



MARKETING IN A DIGITAL AGE

Using Big Search Data to Map Your Market

By **BERND SKIERA** and **DANIEL M. RINGEL**

Recently, we ran a marketing workshop with an experienced business development manager from Amadeus, the tech provider for the travel industry. The workshop focused on Travel Audience, Amadeus' online advertising platform for travel companies. Participants were given four data sets, ranging in size from 770,000 to more than 9 million "observations" – which reveal how users or consumers behave and interact in a specific market category. We asked the participants to analyze the data and answer a few basic marketing questions: Where does the company make the most profit? Where should it spend more money on advertising? What is the return on investment in terms of advertising?

To our surprise, the participants were stumped. First, it took them a very long time to convert the data into a format that would allow them to answer basic marketing questions, because the sample sizes were far too big for a conventional spreadsheet program. They struggled with separating causality from correlation. And finally, they complained about the amount of data they had to analyze – until we reminded them that the data sets they had been given barely covered a week!

This experience illustrates that, unlike yesterday's marketing professionals, marketers today need to upgrade their data-handling skills. It's not enough to ask a junior professional to do the work for you. Marketers who can pull



Big data present distinct challenges. Marketing professionals do not have to become IT specialists. But successful marketers will be those who also have solid math and data-management skills.

and read the numbers themselves are in a much better position than those who have to run to someone else each and every time they need to retrieve key data, waiting days or weeks to get the information they want from the IT department. Being able to read and react faster to market shifts will lend a competitive advantage.

With so many brands and products and so much big data available today, where does one begin? Interpreting the competitive relationships among 1,000 different products is extremely difficult and would entail a gigantic, unwieldy table. As such, we came up with a way of visualizing competitive market structure in a single map. This makes it more manageable to devise effective competitive strategies, pricing policies and communication.

This article explains our approach. Using data of large markets with more than 1,000 competing products, we show how companies can tap into big data and visualize competitive market structure. Various parties can benefit from this method, which reveals valuable insights on customer behavior.

The Big Learning Curve of Big Data

Ever since the advent of the browser, marketers have been trying to get to grips with the

information boom. The internet enabled marketers to measure, on a grand scale, when an advertisement was shown and how consumers reacted to it – how many clicked and how many proceeded to purchase.

With the ubiquity of increasingly affordable and powerful smartphones, the sea of data has grown exponentially. On top of all existing data, companies can now also collect vast quantities of location-based information on consumer behavior around the clock.

While having more and better information is a good thing, big data present distinct challenges – from poor data quality, to incomplete data sets, to mistaking correlation for causality.

There are several signals that marketers are behind the learning curve. It used to be that the chief concerns of marketing were to develop creative slogans and strategies. In fact, many people went into marketing precisely because they didn't want to do math. The digital revolution, however, has changed all that. While coming up with creative content will always be central to the profession, marketers now have to take a data-driven approach to prove that their ideas actually work.

This development doesn't mean marketing professionals have to become IT specialists or set up databases and IT architecture themselves. Such tasks can and should stay with the IT department. However, it does mean that successful marketers will be those who also have solid math and data-management skills. The rise of crowdsourced platforms, such as Kaggle, to carry out predictive modeling and analytics, as well as the growing popularity of online data-analysis courses, are testament to marketing's wider remit.

Market Share and Competition

Recently, a manager at a price-comparison website mentioned that his firm could quickly predict the success or failure of a new product by simply following the outbound clicks to

EXECUTIVE SUMMARY

Markets today are flooded with an increasing number of products and brands, making it difficult for companies to track how their products compete in the market. Big search data allow companies to obtain valuable insights on market share as well as on key competitors. But collecting and analyzing vast amounts of data pose huge challenges.

In this article, the authors describe how they used clickstream data to visualize competition in product categories containing more than 1,000 products. Both manufacturers and retailers can profit from their approach, which yields more meaningful information on customer behavior than traditional market research methods.



Purchase data make it easy to estimate market share, but they won't tell you much about the decision process of consumers before they make a purchase. Our research aims to fill that gap.

retailers' websites. That gave us an idea: why not use clicks to predict market share?

We also decided to do something that market research firms don't do very well. Traditional approaches are largely based on scanner panel data that require multiple purchases within a product category to detect switching between products as an indicator of their substitutability. We sought to identify competition among durable products that consumers typically buy just once. Market share and competition are two important areas in which big data can provide key insights for marketers.

