Women's Clothing E-Commerce Reviews Analysis

Arlene Huang, Cathy C. Hou, Michael Haoxiang Guo, Lanzhen Zhang, Zekun Li

Topics

Topic

E-commerce is gradually becoming **main channel** for distributors to sell products.

Comments largely influence people's shopping **decisions**.

Text is one of the most popular data type for the sell-side.

Dataset

Women's Clothing E-Commerce Reviews dataset revolving around the reviews written by customers. The ultimate purpose of our research is to **approach customers precisely**.

• 23486 rows
• 10 feature variables

Problem statement

• What are the **words** being treated as **useful** information as customers rating the review?

• What are the **relationships** of review text/title and rating, rating and recommend IND, and review text/title and recommend IND?

• Whether **other listed variables** have **potential relations** to title/text?

• Can we **predict with words** alone whether or not a customer would **recommend** a purchased item of clothing?

A Quick Overlook of the Dataset

Age of the reviewer

The number of customers who found this review helpful

	C_ID	Age	Review	Rating	Recommend	PostiveFB	Division	Department	Class
1	1077	60	Some major design flaws I had such high hopes for th...	3	0	0	General	Dresses	Dresses
2	1049	50	My favorite buy! I love, love, love this jumpsuit. it's fu...	5	1	0	General Petite	Bottoms	Pants
3	847	47	Flattering shirt This shirt is very flattering to all due t...	5	1	6	General	Tops	Blouses
4	1080	49	Not for the very petite I love tracy reese dresses, but t...	2	0	4	General	Dresses	Dresses
5	858	39	Cagrcoal shimmer and this in my basket at hte ...	5	1	1	General Petite	Tops	Knits
6	858	39	Shimmer, s ... in ...	4	1	4	General Petite	Tops	Knits
7	1077	24	Flattering I ... t r...	5	1	0	General	Dresses	Dresses
8	1077	34	Such a fun ... e s...	5	1	0	General	Dresses	Dresses
9	1077	53	Dress look: ... run...	3	0	14	General	Dresses	Dresses
10	1095	53	Perfect!!! M ... e r...	5	1	2	General Petite	Dresses	Dresses
11	767	44	Runs big Bought the black xs to go under the larkspu...	5	1	0	Initmates	Intimate	Intimates

My favorite buy! I love, love, love this jumpsuit. it's fun, flirty, and fabulous! every time i wear it, i get nothing but great compliments!
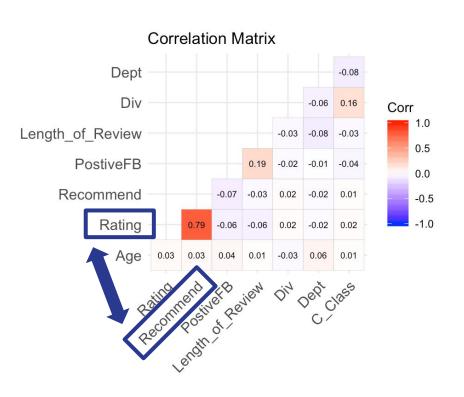
Cloth unique ID

Content of the review (Avg: 63 words)

Rating of 1~5

If they will recommend it

Belong to which division

Belong to which department

Belong to which class

Exploring the data

Correlation Matrix

Rating Distribution
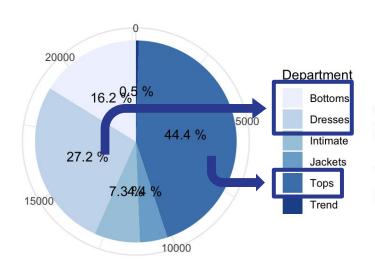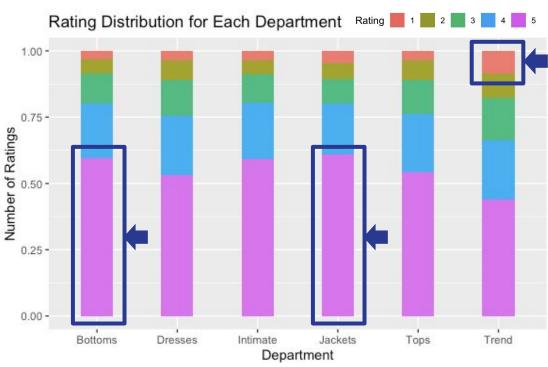
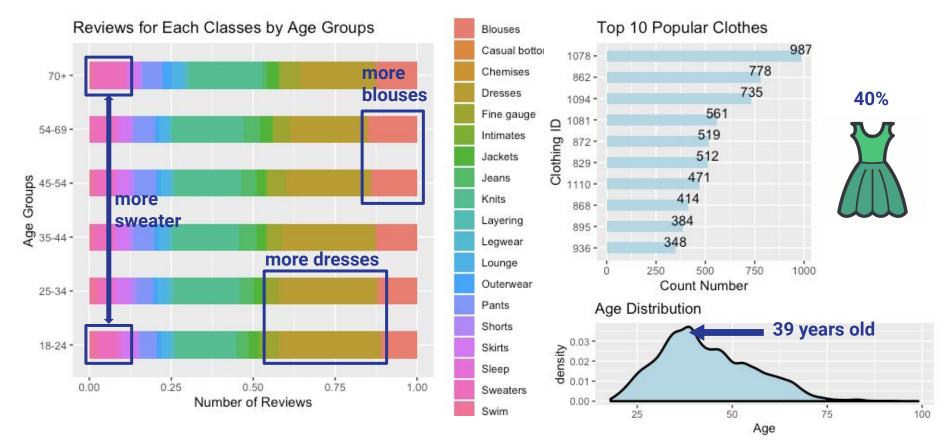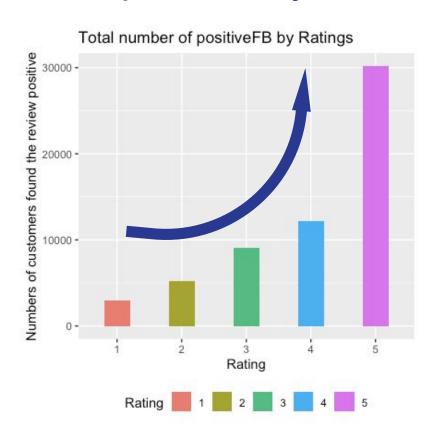# Descriptive analysis - Department

