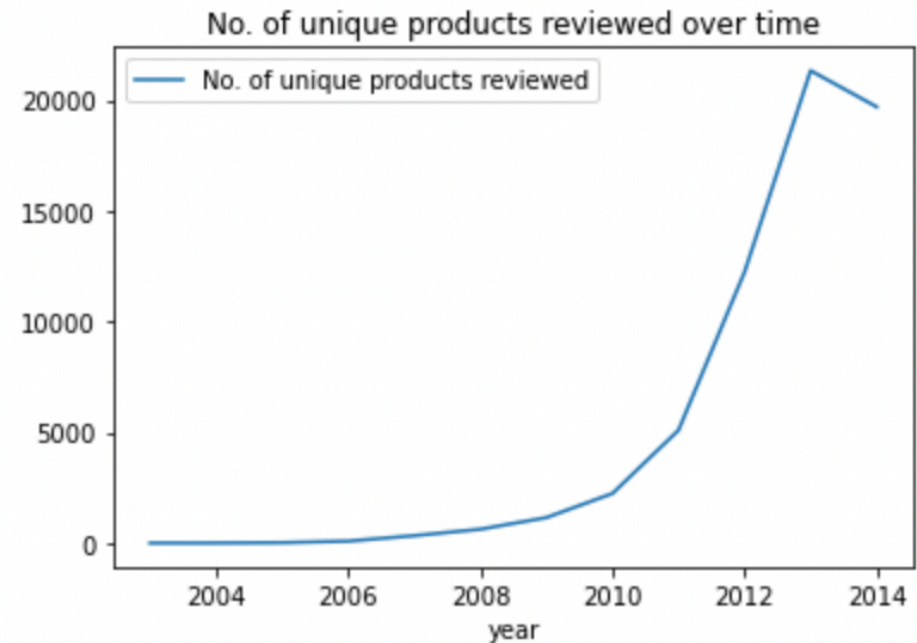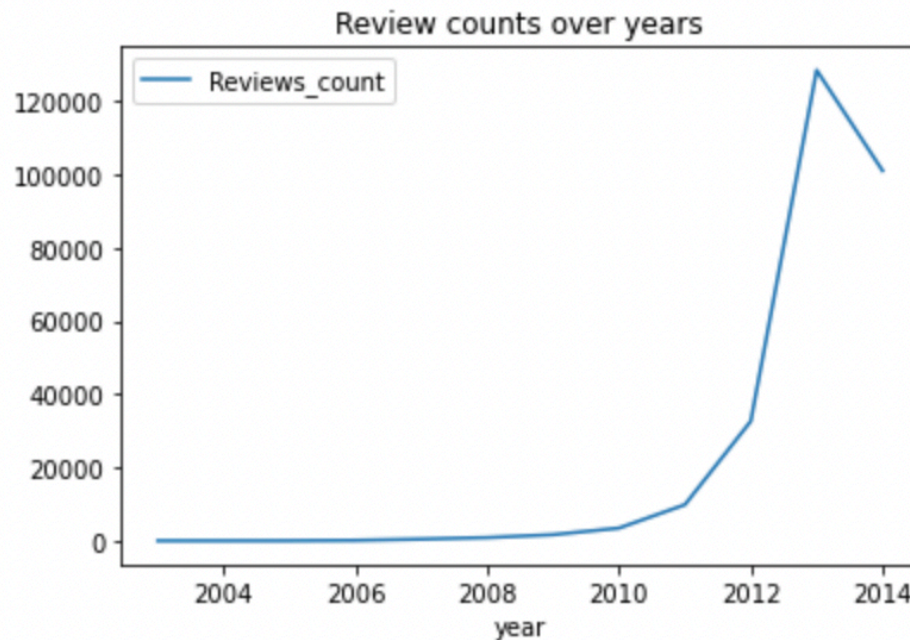# Executive summary

