

# DATA SCIENCE AT SCALE

CREATING A WHOLE NEW  
DEPTH OF UNDERSTANDING

# 8451°

**BY: BRAD BOEHMKE**

- ✉ [bradley.boehmke@8451.com](mailto:bradley.boehmke@8451.com)
- 🐦 [@bradleyboehmke](https://twitter.com/bradleyboehmke)
- 🌐 [linkedin.com/in/brad-boehmke-ph-d-9b0a257](https://www.linkedin.com/in/brad-boehmke-ph-d-9b0a257)

**CONTRIBUTORS:**

JEANETTE MUMBECK  
JOE KELLER  
GIRI TATAVARTY

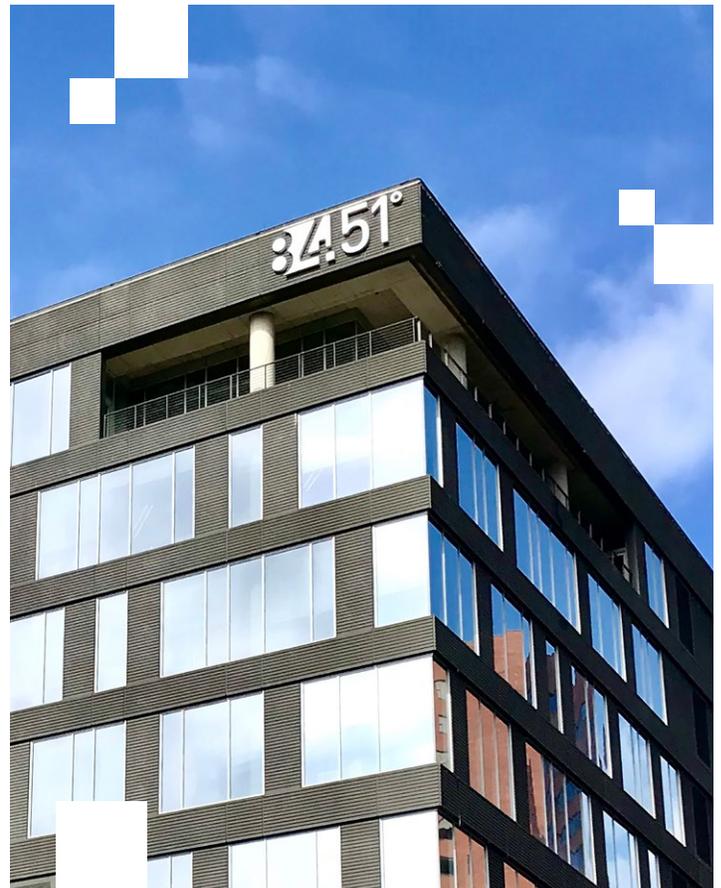
# DATA SCIENCE AT SCALE

## CREATING A WHOLE NEW DEPTH OF UNDERSTANDING

### INTRODUCTION

Searching for new channels to heighten the focus on the customer, organizations are combing for ways to expand the depth and breadth of advanced analytic approaches to be employed in the supply chain. Unfortunately, developing and deploying modern predictive and prescriptive analytic techniques at scale is extremely labor-intensive and, often, a limiting factor in an organization's abilities to create a truly analytically-focused and evidence-based emphasis on the customer. However, a recent article in Forbes<sup>1</sup> highlighted the unique ability of 84.51° to scale its use of advanced analytics across Kroger's supply chain to create a new level of emphasis on the customer.

84.51° is a wholly owned subsidiary of Kroger, which serves over 62 million households across 38 states. Originally part of dunnhumbyUSA, Kroger purchased and renamed the company 84.51° in 2015. The name coincides with the location of its Cincinnati headquarters and is a tribute to the longitudinal analytics the company. 84.51° brings together 10+ petabytes of customer data, marketing strategy, and advanced analytics to drive sales growth and customer loyalty for Kroger and more than 300 consumer-packaged-goods companies in the U.S. using a proprietary suite of tools, technology, and customized data science.