We used clickstream data – that is, a log of user actions on a website – as the basis for our analysis. These data are powerful for marketers, because they provide high-quality information that can be sliced up across various dimensions. Most market research firms only collect purchase data. These data make it easy to estimate market share, but they won't tell you much about the decision process of consumers before they make a purchase.

Imagine you want to buy a washing machine for the first time. The number of options can be daunting. Someone might recommend a product to you, and although you might recall the brand, it is more difficult to remember the exact type or product number of the washing machine.

The failure to remember specific products is the fundamental pitfall of consumer surveys. When market research firms survey consumers on which other products they considered before making a purchase, consumers typically struggle to correctly recall which exact products among hundreds they looked at. One of the goals in our research was to fill this critical gap, since such insight is crucial for marketers to make informed decisions.

A Five-Phase Approach

For the purposes of our research, we examined the LED TV market in Germany. We chose the

LED TV market for three reasons:

- First, LED TVs have become highly popular consumer durables on a global level.
- Second, since the market consists of more than 1,000 products, the use of traditional approaches to analyze competition among products would be cumbersome, if not impossible.
- Third, the market is well known, with easy-to-understand attributes, allowing us to easily check our findings for face validity.

Our approach consisted of five phases:

- Collect data by observing consumer behavior on an online platform that provides information on diverse products.
- Consider products that were viewed together by the same consumer as being products in that consumer's consideration set.
- Identify competitive asymmetry.
- Use a model, called DRMABS, to create a map that visualizes asymmetric competitive market structure.
- Transpose product attributes onto this map.

1 PHASE 1: COLLECT DATA

For the initial step, we collected big search data on individual consumer clickstreams. You can retrieve such data from any website that provides information on a broad range of products, such as online retailers, price-comparison sites and product review sites. For this study, we collected data from idealo.de, a leading price-comparison website in Germany.

To collect consumers' clickstreams, we installed a tracking pixel on each page of the website. This tracking pixel enabled us to collect data on 105,606 consumers who used idealo.de to search for and compare 1,124 products of 56 brands in our chosen product category (LED TVs) in September 2012. A tracking pixel consists of a tiny, usually transparent image that is embedded in the HTML code of a website to record usage by website visitors. This means that each time a consumer opens a new webpage on



the site, the browser downloads the tracking pixel from the web server. This procedure allows the server to log the fact that the shopper has visited that specific page.

During the search process, consumers may revisit sites to review products various times over a period of days. For this reason, we also tracked the behavior of consumers who returned to *idealo.de* several days later to continue their search for products.

Through our research, we discovered that 99.9 percent of all captured clickstreams spanned 16 products or fewer. As such, we chose 16 products as our cut-off, and eliminated clickstreams containing more than 16 products, since such large clickstreams may actually be generated by search bots that systematically collect large amounts of product data but do not provide insight into a consumer's product consideration.

2 PHASE 2: DETERMINE CONSIDERATION SETS

From the data collected in Phase 1, we constructed consideration sets. We looked at which

products consumers perceived as viable substitutes, and then used those sets to uncover competitive relationships among products.

A consideration set was constructed for each individual consumer visiting the website by analyzing individual clickstreams with regard to the products that consumers searched for and compared online. While one shopper may look at five different LED TVs, another might look at two.

We constructed consideration sets for more than 100,000 consumers. In doing so, we were able to determine how often any two products appeared in the same consideration set. The more that any two products appeared jointly in consumers' consideration sets, the stronger their competitive relationship.

We captured all competitive relationships in a 1,124 x 1,124 matrix of joint product consideration. This gave us a very broad view of the marketplace.

3 PHASE 3: IDENTIFY COMPETITIVE ASYMMETRY

Next, we identified two types of competitive asymmetry: across all products in a market (market-share competitive asymmetry) and among pairs of products (local competitive asymmetry).

The idea behind market-share competitive asymmetry is that if a given product, A, is considered by more consumers than products B, C and D, then product A is the stronger competitor overall. In this way, we were able to identify which products were competing against each other, as well as determining how consumers' consideration was distributed across competing products.

