Know Your Customer:
How Retailers Have Used Data Throughout History

Produced for NetChoice

Table of Contents

Executive Summary

This article illustrates retailers' use of data and analytics to better understand customer behavior, improve customer experience and satisfaction, and enhance business operations in the past 100 years. We review academic literature and case studies to demonstrate how retailers have always collected and analyzed customer data throughout history. They further show that the business motivations and outcomes for data analytics are strikingly similar across the different eras, even as data and technologies become better and more complex. As the article moves from "The Early Days, From 1900 - 1980", to "UPC Scanners and the Big Box Era, 1980 - 2000" and to "Modern Omnichannel Era, 2010 - Present", a common thread is that retailers' data analytics use cases and goals, regardless of being brick and mortar, e-commerce, or omnichannel, all follow a straightforward business logic: understand current and potential customers' needs and provide goods and services to satisfy those needs better than the competition. We further discuss how retailers' data usage has benefitted consumers through lower prices, greater convenience, improved shopping experiences, and increased product variety. The article concludes with a contemporary analysis of retailers' data use and shows that among 113 of the largest U.S. publicly traded retail companies, 84% of them use customer data to enhance the customer experience.

Key Insights:

- **Data and analytics have always played a crucial role in retailing throughout history.** Retailers have used data to predict demand, understand consumer behaviors, improve their operations, enhance customer experiences, and align their offerings with customer needs.

- **Despite technological eras and the types of data collected,** retailers ultimately strive to use data to increase sales, enhance efficiency in operations, and improve profit margins. These goals and rationales for retailers' consumer data collection and analysis have been highlighted and advocated by the foundational marketing principles that date back to the early days of the marketing academic discipline and have been followed by retailers of all types throughout the ages.

- **In the "Early Days,"** spanning from the 1900s to 1980, retailers observed shopping behavior through more intrusive methods, long before the advent of modern scanners and online purchase data. For instance, retailers' staff would follow customers and monitor in-store and conduct surveys and interviews to understand their shopping habits. Although these data collection techniques were costly and lacked precision, they enabled innovative retailers, such as the A&P supermarket chain,

to better comprehend regional variations in consumer needs and provide products that catered to and delighted diverse customer preferences.

- **During the "Big Box Era",** which spanned from the 1980s to 2000s and saw the introduction of UPC scanner technology, big-box retailers like Walmart capitalized on scanner data to anticipate consumer demand. They expanded their product categories per store at competitive prices, creating a one-stop shopping experience that significantly enhanced consumer convenience and satisfaction.

- In the "Modern Omnichannel Era (i.e., 2010s to Present)", ongoing advancements in information technology and the availability of new data sources for retailers (such as customer characteristics and interests from social media, geolocation data, and online browsing behavior) have further optimized retail operations through innovative practices. These innovations have led to improved customer experiences, with reduced search time, seamless omnichannel shopping and return experiences, as well as more personalized and relevant product recommendations and marketing messages.

- **The utilization of data to understand** diverse consumer preferences has also contributed to the development of new products that cater to various tastes. The recent proliferation of niche products and private labels is the result of data-driven insights into the diversity of consumer segments. This increased selection has enhanced consumer welfare by offering products that better align with individual tastes and preferences, as well as providing higher quality at competitive prices.

- **Contemporary retailers commonly utilize customer data** to inform their business decisions and elevate the customer experience. Our analysis of the latest annual reports for retailers with a market cap of $1 billion or more revealed that 84 percent of them mentioned leveraging customer data to improve the overall customer experience.

1900   **1980**   **2010**

Early Days   Big Box Era   Modern Omnichannel Era

Concluding Remarks:

Our study demonstrates that despite continuous advancements in technology, data types, and data richness, the fundamental objectives behind retailers' use of data remain consistent - to optimize business operations, enhance customer service, and diversify and improve product offerings for shoppers, ultimately aiming to boost revenue and reduce costs. The sole driving force behind data collection, analysis, and insight generation is profit maximization, as retailers engage in these costly activities with no other motivations in mind. For consumers, retailers' use of data on their shopping behavior leads to greater convenience, better shopping experiences, and products that are better tailored to their needs and preferences.

