Merrell Topic Modeling Project

Topics:

The following are topics from Merrell's reviews. To do comparisons with topics of all Merrell reviews, I extracted a thousand reviews from top and bottom with the sorted sales ranks. Because of the smaller data size of two segments, I only created 20 topics for each one instead of 50.

Topics for All Merrell Reviews

0: great fit highly shoes comfortable
1: boot winter snow foot weather
2: size order half review fit
3: pair shoe merrell jungle mocs
4: toe box room foot shoe
5: water foot rain proof puddle
6: pair year shoe month wear
7: sandal merrell pair style summer
8: style merrell shoe size fit
9: shoe merrill fit great pair
10: shoe foot size width bit
11: shoe trail run mile running
12: shoe balance new merrell foot
13: shoe heel back blister ankle
14: pair price store shoe amazon
15: shoe summer pair great slipper
16: boot hike hiking mile moab

17: foot day shoe hour pain
18: merrell shoe quality comfort fit
19: merrell pair shoe clog encore
20: shoe star foot review people
21: shoe time period box break
22: shoe amazon size return order
23: shoe work day foot hour
24: shoe barefoot vibram toe feel
25: glove shoe barefoot merrell trail
26: quality fit material look thing
27: shoe pain foot barefoot mile
28: pair merrells year love one
29: toe shoe foot top box
30: shoe sole leather month rubber
31: merrell shoe product time quality
32: foot shoe arch support review
33: shoe year pair son foot

34: jean day dress pant work
35: shoe foot problem lot time
36: size shoe woman fit men
37: sock size bit fit thicker
38: shoe weight gym foot workout
39: trip shoe day pair lot
40: traction surface rock trail sole
41: color shoe picture brown blue
42: support shoe arch foot heel
43: shoe sneaker weight foot support
44: shoe lace chameleon tongue foot
45: merrell pair shoe year model
46: boot calf ankle zipper merrell
47: foot strap sandal fit support
48: shoe pair merrell insole sport
49: pair shoe husband wife gift

Topics for Top 1000 Sales Products

0: shoe foot support arch problem
1: shoe sole month merrell lace
2: shoe foot toe box heel
3: foot sandal strap fit support
4: shoe pair merrell year foot
5: shoe foot support fit feel
6: foot shoe day work hour
7: pair merrell shoe jungle year
8: size shoe half order sock
9: shoe glove trail barefoot running
10: shoe traction foot water trail
11: shoe toe size foot box
12: color shoe fit merrell picture
13: shoe merrell pair size quality
14: boot foot time day sock
15: pair shoe year husband time
16: shoe hiking pair day trip
17: shoe pair great fit support
18: sandal pair merrell color style
19: shoe amazon merrell store size

Topics for Bottom 1000 Sales Products

0: shoe foot day support arch
1: shoe foot support problem sock
2: shoe heel foot ankle way
3: pair merrell year merrells moab
4: shoe water foot day winter
5: toe foot shoe box bit
6: color jean clog look fit
7: shoe size merrell amazon product
8: boot pair lace year merrell
9: pair shoe year store time
10: shoe pair merrell year style
11: shoe merrell color style size
12: shoe barefoot glove running merrell
13: great shoe husband shoes size
14: size shoe fit order half
15: shoe pair day merrell foot
16: sandal foot merrell color summer
17: boot winter foot sock weather
18: shoe trail sole rock sol
19: boot day foot mile hike

Topic Description

Comparing the topics of two segments and the topics of original reviews, there is a high level of overlap between them. Therefore, only descriptions for original topics are showing below. The topic descriptions are based on the extended words of each topic and text reviewing from the classified documents.