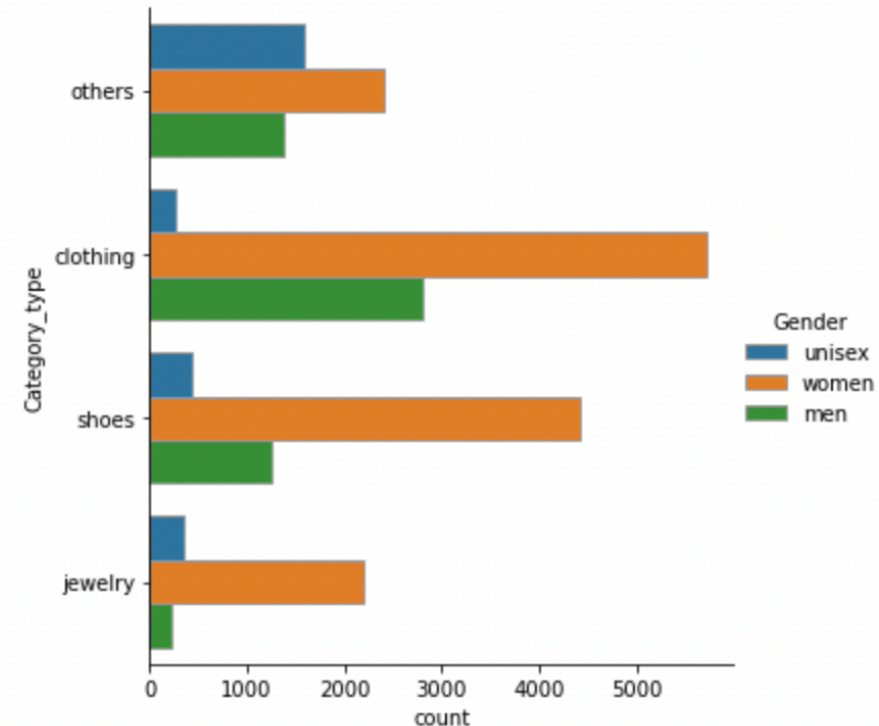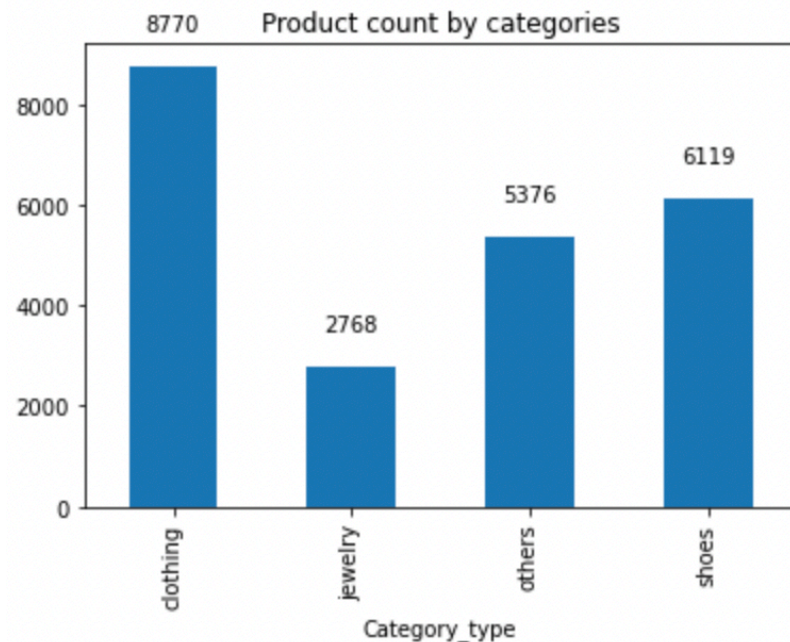Total products listed

23,033

Unique customers

39,387

Total reviews received

2,78,677

Executive summary – Products

- Most of the products listed are for women.

- Clothing category comprises of maximum products.

Executive summary – Products

- More than 60% of datapoints missing on columns description, price and brand.

- Assuming sales rank is available only for 2014 for all products. Sales rank column with blank values are treated as products that are not among the top 1 million top selling products.

- Gender and Broad-level product category were extracted from the category tags listed for each product.

Top 5 most reviewed products

| review_counts	title	Category_type
441	Vintage, Retro Colorful Crystal Owl Pendant an...	others
286	Spalding Women's Capri Legging	clothing
249	Allegra K Women's Long Sleeve Skull Printed St...	clothing
241	Skechers Women's Go Walk Slip-On	shoes
225	Azules Women's Rayon Span Maxi Skirt	clothing

Executive summary – Reviews

- 90% of customers have posted maximum 8 reviews across 2003-2014.

- 75% of reviews are of 350 characters in length.

- Average rating across all product categories are 4+ stars.

Average rating
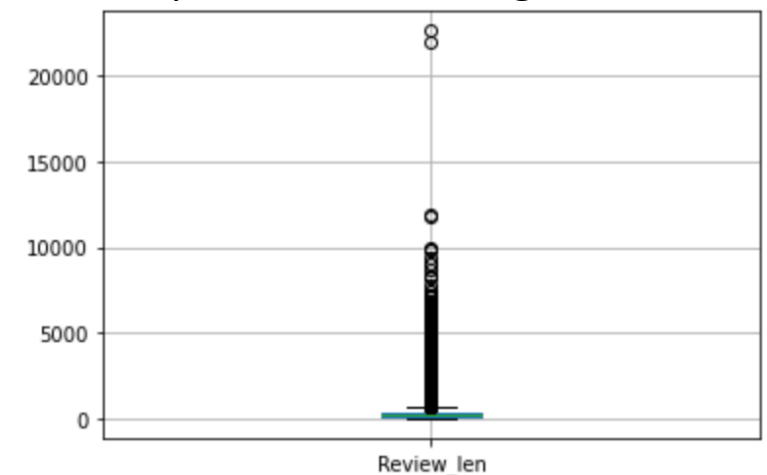
```
Category_type
clothing      4.178824
jewelry       4.314545
others        4.294919
shoes         4.282025
Name: overall, dtype
```

