Making use of data

IN SEARCH OF A BETTER ASSORTMENT

David Waller

Frederic Thomas-Dupuis

KEY TAKEAWAYS

Most food retailers have already begun to collect and leverage enormous quantities of "big data" – which can be used to drive valuable insight about the choices a retailer's particular customers make and how that retailer can optimize its own idiosyncratic assortment.

However, the challenge in harnessing this data is that it necessitates trade-offs between simplicity and value. At the furthest end of the spectrum, the complexity of computing solutions can result in insights emerging from a "black-box," which will not be clearly understood and be mistrusted.

Retailers must instead emphasize solutions that balance the art and science of working with big data, incorporating tools and processes that ensure the insights are actionable.

Big Data 1.0

When considering big data, many retailers may immediately jump to the idea of harvesting Twitter feeds and Facebook histories to serve as proxies for customer preferences. But before looking to outside sources, food retailers should first consider the huge datasets brimming with insight available within their own stores, which the majority have been collecting for years - long before the vogue of Facebook and Twitter.

This data is certainly "big"; the magnitude of information available dwarfs those of comparable industries. (See Exhibit 1.)

For food retailers, the richest data on customer behavior is readily available, and the real question becomes how to best utilize it.

Specifically, data on customer purchases provides a wealth of information about the exact set of customers that shop at a given retailer's stores. This data can be used to deduce a number of buying patterns and preferences at the individual consumer level, including:

- Which items are preferred
- The strength of those preferences
- The role of prices in driving switching and expansion of demand
- Which items are substitutes
- Which items are complements

Exhibit 1: Comparative scale of records: Billions vs. millions

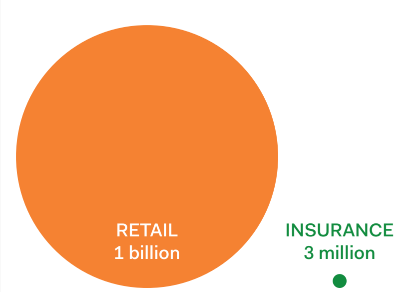
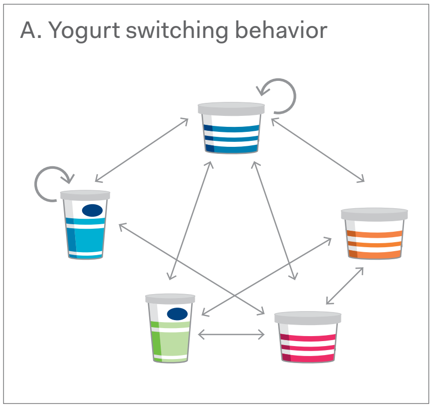
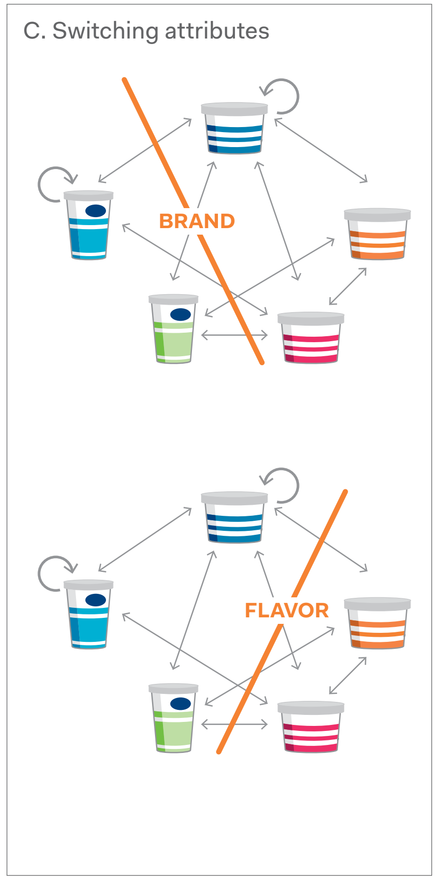
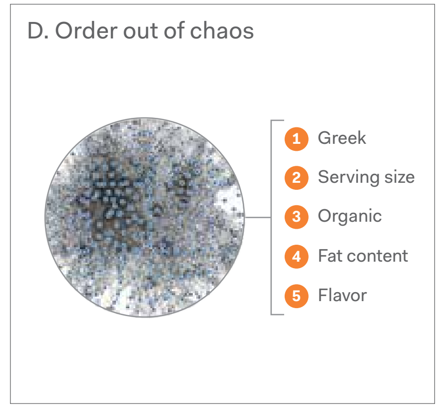


