Uncovering the key drivers of consumer satisfaction in the snack category in China

A case study utilizing big data and natural language processing as a tool in brand insight

As with most things these days, consumer insight is facing disruption…

…and we can trace a similar cause to this disruption, digital. Traditionally companies that are consumer centric conduct research by proactively going to their consumers and asking questions. This can be through an internal CRM system, focus groups or survey methods.

With the advent of the information age, we have seen an exponential growth in the information available for us to look at it. Couple this with a mass migration of consumers to online commerce, as well as the interlink between so many aspects of life being digitized, there is
a wealth of information available digitally to be analyzed for our insight needs.

One of the major issues today is that many brands and companies lack either the right partners, the right tools or the internal capabilities and organizational structures to truly make this insight work for them.

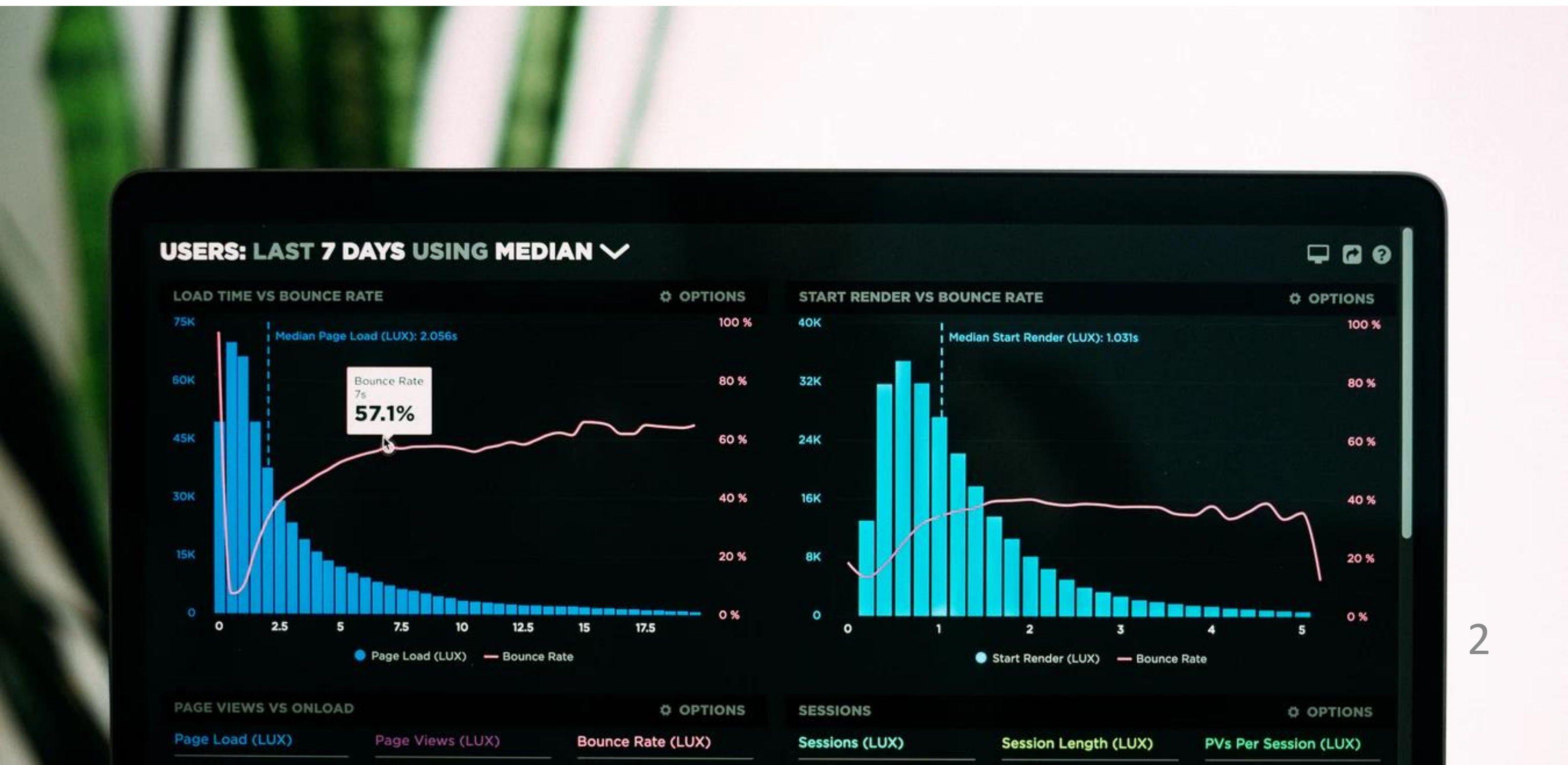
We believe that proactive, survey based, insight will always have a place in the foundations of building strategy. Without asking direct questions to consumers, you can often not find the answers you are looking for. This highlights a big drawback in the use of big data – given its passive nature, it is often very difficult to find answers to specific questions. However, in some senses, its biggest weakness is also its biggest strength.