<sup>1</sup>Davenport, Thomas. "84.51° Builds a Machine Learning Machine for Kroger." Forbes. April 2, 2018.

## A FINELY-TUNED ANALYTIC MACHINE

Although the partnership has a long analytical history, in recent years 84.51° has become known as a finely-tuned analytic machine. Three key ingredients have combined to make 84.51° an analytic powerhouse: data, tools, and people.

### DATA

#### Granular transactions history

Detailed purchase history for millions of Kroger households over 8+ years

#### Integration of data sources

A mountain of information is available, including but not limited to geospatial, cost, primary research, weblog, and weather data.

### TOOLS

#### Developing custom IP

Scientific approaches to developing, testing, and deploying analytic methodologies

#### Leverage modern capabilities

Implement a machine learning and optimization framework to synthesize diverse data. Use both in-house and open source platforms.

### PEOPLE

#### Talented people

Years of domain experience as well as advanced machine learning, statistics, optimization, and programming expertise.

#### Collaboration

Unique internal collaboration approach for education and research along with external collaborations with Duke Univ. and others provide cutting-edge research.



84.51° has access to every transaction made by Kroger customers over the past decade. With over 62 million households served across the 2,800 locations, this generates over 35+ terabytes of customer data per week. Integrating this information with alternative data sources such as primary research surveys, geospatial, weblog traffic, weather data, supplier information, and more has created one of the richest retail data environments available. To leverage this vast information, 84.51° has created a custom data science platform that integrates in-house analytic IP solutions along with open source platforms.

---

With over **62 million households** served across the **2,800 locations**, this generates over **35+ terabytes of customer data per week**.

---

Creating the custom platform solution was the only means to allow 84.51° to apply advanced machine learning techniques and complex numeric and combinatorial optimization in a distributed computing environment. Lastly, the wealth of domain experience combined with analysts, researchers, and engineers with advanced machine learning, statistics, optimization, programming, and data architecture expertise allow 84.51° to develop and deploy a wide variety of techniques and tools to advance customer interactions and truly make the lives of Kroger's customers easier.



## WHAT ARE EXAMPLES OF 84.51° PUSHING THE SCALE OF DATA SCIENCE?

Here are four illustrative use cases.

### 1. REPLENISHMENT FORECASTING AT SCALE

Efficient inventory management processes are required to optimize the use of supply chain resources such as space and shipping while minimizing labor, waste, and out of stock instances. To support Kroger inventory management, 84.51° developed and continues to evolve a machine learning based, replenishment forecasting system that predicts sales for each item in each store for each of the subsequent 14 days. Causal factors such as price, demand elasticities, promotional status, date-time attributes, and other features are explored with tree building algorithms that determine which factors impact demand the greatest for every item and store combination. 84.51° created a custom solution that automates millions of Bayesian option trees to be built every night based on the most recent transaction data and predict the latest expected demands for millions of products along with real-time out-of-stock alerts.



## 2. GRANULAR CUSTOMER LOYALTY

Understanding and increasing customer loyalty is at the forefront of every organization's strategy. Subscription-based businesses use "churn" as a feedback signal for loyalty; however, non-subscription-based environments such as grocery have a difficult time measuring and predicting changes in loyalty. To support Kroger's drive for heightened focus on its customers, 84.51° developed a custom metric to measure customer loyalty. Furthermore, a deep learning model methodology was customized that can produce six-month change in loyalty propensity scores for every customer. This allows Kroger to quickly react to changes in individual customer loyalty by applying a strategic marketing campaign to regain loyalty in high risk customers. The crafted methodology leverages state of the art long short-term memory recurrent neural network models and requires a custom-built GPU configuration to run 10 hours to execute.