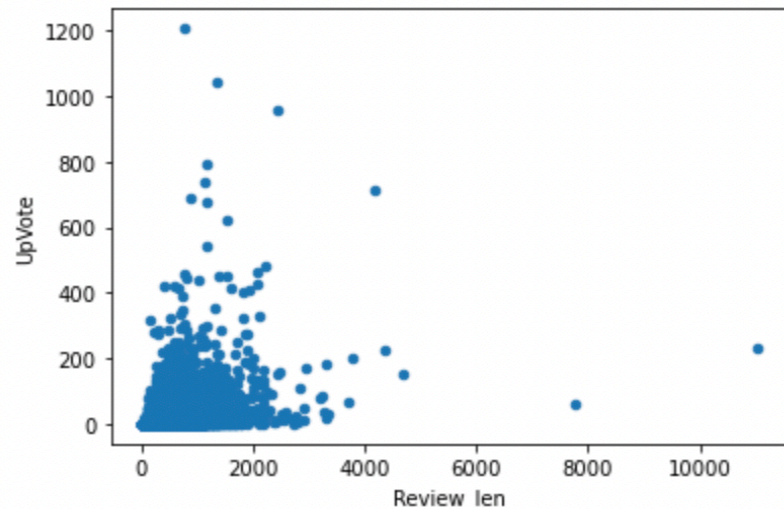Box plot of reviews distribution
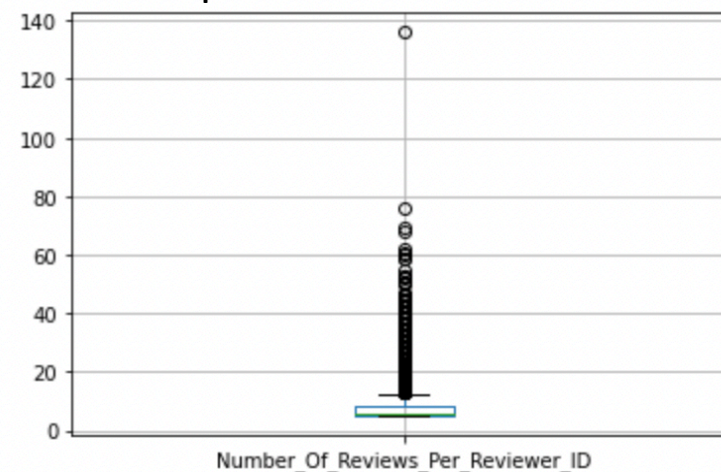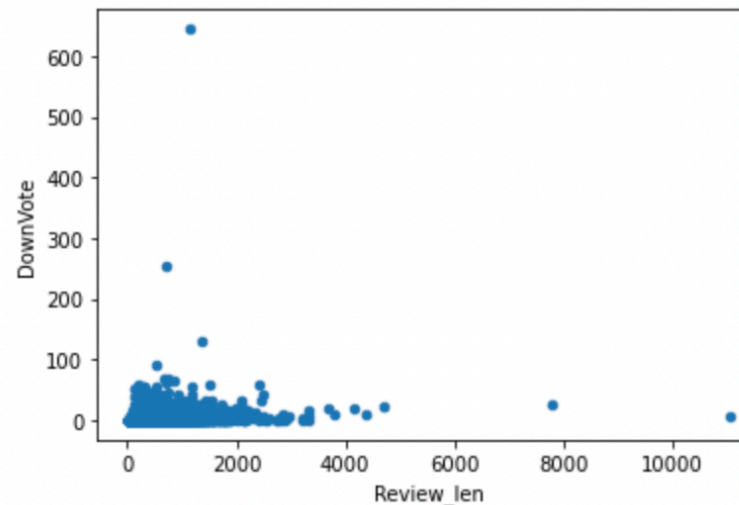