The idea behind local competitive asymmetry is that from the perspective of one product, the other product may be a stronger competitor than vice versa. Take Apple's iPod: from the perspective of an iPod, an MP3 player from *iRiver* is less of a competitor than vice versa.

4 PHASE 4: VISUALIZE COMPETITIVE MARKET STRUCTURE

We then created a map to visualize competitive market structure in large product categories, using a combination of methods drawn from network analysis and graph theory. These methods are often used to analyze and visualize relationships in the fields of communication, biology, medicine and IT applications.

■ ABOUT THE AUTHORS

Bernd Skiera holds the Chair of Electronic Commerce at Goethe University, Frankfurt. In addition, he is a member of the managing board of the e-Finance Lab and a Professorial Fellow at Deakin University in Australia. His research focus is e-commerce, online marketing and interactive media, subjects on which he has published books and articles in *Marketing Science*, *Journal of Marketing* and *Journal of Marketing Research*, among others. He has been a visiting professor and scholar at Stanford University, Cambridge Judge Business School, New York University's Stern School of Business, Duke University's Fuqua School of Business and the

University of California, Los Angeles.

Daniel M. Ringel is completing doctoral studies at Goethe University, and starting in summer 2017 he will be an assistant professor at the University of North Carolina at Chapel Hill. He studied business administration at Baden-Wuerttemberg Cooperative State University in cooperation with IBM, and completed his Executive MBA as Goethe Scholar from Goethe Business School in alliance with Duke University's Fuqua School of Business. He previously worked as a management consultant and founded an e-commerce business.



We were able to identify the top 10 brands and products by share of consumer consideration. We were also able to see how individual products competed, and pinpoint local competitive asymmetry.

To do so, we integrated several existing and newly developed algorithms into a new model called DRMABS, which stands for Decomposition and Reassembly of Markets by Segmentation. The underlying concept of DRMABS is that it decomposes the total market into submarkets, maps submarkets individually, and then reassembles all local submarket maps into a global map of the whole market.

Finally, we introduced global competitive symmetry to the global map by using bubble size to represent the frequency of consideration as a proxy for market share, and weighted arrows to indicate local competitive asymmetry. See **Exhibit 1**.

5 PHASE 5: TRANSPOSE PRODUCT ATTRIBUTES ONTO THE MAP

In the fifth phase of our research, we transposed product attributes, such as brand or display size, onto the competitive market structure map. We visualized these attributes by using different colors and bubble sizes.

To create this map, we first removed all competitive asymmetry information (arcs and bubble size) as well as cluster membership (bubble color) from the map. Next, we depicted product attributes by varying bubble colors and sizes that represent individual products on the competitive market structure map. The final visualization provides a clear view of what consumers searched for and the drivers of competitive market structure. See **Exhibit 2**.

Using our approach, we were able to identify the top 10 brands and products by share of consumer consideration. As **Exhibit 3** shows, Samsung dominated the market, with a share of 43 percent. Overall the top 10 brands held a joint share of roughly 97 percent of the German market, while the remaining 46 brands accounted for just 3 percent of the market.

We were also able to visualize how LED TVs were organized into submarkets, to see how individual products competed with each other and to pinpoint when there was strong local competitive asymmetry.

Since price-comparison sites generate revenue with every click on any retailer, regardless of what the product is, they are not affected by which products are viewed by consumers, making them an unbiased data source for consumer consideration.

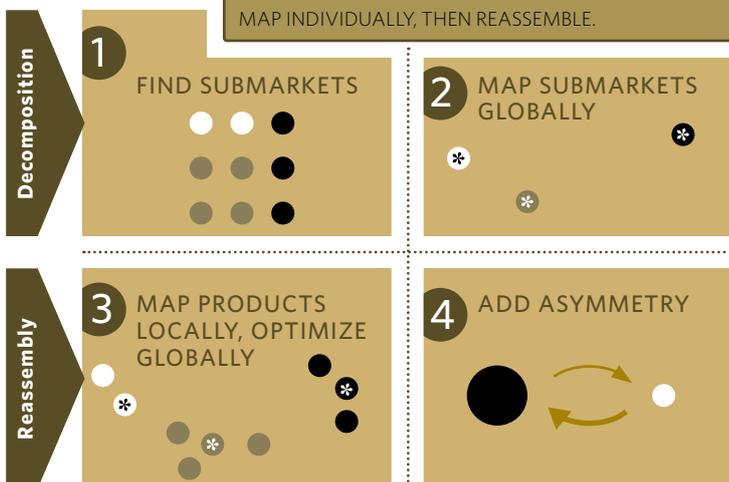
Another benefit of price-comparison sites is that they provide low-cost clickstream data generated by millions of consumers who are comparing thousands of products. They capture individual consumer searches in real time, unlike other sources – such as Google – which only provide summary information, such as total keyword searches.