Given their passive nature – the reviews are primarily based on an actual purchase, with no applied screener and show a lower degree of bias than surveys.

Additionally, the high frequency and high volumes of data lead to ability to lend more credibility (for those put weight on sample size) and the ability to track at more regular intervals. Insights and strategy departments within companies today must move beyond the current thinking of ***either/or*** and uncover the benefit of taking a ***both/and*** approach when it comes to using these two types of insight.

To highlight the power of these solutions Kantar and Re-Hub teamed up to investigate the snack category in China. Using the power of Revuze, a natural language processing tool, applied to over 400,000 eCommerce reviews, each review is automatically analyzed to identify linguistically which different language aspects emerge and assign a sentiment to each of them given the context (positive/negative). As a result, it is possible to look into the breakdown of each of the aspects and their sentiment for any given brand or product.

In the following paper we will look at the findings across the category and some representative brands, highlighting what drivers consumer satisfaction in the category, which brands are performing well and what products are driving the brand satisfaction.



Overall Satisfaction

The first area we want to understand is what drives overall satisfaction in the category.

A good or bad product experience can be the difference between trial and drop-out, or trial leading to true brand loyalty.

We often tell ourselves that the complete product experience will drive satisfaction, and it is important to focus on improving everything, but what is a complete product experience? The perfect blend of different aspects that all roll up to create an overall experience for consumers; whether its sensory (taste, smell, touch etc.) or service orientated (friendly staff, great shipping experiences etc.), all these aspects play a crucial role (with different weightings across categories) in building overall experience.

However, with finite resources, come difficult decisions, and given a limited investment where should we look to improve our overall satisfaction to increase ROI? The first step in that journey it to break down the experience into manageable aspects. These aspects should be clear enough to be able to address on their own, and create action plans against them.

By taking this approach it is possible to improve your overall experience by identifying which element of that experience is currently underperforming and addressing it directly.

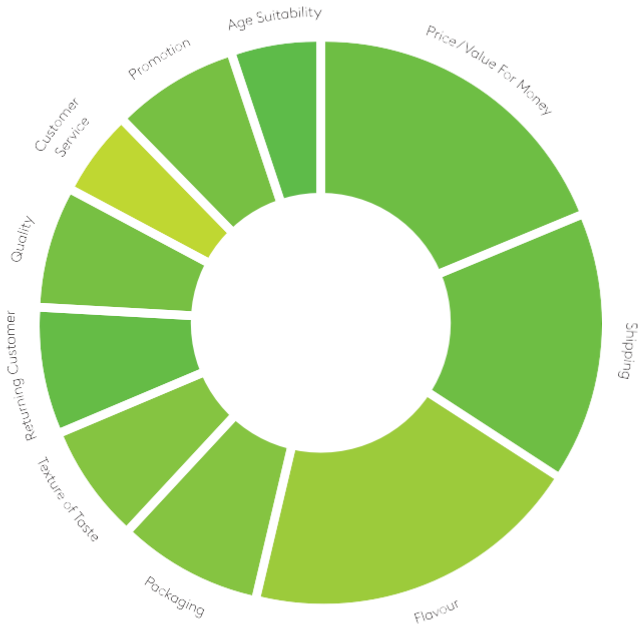


Figure 1

Overall Satisfaction aspect is mentioned **positively in 32.31%** of total reviews.