Box plot of review length distribution

5

Executive summary - Reviews

- Most of the up votes and down votes are observed for reviews less than 1500 characters in length.
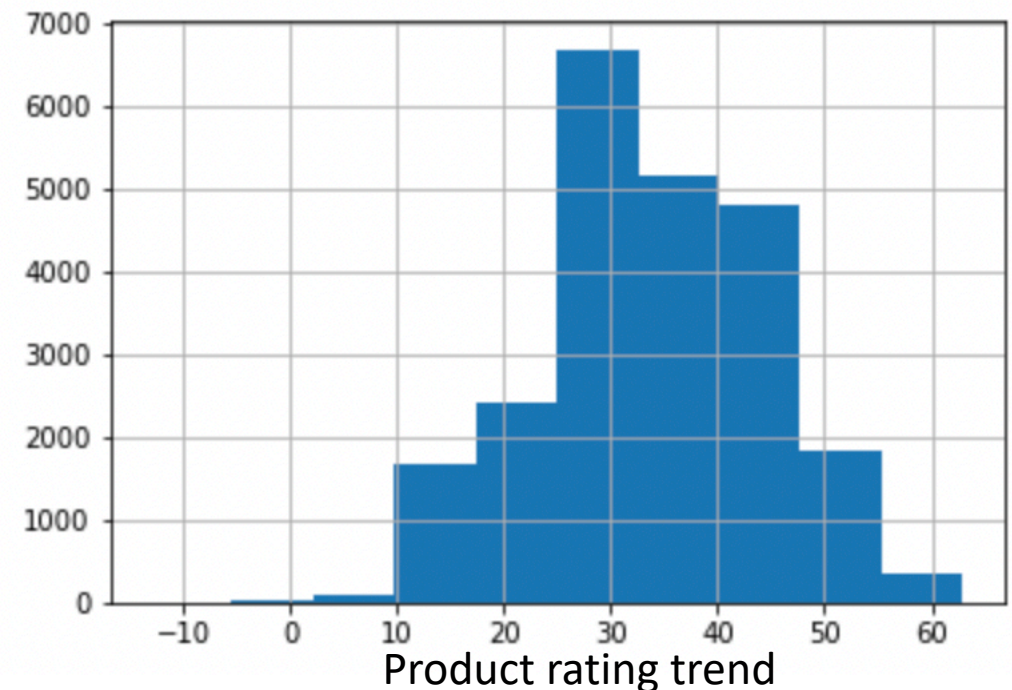
Review upvotes Vs Review length

Review downvotes Vs Review length

1. Product rating trend analysis

- On analysing the year-on-year changes in average customer rating, more than 99% of products displayed positive rating trend.

- 18 products shown consistent negative trend.

Distribution of product rating trend across 10 years

Product rating trend

Top 5 products with negative rating trend

| Product	Rating trend
Nine West Women's Ealene Bootie,Black Leather,7 M	-13.11
Suunto X-Lander Wrist-Top Computer Watch with ...	-11.36
Columbia Sportswear's Women's Benton Springs S...	-9.79
Timex Men's T56371 Ironman Triathlon 42 Lap Co...	-7.93
Olga Women's Sheer Tapestry Underwire Minimize...	-7.93

2. Correlation analysis

- There is no significant relationship identified between the following:
    - sales rank vs average product rating.
    - sales rank vs overall count of product reviews.
- There is a weak positive relationship between count of product reviews and positive product rating trend.

Correlation table

	review_trend1	converted_salesrank	overall	review_count
review_trend1	1.000000	-0.027874	0.392938	0.274183
converted_salesrank	-0.027874	1.000000	-0.006537	-0.060010
overall	0.392938	-0.006537	1.000000	0.003073
review_count	0.274183	-0.060010	0.003073	1.000000