Percentage of Reviews from Each Department

87.8% reviews are for
Tops, Bottoms, Dresses
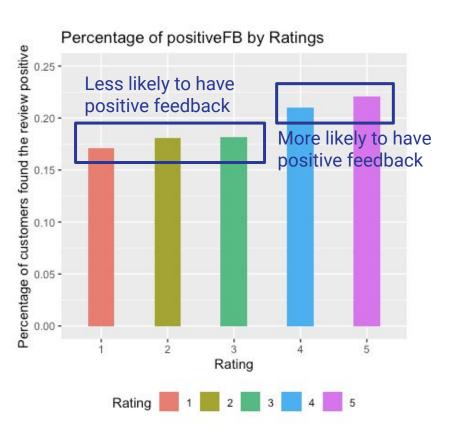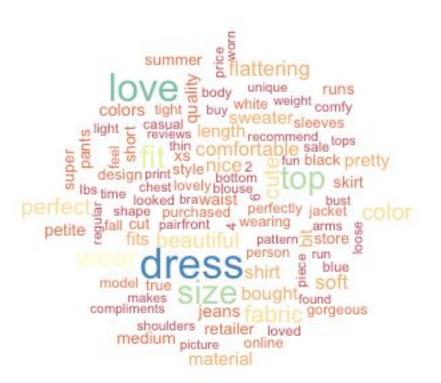
Descriptive analysis - popularity analysis

Descriptive analysis - Positive Feedback

Total number of positiveFB by Ratings

Percentage of positiveFB by Ratings

Less likely to have positive feedback

More likely to have positive feedback

NLP - High rating words VS Low rating words

NLP - What factors led to low ratings?

Top 8 Bigram Terms for Each Rating

Top 30 Words in Reviews with Low Rating

Targeting low-end mkt:
Improve fabric quality;
be careful of arm holes;
use loose waist design

Targeting high-end mkt:
cute design, perfectly fit size.

Predictive model - tree

- Typical score of review: 4.2/5.0

- reviews that contain "perfect", "comfort", "great", "love" are rating higher than if not; reviews that contain "return" are rating very low as around 3.2;

- If the review only says "perfect" without much other comments, the rating would be very high as around 4.7;

- RMSE: 1.00645.

Insights & suggestions

Suggestions

- **Bottoms department can become the next cash cow**, as there the volume indicates constant high demand of Bottoms and analysis shows that customers are less picky for Bottoms.

- Products in **Trend department are risky** in getting lower ratings. Only develop product in Trend department when there's higher return.

- Sweaters: Focusing on under 18 and 70+ people

- Blouse: middle-age people, need a bit fashion.

- Dresses: young people, more fashion and design, **dresses can become very popular suddenly,** so factories must be prepared.

- **A good fit can be extremely important for getting high ratings.**

Insights

- The majority of the cloth sales are in Tops, Bottoms and Dress department.

- Clothes designed for Trend department are getting far more negative reviews than others, meanwhile it seems that customers are less picky in Bottoms and Jackets departments.

- Children under 18 and seniors above 70 love sweaters, middle-age people are focusing more on blouses; young people favor dresses more.

- Nearly half of the top selling clothes are dress.

- People found 4 or 5 star reviews much more helpful than others.

- Typical score of review: 4.2

- Complaints are mainly regarding poor quality (especially in arm holes), or the cloth has odd fit.

- Positive reviews typically mentioned not only design are cute, but also the cloth fit perfectly.

Thank you