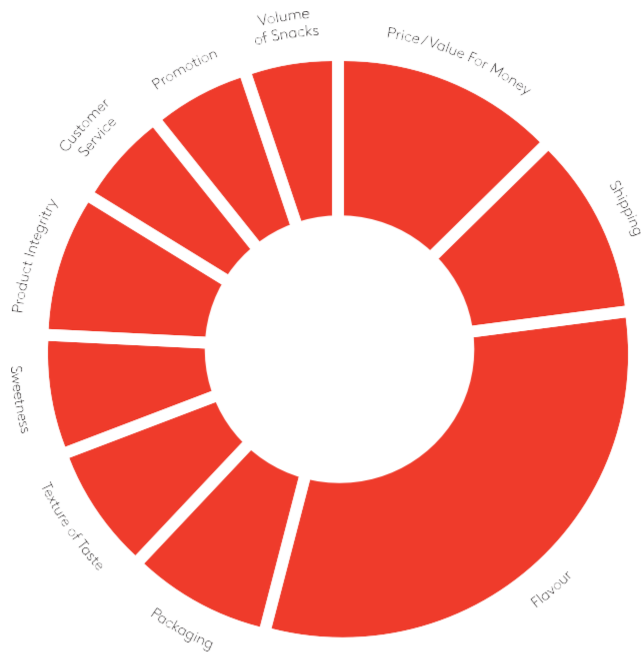


Figure 2

Overall Satisfaction aspect is mentioned **negatively in 3.48%** of total reviews.

Aspect breakdown that corresponds to a positive or negative overall experience

When we look into the snack category in China, we are able to unlock insight as to what is driving both overall positive and negative experiences to real world consumers.

Our data shows that ~66% of what makes up an either an overall positive or negative experience comes down to 4 factors – **Price, Shipping, Flavor** and **Packaging**.
This means that getting these four elements right gives a higher chance of driving customer satisfaction.

Looking deeper at the relationship between positive and negative, we can see that a bad flavor experience is much more likely to drive dissatisfaction than a good flavor experience drives a positive experience. This means getting flavor wrong is the potential biggest pitfall of the category.

The inverse is true for price – meaning that often the price or value is seen as an unexpected nice to have, rather than an element that defines a fully negative experience.

Outside of these factors we start to see differences emerge between positive and negative product experiences. Quality, customer service, and age suitability all come into play for positive experiences, highlighting the heavy shopper nature of the category whereby the purchaser is not the consumer. Within negative experience we see two elements, sweetness – potentially highlighting that across the category there is a sweet issue, and volume – also potentially highlighting that there is not an appropriate pack format available to fulfill the specific needs of the consumer/shopper.

This interlinkage between satisfaction and sentiment is important to identify the jobs to be done on a brand level but also to identify the white space opportunities on a category level.