3. Common issues with shoe category

- Size chart mismatch
- Hurts heel

Shoes - 10 Most Frequently Occuring Tri-grams - Negative reviews
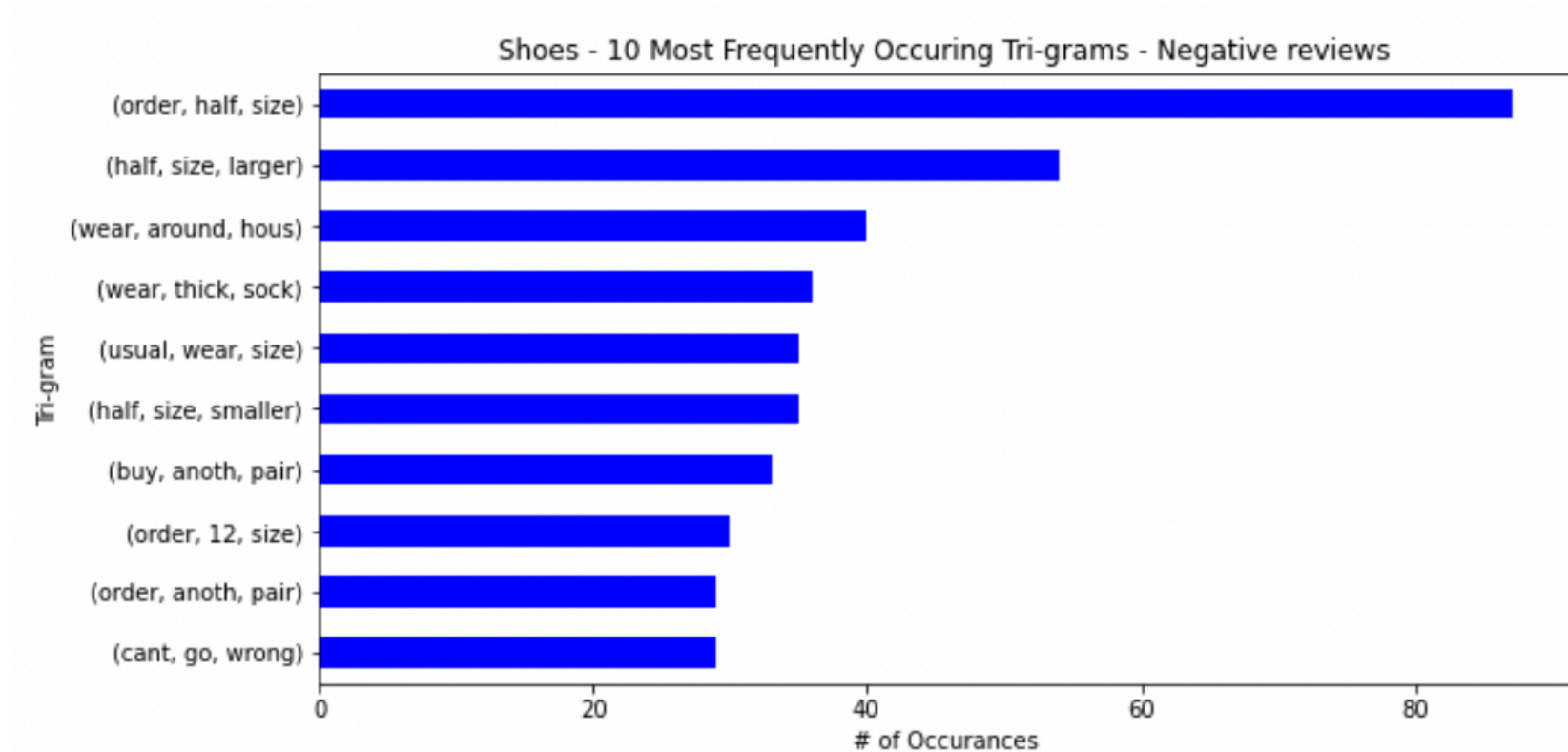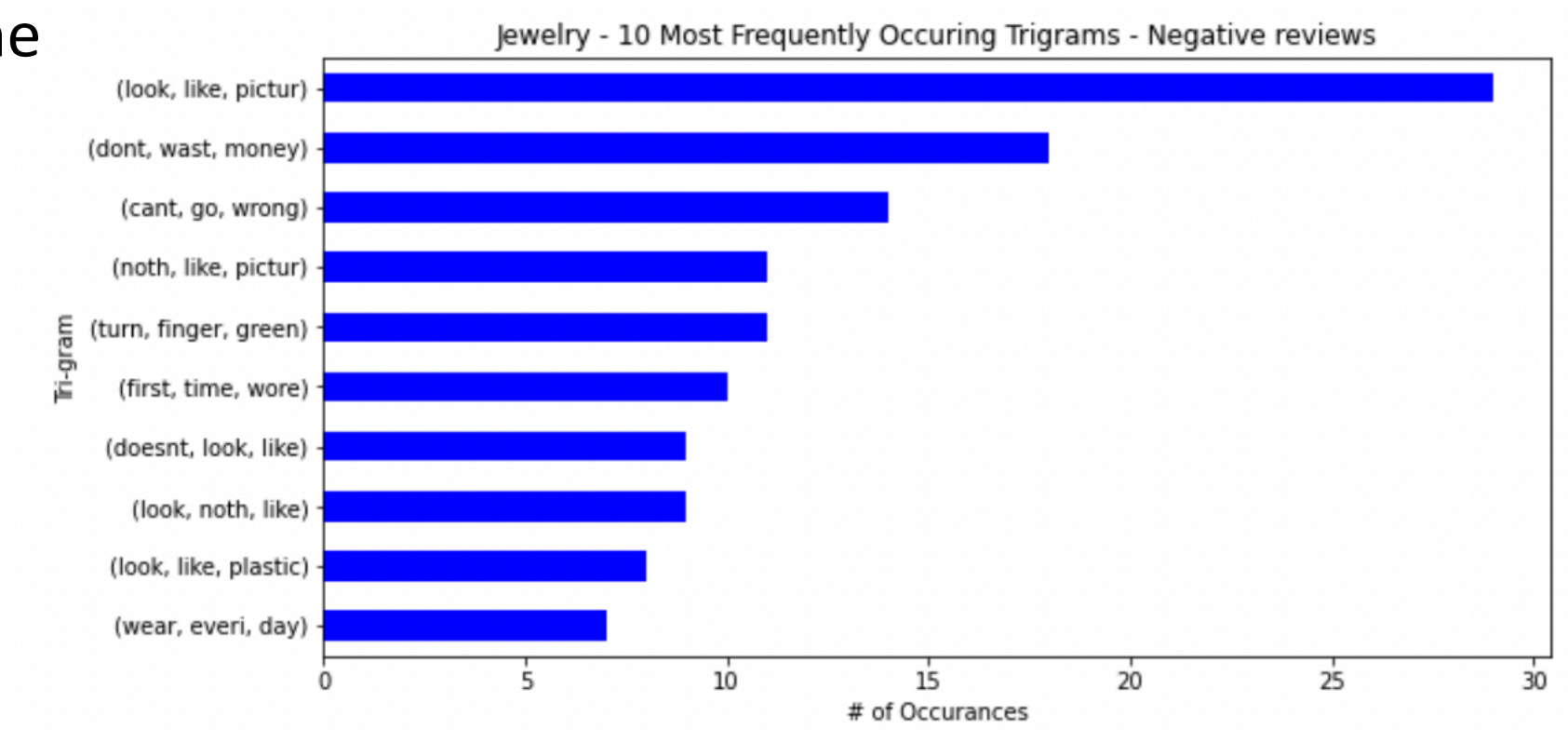
3. Common issues with clothing category

