amazon

Customer Review Analysis

A quantitative approach to identifying valuable customers

Prepared by Yusi Chen

January 14, 2021

amazon

Survey Findings | Why customer review matters

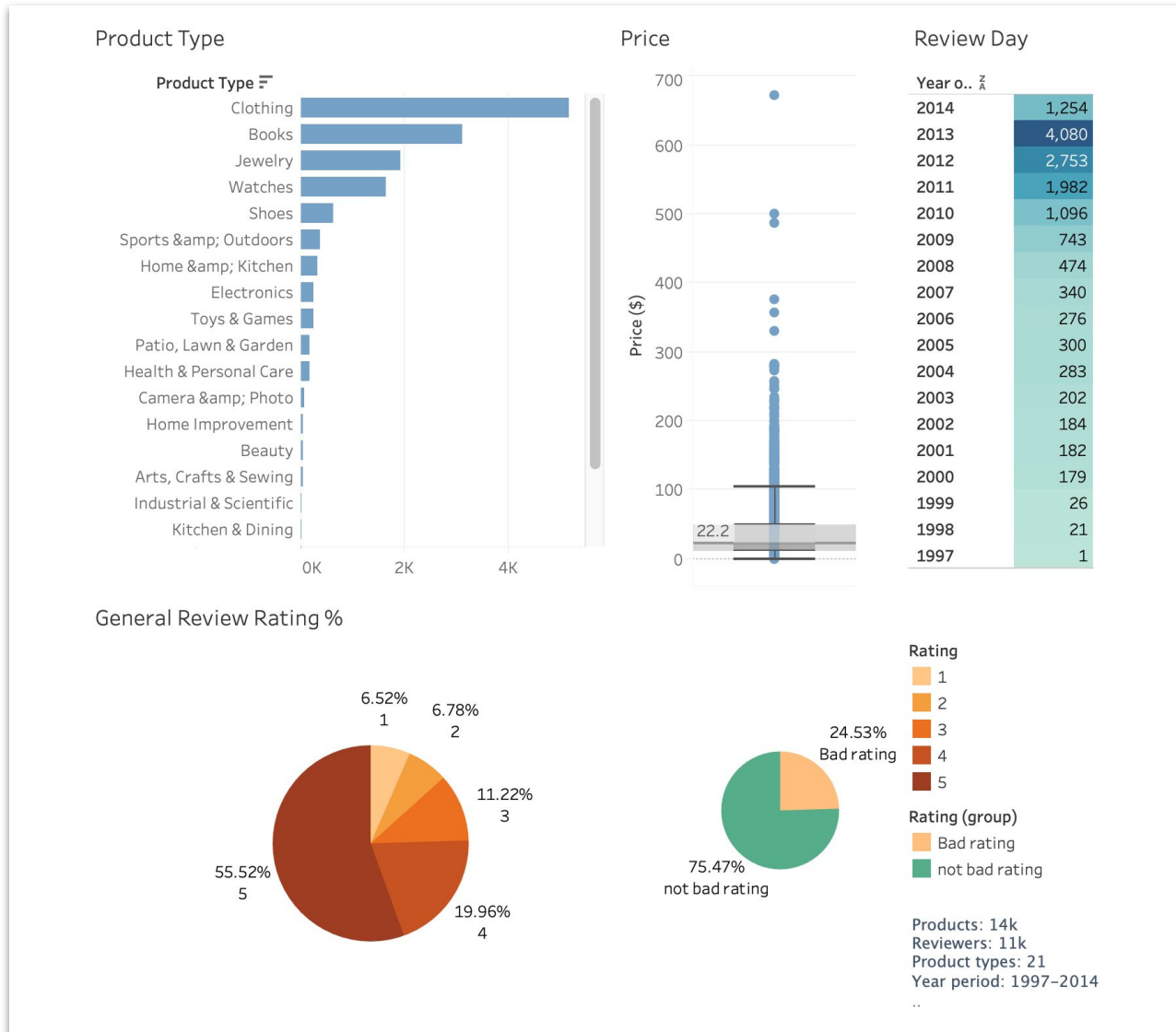**270%** 5 reviews increase purchase likelihood

5-9% Every 1-star brings revenue increase

50% Need at least a 4-star rating

40% Negative reviews can stop buyers

73% Value more on the written review

Data distribution │How does the data look like

Product Type

Product Type	
Clothing	
Books	
Jewelry	
Watches	
Shoes	
Sports & Outdoors	
Home & Kitchen	
Electronics	
Toys & Games	
Patio, Lawn & Garden	
Health & Personal Care	
Camera & Photo	
Home Improvement	
Beauty	
Arts, Crafts & Sewing	
Industrial & Scientific	
Kitchen & Dining	

0K   2K   4K

Price

Price ($): 700, 600, 500, 400, 300, 200, 100, 0

22.2

Review Day

Year o..	
2014	1,254
2013	4,080
2012	2,753
2011	1,982
2010	1,096
2009	743
2008	474
2007	340
2006	276
2005	300
2004	283
2003	202
2002	184
2001	182
2000	179
1999	26
1998	21
1997	1

General Review Rating %

6.52% 1
6.78% 2
11.22% 3
55.52% 5
19.96% 4

24.53% Bad rating
75.47% not bad rating

Rating
- 1
- 2
- 3
- 4
- 5

Rating (group)
- Bad rating
- not bad rating

Products: 14k
Reviewers: 11k
Product types: 21
Year period: 1997–2014
..