Aspect/Sentiment Breakdown

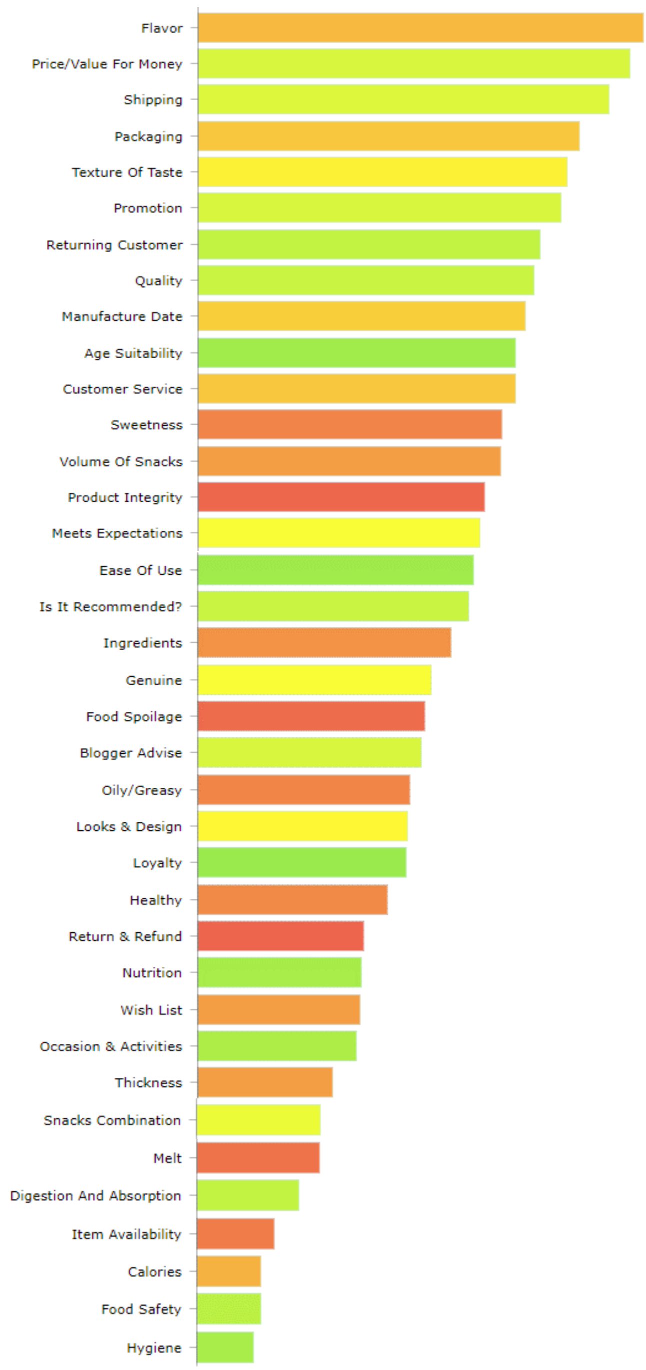


Figure 2: Total category aspects; bar length represents importance (frequency of mention) while the color indicates positive/negative (green/ orange/red) sentiment as indexed against overall category sentiment

While we have understood what makes up overall positive and negative experiences, there are multiple other aspects that are mentioned within the category.

The category analysis is made up of two elements; the first is the frequency of mentions. For a given aspect as decided by NLP, we calculate how many times it is mentioned in the pool of reviews analysed. This represents the importance of that aspect to the category.

The second analysis metric looks at whether the language used in each of the reviews, with regard to that specific aspect, is positive or negative, giving an overall percentage score to the sentiment. This is then indexed against the entire category to understand its performance.

Using this analysis, we can see the different aspects that perform well in the category and those that do not. Our aspects cover a wide range of implications that spread across multiple functions within any business from **sales – price; RTM – manufacturer date, returns/refund; innovation – flavor, sweetness; marketing – packaging, occasion.**

Across the category, drivers such as age suitability and ease of use are overall important and positive – this shows that these are areas that most players within the category are performing well at. The implication is that within these specific aspects there is little room to create true category differentiation, however, when looked at on a brand level, may instead offer room for incremental improvement.

Instead, it is perhaps more interesting to look at areas which fulfil two criteria; absolute importance to the category; and a mid/low (orange/reed) sentiment level. This indicates that most brands are not satisfying consumers in this area, and if an individual brand is able to provide solutions that overcomes the consumer pain-point they stand in a good position to build positive equity with consumers.

In order of importance these areas of potential differentiation are; manufacture date; customer service; sweetness; volume of snacks (primary and secondary packaging); product integrity; ingredients; spoilage; oily/greasiness; health; return and refund; thickness; melt and item availability – this means that if snack brands are able to put teams and resources against fixing these consumer facing issues through working with their internal or external partners they will be able to create category level differentiation.

Four Drivers

While it is important to understand the sentiment and satisfaction drivers on a category level, the ability to compare performance on a more granular/brand level creates a greater degree of actionability for companies.

The benefit of consumer eCommerce reviews is that they are based on a specific product as taken from eCommerce platforms, giving us the ability to look at a product and brand level in order to understand current performance.

We have seen that four aspects drive the majority of positive consumer experience within the category. It is possible to break these down to a brand level (product level analysis is possible, but not covered in this paper).

We took five representative brands from our brand list to deep dive into, however any selection of brands and/or aspects is possible to analyze.