- Size mismatch

- Doesn't look like the one displayed in picture

- Quality issue

Clothing - 10 Most Frequently Occuring Tri-grams - Negative reviews

11
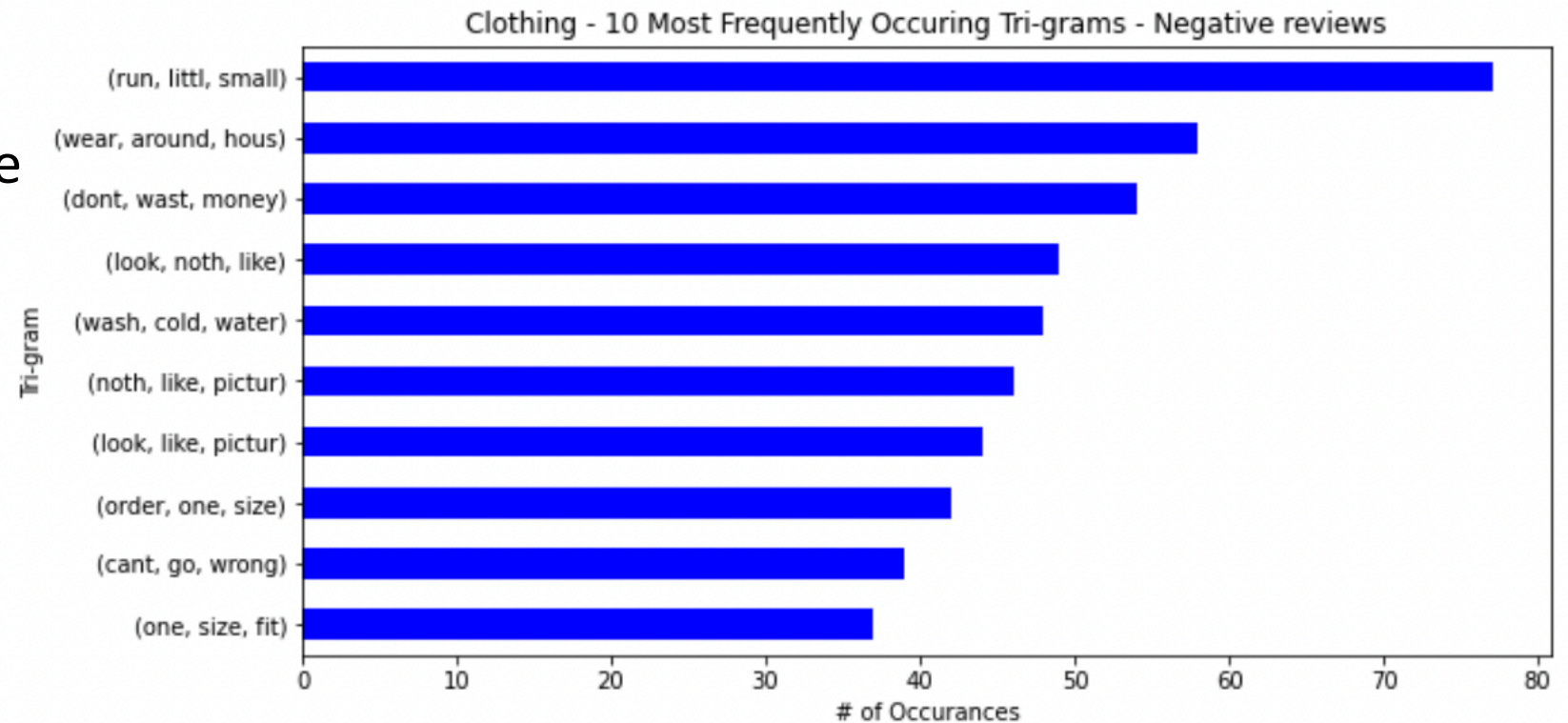
3. Common issues with jewellery category

- Doesn't look like the one displayed in picture.