- An unevenly distributed dataset

- Half of reviews = 5

- ¼ bad reviews rating below 3

- Review rating interpretation

★★★★★ **Loved it**
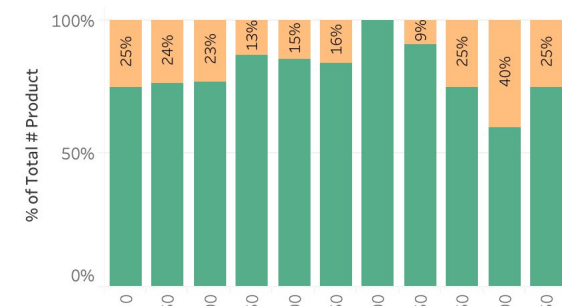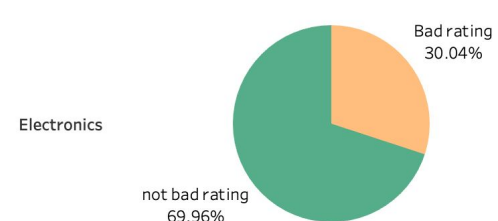★★★★☆ **Liked it**
★★★☆☆ **It was ok**
★★☆☆☆ **Disliked it**
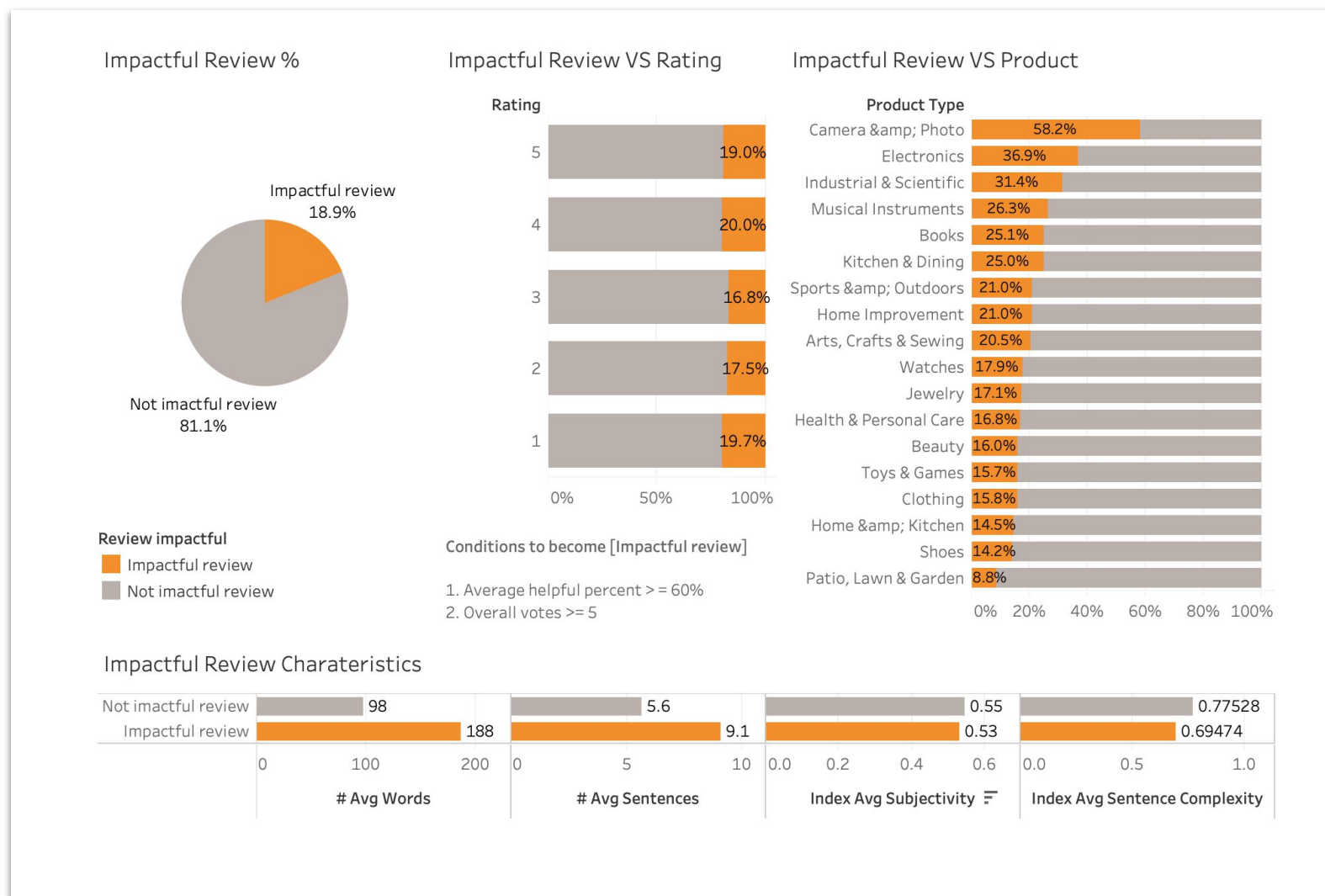★☆☆☆☆ **Hated it**
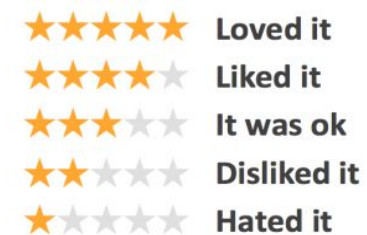
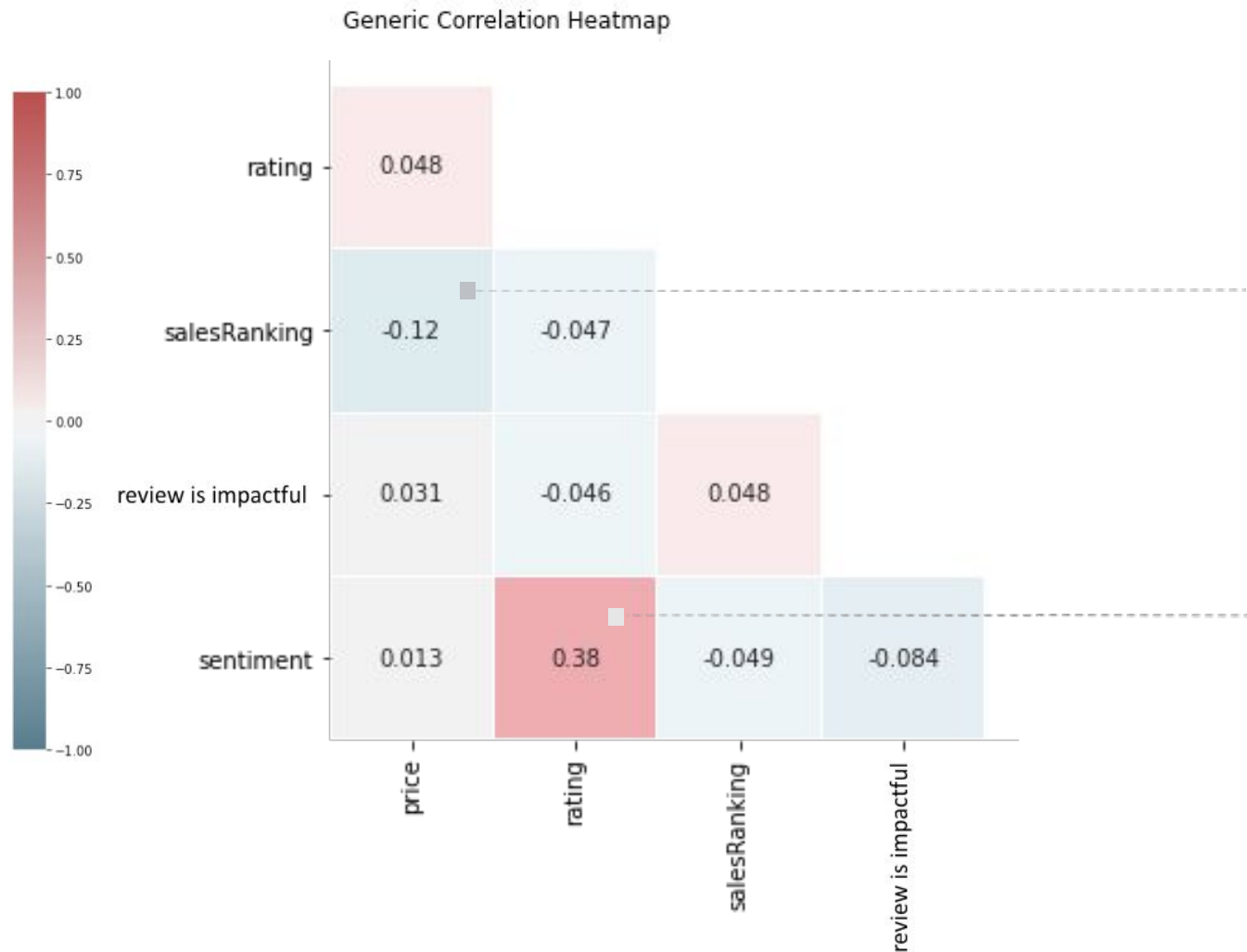# Rating | What is the relation to other features
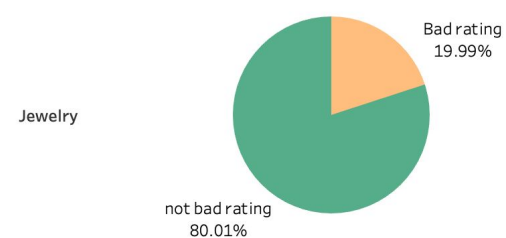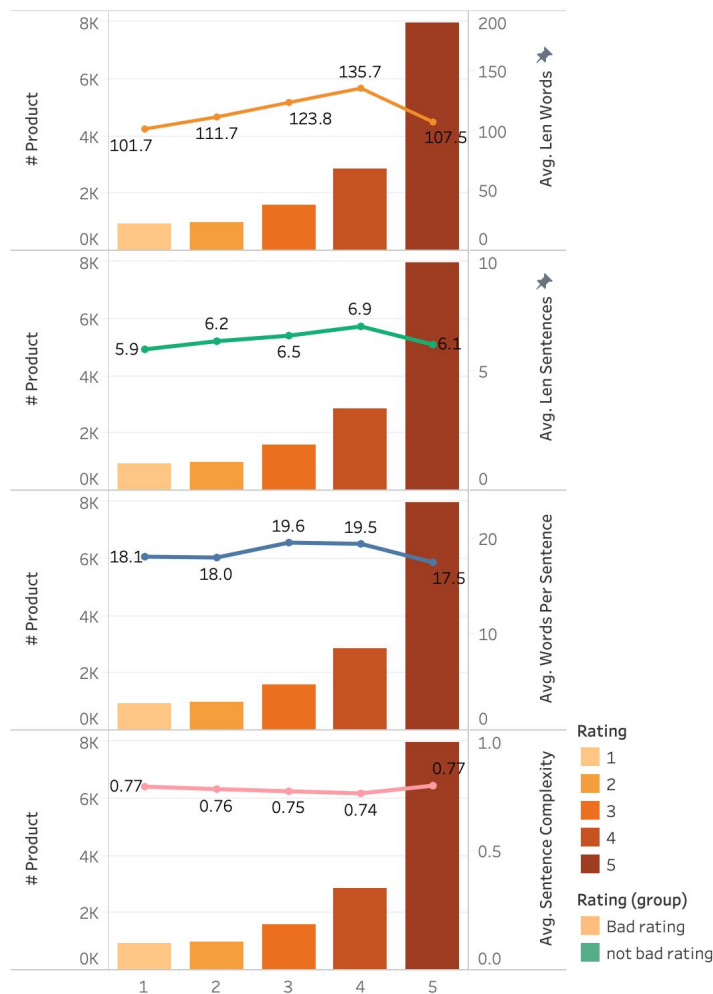
- The bad rating rate ranges from 20% to 30% among 21 product types.

