

Tonie Enriquez Caponpon

2000985H

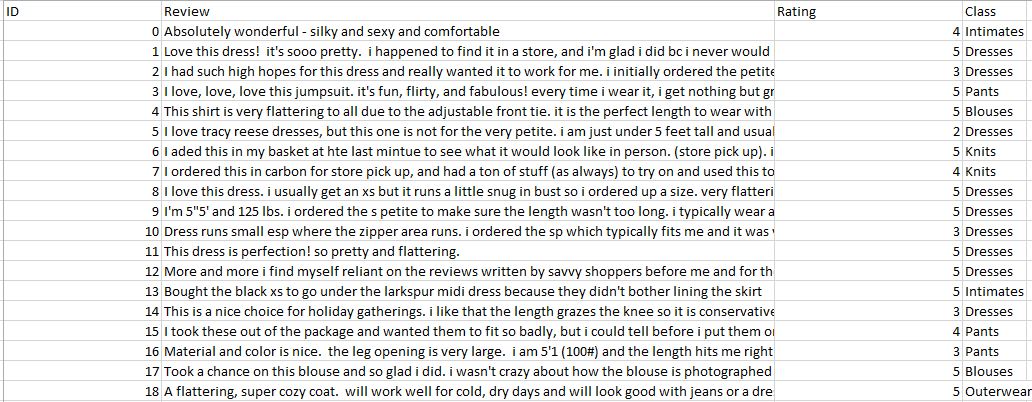
P01

Group 1

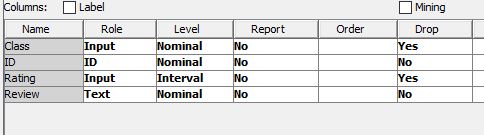
6 June 2022

1. Dataset

I chose this dataset from Kaggle, named “Womens Clothing E-Commerce Reviews”



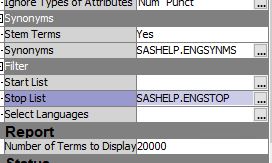
1. File Import Node - Importing the dataset into SAS



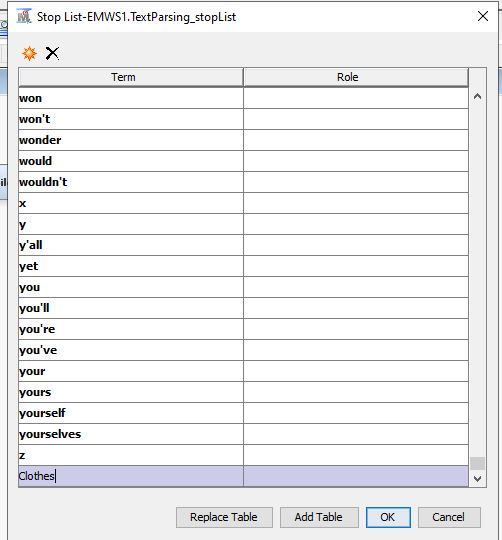
Dropped ID and Class

1. Text Parsing Node

* Formed a stop list to remove unnecessary English words.
* Also set “Stem Terms” to “Yes”



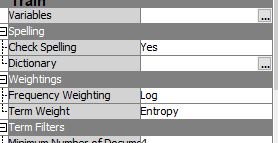
* Also added “Clothes” to set of “stop list”



1. Text Filter Node

Term Weight - Mutual Information because categorical target/Class

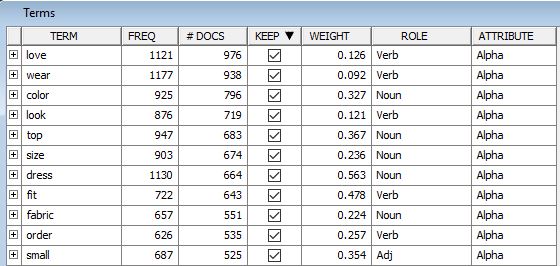
Frequency Weighting – Log



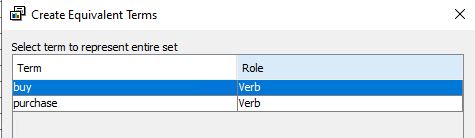
FINDINGS:

* 1. Top 10 terms frequently used in the reviews!