Exhibit 2: Customer constellations







Visually linking these decisions for an individual customer, we can create a unique constellation that summarizes the switching behavior of that customer over time for a specific product. (See Exhibit 2A.)

These constellations encode the crucial components of customer behavior into an easily digestible pattern; the aggregated picture for all customers, however, is significantly more complex. (See Exhibit 2B.)

The "art" of big data is making sense of the complexity and gleaning valuable insight about your customers – with the ultimate goal of translating these insights into implications for your assortment.

Make it simple

A familiar approach is to organize aggregated choices into a customer decision tree (CDT). Typically, these trees are constructed by vendors with huge research budgets who analyze focus groups to understand the prioritization of factors influencing their decisions.

The process, however, is not totally objective; vendors may have their own agendas which can cause retailers to question, for example, the reported significance of brand in influencing decisions. An even subtler difficulty with the typical CDT approach is that focus groups are not necessarily representative of the distribution of customers who shop at a given retailer, which is really the population retailers should be interested in observing.

This is the type of mental shortcut that Daniel Kahneman, a Nobel laureate in economics, warns us to avoid in his book, *Thinking, Fast and Slow*. The real question retailers are hoping to answer with CDTs is "What do customer behaviors reveal about their buying preferences?" Retailers must not be lured into substituting the answer to the simpler question: "What do a few customers in a focus group say about what they buy?" and assuming that this is sufficient to solve the more complex problem. Relying on what a small group of people tell us about their behavior is very different from observing what the sum total of all customer actions indicates.

This is one area where the data retailers already have available can be leveraged to improve the existing CDT process.

From the switching patterns of a retailer's customers, we can deduce the important factors driving these changes. (See Exhibit 2C.)

These could include any number of attributes, such as, for yogurt:

- Sweetener
- Calories
- Size
- Flavor
- Organic
- Single / Multi
- Blend or FOB
- Packaging
- Brand
- Fat Content

By determining the most important factors, those that are the largest explanatory variables for customer switching at the aggregate level, we can create order from what was originally chaos. (See Exhibit 2D.)

The final output is a traditional CDT, but one that is customized to a particular retailer and reflective of actual customer decisions. Data derived from the observed customer switching behavior indicates which attributes are most important to customers – that is, which factors are most likely to predict switches in customer buying patterns. This type of analysis can drive new insights about the relative importance of different attributes; in one real world example for a particular retailer, we saw that whether or not the orange juice has pulp is just as important as the juice brand – and other factors, like serving size, are even more important.

"All models are wrong but some are useful"

The outcomes of the type of switching analysis described above are more unimpeachable than the experimental design-dependent focus groups typically employed, as the insights are derived directly from the actual observed behavior of a given retailer's particular set of customers. However, as the famous statistician George Box once wrote: "All models are wrong but some are useful."

So rather than questioning whether such models are right, one could ask how useful the distilled observations are in creating a reliable picture of the world. The answer is that they are extremely effective.

Just understanding switching behavior alone - that is, having a global view of what amounts to a collection of probabilities about how likely a consumer is to switch between items – allows one to almost perfectly reconstruct the relative sales distributions of those products. (See Exhibit 3.)

As Box reminds us, this correlation doesn't mean the model is right – but it does mean that there exists a deep underlying linkage between the observed customer switching behavior and real world outcomes.

Which one is best?

Beyond gaining a general understanding of customer behavior, big data can also help retailers make decisions about the specific products that will optimize their shelves.