- The bad ratings give more to two ends of products- either very cheap or very expensive.

- The avg number of words & sentence in review texts spikes in rating 4.

Review Helpfulness | What is the relation to other features

Impactful Review %

Impactful review
18.9%

Not imactful review
81.1%

Review impactful
- Impactful review
- Not imactful review

Impactful Review VS Rating

Rating

5	19.0%
4	20.0%
3	16.8%
2	17.5%
1	19.7%

0%   50%   100%

Conditions to become [Impactful review]

1. Average helpful percent > = 60%
2. Overall votes >= 5

Impactful Review VS Product

Product Type

Camera & Photo	58.2%
Electronics	36.9%
Industrial & Scientific	31.4%
Musical Instruments	26.3%
Books	25.1%
Kitchen & Dining	25.0%
Sports & Outdoors	21.0%
Home Improvement	21.0%
Arts, Crafts & Sewing	20.5%
Watches	17.9%
Jewelry	17.1%
Health & Personal Care	16.8%
Beauty	16.0%
Toys & Games	15.7%
Clothing	15.8%
Home & Kitchen	14.5%
Shoes	14.2%
Patio, Lawn & Garden	8.8%

0% 20% 40% 60% 80% 100%

- Impactful reviews are least likely to go together with rating 3.

- Impactful reviews favor technical and electronic gadgets.

- Impactful review characteristics
  - more words
  - more sentences
  - Less language complexity
  - more fact-based

Impactful Review Charateristics

	# Avg Words	# Avg Sentences	Index Avg Subjectivity	Index Avg Sentence Complexity
Not imactful review	98	5.6	0.55	0.77528
Impactful review	188	9.1	0.53	0.69474

Avg Words   # Avg Sentences   Index Avg Subjectivity   Index Avg Sentence Complexity