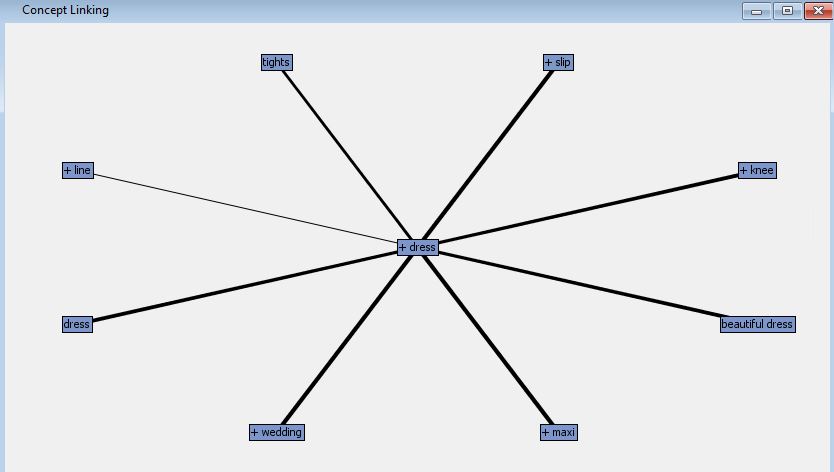
From here I can infer that most of the reviews are positive as the word ‘love’ appears. The customers mostly love the products they have purchased from the shop and most of the reviews talk about the products. Most are neutral words but the word ‘love’ suggests that many would be positive reviews.



* 1. This word has been removed.
  2. Created a synonym for purchase/buy

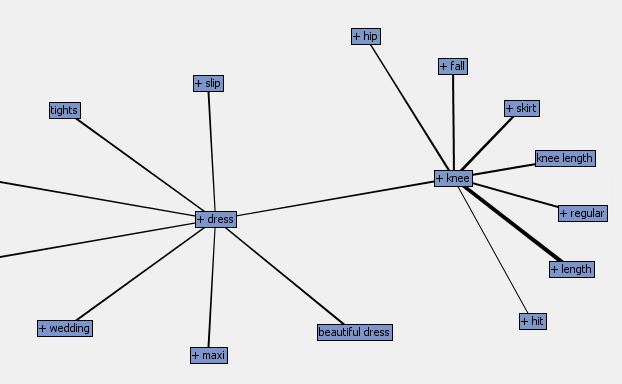


* 1. Concept links on the word ‘Dress(noun)’

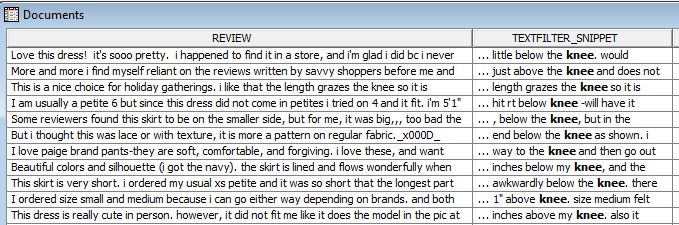


As we look at the bigger picture, we can see that the reviews about “dress” talk about how the dress looks like and how they feel about the product.

* 1. Further concept links on dress – expanded on “knee”

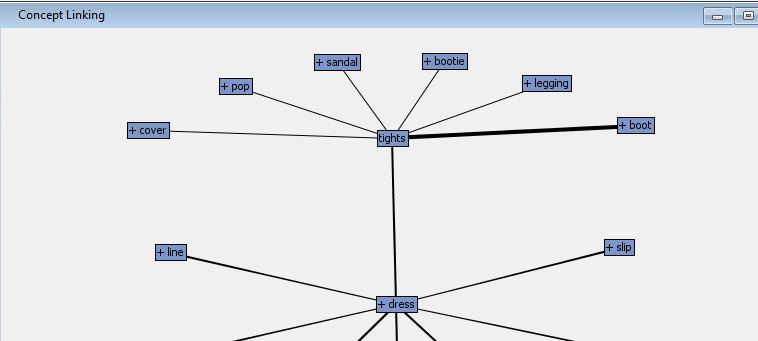


Explanation on link between knee and DRESS



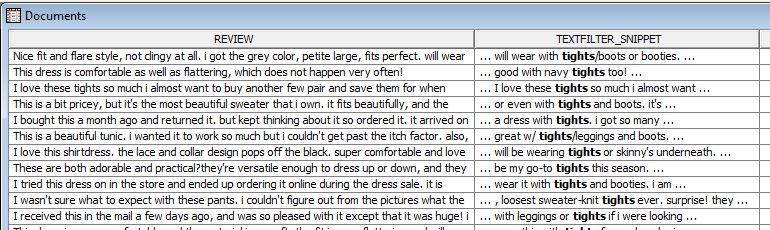
Under “FALL” there are words like “knee” and “ankle” I believe the reviews are talking about the length of the dresses that they bought and love. For example, the first one talks about the dress ending a little below the knee, and the second review talks about the dress ending just above the knee. People are talking about its length.

