

Case: Predicting Consumer Tastes with Big Data at Gap - The Art and Science of Fashion

Team 6

Avishek Dasgupta
Roshan Raj Singh
Sai Mona Duvvapu
Sohan Kumar Sahoo

Business Problem

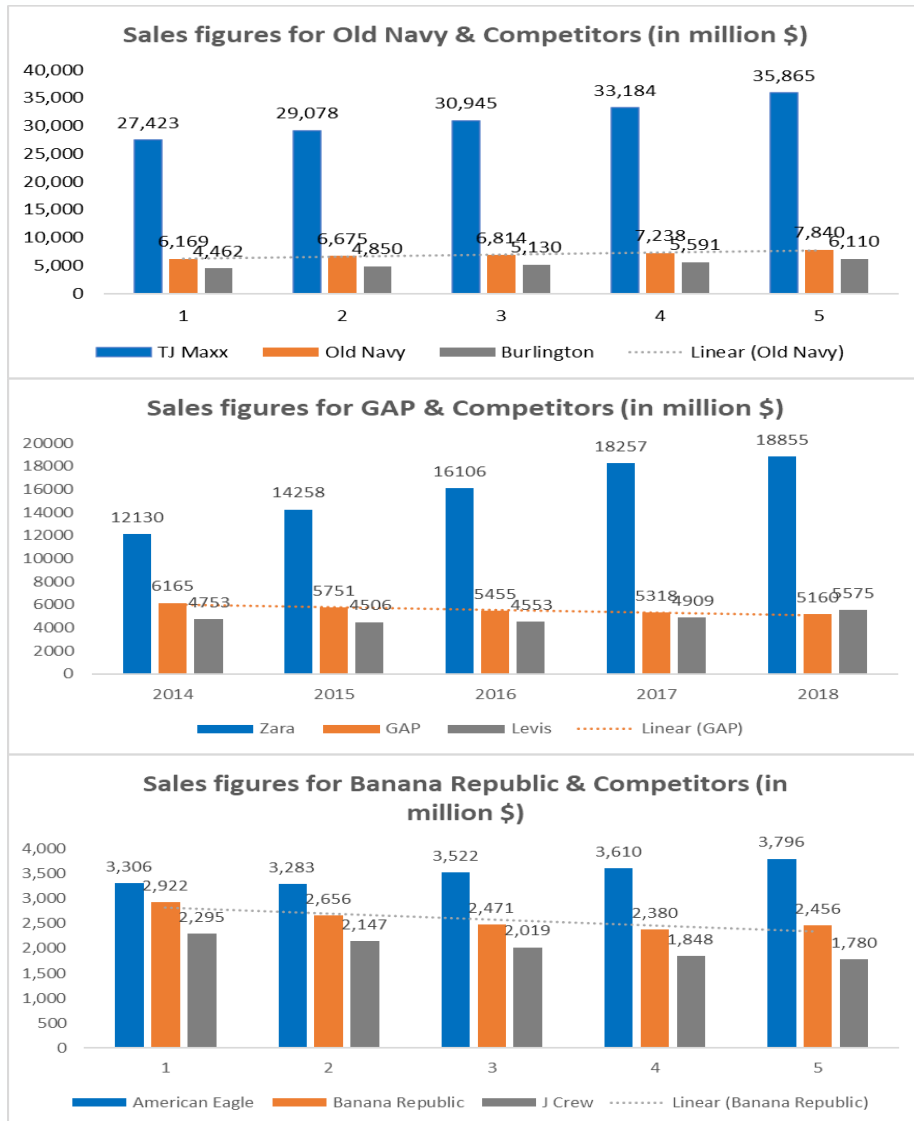


**Can machine intelligence rule
over human creativity in the
fashion designing process?**

OLD NAVY



Was Peck Right?



Sales Figure Analysis

1

Old Navy's sales grew between 2014 and 2018, but its direct competitors grew faster

2

Gap's and Banana Republic's sales continued to decline, but it was an industry-wide trend

3

Only a few competitors, such as Zara and American Eagle, showed signs of success

Limitation: Revenue is dependent on multiple factors and not just big data.