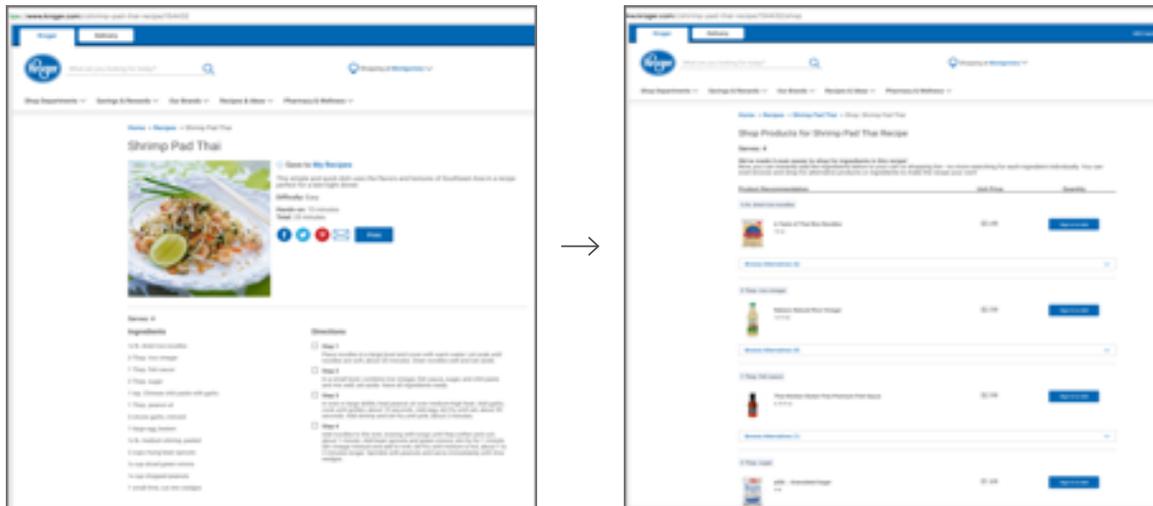
## 3. BASKET OF OPTIMAL GOODS

What items should Kroger stock in each of its stores? The optimal answer is not as simple as identifying the items with the highest sales in each category; rather, the objective is to estimate the combination of items that most satisfies the customers' needs in total. Furthermore, the combination of optimal goods changes based on store demographics and product substitute options. To help Kroger deliver the optimal basket of goods for its customers, 84.51° leverages machine learning to understand how the relationship between products effect their relative values based on demographic and product attributes. This provides insight into how customers in different regions prefer to fill their baskets based on product availability. 84.51° then applied greedy and non-greedy optimization techniques to choose the mix of products that maximize the cumulative expected value to the customer. This allows Kroger to personalize the optimal combination of products to the different attributes and needs of customers visiting its 2,500 stores.



## 4. SHOPPABLE RECIPES

The Kitchen is becoming increasingly digitized with shopping lists, recipes, meal planning, ingredient and recipe personalization and pantry management moving to digital platforms. Kroger is very passionate about putting the customer first and leading the customer in every step of the meal preparation. To help Kroger evolve this capability, 84.51° developed a new feature where users can not only look at recipes but efficiently purchase the ingredients through the online ordering system called clicklist. To make this feature possible, 84.51° applied advanced natural language processing to recognize the key ingredients, quantity, and measurement descriptors. For example, given a recipe that includes the ingredient: *4 cups cauliflower florets, boiled*, the model will extract the required information and supply the customer with the products and quantity that will fulfill the requirement. The customer can then add all necessary recipes with a single click of the mouse.



## CONCLUSION

Developing and deploying modern predictive and prescriptive analytic techniques at scale is extremely difficult. Many companies today are experimenting with analytics, but 84.51° and Kroger have managed to scale data science to the next level. Combining a rich data ecosystem, modern analytic platforms, and the brightest analysts, researchers, and engineers have all helped to create a finely-tuned analytic machine. Advanced analytic methodologies are framed, developed and deployed in the same innovative and methodological way that Henry Ford approached in revolutionizing the assembly line. As Davenport eludes to in his Forbes article, the future of data science lies in strategically applying a factory-style approach to developing and deploying analytics. What sets 84.51° apart is that they are practicing it today.