Topic Descriptions for All Merrell Reviews

Topic 0 – Shoes are highly comfortable with great fit
Topic 1 – Boots are great for foot when snowing
Topic 2 – Shoes ordered are half size bigger or smaller
Topic 3 – Merrell jungle mocs
Topic 4 – Toe box area has plenty of room
Topic 5 – Shoes are waterproof in raining day and in puddles
Topic 6 – Shoes can be wore daily for months and years
Topic 7 – Merrell sandal pairs'summer styles
Topic 8 – Styles and colors of Merrell's fit and comfortable shoes
Topic 9 – Merrell shoes have great quality and fit anyone
Topic 10 – Merrell shoes' fit width makes consumers feel they are in bare feet
Topic 11 – Shoes are used to running miles on rock road
Topic 12 – Merrell shoes are like the mix of Nike and New Balance
Topic 13 – Blister problem on ankle caused by heels
Topic 14 – Sale time of one pair of shoes on Amazon's store
Topic 15 – Slippers greatly support walking around house in summer and winter
Topic 16 – Moab shoes are not waterproof but great for hiking without rain and puddles
Topic 17 – Pain of wearing shoes for hours at the end of a day
Topic 18 – Price and quality of comfort and fit Merrell shoes
Topic 19 – Merrell Encore Breeze and Merrell Clog
Topic 20 – Reasons of stars and reviews about shoes over time
Topic 21 – Shoes need no breaking-in time right out of box
Topic 22 – Return shoes bought on Amazon because of size issue
Topic 23 – Walking in shoes for hours a day for job
Topic 24 – Shoes with vibram makes barefoot feeting
Topic 25 – Barefoot feeling from Merrell Pace and Merrell Trail Glove
Topic 26 – Quality of material affect fit and look of shoes
Topic 27 – "Barefoot" shoes cause pain if walking or running for miles
Topic 28 – Love Merrell pairs for more than one year
Topic 29 – Ample of room in toe box makes toe box comfortable
Topic 30 – Consumers worry about the connection between rubber sole and upper leather
Topic 31 – Merrell's quality of products and customer service over time
Topic 32 – Reviews of shoe arch support for feet
Topic 33 – Condiiton of shoes for kids after a year
Topic 34 – Wear with jeans for daily work
Topic 35 – Shoes for people with foot problem and people waking a lot
Topic 36 – Fit of women and men shoes
Topic 37 – Shoes fit even with a bit thicker socks
Topic 38 – Shoes support weight tranning and workout in gym and provides barefoot feeling
Topic 39 – A pair of shoes for walking a lot per day during trip
Topic 40 – Traction surface sole for walking on rock trail

Topic 41 – Picture of shoe with color brown and blue
Topic 42 – Shoe with arch spport for foot with problems caused by heels and fasciitis
Topic 43 – Light-weight sneakers with good food support
Topic 44 – Lace and shoe tongue of Merrell Chameleon
Topic 45 – Some consumers are fans of Merrell brand for it's shoe models
Topic 46 – Merrell calf boots and ankle boots with zippers
Topic 47 – Foot strap supports sandles to fit
Topic 48 – Sport shoes and hiking shoes' insoles and inserts
Topic 49 – Shoes as Christmas gift for husbands, wives, and daughters

Preprocessing steps:

Before formally starting the preprocessing, I loaded the merrellreviews.json file that is created from previous steps, then use for loop to extract content of the *reivewText* variable for every review and append each into an empty list called *merrelltext*. After having a list of all review texts, I converted it into a dictionary and then transform it into a corpus for further use. Next, I did a histogram to show how the length of texts is generally distributed. I saw that most reviews are short and less than 1000 words. This means I can do some text preprocessing to shorten those long reviews instead of building another topic model for them.

There are several steps needed for the text preprocessing in order to improve the quality of topics. The first step is to create a new variable containing a new data structure the initial corpus for preprocessing. The second step is to generate tags of part of speech for each word because there are some parts of speech tend to be more informative than other, such as nouns. The third step is to unify the format of words using lemmatization and lowercase transformation, which makes words cleaner. To cleaner the texts, I removed all special characters and stop words. After the first round of preprocessing, I found that the texts were not as clean as expected. Therefore, data needed more processing, including extracting only nouns, diminishing numbers and words with length shorter than two, and removing common words. The two last steps before creating the model are generating vocabulary array and document-term matrix (DTM).