Our Approach



Build a methodology to predict fashion trends



Understand the internal/external data available to Gap.



Insights that can be drawn from such data



How big data can help individual brands



Data-backed inferences on whether big data was a good way forward for all 3 companies



Inferences on creative designer vs big data

Data-Driven Creative Process At GAP

Data model should give us insights on:

01

Design of product: color, fit type, pattern, fabric composition, silhouette

New product launch time

02

03

New Product Launch Locations

Quantity and Price range for the targeted customer segment

04



Data to build predictive model

In- House Data



Website Visits, Wish-Lists



SKUs sold via Online/offline



Time from Shelf to POS



Review & Sentiment Analysis



Online/ Offline Surveys

External Data



Competitor Prices & New Offerings



Fashion week/Catwalks Image Analysis



E-Commerce Bestsellers



Google Trends



Social Media



Social Events

Competitor Analysis

Zara vs. GAP

1

Hypothesis: Mean of the prices of the product from Zara and Gap are different for Jeans

2

Run t test to determine if diff in mean prices are statistically significant

3

GAP's jeans sample mean= \$69.06

4

Zara's jeans sample mean= \$62.1

5

From t test results we can infer that difference in mean of prices on offer across GAP and Zara statistically insignificant

6

As Zara and Gap have similar prices, we can extract design patterns of Zara's new offerings and use that data to design new products for GAP

t-Test: Two-Sample

	Zara	GAP
Mean	62.10	69.74
Variance	293.73	72.10
Observations	10.00	36.00
Hypothesized Mean Difference	0	
P(T<=t) two-tail	0.20	

Insights on Zara's New Offerings

1

Analysis of its new product offers gives us valuable insight into price, fabric, fit type and color

2

Zara has forecasted baggy jeans to be the norm

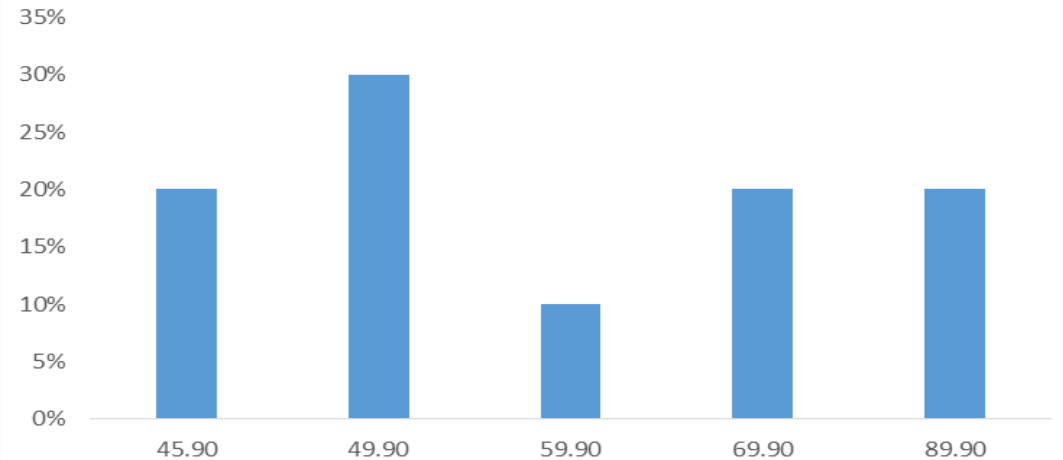
3

Offering products across entire price range, giving sense that demand to remain same across segments