- Quality issue

Jewelry - 10 Most Frequently Occuring Trigrams - Negative reviews
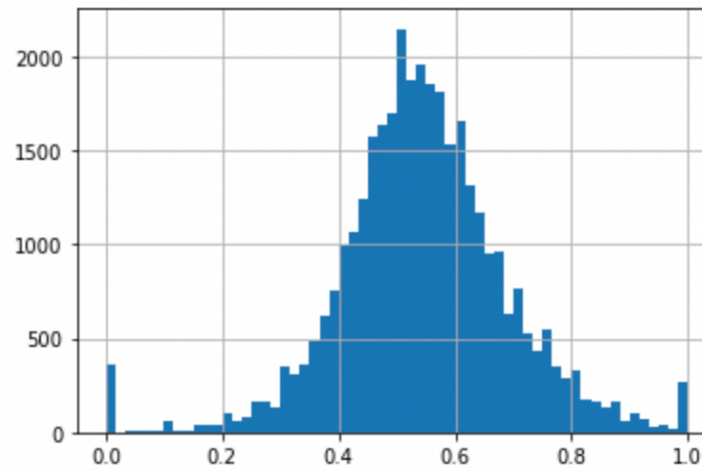
4. Polarity analysis of reviews

- Overall polarity of reviews are heavily inclined towards positive tone.

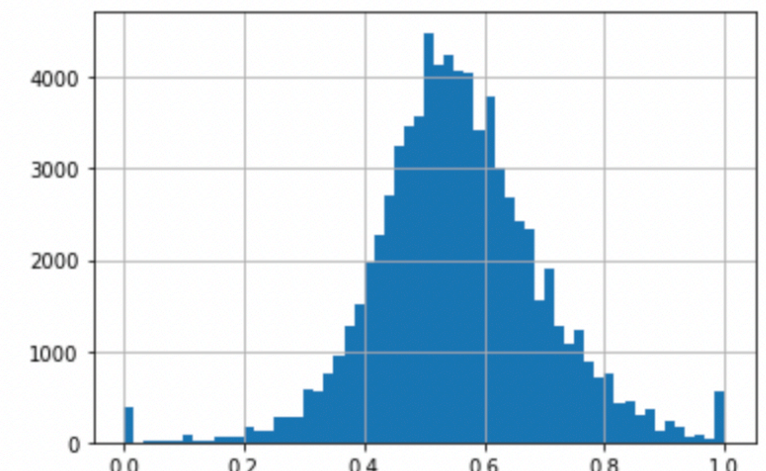- Both upvoted and downvotes comments are observed to be neither subjective nor objective
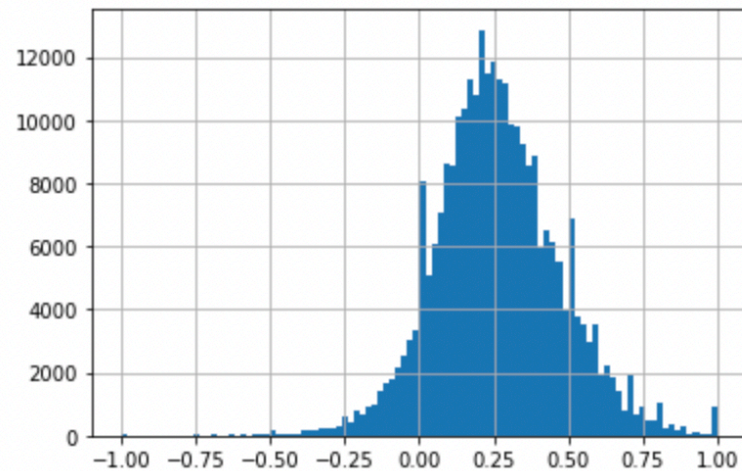
Overall polarity of reviews   Subjectivity of down voted reviews   Subjectivity of up voted reviews

5. Rating a customer based on reviews

- 24.2% of reviewers are posting helpful reviews.

- 0.14% of reviewers are posting consistent negative reviews.

Score (out of 7) = Average rating (out of 5) +
       1 if (Number of upvotes – number of down votes) > 0 else 0 +
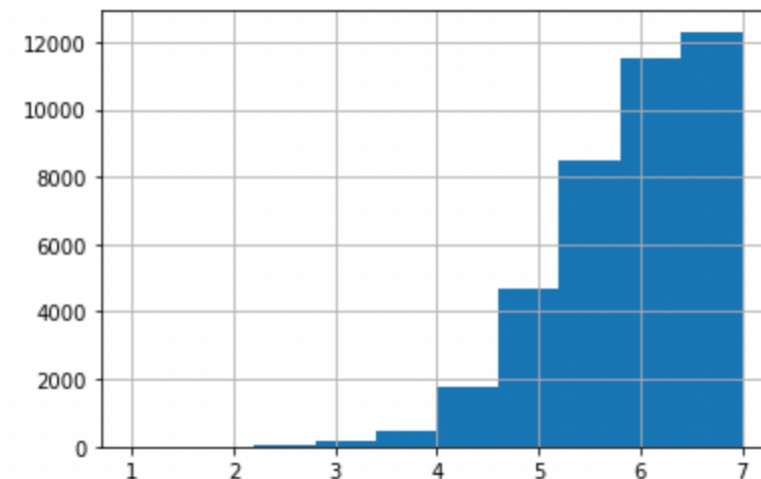   1 if review length between 50 and 706 else 0

```
reviewer_df['Reviewer_Score'].describe()

count      39387.000000
mean           5.878635
std            0.786865
min            1.000000
25%            5.400000
50%            6.000000
75%            6.500000
max            7.000000
Name: Reviewer_Score, dtype: float64
```

Distribution of reviewers' score

6. Product recommendation effectiveness

- 15% of people have purchased at least 1 product from the recommended items.

