

Generation Luxe Analysis Report

XTOL: Data Science & Big Data,

Task 4: Evaluate information needs for a new project

Author: Esteban Andrés Villalobos Gómez

Date: February 6th, 2020

Synopsis

Generation LUXE is a high-end apparel company focusing on young women. The company is committed to offer the best product from well establish as well as emerging designer brands.

Alert! Analytics has been asked to look at a set of questions the customer would like to answer using data analytics and machine learning (ML) techniques, and we have done an analysis of the requirements as we all a draft proposal on the information needs and how to address the customer's questions.

Questions analysis

Pre-requisite: In order to answer some of the questions below, we will require access to Generation LUXE's sales records, inventories and customer demographics.

Limitations: please be aware that any question we might be able to approach an answer to, will use data that might contain biases, and that doesn't cover the entire population. There is a digital divide that excludes usually elder and low resources people. However, this should not be of great concern, as the target population of LUXE is wealthy and young women, who have access to the internet and have very active digital lives.

Which brands are the most popular?

A brand could be popular in a positive or negative way. **We can answer** this question by looking at brand references on popular shopping sites, both from "Generation Luxe" as well as other resellers/retailers, blog posts, and social media.

Once we identify the mention of the brand, we will try to classify the sentiment to it, by performing a sentiment analysis based on adjectives used about the brands, and/or their products. We would not search the entire internet for reviews, we would focus on specific social media networks and influencer accounts.

Some of the features (adjectives and close terms) we would look for regarding the brands and apparel:

- **Classification:** brand name
 - **Sub-classification:** top products from each brand.
- **Positive** terms (initial draft):
 - Famous, aggressive, proud(ly), appreciated, nice, good | great (looking), elegant, classy, cool, great...fit, deserved, worth(s)...it, elegance, fashion, premium...materials, perfect, beautiful (collection), historical (identity), luxury, minimalist(m), high (end), appreciated.
 - Examples taken from actual reviews:
 - *"But when you are making **classic**, detailed clothing for the **perfect** urban look, you let your work speak for you"*
 - *"The Danish company has a finger on timeless style and simple lines that work in any season. **Minimalist** is a word used a lot in the fashion world, but Norse Projects are one of the few who have **minimalism** down-pat. Wear it as streetwear, techwear, or just as **great-looking** clothes that work – the label will lead urban contemporary fashion for years to come."*
 - *"the French brand Lacoste has a well-established reputation in the world of **high-end** ready-to-wear apparel. Made **famous** for its polos, Lacoste was founded in 1933 by the tennis player René Lacoste, whom an American journalist called at the time "The Alligator" for his **aggressive** playing style. Since then, the **famous** animal has been adopted as the logo **proudly** seen on all Lacoste products, which are now **appreciated** internationally."*
- **Negative** terms (initial draft):
 - Terrible (company | <brand-name>), awful , wrong, too...large, too...small, low...quality, expensive, cheap(ly)..made | materials | fabrication | quality, regular, poor, never...purchase, tired, ripped, repaired, not...expected, stiff.
 - Examples:
 - *"**Terrible** company and customer service! I ordered this jacket 3 times (XXL) and they sent me the **wrong** size 3 times (L). I am not sure why they do this! I talked to them on the phone to make sure that they should send the correct size. NO luck! I will **never ever purchase** anything from Lacoste anymore. I am really **tired** of dealing with them."*
 - *"**Cheaply made?**
Color: Misty Rose
Not sure what happened but both my pockets **ripped** along them seams and I spent over \$80 having it **repaired**."*
 - *"**Not** the color as **expected**
Color: Medium Grey Heather
Was **stiff** and the gray was shiny and bright."*

- **Neutral** terms (initial draft):
 - Fabric, color, not exactly, etc.
 - Examples (from 3 and 4 star reviews):
 - *“Not the **color** as expected”*
 - *“The **color** was **not exactly** to me as it appeared in the pictures and was a bit smaller than I like.”*

Which brands are our customers more likely to buy based on their attitudes toward the economy?

We believe **this question should be taken out of the scope** of the automatic analysis, since the target population of the brand is wealthy young women, who tend to consider that the general attitude of the economy is not affecting their shopping habits.

Our hypothesis is that there is no correlation between consuming habits of the wealthy and the state of the economy (unless they go bankrupt of course). If the customer considers this is still important, we would focus on performing hypothesis tests.

Which brands are trending upward in popularity?

In order to answer this question, we will collect historical Internet information, from sites like common-crawl, and social media sites, by acquiring the required licenses to crawl public information.

We will look for the patterns defined in the first question, and assess popularity given specific time ranges (by summarizing daily, weekly and monthly) information.

Once the information has been summarized, we will use it to do a time-series analysis, in order to assess each brand's popularity over time.

Which brands are trending downward in popularity?

We will perform the same technique describe the previous section, in order to answer this question.

Which brands have an inconsistent trend in popularity?

We need to circle back this question to Luxe, since it is not clear for us what the customer is looking to for. We believe the question is ambiguous.

Is LUXE trying to find brands/products with very positive popularity and low sales, or products with low or negative popularity, but extremely high sales?

Which of these brands look best on our customers?

We won't be able to answer this question using only machine learning techniques. The question refers to a concept that is quite subjective and dynamic, as trends in fashion change very fast. This question would be better analyzed with social analysis techniques, that are focused in specific groups and gather more context than looking at reviews, or sales information.

Which brands are the most popular among fashion critics?

We will answer this question by reducing the focus of our content acquisition (content crawl) to specific fashion critics pages, and social media accounts. Then we will do a sentiment analysis, using the terms defined in the first question, and compare that to a profile created for each one of the main fashion critics.

Another question we can try to answer with this analysis is: **"Are fashion critics reviews biased towards a specific brand?"**

We suggest not only focusing on fashion critics but looking into the main influencers over the social networks mostly used by the target population.

For each product category (dresses, skirts, tops and pants) which brands carry the most popular items?

In the first question we looked for the most popular products for each brand.

In order to predict the size of the inventory, we would first need to correlate historic sales information from Luxe of **similar items** with the most popular new products. With this data we would then create a prediction model, which can help LUXE determine the quantity of each item that should keep in the inventories.

Which retailer will be carrying the most exciting color combinations next season?

This is a very ambiguous question to try to answer with ML, for instance, what does it mean an existing color combination?

Usually only looking at data would not be useful to answer this question, since this is a very subjective question, which requires a lot of contextual (social) information.

This question will be discarded from the automatic analysis and we recommend using social analysis techniques to try to approach this.

What qualities do customers like about each brand?

We can answer this question by taking the positive reviews and comments of each brand, and extending our sentiment analysis to look for specific qualities in the product like:

- Warranty, fabric, color, size, durability, material, price

What qualities do customers dislike about each brand?

We will answer this question by looking at negative reviews and correlate them with the features defined in the previous question.

Which of the brands are popular because they are well liked by celebrities?

We will not be answering this question using machine learning only models, but we can do a hybrid approach in which we look at the results on the time series analysis on brands that are upward in popularity and trying to correlate which brands have products that were used by celebrities during the same periods of time.

The results of the research can then be used by marketing analyst to decide whether or not to sponsor specific celebrities at red-carpet activities, in order to boost sales over a specific product.