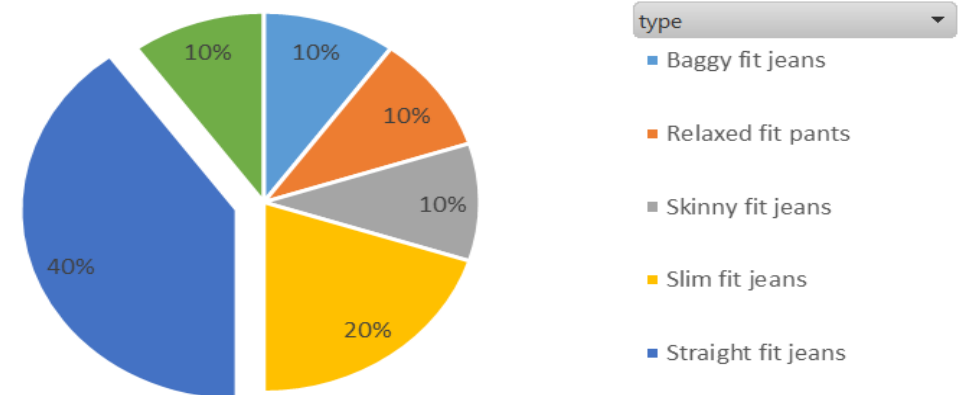
4

Under the hypothesis that Zara has finest predictive model, directly copying the trends in new offerings can really work for brands GAP and Old Navy

Skus offered under diff prices



Fit type



Amazon's Bestsellers

1

Amazon lists out its best sellers with rankings and associated details of product

2

Data can be utilized in finding the relation of the rank(dependent variable) with other variables such as price, fabric, color etc specific to fashion industry

3

This data can be used to build an inferential model

Below is a sample collected from Amazon best seller website

Name	Color	Price	Info	Ratings_count	Ratings
Levi's Men's 505 Regular Fit Jeans	Medium Stonewash	41	100% Cotton Impc	109 796 ratings	4.5 out of 5 stars
Wrangler Authentics Men's Regul	Dark Stonewash	30	88% Cotton 10% Pc	68 548 ratings	4.6 out of 5 stars
Levi's Men's 559 Relaxed Straight	Timberwolf Twill	20	100% Cotton Impc	44 039 ratings	4.5 out of 5 stars
Signature by Levi Strauss & Co. Go	Headlands	28	99% Cotton 1% Ela	40 321 ratings	4.5 out of 5 stars
Wrangler Authentics Men's Classi	Carbon	32	100% Cotton Impc	83 659 ratings	4.5 out of 5 stars

Inferential Model:

$$\text{Rank} = C + B_1 * \text{Color} + B_2 * \text{Price} + B_3 * \text{Color} * \text{Price} + B_4 * \text{Material composition} + B_5 * \text{Brand image/value} + B_6 * \text{Brand image} * \text{Price}$$

Interaction term of color*price takes into consideration that certain colors demand a small premium from other colors and similar is with brand image*price, certain brands enjoy premium based on popularity

Fashion Week/Catwalk Image Analysis

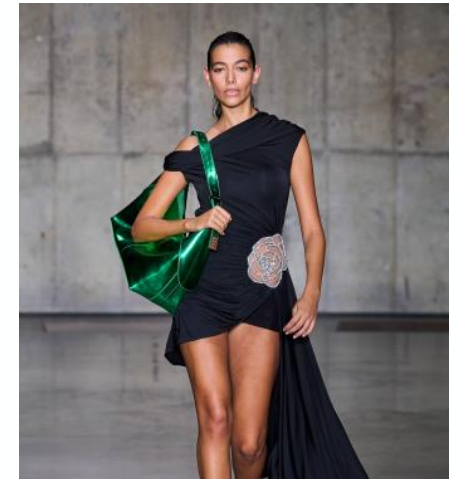
Images scraped from Vogue

1 Images can be collected from weekly or monthly fashion week and catwalks

2 Image recognition technology can be incorporated to identify the fabrics, pattern, fit and color from these images

3 Secondary validation of the analysis can be done from designers

4 Methodology can be applied to social media platforms like Instagram, whereby outfits of influencers can be tracked



GAP Website & Store Level Data



Website Visits, Wish-Lists



SKUs sold via Online/offline



Time from Shelf to POS



Store Manager Feedback



Online/ Offline Surveys

1

Present sales information gives insights into customer preferences and current inventory