The first step in this process is to re-think the traditional method for identifying a "good" product. Sales and margin are no longer sufficient to determine whether a product should be stocked; retailers must now consider a range of additional product factors, including: space, facings, funding, customers, strategy, incrementality, trends, halo, JBPs, vendor support, vendor strategy, case packs, etc.

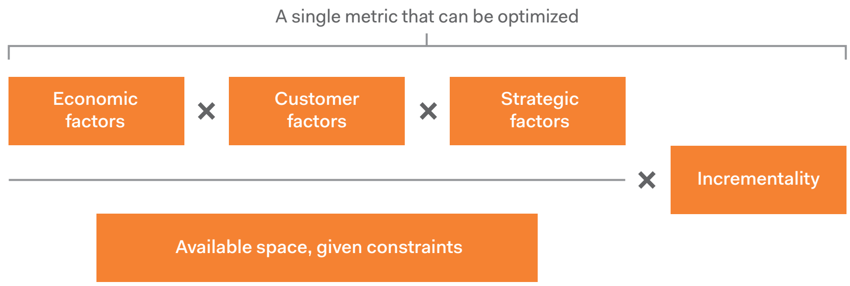
As it can be difficult to balance the impacts and determine the final net effect from this multitude of competing factors, one approach is to distill the relevant components into a single metric that can be optimized across products and categories. This metric should capture the relationship between all of the economic, customer, and strategic factors at play, as well as the limitation of available space and the incremental impact of adding the product relative to the existing assortment. (See Exhibit 4.)

However, assessing the incremental impact of a new item is a difficult problem to solve. Not only does it require an understanding of aggregate customer behavior, but it is also dependent on the items already on a given retailer's shelf. Said differently, adding or removing a single item in the assortment will necessarily change the value of the remaining items and the utility of adding new ones.

While this incrementality creates significant complexity, it is crucial to get it right. Exhibit 5 shows total utility, relative to the starting point utility, as products are removed one-by-one (an "iteration step"), across three different strategies

Exhibit 3. Product sales distributions: Actual vs. predicted by switching behavior

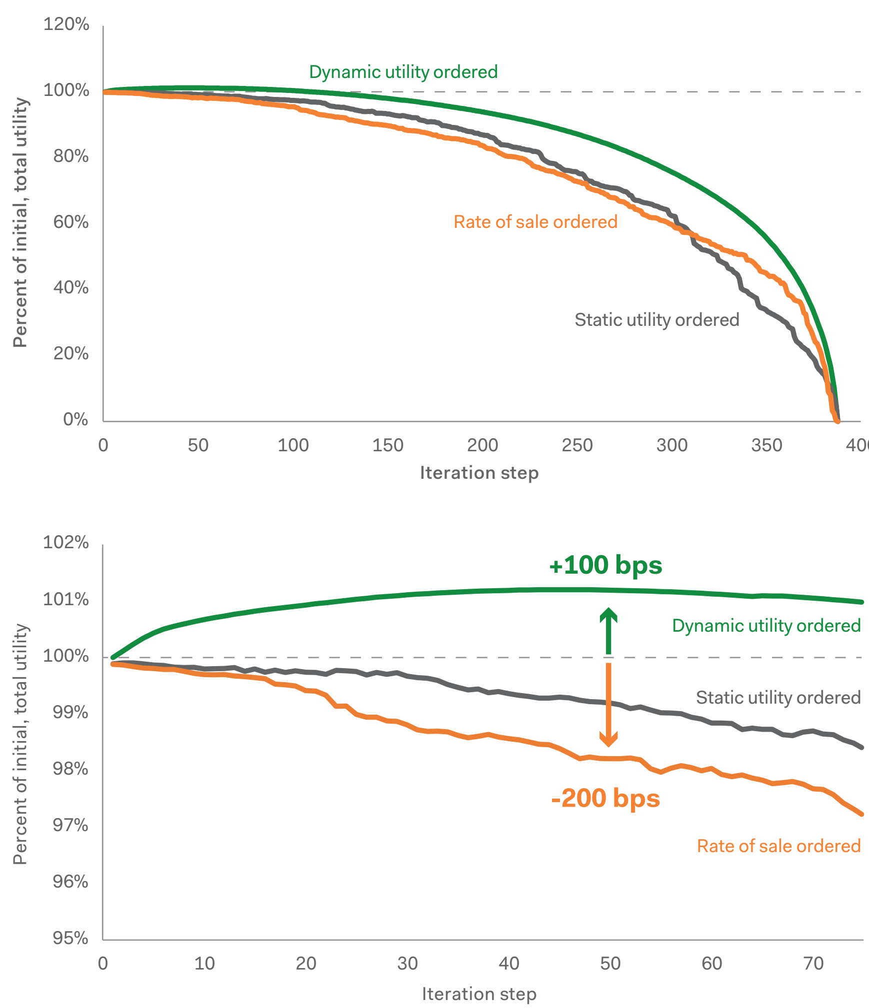
Exhibit 4: Defining utility: One metric to rule them all



