1. Task Description

Hello,

Thank you for your efforts on the Helio project.

We have a potential project with a new client that I would like you to start working on right away. To get you started, let me share with you what I know about the client's needs and explain what you need to do to get the project off the ground.

Our client, Generation Luxe, is a retailer of high-end, fashionable apparel targeted at young, affluent women. They have a number of stand-alone boutiques within selected department stores.

Julia Chan, the buyer for Generation Luxe, has approached Alert Analytics to explore how an analysis of web data may enable her to make better purchasing and marketing decisions for some emerging brands she is considering adding to next season's collection.

In the Generation Luxe Analysis Request (attached to this email), Julia explains her initial thinking on the business decisions she needs to make and for each decision, a list of questions she believes need to be answered to make those decisions.

For this project to be successful, we need to ensure that the questions we set out to answer can actually be answered through data analytic methods. Clients often view data analysis as a magical process that can answer any question they dream up---however, it has definite limitations and one of our responsibilities to our clients is to set accurate expectations about what question data analysis can and cannot answer and help our clients frame feasible and meaningful questions.

This is where you come in— your task is to review the business needs and assess the associated questions in the Generation Luxe Analysis Request to identify any questions we can't accurately answer using a sentiment analysis of data on the web, either because it is not feasible to answer the question using data analytic methods or because we need additional information to better define the data we will capture and analyze to answer the question.

Please explain your assessment of the questions she has enumerated in a brief 3-5-page report that I can share with Julia.

For each of the questions that you believe can be answered by a sentiment analysis, sketch your approach to answering the question. At this stage, I don't need a specific enumeration of all the features that would be required, but I would like to review your thoughts on the types of data patterns you plan to capture. Please provide a few examples for each type. Also clearly state which questions can't be answered, why, and what information is needed to clear up any ambiguous questions. If you have suggestions to better address, her business needs by adding new questions or rephrasing the original questions please include those suggestions in your report.

Best,
Michael Ortiz
Senior Vice-President
Alert! Analytics Thanks,
Danielle

2. Task Solution

The document "Request for an Apparel Brand Analysis" was provided to answer some questions submitted by Generation Luxe; the main goal of this document is to address their objectives for improving business for the next fashion season.

In order to achieve this goal, the response for each questions / business context will be listed in section 3 – Information Needs Proposal.

3. Information Needs Proposal

The information needs proposal is divided in five sets of questions (Objectives), in which each set of questions has his own business context.

1- Objective I
2- Objective II
3- Objective III
4- Objective IV
5- Objective V

The tables below (Table I to V), contains the business context, question set and the proposal to address each question, along with a brief summary of the objective.

Business Context	Questions	Proposal
Objective I: We need to know which of the candidate brands will be the most popular in the coming season and which of them will have the highest likelihood of remaining popular, so we can decide on the brands to add to our collection next season.	Which brands are the most popular?	Greatest quantities of positive reviews across multiple sources, greatest number of sales
	Which brands are our customers more likely to buy based on their attitudes toward the economy?	Buyer confidence values covering previous 5 years, Brand purchase analysis, Market Basket Analysis
	Which brands are trending upward in popularity?	Positive review scores, positive sentiment social media (FB, Tweeter, web pages)
	Which brands are trending downward in popularity?	Sentiment analysis (FB, Tweeter, web)
	Which brands have an inconsistent trend in popularity?	Review scores, date
	Which of these brands look best on our customers?	Sentiment text analysis (FB, Tweeter, web pages)
	Which brands are the most popular among fashion critics?	Product decision engine for retail such as Maker / Sights.

Table 1 – Objective I

Which brands are the most popular? To determine the most popular brands I would want to correlate the greatest number of positive reviews across multiple sources with the greatest number of sales for each brand.

Which brands are our customers more likely to buy based on their attitudes toward the economy? My approach would be to track consumer buying habits during the last 5 years while comparing it to the "Índice de confianza sobre la actividad económica" from BCCR, which expresses consumers' optimism with the current state of the economy. Additionally, the digital print of the recommendation engines shall be correlated with Confidence Index.

Which brands are trending upward and downward in popularity? To determine the most popular brands I would conduct a sentiment analysis in social media such as Facebook, Tweeter and web pages. **Which brands have an inconsistent trend in popularity?** Similar to finding the downward and upward popularity trend analysis the task of finding inconsistent trends will use the same process of charting the positive, neutral, and negative sentiments of consumers while also factoring in the change in upward and negative trending of the data over time.