2

Stockouts and markdown information give us performance of products

3

Surveys and reviews collected at stores gives us valuable sense of customer's fashion sense and needs

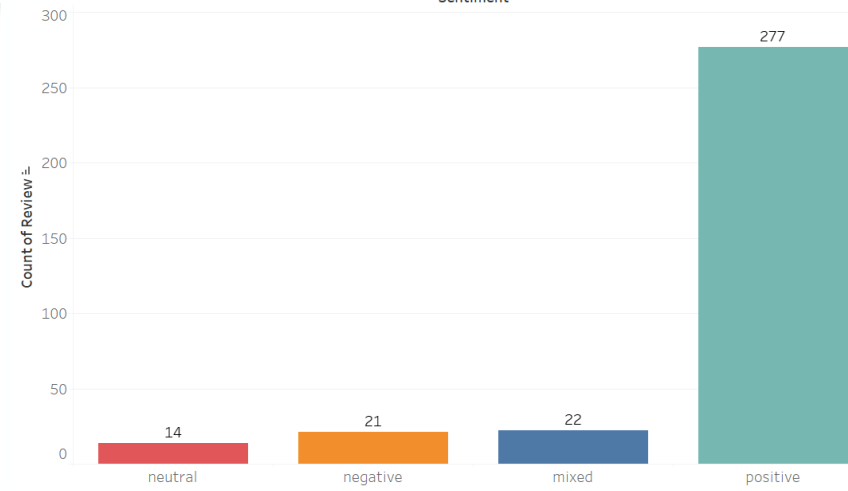
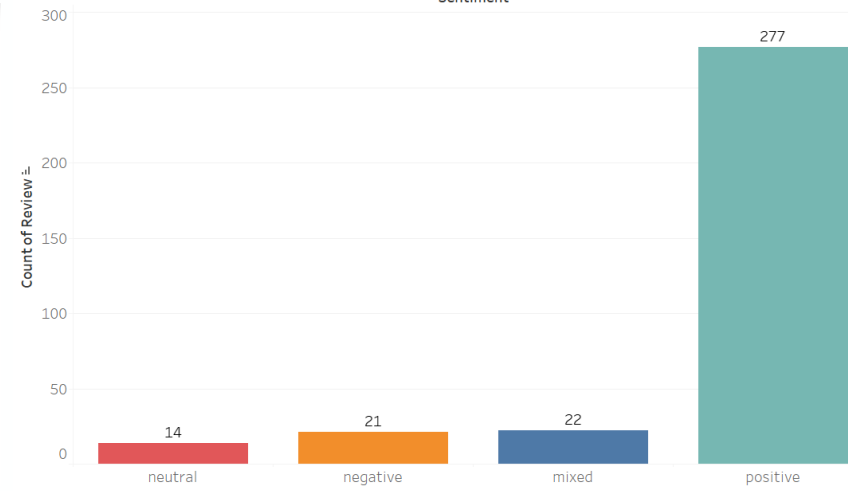
4

Allows efficient management of sourcing and stocking merchandise

[illegible][illegible]

Sentiment

Sentiment	Count of Review
negative	55
mixed	58
neutral	65
positive	1,033



Google Trends

Related topics ?

Rising ▾



1	Levi's - Topic	Breakout	⋮
2	American Eagle Outfitters - Topic	Breakout	⋮
3	NewJeans - Girl group	Breakout	⋮
4	Wrangler - Topic	Breakout	⋮
5	fit - Topic	Breakout	⋮

Related queries ?

