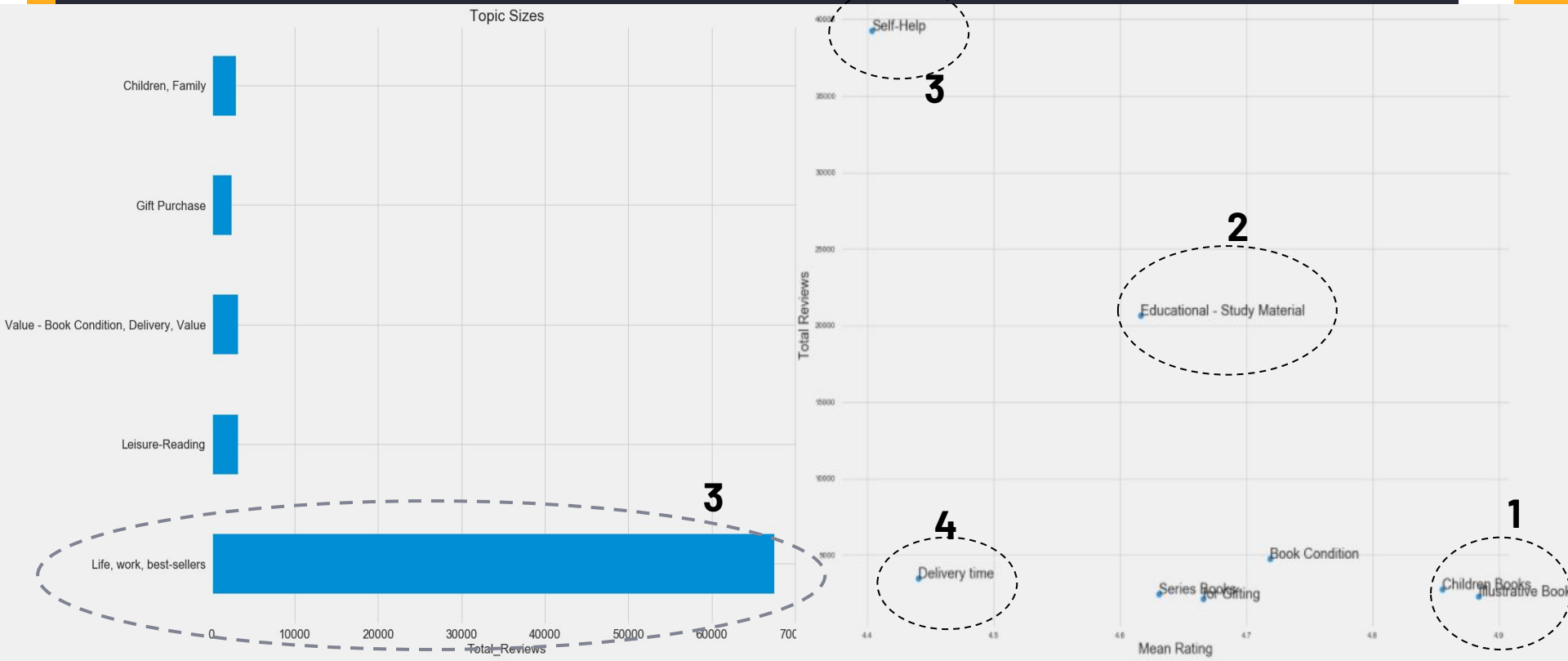
Advice for Authors / Sellers



Advice for Authors / Sellers



Advice for Authors / Sellers





Results – Vine Reviews

Identify the Role of Vine Reviewers



Vine Reviewers vs Regular Users (< 5 Star Reviews)