that incorporate incrementality to a varying degree. The orange path shows the change in utility as products are removed simply based on which have the lowest sales; incrementality is not considered at all in this strategy. The grey line describes a slightly more sophisticated approach, which considers incrementality in the initial ranking of which products to remove first, but doesn't recalculate this ranking as items are subsequently removed. The green curve shows the impact of dynamically re-calculating incrementality based on the new assortment at each iteration step. Recall that the incrementality of an item depends on the existing assortment, so that as the assortment changes, the incremental value of an item also changes – the green line takes this evolution into account.

As demonstrated, it is clear that the strategy that delivers maximum utility is that which dynamically adjusts for the incremental impact of each item (the green line). In fact, in the early iteration steps, this strategy can actually increase utility relative to the starting point, as shown below.

Exhibit 5: Impact of dynamic incrementality: Getting it right matters



The calculations to perform this type of analysis are huge. But we must remember that, these days, it is cheap to do huge calculations and the impact can be significant.

Unique: Just like everyone else

In theory, the switching behavior and utility analyses described above could be carried out at the individual store level, given that each store, and its population of customers, is unique – but it's not necessarily true that this degree of convolution will pay out sufficiently compensating dividends for a large retailer.

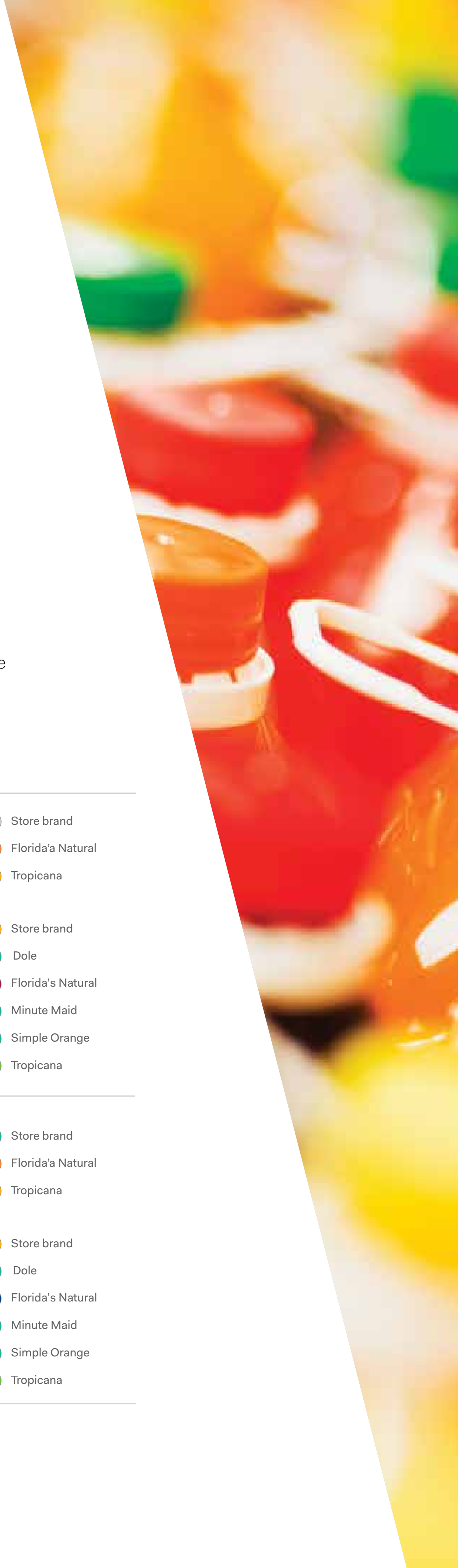
The crucial trade-off in the problem of localizing assortment balances the tension between simplicity and value. As assortments become increasingly localized, managerial complexity will grow in tandem, and at some point, the incremental value will not justify the additional complexity.