A Typical Consumer Clickstream

To understand the value of clickstream data, consider a typical scenario. Imagine a shopper – let's call her Susan – is looking for an Apple smartphone. She uses Google to search for this product, and one of the first results is a link to a

This document is an authorized copy for personal use of Mr. Skiera, 25/03/2017

DRMABS Model for Visualizing Market Structure
EXHIBIT 1
WE DECOMPOSE THE MARKET INTO SUBMARKETS, MAP INDIVIDUALLY, THEN REASSEMBLE.





Less nuanced methods would only register if Susan clicked on an HTC product page and then purchased a smartphone there. Clickstream data offer a more detailed picture of the customer journey.

product and price comparison site.

Susan clicks on the site and is taken to a product detail page, which includes technical details, pictures, other consumers' opinions and ratings. She studies the information and decides to have a look at an alternative phone. She uses the search bar to search for "Samsung smartphone." She clicks on the phone she is interested in, but since she's running late, she decides to continue her search another time.

The next day, Susan goes directly to the comparison site to search for "Apple phone." She likes the phone, but wants to make sure it is the right choice. To reduce the list of several hundred smartphones, she applies filters of attributes important to her. She clicks on a model from HTC and is taken to its product detail

page. She likes the phone, but is still not sure.

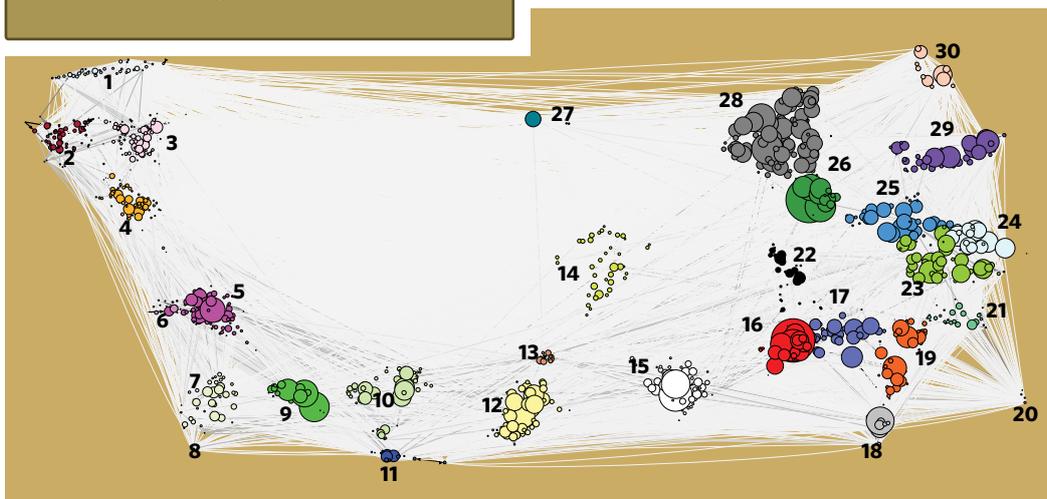
After revisiting the Apple page, she returns to the HTC page. She clicks on the one with the lowest price and a new browser opens, displaying the website of the corresponding online shop.

In this case, Susan's consideration set size is three. The products she weighed are referred to as being "considered jointly" with each other, in that Apple was considered jointly with Samsung, Samsung was considered jointly with HTC, and HTC was considered jointly with Apple.

This is an example of the kind of information that clickstream data can yield that traditional marketing research methods cannot. Less nuanced methods would only register if Susan clicked on an HTC product page and then purchased a smartphone there. Clickstream data offer a more detailed picture of the customer journey.