Vine Reviewers

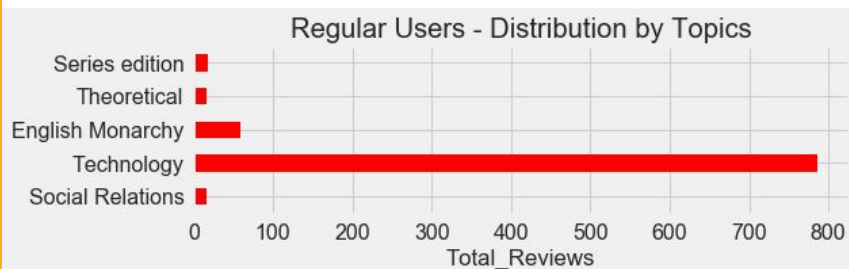
Topic 0: Race

Topic 1: Technology

Topic 2: English Monarchy

Topic 3: Theoretical

Topic 4: Series Edition

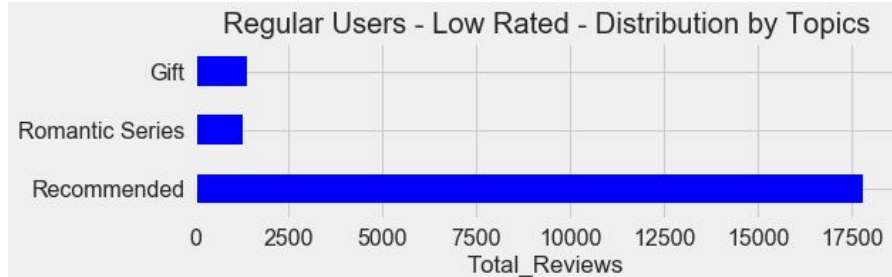


Regular Users

Topic 0: Recommended

Topic 1: Romantic Series

Topic 2: Gift



Vine Reviewers vs Regular Users (< 5 Star Reviews)

Vine Reviewers

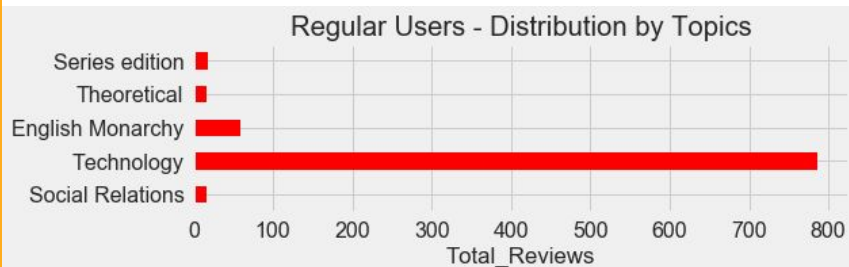
Topic 0: Race

Topic 1: Technology

Topic 2: English Monarchy

Topic 3: Theoretical

Topic 4: Series Edition

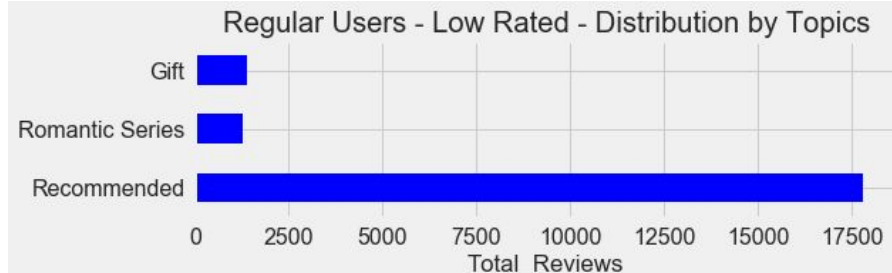


Regular Users

Topic 0: Recommended

Topic 1: Romantic Series

Topic 2: Gift



Nothing in Common

Vine Reviewers vs Regular Users (5 Star Reviews)

Vine Reviewers

Topic 0: American values

Topic 1: Educational books

Topic 2: History

Topic 3: Social books



Regular Users

Topic 0: Self-help

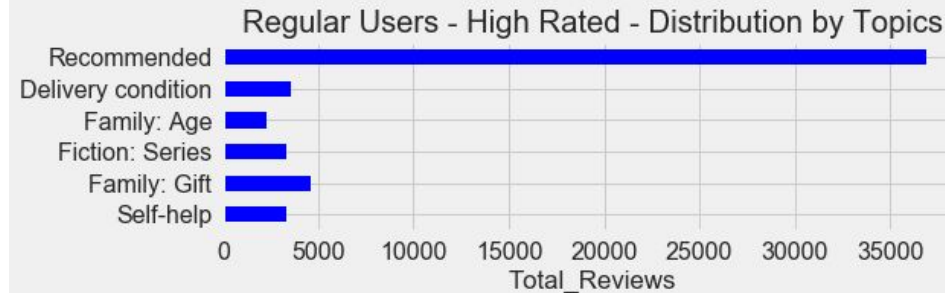
Topic 1: Family - Gift

Topic 2: Fiction-Series

Topic 3: Family - age appropriateness

Topic 4: Delivery condition

Topic 5: Recommended



Vine Reviewers vs Regular Users (5 Star Reviews)

Vine Reviewers

Topic 0: American values

Topic 1: Educational books

Topic 2: History

Topic 3: Social books

Again, Nothing in Common.