Correlation Heatmap |What is the correlation of multiple features

Generic Correlation Heatmap

Slightly negative coefficient

The higher the price, the sales raking number is smaller

The higher the price, actual sales ranking is better.
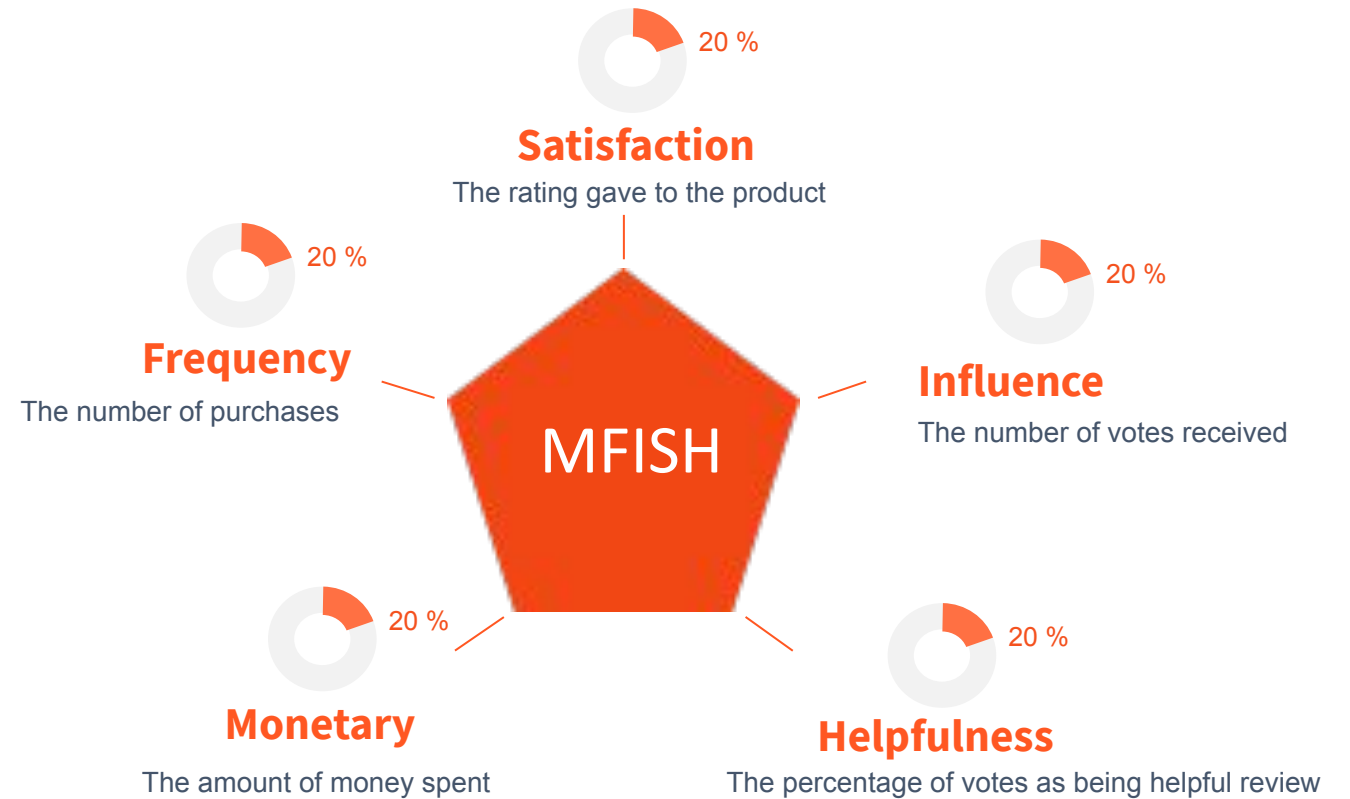
Moderately positive coefficient

The higher the rating score, the more positive the review texts.

Customer Segmentation | Which group of reviewers is valuable

Invent the conceptual framework MFISH to assess the value of each customer

	Frequency	Monetary	Satisfaction	Influence	Helpfulness
reviewerID					
A0011146133O3F18Q5NVR6	1	52.99	5.0	1	100.0
A00338282E99B8OR2JYTZ	1	11.55	3.0	0	0.0
A00354001GE099Q1FL0TU	1	18.02	3.0	0	0.0

20 %

Satisfaction

The rating gave to the product

20 %

Frequency

The number of purchases

20 %

Influence

The number of votes received

MFISH

20 %

Monetary

The amount of money spent

20 %

Helpfulness

The percentage of votes as being helpful review

Customer Segmentation |Which group of reviewers is valuable

Apply unsupervised machine learning algorithm to automatically divide the groups

Algorithms KMeans

Silhouette Score suggest 3-clusters is optimal

Fancy…

But, shall I trust it ?
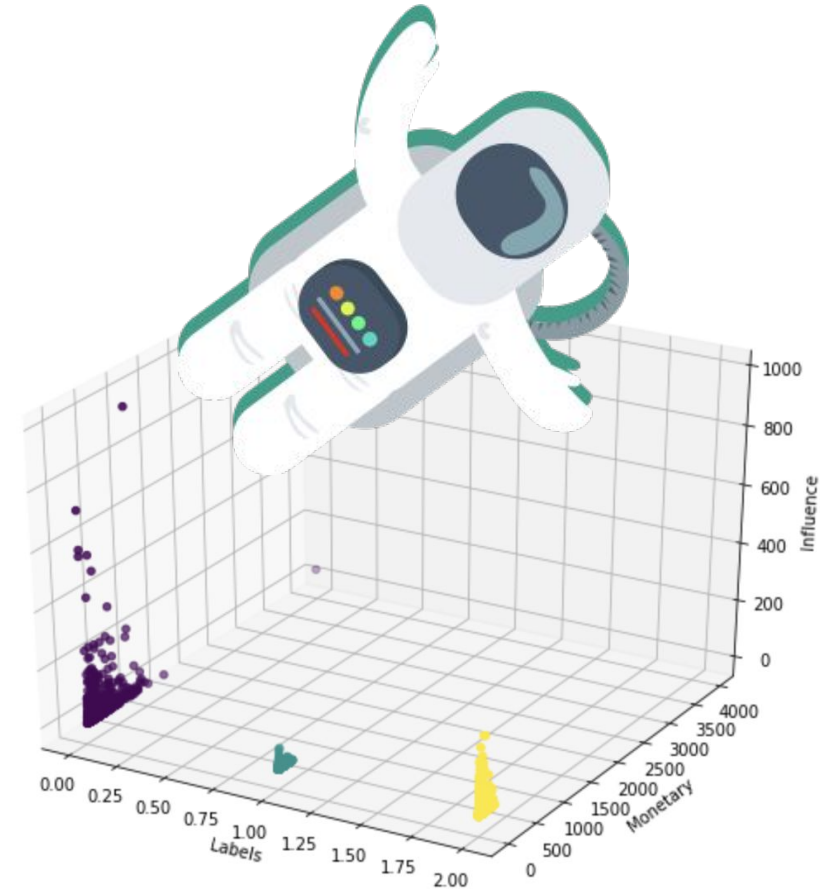
Customer Segmentation 3D Plotting Model