Five number summary of
orders from recommendations

```
:  count    6098.000000
   mean        8.952115
   std         6.033720
   min         5.000000
   25%         5.000000
   50%         7.000000
   75%        10.000000
   max       136.000000
Name: order_count, dtype: float64
```

Distribution of orders from recommendations

Order count
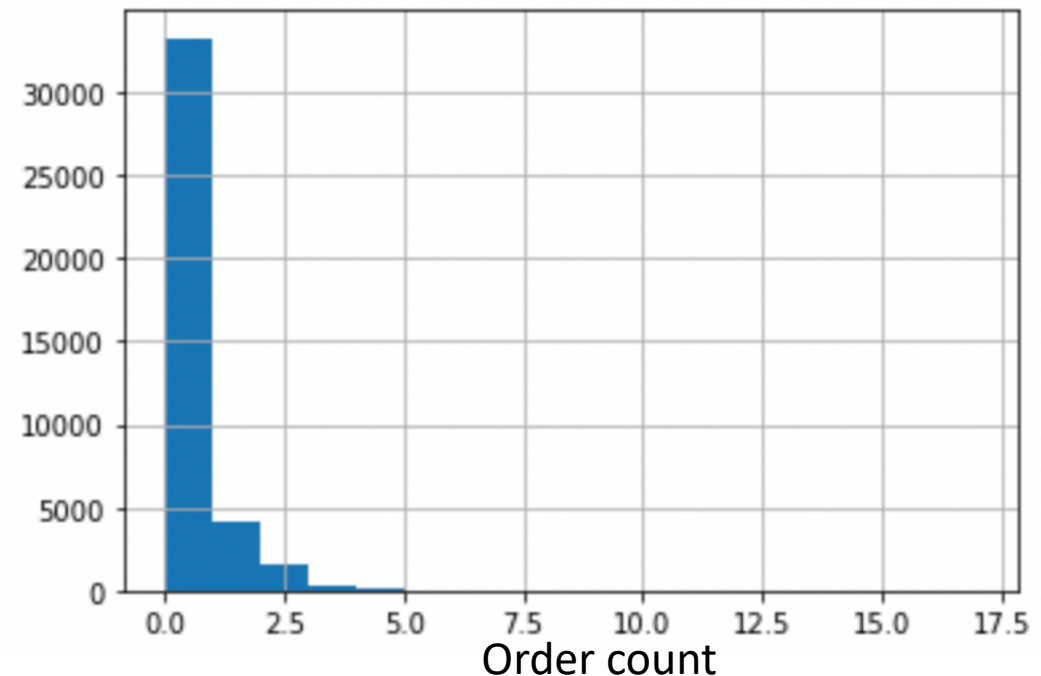
Identifying helpful reviews

- Assumption based on distribution: Minimum 5 upvotes/ downvotes are required to classify a review as helpful or not helpful

- Total 5% upvoted reviews are found helpful overall

- 0.4% downvoted reviews are not helpful

```
reviews['UpVote'].describe()

count       278677.000000
mean             1.145936
std              6.579834
min              0.000000
25%              0.000000
50%              0.000000
75%              1.000000
max           1200.000000
Name: UpVote, dtype: float64
```

```
reviews['DownVote'].describe()

count       278677.000000
mean             0.198664
std              1.041042
min             -2.000000
25%              0.000000
50%              0.000000
75%              0.000000
max            248.000000
Name: DownVote, dtype: float64
```

Action Plan - Products

1. Products showing negative rating trend should be black-listed/ moderated before further listing as they affect he reputation of Amazon/ seller.

2. The rating trend of the products can be utilized to prioritize the products in recommendation list.

Action Plan - Feedback

1. For issues related to size mismatch, the seller should provide a brand specific size chart or information like "order a size up than usual".

2. For issues related to quality, if it's a repetitive occurrence, then the product should be black-listed.

3. For issues related to differences with the catalog displayed, the seller should list a recent image of the original product in the site instead of a catalog picture.

4. Positive feedbacks are found most helpful than negative feedbacks. Hence the feedback section can be loaded with comments having positive polarity to influence purchase decision.

Action Plan - Customers

1. Top rated customer who post helpful reviews should be given a tag like 'trusted reviewer' as their reviews can influence the purchase decision of others.

2. Bottom rated customers who post consistent negative reviews should be moderated as they aren't found helpful.