Retailers should approach this problem by trying to reduce complexity, while still capturing real and meaningful differences in customer demand. One obvious technique for this strategy is clustering. Historically, there are three, not necessarily mutually exclusive, approaches to clustering: demographic, geographic, and behavior-led clustering.

For a behavior-led approach, we can use a typical CDT to summarize customer demand in a simple way that allows us to recognize clusters across multiple dimensions. Each store has its own CDT "fingerprint" – how each branch of the tree, that is, each facet of a product type, indexes as compared to other stores. (See Exhibit 6.)

Comparing these fingerprints, retailers can determine which stores have similar fingerprints, and can be clustered together. This analysis will also show just how different a population of stores is, and thus, help determine how many clusters are appropriate, as guided by the relative intensity of customer demand across stores. (See Exhibit 7.)

Exhibit 6: CDT fingerprints



How does this method compare to the other clustering approaches, namely demographic and geographic clustering? Because it is based on customer behavior, this approach captures nuances of each of the other two approaches. For example, as shown in Exhibit 8, the demand for corn & peas versus greens & carrots seems to be driven largely by geographic taste preferences, a distinction that would have been lost in an assortment based solely on demographics. However, in the brand dimension, demographics play a much larger role and thus cannot be ignored. Clustering by behavior is the best of both worlds, highlighting the most powerful implications of each of the other two approaches.

This approach allows us to determine how different demand truly is across stores and then calculate what the corresponding assortments would be at various degrees of clustering. But how many clusters optimize the tension between simplicity and value? Computers can analyze the additional value from each increased level of granularity, but it is ultimately up to a human to consider whether this justifies the corresponding increase in managerial complexity.

Making it stick

Having interpreted the data, translating these insights into actionable processes within an organization presents another obstacle to consider. Without proper procedures to implement the results, all of the benefits of the analysis are at risk. Because of the complexity of these types of problems, it can be tempting to plug the data into a black box which spits out a final answer, abandoning any familiarity with the underlying logic.

This type of black-box approach, however, is ultimately doomed to failure. People reject what they don't understand – and they stop feeling accountable for the results. Senior leadership will expect merchandisers to be able to answer questions about the outputs and will