Customer Segmentation |Which group of reviewers is valuable

Manually divide the group based on the data distribution of each factor

reviewerID	Frequency	Monetary	Satisfaction	Influence	Helpfulness	AvgScore	Label
A7NJ6BVT92TJS	2.0	5.0	4.666667	1.0	1.0	2.733333	2-star
A3JOWXRNYC6UHL	4.0	5.0	4.545455	4.0	4.0	4.309091	4-star
A7O3870BUEO4K	4.0	5.0	4.333333	1.0	1.0	3.066667	3-star
A1247LFDYISISU	3.0	5.0	3.000000	3.0	4.0	3.600000	3-star
A2G4MGVWR20OOI	3.0	5.0	5.000000	2.0	3.0	3.600000	3-star
...	...	...	...	...	...	...	...
A35WGVT8AUEODN	1.0	1.0	4.000000	1.0	5.0	2.400000	2-star
A35WQNQID00GBW	1.0	1.0	3.000000	1.0	1.0	1.400000	1-star
A1URRWRCDRHHWZ	1.0	1.0	5.000000	1.0	1.0	1.800000	1-star

Divide the value on a scale of 1-5 through the percentile cut. Then sort them out…

Score	Size
1.0 - 2.0	44.28%
2.0 - 3.0	39.67%
3.0 - 4.0	15.37%
4.0 - 4.5	0.64%
4.5- 5.0	0.05%
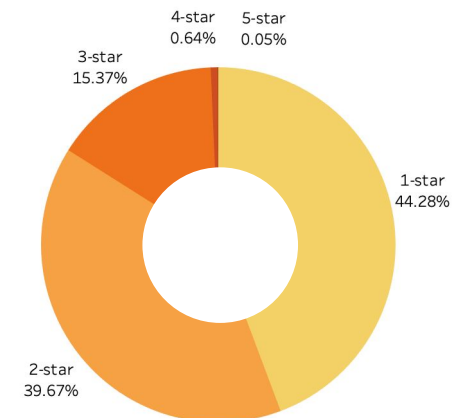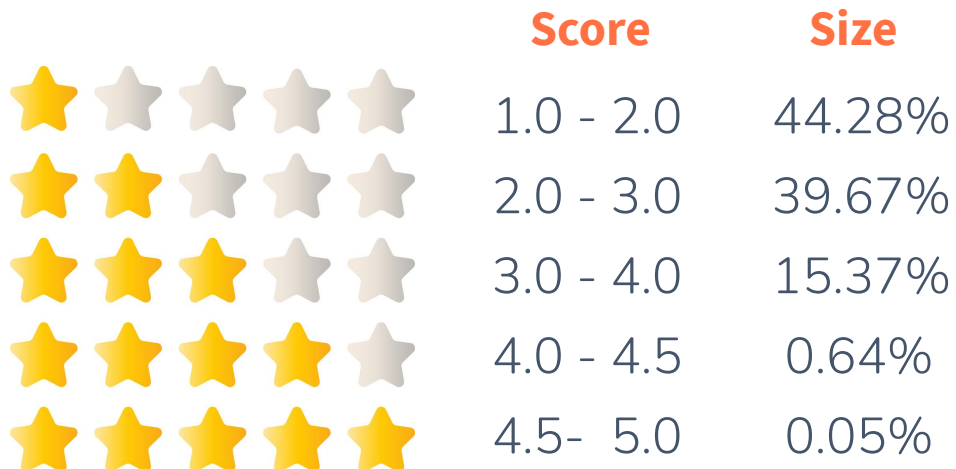
Customer Persona │How do our customers look like
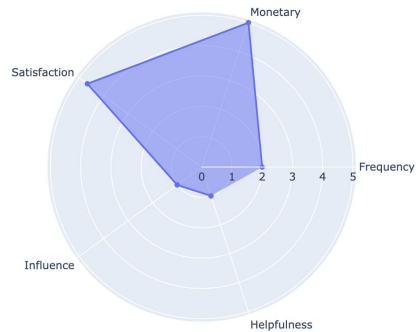
A. Bell