* 1. Further concept links on dress – expanded on “tights”

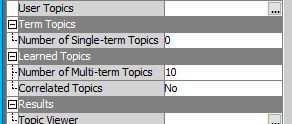


Explanation on link between TIGHTS and DRESS

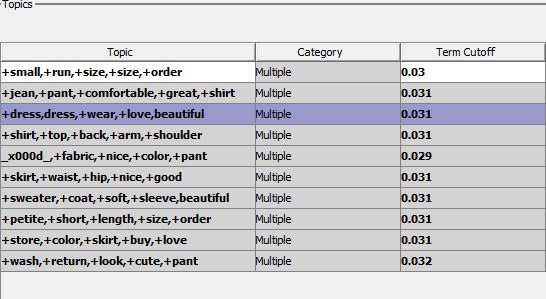
As we can see, there are boots and leggings that show up in the expanded link. This suggests that people may have been talking about how the dresses they have purchased look good with tights, and/or boots! We have evidence of this:



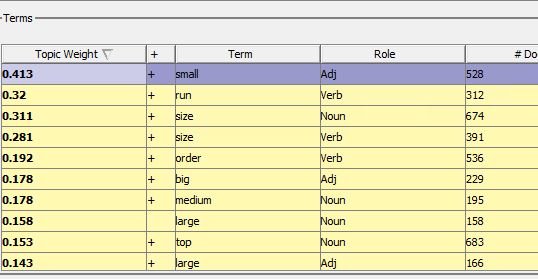
1. Text Topic



* 1. These are the 10 topics! I will be looking into 3 topics.

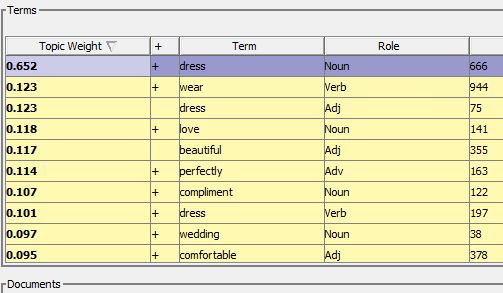


* 1. FIRST TOPIC:



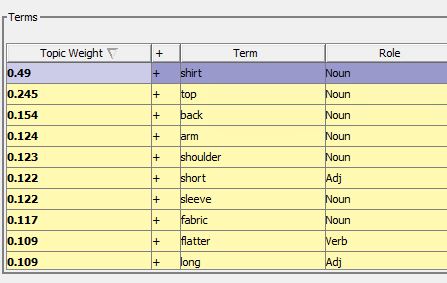
From the terms in TOPIC 1, I can infer that the reviews under this cluster talks about the sizes of the dress, tops or bottoms that they purchased. Most of the words fall under the size category like “medium” or “large”. I believe this topic mostly includes reviews of people talking about how the products they bought fit them!

5.3 SECOND TOPIC:



From the terms in TOPIC 2, I can infer that the reviews probably talk about the design of the clothes, such as its quality and how it can be worn. The terms “beautiful” suggests that people are talking about how beautiful the dress is, for example. The reviews under this topic talks about how the clothes look on them and their quality, as suggested by the words “wear”, “beautiful” and “perfectly”!

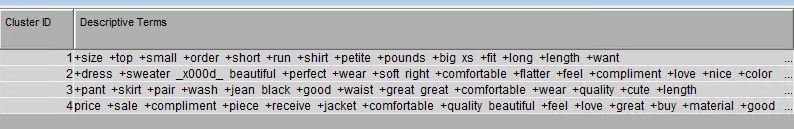
* 1. THIRD TOPIC:



From the terms in the third topic, I can infer that the reviews under this topic are probably talking about tops/blouses/shirts! As suggested by “shoulder” and “short” it could be talking about the length of the “sleeves” and also the fabric of these tops!

1. Text Clustering Node

I used the text clustering node and created 4 clusters.



CLUSTER 1:

The first cluster seems to be talking about the size/fit of the clothes the customers have purchased. This cluster may also talk about the length of bottoms as suggested by the terms “length” and “short”. As suggested by “big xs”, this could also mean that some people think that the sizes do not highly match the expectations of sizes by the customers, meaning that an XS may feel like an S, or maybe and M feels like a size S!

CLUSTER 2:

The second cluster seems to be talking about the tops/dresses and how pretty or beautiful they look on the customers. The designs of the clothes being sold are probably beautiful, or perfect or may even feel comfortable, as suggested by the terms in this cluster! It may have a nice colour or may compliment the customer’s body perfectly, as suggested by the terms “compliment” and “flatter”!