Aspects	Brand > Snack Category Average	Three Squirrels	Lay's	Oreo	HsuFuji	Pocky
Flavor	70%	66% (94)	68% (97)	78% (111)	68% (97)	74% (106)
Shipping	85%	76% (89)	89% (105)	90% (106)	74% (87)	82% (96)
Price Value for money	86%	73% (85)	91% (106)	92% (107)	87% (101)	87% (101)
Packaging	72%	78% (108)	66% (103)	74% (103)	70% (97)	72% (100)

Table 1: Brand performance across four key drivers to overall brand satisfaction for top 6 key brands. Bracketed number indicates the index against the overall category. Index is calculated by dividing the brand level sentiment to overall category sentiment and expressing as a whole number, where 100 is equal to the average category sentiment.

6

Flavour

Flavor has proven to be one of the most important factors in driving a positive overall experience for consumers and the largest driving factor for a negative overall experience. It also represents a low hanging fruit for adaptation from an innovation/R&D perspective, with additional implications to insights and marketing function within organizations.

With regard to flavor, Oreo comes out on top (111 index) against the category, while we see that Three Squirrels (94 index) underperforms against the category.

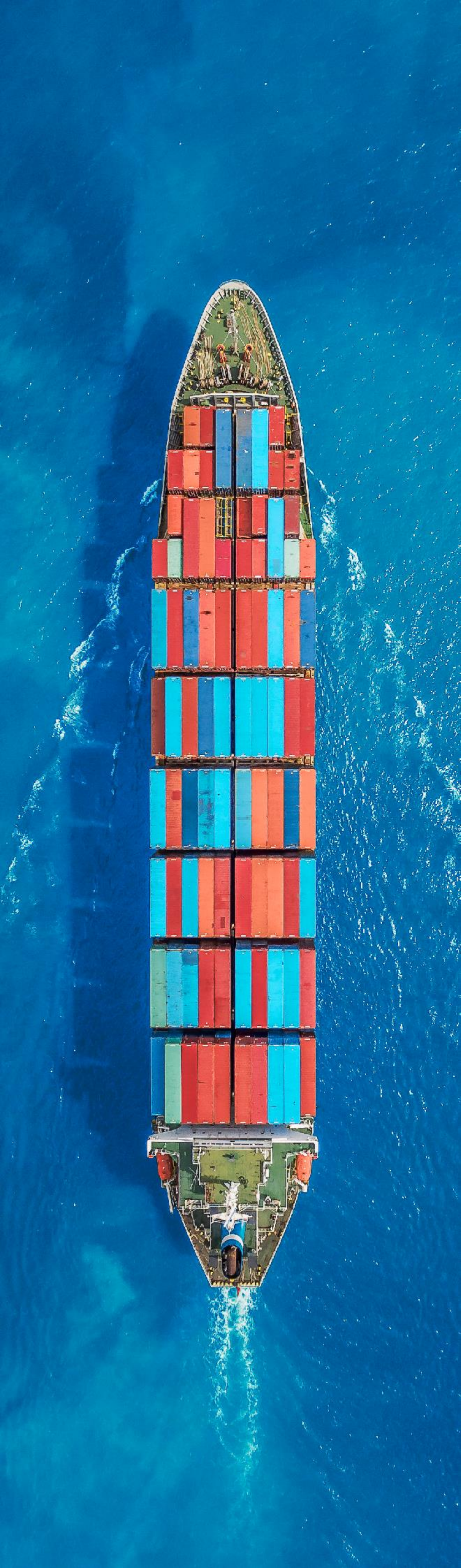
In order to understand what is driving this we must take a full portfolio view of all SKUs under each brand. Lay's as a brand is made up of a collection of different flavors, and according to consumers, there is significant polarization in attitude toward flavors – original, cucumber and finger licking braised pork flavors are received positively, with high overall satisfaction scores in flavor above the brand average (68% from Table 1). However, caramel pumpkin chips, and 'French fries' flavors have negative feedback, scoring significantly lower than the brand average.

Using this level of analysis, it is possible to identify specific flavors that may be considered for discontinuing or putting back into formulation. In this way, this type of analysis can be a useful tool in the innovation toolkit to both assess current portfolio, as well as being able to track the progress of new releases in the market.