"I shop online a few times and I spend a lot to buy nice stuffs. I am lazy to write reviews."
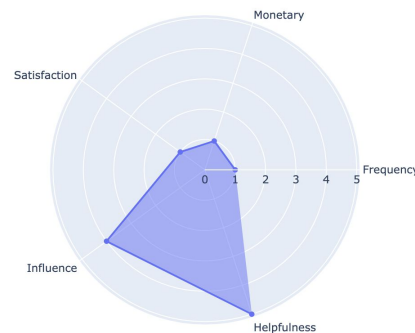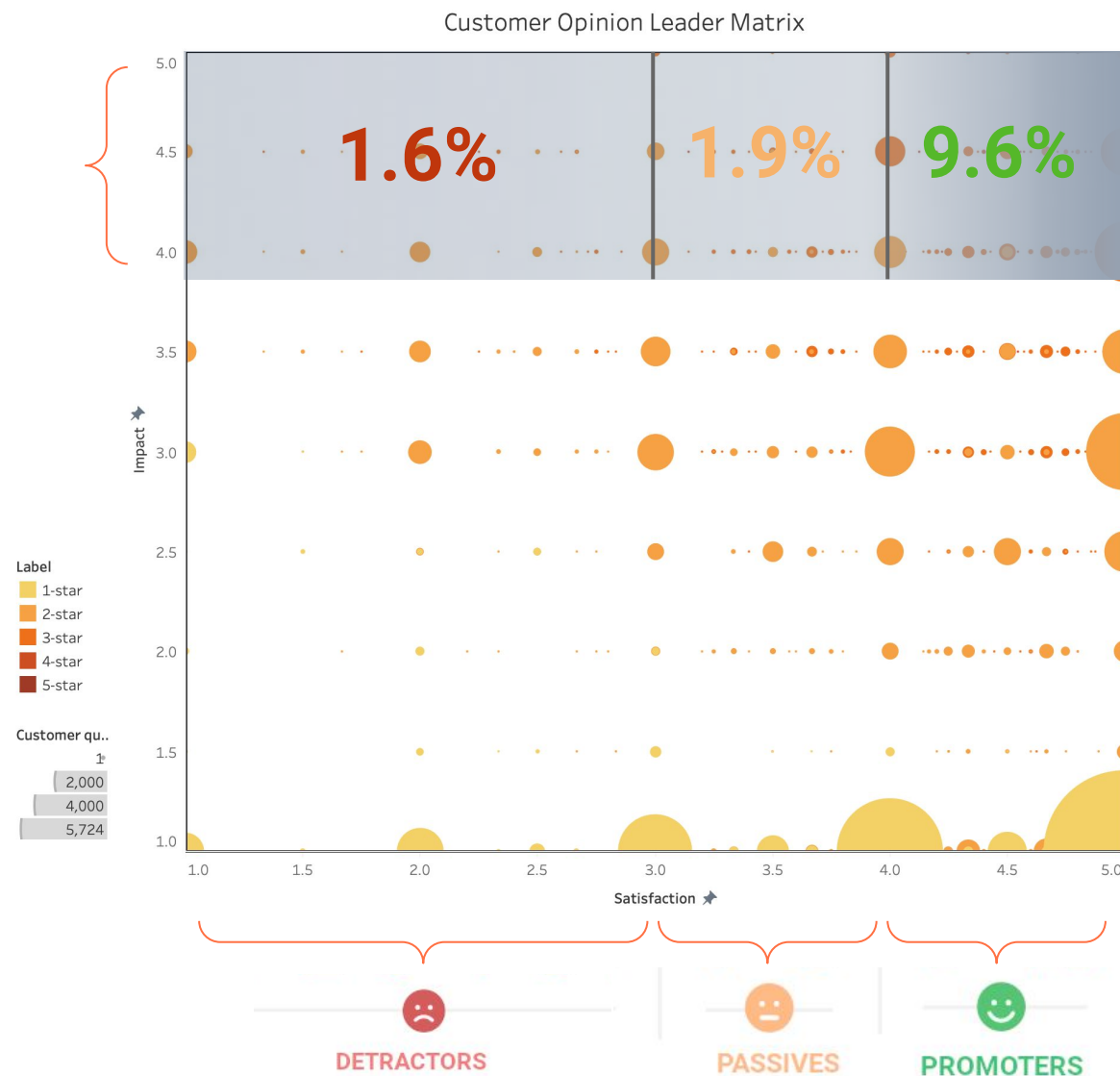
Tario

"Although I don't shop a lot on Amazon, I make sure my review is a bomb when the product is disappointing."
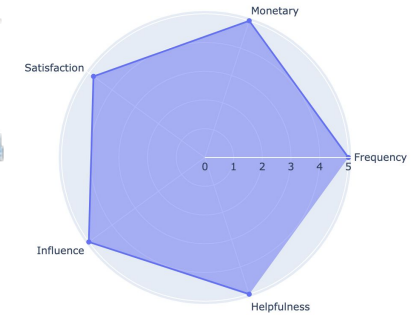
Matthew