Our study serves as both a historical overview and a testament to the enduring significance of data and analytics in the retail industry, highlighting their importance in product innovation, business innovation, and enhancing consumer welfare. As innovative retailers adapt to the ever-changing technological and data landscape, they gather insights into evolving consumer tastes and behaviors, enabling them to develop new products and services that delight and benefit consumers. Given the significant improvements in customer experiences that retailers of all types have made in recent years, consumers now expect retailers to possess the necessary information to provide seamless and enjoyable shopping experiences.

Introduction

Data are ubiquitous in retailing. In both online and brick-and-mortar settings, data and analytics enable retailers to better understand their customers' behaviors, which can help retailers improve decision- making regarding operations and product and service improvements that satisfy consumers. From retailers' perspectives, these efforts result in more efficient use of marketing resources and improved sales and profit margins. From the consumers' perspectives, retailers' use of data and data analytics can provide them with more convenience, an improved shopping experience, and better products that more closely fit their needs and tastes.

In this article, we discuss how retailers, throughout history, have collected and analyzed customer data to improve their offerings and services and enhance customer delight. To do so, we draw upon decades of rigorous academic literature and case studies. The literature review will explain what data was used by retailers, how it was obtained, and how it increased sales or reduced costs. The case studies will illustrate data collection processes and data usage that have already been chronicled in the top academic literature, trade publications, or other reputable sources.

The rest of the article is organized as follows: we describe Why Retailers Use Data, as illustrated and advocated in foundational marketing principles that go back to the very early days of marketing academic discipline. We then review how data types, data collection efforts, and retail use cases evolved over the past 100 years during three periods:

1. The 1900s to 1980 (before UPCs became prevalent),

2. The Era of UPC Scanners and Big Box Stores (1980s to 2000s), and

3. The Modern Omnichannel Era, 2010s to Present.

A common thread throughout these eras is that although data and analytics have become increasingly sophisticated over the years due to technology, retailers' data analytics use cases and goals, regardless of being brick and mortar, catalog, or e-commerce, all follow a straightforward business logic:

understanding current and potential customers' needs and providing goods and services to satisfy those needs better than the competition.

Finally, to provide a more generalized background and demonstrate how extensively contemporary retailers use customer data to enhance the customer experience, we reviewed the annual reports of 113 of the largest U.S. publicly traded retail companies by market cap. Our analysis of the most recent annual reports for retailers with a market cap of $1 billion or more showed that 84 percent mentioned using customer data to enhance the customer experience.

Why Do Retailers Use Data?

In answering the fundamental question "Why Do Retailers Use Data," it is worthwhile to first define retailers' main functions and goals and assess how the use of customer data helps them achieve these goals. Retailers are the final link in the supply chain from producers to consumers. Retailers make money by attracting shoppers and converting them into customers. Retailers only profit by making customers happy, so the interests of both shops and shoppers are aligned. Only by consistently providing compelling value to consumers can retailers retain existing customers, acquire new customers, and earn a profit.

Retailers need to know:

1. What products consumers might want to buy and how much they might want to buy;

2. What price to charge for the products;

3. How to use their services, promotions, and shopping environment to attract new customers; and

4. Increase loyalty among current customers.

Fundamentally, retailers need to be able to predict demand and to understand consumers' behaviors. These tenets are set forth by early marketing scholars and published in the Journal of Marketing (the oldest and one of the most prestigious academic journals in marketing) which has greatly influenced marketing frameworks, business education, textbooks, and generations of retail managers and executives.

In "Marketing Principles" (Bartels, 1944), one of the earliest academic papers to list guiding principles in how marketing should be conducted, Robert Bartels discusses **these principles:**

1. **Properly understanding demand is fundamental for retailers to run their business** ("the initial stock for a new store should consist approximately of the proportions of stocks in going stores. Subsequent stocking should reflect demand and supply peculiarities of the locality," pg. 153).

2. **Pricing should be set based on demand data** ("The pricing policy for a store cannot be based upon the gross margin of any one profitable item. When a number of items are handled, the general policy must be composite, reflecting the nature of demand for the good..." pg. 153).

3. **Consumers' needs and wants vary with income** ("as income for family increases, the percentage of income expended for food and housing decreases, and the percentage ex- pended for clothing and miscellaneous items increases." pg. 154).