Star Product

Lay's Cucumber Flavor 104g

Overall Sentiment – **94%**



Shipping

Shipping is an aspect that covers much more logistic focused concerns. The fundamental nature of this aspect means that its control sits outside the brand itself, however a deep understanding of this aspect can arm businesses with the right information and storyline with which to speak to logistics partners with.

Within shipping we see Oreo again coming out on top (106 index), many citing the very quickly delivery times they would receive even from multiple purchases, while 徐福记 (HsuFuJi) comes out lower (87 index) with consumers citing areas such as the delivery method (postal vs express) that were not clear to them on the platform and therefore undelivered versus expectation.

The overall score can provide an initial check if there are shipping issues, however a further analysis within Shipping can be conducted on a platform level (JD vs Taobao vs TMall etc.) to understand if there is a specific issue on a specific platform.

Star Product

Oreo 349g 12 pack

Overall Sentiment – **98%**

Price / Value for Money

Price plays an important role in business strategy, the right pricing will attract consumers, but the wrong pricing may scare them away. It is important to remember that this can be seen for both being too expensive or too cheap depending on the current image of your products and brands. Interestingly when it comes to pricing, the idea of 'Promotions' is separated out in our aspects and plays a smaller role. Thus, we can consider pricing as a true indication of the non-promotional price, rather than something that is seasonal or affected by promotions.

From a brand level, Three Squirrels is underperforming against the category (85 index). Within this aspect consumers offer feedback such as a sense that the price has been rising over the years, and over use in packaging is driving 'wastefulness'.

Similarly to flavor it is possible to conduct this analysis on a product level which can offer specific insights into portfolio pricing management for sales teams.

Star Product

Orion 1020g 30 pack

Overall Sentiment – **95%**

Packaging

Packaging incorporates the overall packaging experience of the product – from primary to secondary packaging.

Interestingly this is the one area within our top four drivers where Three Squirrels out performs its competitors (108 index) – with the main drivers for this success being activations such as packaging the fruit and nuts separately, as well as packaging items in smaller portions.

Lay's on the other hand is underperforming in this aspect (90 index) against the category. Consumers offered feedback that is intrinsic to the product such as the propensity for crisps to break and become crumbs before consumption. While others would cite receiving bags that had burst and spilled before being able to enjoy the product.