Exhibit 7: Clustering with CDT fingerprints

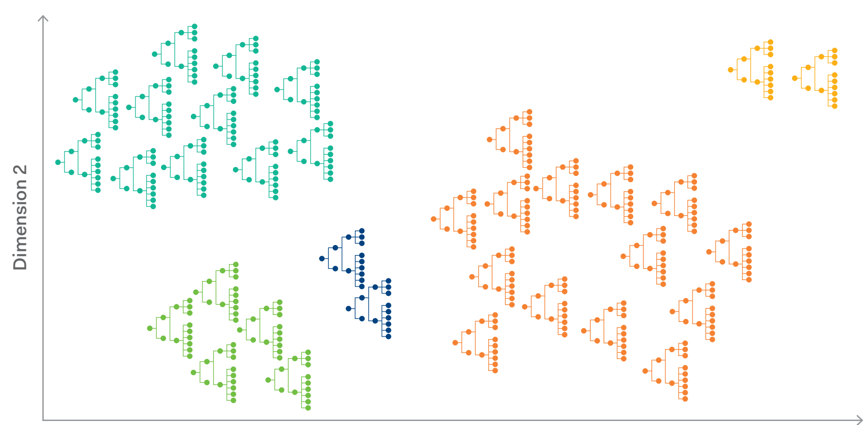
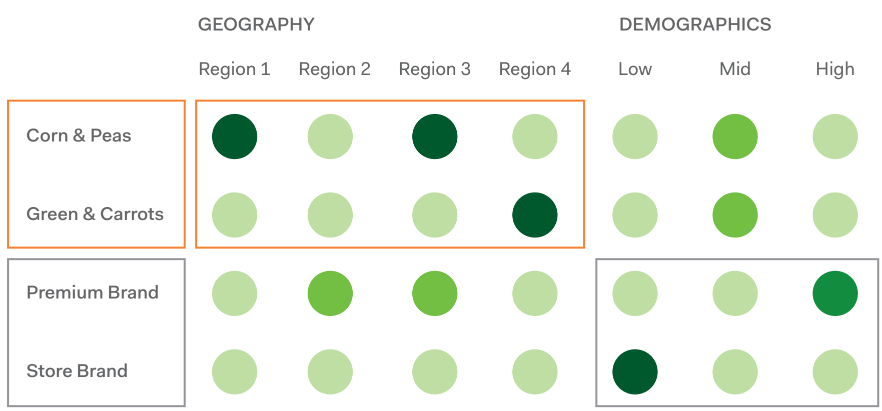


Exhibit 8: Behavior-led clustering: Best of both worlds



mistrust and reject solutions that aren't accompanied by explanations.

Because these analyses rely on human interpretation as well as purely computed outputs, the "art" of this approach necessitates a seamless interaction between the two components. When this type of coordination is successful, the outcomes can be ground-breaking.

Take as an example a series of famous chess duels that illustrates the power of a partnership between tools and their users. (See Exhibit 9.)

In 1997, the chess Grandmaster Garry Kasparov faced off against IBM's supercomputer Deep Blue and was defeated, in a dark day for mankind. Several years later, an enhanced supercomputer, Hydra, with even more computing power, was paired against various chess Grandmasters who were given access to personal computers. Working as a unit with their computers, a number of these Grandmasters were able to defeat Hydra. Even more compelling, when later matches set chess Grandmasters, with their personal computers, against amateur chess players also using personal computers, the amateurs fared better!

The implication of these tournaments is that the best outcomes occur when people and their computers work as a unit – outcomes that strictly dominate pure computing or pure people-driven solutions. The intuition of a chess Grandmaster can't be fully encapsulated into a supercomputer, but nor can a chess Grandmaster equal the computing power of Deep Blue or Hydra. It takes both to defeat the supercomputers.

Furthermore, an amateur and his PC can outmaneuver a chess Grandmaster and his PC, even as this combination was able to beat Hydra. This is because the amateurs were more willing to rely on the information their computers gave them and to accept the limits of their own knowledge, whereas Grandmasters were tempted to override computer results to their own detriment.

Analogously, the best implementations of big data solutions are those that enable their users to work in partnership with the tool. The users must be well-trained, and the tools must also be designed with the goal of interacting and communicating with users.

User training cannot stop at tool documentation, but should require further certification tests. One successful technique requires future tool users to give a presentation explaining the tool to a panel, which necessitates a far greater depth of knowledge than does simply answering multiple choice questions. Training in this way also assures that the tool outputs can be clearly communicated to senior management, backed by a deep understanding of the underlying logic.

Similarly, customized tools that prioritize intuitiveness will result in better informed users. Design your tool to complement the amateur chess player from our chess example above, as that unit will be the overall tournament winner.

CONCLUSION

When it comes to optimizing assortment using big data, there are a few key points to remember:

- You have the most informative data already (in all probability).
- The tools need to fit your processes, not the reverse.
- Your processes need discipline – don't miss the obvious stuff.
- With today's technology, fast ≠ lame and cheap ≠ fragile.
- Faced with a choice of art or science, choose both.

Exhibit 9: Chess duels