4. **The need to understand the variation in consumers' tastes in order to carry and develop proper product assortments to satisfy those tastes** ("So long as tastes vary it will be impossible to standardize consumer goods in the same way as paving-bricks or steel rails can be standardized." pg. 154).

These points demonstrate that critical retailer decisions require data to gain insights into consumer demand and consumer behavior. Acknowledging such needs, the early marketing academics then proceeded to publish papers on the associated data types and data collection methods to achieve retail objectives. In the next section, we review the early days of data collection and analytics in retail.

The Early Days (1900 – 1980)

This section discusses the foundational approaches to retail data and goes back to the very early days of systematic marketing and retail knowledge. This foundational knowledge influences retailers, to this day, in how they approach data collection, analysis, insights, and business improvements.

Before the 1980s, there were no UPC scanners and no point-of-sales (POS) data – technologies that every retailer nowadays takes for granted. So, retailers (as advocated by marketing academics) relied on grocery diaries where some customers were paid to keep track of what, when, and how much their households made grocery purchases. Despite its low accuracy and low reliability, this data source gave market researchers some basis in data to look at household purchase patterns and understand simple yet important descriptive statistics such as category demand and brand shares. Retailers also surveyed consumers to understand taste variations in local markets, why some brands are preferred over others, and the factors that made it easy for people to shop. These data collection efforts were labor-intensive and often used mail or in-person surveys, in-store observations, and pen-and-paper recording by retailers' staff.
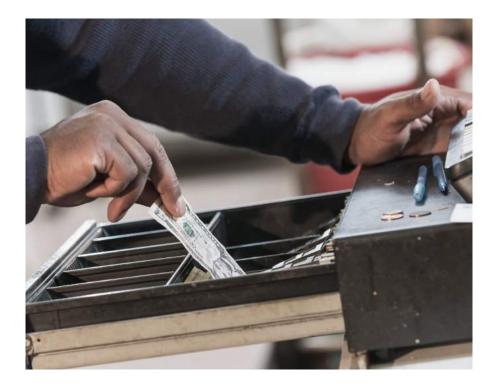
To illustrate the various approaches, below are some of the most influential and seminal academic articles from the 1950s and 1960s that provided very early roadmaps and examples of why consumer data are valuable to retailers and how retailers should systematically collect and analyze consumer data.

By doing this, marketers and retailers could focus on improving the attributes that led to more sales and assess the effectiveness of price promotion for each product category.

Given the strong connection between people's survey preferences and their actual purchases, Banks advocated for brands to undertake routine consumer-based market research efforts to increase the success rate and benchmarking across the competition. The following quote wonderfully summarizes how the data can improve decision-making:

> *In the views of the writer, preference measurement can come to have a wider application than the prediction of the relative purchase of a series of competing alternatives. It can offer basic information as a guide to the formation of marketing policy. Market by market, the marketing manager can determine how well his product or brand is known, what people who have used it think of it, and the purchase intentions of users and non-users. These findings will serve to point directly to basic marketing weaknesses.*

Using Consumer Surveys to Measure Product References

For example, in Banks' (1950) article "The Relationships between Preference and Purchase of Brands", the researcher surveyed 65 housewives in Chicago from April to June 1948. He linked the survey of consumers' stated preference to actual purchases, to predict the success of brands in various household categories (e.g., cleanser, coffee, ice cream, peanut butter, potato chips, catsup) and to see which product attribute (e.g., cleansing ability and price for cleansers, and flavors and prices for coffee) was the most important.

Remedies can be proposed for the weaknesses revealed by these diagnostic tests – advertising and sampling to increase knowledge level, product or package improvement to raise preferences for the product, and the use of price reductions or special promotional deals to increase purchase intentions if preference is high and purchase is low. (Banks, 1950, pg. 156)

Merging Industry Trends And Consumer Survey Data To Inform Store Location Decisions

Kelley (1958) showed that consumers' "convenience cost" can drive behaviors and has implications for how shopping centers and stores of various sizes should be located relative to population centers to maximize the selling rates of goods. Accordingly, he developed a method for retailers to estimate customers' travel costs in terms of time and distance.