Star Product

HsuFuJi 1596g Giftset

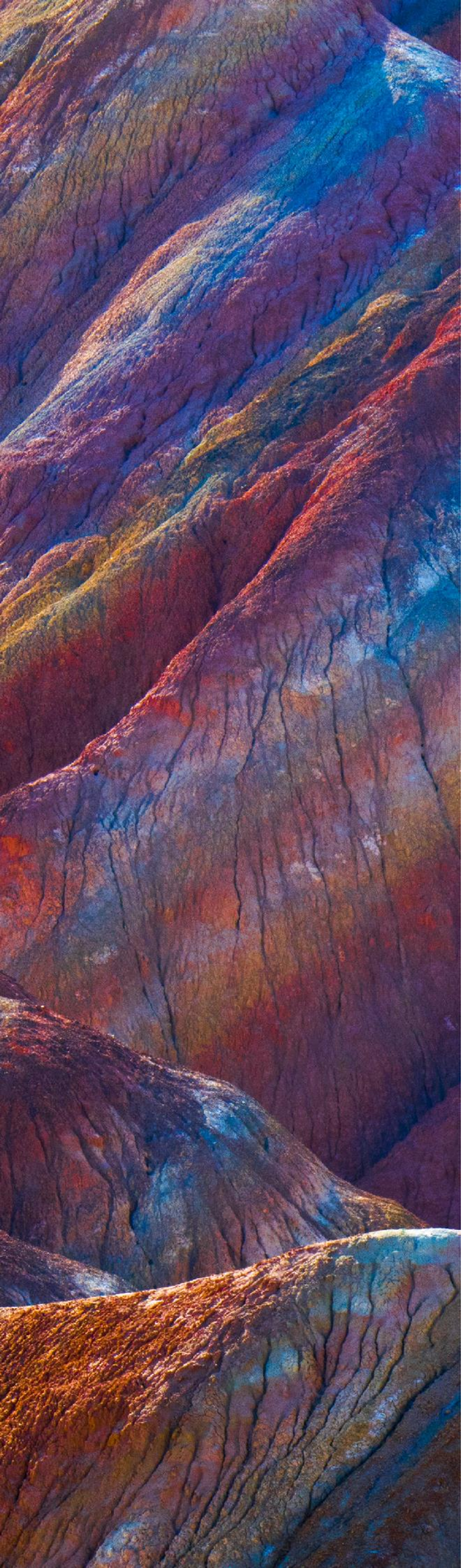
Overall Sentiment – **87%**

The Trend – The impact of COVID

We know that 2020 has been a tumultuous year for many categories, however what has been the impact on the snack category? Looking at the overall trend in the category, we have found that many more users are leaving reviews on eCommerce categories within this category. We do not cover an in-depth analysis of the changes of sentiment throughout this period – although it is possible.

From a high level, this increased engagement with the category and with eCommerce reviews shows the importance of understanding the insight behind them. With the volume of such reviews growing at such a strong rate, it is no longer possible for an individual or even a team within a company to fully extract the insight behind such a large amount of data.

This is another example of COVID acting as a catalyst for already existing trends, accelerating the digital transformation of not only consumers, but additionally the required digital response of companies to counteract for the trend.



Conclusion

The wealth of insight that is available to us through eCommerce review analysis reaches from category, to brand, to product specific insights. While not covered in this paper, it is also possible to look at time bound analysis (year on year, month to month, week vs week), as well as platform level analysis for eRetailers. We have presented a representative view of a specific category in China, but it is possible to extend to larger collections of brands, as well as analyze different categories.

While passive collection of big data lacks the tailormade nature of survey data, the tools and the size of raw data available to brands to use as they see fit is growing exponential as markets move towards greater degrees of digitization. The data exists, can be refreshed up to real time, and contains a significant enough sample size to build confidence in decisions made on the back of it. The result is that, while currently underutilized by many brands who have not digitized their insights functions, the benefit of such analysis is clear and should be considered as an additive tool to the currently existing insights toolkit.

If you are interested in how big data solutions such as this applying natural language processing to the big data of eCommerce reviews can be utilized for your brand and business questions, contact us today.

For further advice on building big data insight capabilities and crafting big data enabled strategy, please contact:

Dr. Thomas Piachaud
Director of Consulting

Kantar

Thomas.Piachaud@kantar.com

Max Peiro
CEO

Re-Hub

max@rehub.tech

Yoyo Liang
CCO

Re-Hub

yoyo@rehub.tech

Appendix

Category Definition

We automatically build a unique taxonomy for each category analyzed based on users' comments, so it is crucial that the category analyzed includes products of same form, need and usage to automatically extract the most relevant topics with the highest accuracy. For this project we selected “snacks” as a category and included 4 sub-categories: nuts, chips, cookies and pastries. We analyzed 13 brands on Tmall and JD.com with around 400,000 user comments.

Machine Learning

Our machine learning algorithms automatically & independently build a unique taxonomy for each category analyzed, allowing to discover unique aspects in each product category without the need to use predefined keywords. This self-learning technology learns by itself from eCommerce reviews to understand how customers talk about any given product category, delivering higher accuracy and granularity than traditional methods

Sentiment extraction

Our sentiment analysis combines the power of computational linguistics, text analysis, and natural language processing to clarify subjectivity in customer perceptions. We filter customer attitude, recognizing contextual polarity and interpolating judgement, affective state, and intended emotional communication to create easy-to-understand and usable analysis.

Contextual understanding

Our insights engine automatically categorizes, identifies, and extracts trends and topics from unstructured data – understanding context with exceptionally high precision and delivering truly actionable business insights. Our contextual intelligence understands topics and sentiment, regardless of the actual words customers use, making manual keywords definition a thing of the past

Brand/Company List

妙芙 – MiaoFu
徐福记 – HsuFuJi
好丽友 - ORION
乐事 - Lay’s
上好佳 - Oishi
旺旺 – WantWant
奥利奥 – Oreo
健达缤纷乐 – Kinder
脆脆鲨 – Crispy Shark
3+2
三只松鼠 - Three Squirrels
绅士坚果 - Planters
格力高 - Glico