Regular Users

Topic 0: Self-help

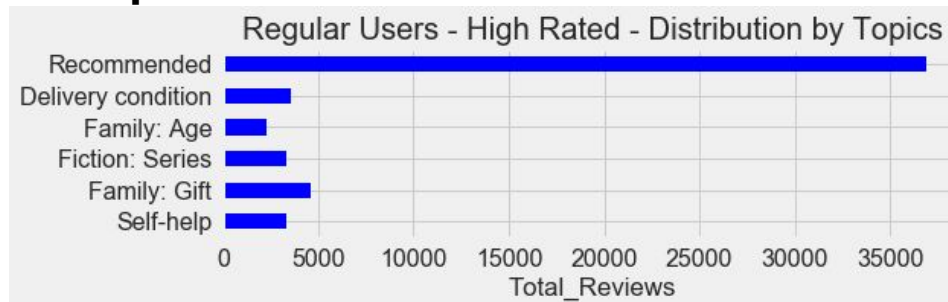
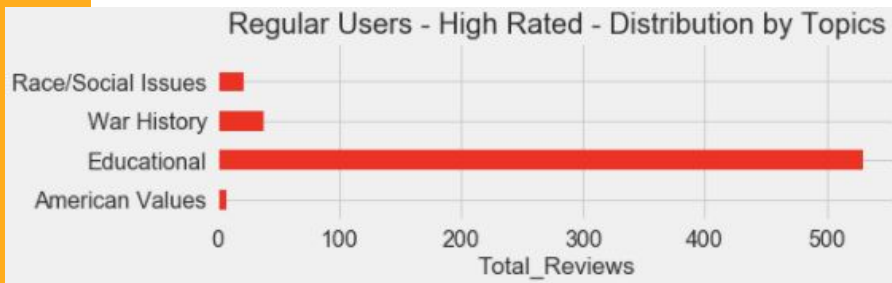
Topic 1: Family - Gift

Topic 2: Fiction-Series

Topic 3: Family - age appropriateness

Topic 4: Delivery condition

Topic 5: Recommended



Impact of Vine Reviews on Sales

1. Vine reviews were more detailed and thoroughly described the product.



Impact of Vine Reviews on Sales

1. Vine reviews were more detailed and thoroughly described the product.
2. Products recommended by vine reviewers resulted in more buys.



Impact of Vine Reviews on Sales

1. Vine reviews were more detailed and thoroughly described the product.
2. Products recommended by vine reviewers resulted in more buys.
3. User reviews shows that they very concerned about topic '**recommended**' and occurred the most on both highly and poorly rated reviews.





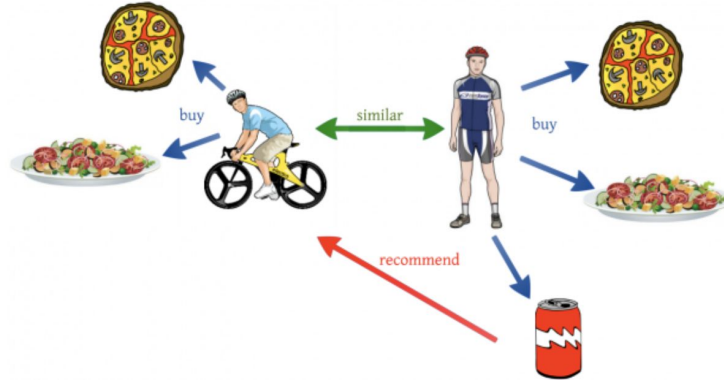
Results - Recommender Engine

Giving the customer what they need



Collaborative Filtering - Based Rec Engine (SVD)

For Repeat Users:



For First-time Users:

**From popularity bases library + user cosine similarity
(to input product)**

Collaborative Filtering - Based Rec Engine (SVD)