Which of these brands look best on our customers? The solution to this question will rely heavily on customer review text as the source for the data we will need to tabulate sentiment scores. A major difference I see compared to previous solutions, such as the one for "Which brands are the most popular?", is that the word usage we will be tracking will need to be specific to the topic at hand.

Which brands are the most popular among fashion critics? I will use a product decision engine which uses data analytics combining factors such as search queries, social media activity, e-commerce sell-throughs and consumer feedback to provide clues into what is most likely to become a trend.

Business Context	Questions	Proposal
Objective II: For each of the product categories we carry, we need to know which items will be the highest sellers, so we can order the right inventory of each brand for each product category.	For each product category (dresses, skirts, tops and pants) which brands carry the most popular items?	Count of purchases based on brand and item (Tops are popular, Brand X carries the most popular tops)

Table 2 – Objective II

For each product category (dresses, skirts, tops and pants) which brands carry the most popular items? To find the brands that carry the most popular items we would need to make a tabulated sentiment score, tracking the different dresses, shirts, tops, and pants that we find through product analysis to be most popular.

Business Context	Questions	Proposal
Objective III: We need to know which retailer may give us additional insight on what color combinations to consider carrying.	Which retailer will be carrying the most exciting color combinations next season?	Color options for each item and each brand

Table 3 – Objective III

Which retailer will be carrying the most exciting color combinations next season? What we could do is chart the use of different color combinations across multiple high fashion retailers. The challenge here would be determining what combinations are the most exciting for the next season. Per fashion trend tracking sites.

Business Context	Questions	Proposal
Objective IV: We need to know what characteristics customers value most about each brand, so that we can design advertising and displays to emphasize those characteristics.	What qualities do customers like about each brand?	Fit, color, price, availability, material.
	What qualities do customers dislike about each brand?	Fit, color, price, availability, material.

Table 4 – Objective IV

What qualities do customers like about each brand? Unlike measuring the popularity of different brands, the charting of what customers like will require a more tailored topic set.

What qualities do customers dislike about each brand? Similar to the process I took for finding what customers like about each brand. The task of finding what customers dislike will require us to make tabulated sentiment scores based on negative traits.

Business Context	Questions	Proposal
Objective V: We need to know if celebrities have influence on the popularity of a brand so we can consider special promotions linked to red-carpet events.	Which of the brands are popular because they are well liked by celebrities?	Match celebrity mentions on social media and/or commercials to sales

Table 5 – Objective V

Which of the brands are popular because they are well liked by celebrities? In order to correlate the relationship between celebrity mentions and endorsements of brands we will need to track their activity on social media platforms (e.g. Twitter, Instagram, Facebook, YouTube, specialized websites) for brand awareness. The other half of the correlation work will be easier as it will be like what will be done in previous tasks by recording when purchases are made of various branded fashion items.

4. Recommendations

The processes described above requires additional work, but these just a portion of a typical data science project. As IBM described in Figure 1, the Foundational Methodology for Data Science consists of 10 stages that form an iterative process for using data to uncover insights. Each stage plays a vital role in the context of the overall methodology.
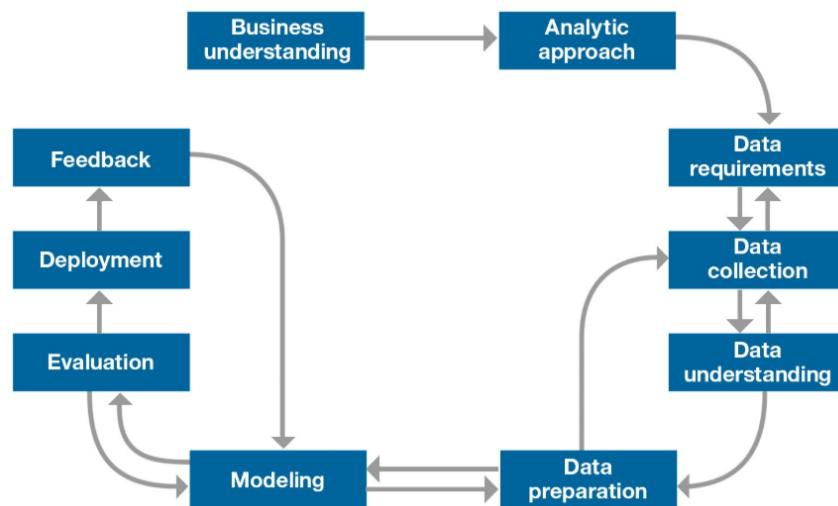

Figure 1 – Foundational Methodology for Data Science.

In this section, we only covered the "Data Requirements" stage. Therefore, I strongly recommend reviewing the other stages to get a clear picture of the iterative process to solve business problems with data science.