Top ▾



1	black jeans	100	<div><div></div></div>	⋮
2	mens jeans	98	<div><div></div></div>	⋮
3	men jeans	77	<div><div></div></div>	⋮
4	blue jeans	76	<div><div></div></div>	⋮
5	skinny jeans	75	<div><div></div></div>	⋮

Gives insights on rising, trending queries related to a search term. For our chosen product "Jeans", it revealed trending styles, colors and brands that people are searching

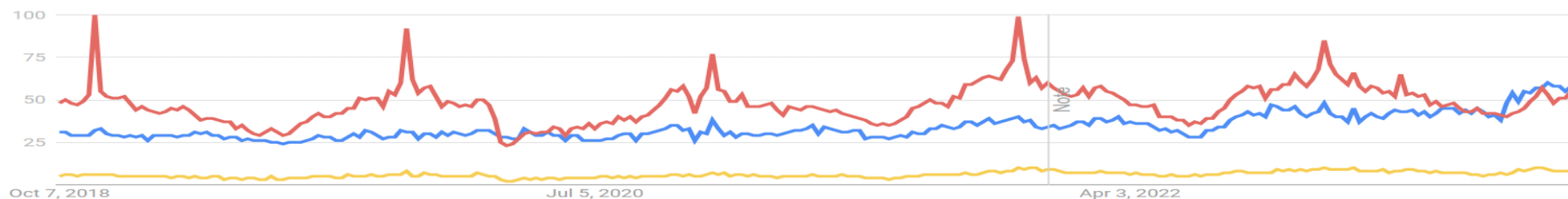
Related queries ?

Rising ▾

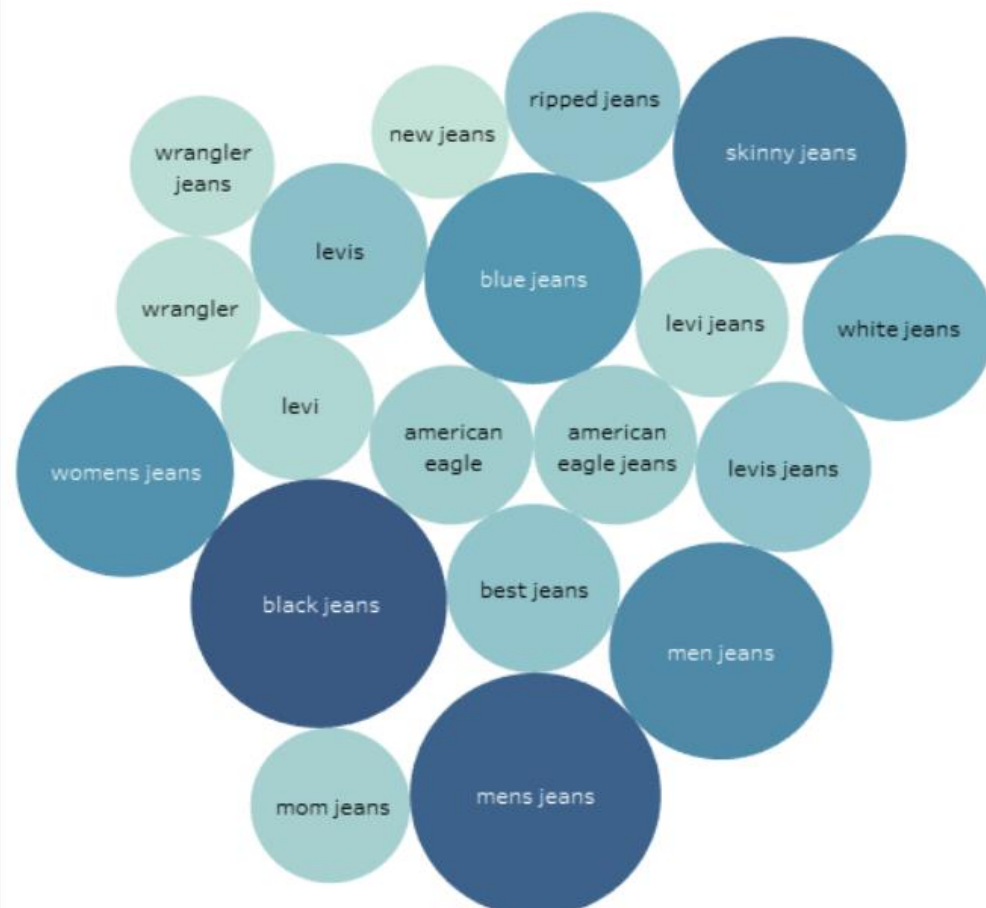


11	agolde jeans	+1,350%	⋮
12	stacked jeans	+1,300%	⋮
13	amiri jeans	+800%	⋮
14	ksubi jeans	+800%	⋮
15	rta jeans	+650%	⋮

Interest over time ?