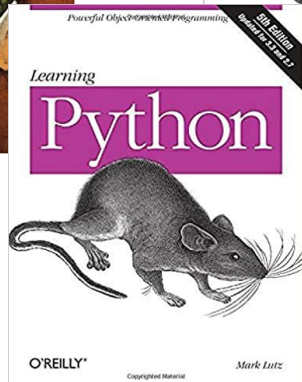
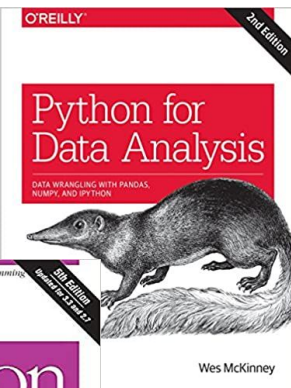
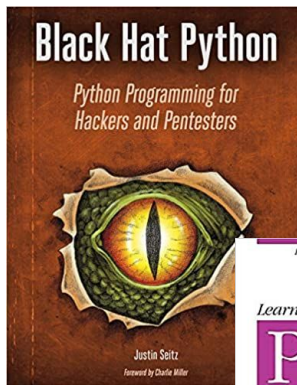
**Feature Engineering
(Adjusted Rating):**

$\text{unhelpful_votes} = \text{total_votes} - \text{helpful_votes}$
 $\text{adf_factor} = 1 + \text{helpful_votes} - \text{unhelpful_votes}$
 $\text{adjusted_rating} = \text{actual_rating} + 0.1 * (1 + \text{adf_factor})$

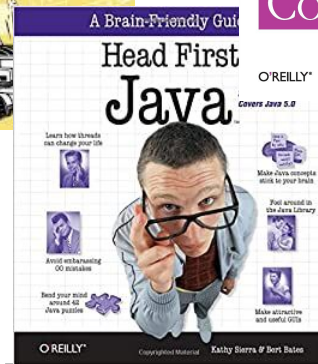
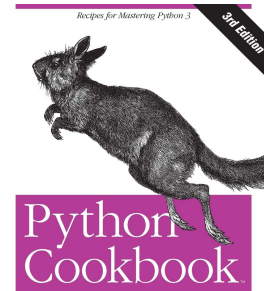
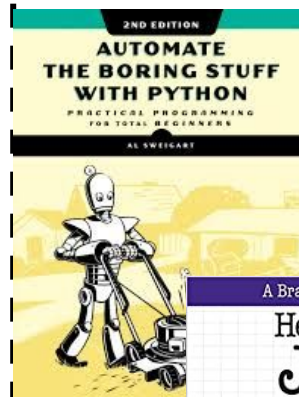
Rating-Based Rec Engine – Results

Method: Matrix Factorization with SVD

Input



Output



Conclusion / Future Work

1. **Reco for authors / sellers (from user feedback):**

- a. Market for children's books and illustrative books is saturated
- b. Educational books (Technology) → some room for improvement
- c. Self help / recommended books → lot of room for improvement (probably due to over hype and lots of poor quality books.
- d. Delivery time / book condition - needs improvement, not very pressing.



Conclusion / Future Work

1. **Reco for authors / sellers (from user feedback):**

- a. Market for children's books and illustrative books is saturated
- b. Educational books (Technology) → some room for improvement
- c. Self help / recommended books → lot of room for improvement (probably due to over hype and lots of poor quality books.
- d. Delivery time / book condition - needs improvement, not very pressing.

2. **Vine reviews had a huge sway in customer decision – helped sell more books.**

- very detailed. Very different from reviews by regular users.

3. **Recommender works fairly well.**

Conclusion / Future Work

1. **Reco for authors / sellers (from user feedback):**

- a. Market for children's books and illustrative books is saturated
- b. Educational books (Technology) → some room for improvement
- c. Self help / recommended books → lot of room for improvement (probably due to over hype and lots of poor quality books.
- d. Delivery time / book condition - needs improvement, not very pressing.

2. **Vine reviews had a huge sway in customer decision – helped sell more books.**

- very detailed. Very different from reviews by regular users.

3. **Recommender works fairly well.**

1. **Future work:** Develop Flask App with Heroku (half-way there)

“

The best customer service is if the customer doesn't need to call you, doesn't need to talk to you. It just works.

- **Jeff Bezos**



Appendix – 2d representation of features using PCA