Visualization of Asymmetric Competitive Market Structure Map of 1,124 LED TVs

EXHIBIT 2



LEGEND

- Bubbles represent individual products (SKUs).
- Bubble color indicates submarket membership.
- Bubble size indicates global

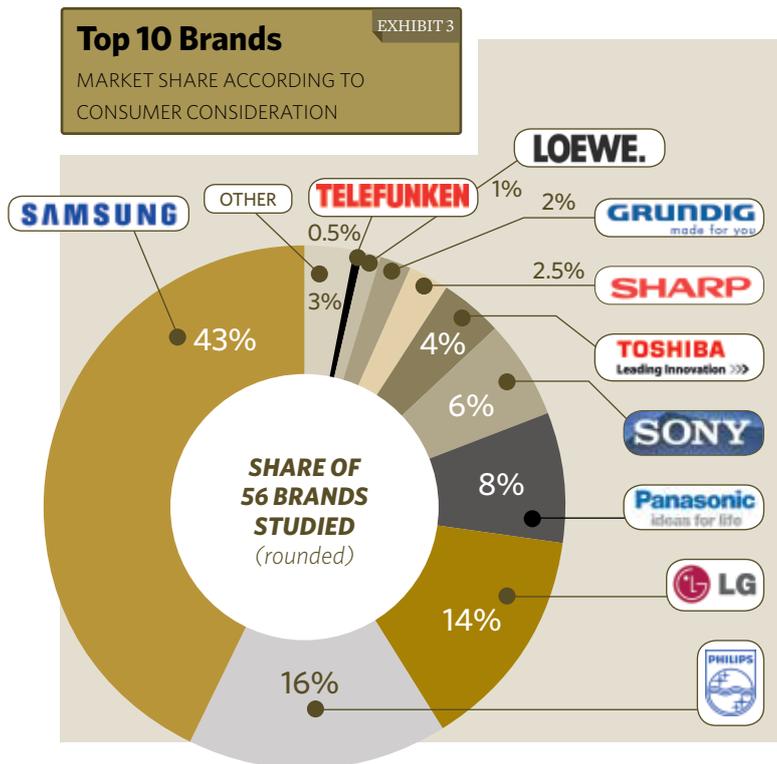
- competitive asymmetry (consideration frequency).
- Arrows represent local competitive asymmetry and point at competitors of the product they originate in.

- Arrow weight indicates how intense a competitive relationship is: the darker and thicker the arrow, the more intense the relationship
- Submarkets are numbered 1 through 30.

This document is an authorized copy for personal use of Mr. Skiera, 25/03/2017



If our method seems complex, let it serve both as an eye-opener and an action agenda for how marketing professionals need to start reorienting their future activities and capabilities.



Reaping the Benefits

Our approach for visualizing competition can yield valuable insights, not only for manufacturers and retailers but also for legislators and public policymakers, which cannot be obtained from other sources.

MANUFACTURERS can use asymmetric competitive market structure maps to quickly see how a market is organized, how many submarkets exist, which competitors they face in each submarket, and how strong these competitors are.

RETAILERS can use the maps to improve purchasing and inventory management decisions. Poor decisions regarding inventory can leave retailers with overstock that cannot be sold or shortages of hot products, leading to unhappy

customers. Both situations cut retailer revenue. When retailers have a clear picture of submarkets and their most popular products, they can serve a broad spectrum of consumer needs with a relatively small number of products.

LEGISLATORS AND PUBLIC POLICYMAKERS need to keep an eye on competitive market structures to spot early signs of problems, such as the consequences of a merger or acquisition on free markets or potential monopolies forming. In our case (**Exhibit 3**), they would be able to see that Samsung, Philips and LG have strong presences in most submarkets, and that the potential merger of any of these three companies would lead to undesirable levels of market power.

Our approach can be used to process big search data taken from any website where consumers search for and compare products in large product categories. If our method seems complex, let it serve both as an eye-opener and an action agenda for how marketing professionals need to start reorienting their future activities and capabilities. As the amount of available clickstream data continues to grow, the need for sophisticated analysis methods will become more acute. Marketers who give proactive attention to this area will have an edge on their competitors.

Overall, our approach opens the door to a deeper and more expansive analysis of today's large markets. It also demonstrates how marketers – when equipped with the right information management skills – can harness the power of big search data. □

TO KNOW MORE

- Ringel, D.M. and B. Skiera. "Visualizing Asymmetric Competition Among More Than 1,000 Products Using Big Search Data." *Marketing Science* 35, no. 3 (May-June 2016): 511-34.