"I am a happy customer, I make online shopping as my lifestyle. I also love sharing my experiences with other people, seriously."
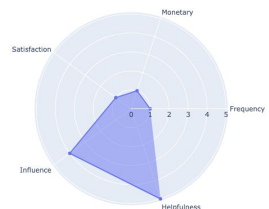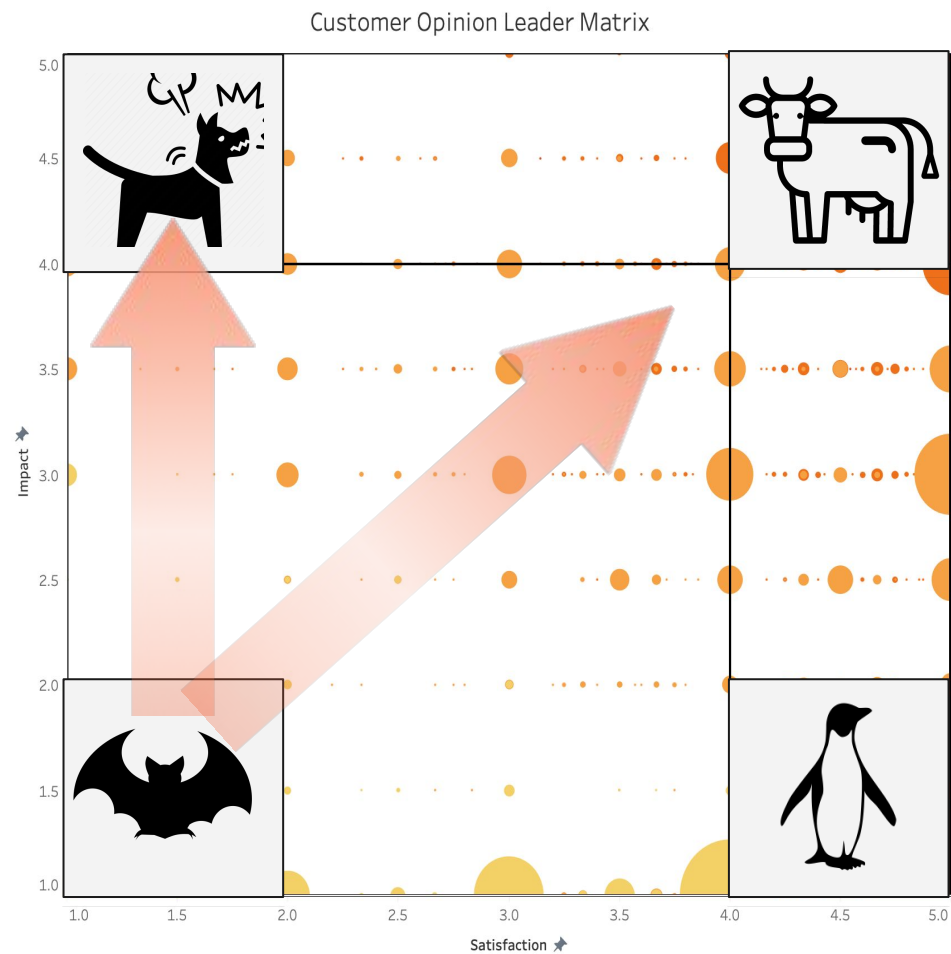
Reviewer Matrix | Who is the opinion leader

Opinion leaders

Customer Opinion Leader Matrix

1.6% **1.9%** **9.6%**

13.1%

of customers are opinion leaders

Label
- 1-star
- 2-star
- 3-star
- 4-star
- 5-star

Customer qu..
- 1
- 2,000
- 4,000
- 5,724

Impact

Satisfaction

DETRACTORS   **PASSIVES**   **PROMOTERS**

Note: Impact is a new feature further created from the mean value of Influence and Helpfulness

Reviewer Matrix |Who is the opinion leader

Customer Opinion Leader Matrix

I am not happy and I am barking

I am happy and giving
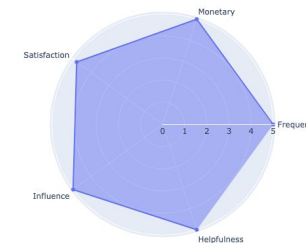
I am not happy and I silence myself in the cave

I am happy and I distance myself in the island
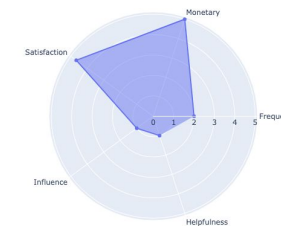
Conclusion

It is the implementation of the strategy to make a company great, not the strategy itself

Provide exceptional customer experience

◎ Respond to customer support queries faster

◎ Deliver personalized support interactions with context

◎ Establish better customer service workflows

◎ Offer omnichannel support to reach customers where they are

◎ Gather customer feedback on an ongoing basis

◎ Incentivize loyalty

Focus the prioritized resources on the prioritized customer groups

Example:

In face of resource constraints, the customer service team is recommended to guarantee the most valuable customer group (5-Star or Opinion Leaders) to always have the

-fastest first time response time

-fastest issue handle time

-deepest personalized communication

Empowering businesses through data analytics

Thank you

Keep in touch
yusi0928@gmail.com