He suggested that retailers should calculate the convenience cost by merging aggregated industry data on retailing dynamics (which was first collected in 1927) with population trends and actively collected data on convenience costs (e.g., driving distance or travel time to the store) from consumers either through surveys or observations.

In-Store Data Collection To Understand "Who Buys Where, What, When, And How"

Applebaum (1951) offered an early roadmap for consumer data collection in-store. In the opening paragraph of the paper, he wrote about the fundamental importance of these consumer-level data for insights and predicted their popularity in retail for years to come:

Studies of customer behavior in retail stores usually deal with (i) identification of customers and (2) their buying behavior patterns. The aim of such studies is to ascertain who buys where, what, when and how. In addition, such studies endeavor to learn about customer response to sales promotion devices. The results of these studies are useful in the solution of an array of marketing problems. These studies are gaining importance in marketing research. (Applebaum, 1951, pg. 172).

He observed that customers' buying behaviors can be grouped according to where they go to purchase, quantity and types of items purchased, timing and frequency of purchase, method of purchase (e.g., shop on cash vs. with credit, alone or with someone else, walks or drives to store), and response to sales promotion devices (e.g., displays, pricing, product demonstrations).

Furthermore, he offered data collection and analysis techniques for acquiring this information, via 1) Analysis of Records, 2) Observation, 3) Interviewing, and 4) Controlled Experimentation.

These data goals, data collection methods, and analyses set the foundation for the present-day attitudes and approaches toward data of all retailers. Below are some of his best practices for each data collection technique:

Analysis of Records

Applebaum advocates the importance of transaction records for retailers, noting that they naturally possess such data during daily operations, and if not, they can collect it using existing technologies. He asserts that sales records, including details of products sold, timestamps, and prices, can reveal customer buying behavior patterns. Valuable records include commodity movement by brand, size, and price; dollar sales by departments and time; and customer sales transactions by departments, purchase amount, and items bought. Comprehensive and accurate records spanning a long period are more valuable. If desired records are unavailable, retailers must implement procedures to obtain them.

In-Store Observation

Retailers can easily collect demographic information such as gender, race, and age through in-store observations. Customer behavior can be observed discreetly, with data recorded using codes in notebooks. Checkout stands are ideal for observing purchased items.

Interview

In-store customer interviews can yield information on their characteristics and buying behavior. Questions about travel, visit frequency, and length of patronage can be easily asked and answered, ideally after the purchase. Customers' addresses help determine their origin and the store's trading area. This information, combined with other data, can reveal customers' income, education, and national origins. It is essential to avoid asking too many questions at once.

Controlled Experimentation

Retail stores are an invaluable setting for conducting experiments to study customer behavior and responses to products, methods, and marketing devices. Controlled experiments involve selecting two comparable groups of stores, with one serving as the control group and the other as the test group. In the test group, one condition (product, method, or sales promotion) is altered, and its impact on customer buying behavior is measured through sales. The results are then compared between the two groups to determine the effect of the changed condition.

Conducting Experiments

Applebaum (1951) briefly discussed the value of running in-store experiments to understand customer response to price promotions and other promotion methods. Pessemier (1959) further developed this technique in his paper "A new way to determine buying decisions": "By analyzing how shoppers behave under simulated market conditions, brand-switching and brand-loyalty patterns can be analyzed. Further development of the technique described could be particularly valuable in shaping a firm's marketing strategy." (Pessemier, 1959, pg. 41).

To illustrate the effectiveness of this research method, he conducted price experiments on 103 subjects in toothpaste and cigarette categories, by varying the price for the focal brand, changing the number of options, and changing prices for other brands. The results allowed the retailer to understand how customers make tradeoffs between prices and brands.

He then urged retailers to use experiments on a smaller scale, as his study "demonstrates that large amounts of significant

data about consumer behavior can be gathered rapidly at low cost by controlled experiments and the experiments show that the type of information obtained could be especially helpful in making marketing decisions." (Pessemier 1959, pg. 41).