Model steps

The method I use is LDA modeling from *tmtoolkit* package. Before building the model, I did several settings to varnish the output: using *logging* package to prevent verbose topics and *warnings* package to not display warning message. Besides, setting DTM and model parameters upfront makes them convenient to manipulate. After I built the model, I tried to evaluate it for several times with different parameters and decided to create 50 topics for

Merrell. In order to flexibly view words for each topic, after printing topics out, I transformed all topics into a data frame, and then converted it to a more viewer-friendly version as the given sample.

Marketing & Product Insights

Positive Attributes
Topic 0 shows that people love Merrell shoes that are comfortable and fit well. Topic 2 shows that Merrell boots are great choices for snow day. 5 shows that people confirm some models are really waterproof during rainy day. Topic 21 shows that people do not need to spend time to adapt Merrell shoes. They are comfortable from the beginning. Topic 37 shows that people are satisfied with the elasticity of shoes that can fit with thicker socks. Topic 42 shows good arch support of Merrell shoes. Furthermore, Topic 43 shows that people think those light-weight sneakers give good support to their feet. All positive attributes including not mentioned ones should be considered to emphasize in product promotion.

Negative Attributes
Topic 2 and Topic 22 shows some people comment that shoes are not true-to-size, though some other topics illustrate good fit of Merrell shoes. Merrell can do further investigation on what models collect the most negative review about size. Topic 30 extracted another concern that consumers are worried about the connection between rubber sole and leather is not firm enough. Those negative topics can give Merrell a good reference of what needs to be improved or solved.

Purchase Occasions
Topic 39 shows that Merrell shoes are suitable for people who need to walk for long time. Topic 34 shows that the products are often paired with jeans and used for daily work. Topic 49 shows that Merrell shoes are good choices as Christmas gifts for families. Topic 42 shows that the products are good substitution for people with fasciitis and other foot problems caused by heels.

Product Development/Improvement Ideas
In general, consumers' feedback on quality is positive. However, there are also negative reviews such as the Topic 30. This show that the quality of Merrell's products is not stable. The brand can improve the process of quality inspection to eliminate the quality concern. Besides product quality, Topic 31 also shows the concern of customer service. Also, there are not much positive reviews for product styles of this brand. Furthermore, the keywords

style and *color* appear relatively frequent in the topics of reviews for products with bottom sales ranks. Therefore, Merrell can think about putting some fashionable features into their new products, or at least creating some attractive color designs.

Pricing Suggestions

Only Topic 18 mentions price. However, price does not appear in the top five words. This mean that price is only a small concern of people. Some people may think prices of Merrell's product is somewhat expensive. However, it is not a main problem for the brand. As I know from the topics, Merrell shoes more aims at comfortability and practicality instead of fashion and luxury. Thus, I recommend Merrell to keep the pricing at levels that most people can afford than too high.

Models Mentioned

There are several models are directly mentioned in the original topics, such as Jungle Mocs, Encore Breeze, Moab, and Clog. Some of them also appears in the segment topics. Jungle Mocs(Topic 7) is in the segment of top thousand. Moab(Topic 3) and Clog(Topic 6) are in the segment of bottom thousand. Though the topic may not necessarily show the drawbacks of these two models. Digging deeper into the data, I found that some people think that Moab is poor on supporting ankles; the sole of it is too hard; also, it makes people easier to fall on the trail. For Clog, some people feel that it's the inferior of Jungle Mocs; it does not well support foot problem. In order to increase the sales of these two models, Merrell can consider eliminating these problems while thinking of target audience. In my opinion, the entry points could be design and material. However, these models do have advantages. For example, Clog is a good model for people with no food issue to stand for a long time. Maintaining competitive features is significantly important.