Most Searched in Google



Brand Positioning

OLD NAVY

- Budget conscious shoppers' choice
 - More affordable than Gap
 - Constitutes mainstream fashion
 - Consistent promotions & sales to attract more customers
-















- Casual, every-day clothing
 - Constitutes trendy fashion
 - Affordable – mid-range pricing (higher than Old Navy but cheaper than Banana Republic)
-



- Exclusive and premium brand of Gap Inc.
- Constitutes edgy and classic fashion
- Emphasis on quality, style and exclusivity

Only Big Data: Right Approach?

	Old Navy	GAP	Banana Republic
 In-house and external data availability			
 Complexity in design			
 High customer expectation			

With limited and exclusive collection with prolonged sales cycle, availability of in-house data for Banana Republic is a concern

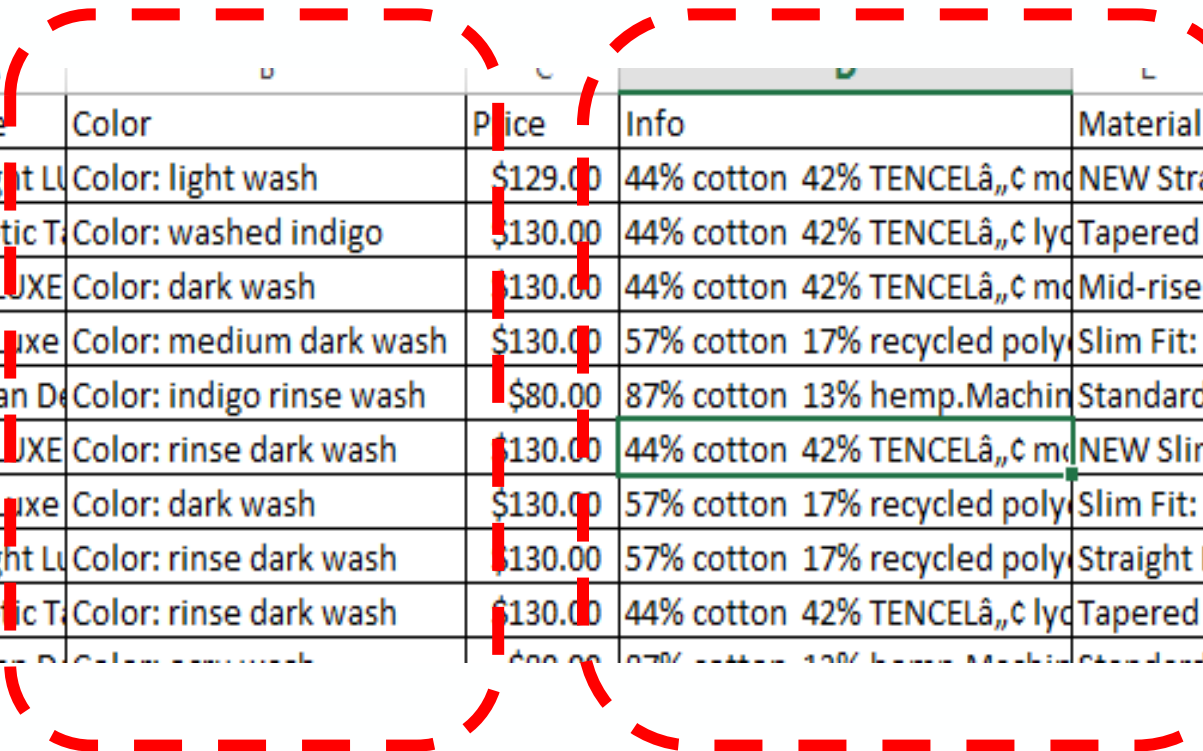
Have to rely heavily on external data from fashion magazines, catwalk images, detection of cultural shift

Level of complexity in production and fabric composition makes it much more difficult to predict

Customer expectations of this niche segment is higher in terms of design, exclusivity, personal experience

High opportunity cost: Loss of customer's loyalty, switching of brands, loss of production time, loss of higher quality fabrics from unsold products

Fabric & Material Complexity: Banana Republic



Name	Color	Price	Info	Material						
Straight Luxe	Color: light wash	\$129.00	44% cotton 42% TENCEL™, 14% modal	NEW Straight Fit: Mid-rise. Slim through the thigh. Straight leg opening						
Athletic Tapered	Color: washed indigo	\$130.00	44% cotton 42% TENCEL™, 14% lyc	Tapered Fit: Mid-rise. Extra room through the seat and thigh. Tape						
Slim Luxe	Color: dark wash	\$130.00	44% cotton 42% TENCEL™, 14% modal	Mid-rise sits at the waist slim through the thigh slim leg opening						
Slim Luxe	Color: medium dark wash	\$130.00	57% cotton 17% recycled poly	Slim Fit: Mid-rise. Slim through thigh. Updated slim leg opening m						
Keegan De	Color: indigo rinse wash	\$80.00	87% cotton 13% hemp. Machin	Standard fit. Long sleeves. Hip length. Body length (size M): Regula						
Slim Luxe	Color: rinse dark wash	\$130.00	44% cotton 42% TENCEL™, 14% modal	NEW Slim Fit: Mid-rise. Slim through thigh. Updated slim leg open						
Slim Luxe	Color: dark wash	\$130.00	57% cotton 17% recycled poly	Slim Fit: Mid-rise. Slim through thigh. Updated slim leg opening m						
Straight Luxe	Color: rinse dark wash	\$130.00	57% cotton 17% recycled poly	Straight Fit: Mid-rise. Slim through the thigh. Straight leg opening						
Athletic Tapered	Color: rinse dark wash	\$130.00	44% cotton 42% TENCEL™, 14% lyc	Tapered Fit: Mid-rise. Extra room through the seat and thigh. Tape						
Keegan De	Color: medium dark wash	\$80.00	87% cotton 13% hemp. Machin	Standard fit. Long sleeves. Hip length. Body length (size M): Regula						

Type of color is not regular which is found in other brands but are blended

Different composition of fabrics

Complexity involved here will be difficult for any data model to predict

Data Limitations

- Data available from Zara new collection is limited, t-test on sample applied for comparing mean of prices against GAP
- Unavailability of store level and online sales, reviews data points for finding correlation between demand against product variables like price, color, pattern, fabric, silhouette.
- Amazon's best-selling data scraped is across the entire product category but not specific to any particular customer segment.

Big Data in Marketing

Product

Science more useful

- Identifying market opportunities
- Tracking customer journey in the product and identifying drop off pages

Art more useful

- New Product Development
- Designing and enhancing user experience
- Packaging

Price

Science more useful

- Tracking competitor prices
- Building a dynamic pricing model taking demand into consideration

Art more useful

- Developing value based pricing strategy by identifying customer's perception of value

Place

Science more useful

- Online vs offline presence
- Identifying right place for right product

Art more useful

- Store design
- Enhancing in-store experience

Promotion

Science more useful

- Targeted Advertising
- Content Personalization
- Measuring Campaign Effectiveness

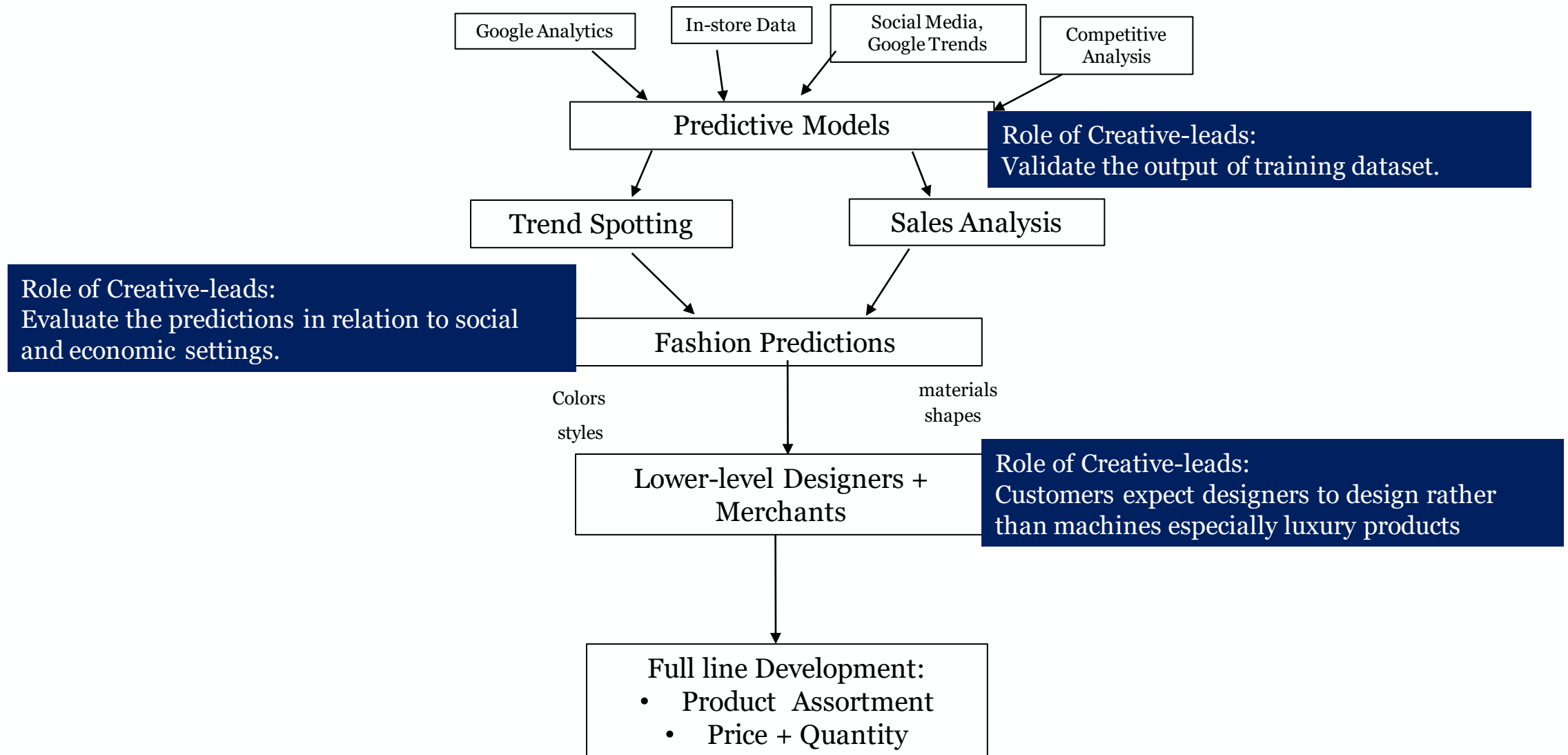
Art more useful

- Branding
- Advertising and Creative Campaigns
- Content Creation

Recommendations

- Visibility and creating brand image with respect to the audience a company is catering to. For example, Banana Republic should prioritize being part of high-end fashion shows to associate with such kind of customers.
- Employ a creative director for:
 - data validation of output of the training data set before being used in predictive model
 - enhance brand image by employing reputed designers, higher segment customers expect clothes to be designed by designers rather than machines

New Big-Data + Creative-led Model



Thank You