Data Fusion – Combining Multiple Data Sources For Insights

These data use techniques were further enhanced and creatively combined by Granbois, in the Journal of Marketing article "Improving the Study of Customer In-store Behavior" (Granbois, 1968), through a research collaboration with a retailer. In the opening paragraph, he laid out the motivation behind his research: "Successful decisions about price, display, layout, and other variables controlled by retailers depend upon reasonably accurate predictions of customer response to variations in these factors. Such predictions may be substantially improved by continuous or periodic attempts by retailers to measure variations in customer in-store behavior as display, price, or layout factors are varied." (Granbois, 1968, pg. 28).

He then showed how a combination of entrance interviews and direct in-store observation can give management more meaningful information about their customers. In particular, he proposed that retailers should first collect, via interview, customer characteristics such as shopping party (i.e., how many people are in the party) and household characteristics (age and gender of each person in the shopping party, household income). He believed that those "characteristics provide useful knowledge of the types of customers attracted to the stores being studied, and when analyzed in terms of shopping behavior they provide insights into possible explanations of observed variations in behavior." (Granbois, 1968, pg. 29).

Then, retailers' staff should observe the shopping party's store traffic patterns to understand the types of products purchased, the amount of time shoppers lingered in areas, the decision-making dynamics within the shopping party, whether the shopping trip was planned or unplanned, and shoppers' responsiveness to promotions.

He provided best practices for observing customer behaviors in-store, stressing the thoroughness of data collection but also

non-intrusiveness: "As the shopping party moved through the store, the observer followed discreetly and traced its path on a map of the store's layout, marking each point where an item was handled or purchased. The exact composition of items purchased, the dollar transaction size and the time of entering a checkout lane were visually observed and recorded. Of a total of 388 shopping parties studied, only one respondent appeared to be aware of this observation. Customers either failed to notice the observers or presumably mistook them for inventory takers." (Granbois, 1968, pg. 29).

This approach of using multiple data sources (retailer's transaction records, interview/surveys, and observation) allows the retailer to understand the mechanisms of how and why customers shop and how these behaviors are linked to sales. Understanding these underlying mechanisms allow the retailer to open up this "black box" and design promotions and optimize store layout to enhance total sales:

> *"Combining conventional audit procedures with the interview-observation technique facilitates estimates of the impact of an experimental promotion or price for a single product on purchase timing for that product and purchase rates for other products, An additional benefit is that the characteristics of shoppers responding to the experimental variable can be determined, so that experimental promotions or prices which attract specific customer sub-groups can be distinguished from those with more general appeal." (Granbois, 1968, pg. 33).*

> *"The method [also] permits the assessment of the effects on traffic patterns of layout or display changes that are not limited to single products but that are expected to have broad effects. By studying customers both before and after the change, improvements in traffic as well as the value of additional traffic in terms of unplanned purchasing (and, therefore on transaction size) can be estimated." (Granbois, 1968, pg. 33).*

Setting The Foundation For The Future

The approaches to data by these early academics created a foundation for the future. For example, "analysis of records" set the foundation for collecting and analyzing data from UPC-

scanner, point-of-sale (POS) systems, customer relationship management (CRM) systems (e.g., loyalty card programs), and consumer online clickstream, to inform on overall demand patterns and consumer behaviors. "Observation" set the foundation for practices such as in-store video tracking, geo-tracking using mobile GPS, and understanding market needs and market structures through social media and forum discussions. "Interviewing" provided best practices for survey designs and focus groups. "Controlled experimentation" is used routinely by contemporary brick-and-mortar retailers to vary prices and promotion practices and store layout, understand what products sell best where and when, and optimize accordingly.

The experimentation approach is the same concept that e-commerce companies, and companies in general, use in A/B testing – for different customers, changing the design of the webpage, and sending out different promotion offers, to assess which design and offers provide the best response. The concept of "Data fusion" illustrated by Granbois (1968) is the underlying philosophy behind big data in retail, where retailers combine and correlate multiple data sources – transactional, survey, observational, and 3rd party purchased data from data aggregators, to better understand and serve their customers.

As we will see in later eras, while technologies improved the efficiency of data collection and the quality of data, the ultimate insights as laid out by Applebaum's 1951 article "...to ascertain who buys where, what, when and how... and to learn about customer response to sales promotion devices", and more broadly, Bartel (1944) 's "Marketing Principles", remain unchanged and are central to all retailers' approaches to data use.

Case Study – A&P As A Pioneering User Of Data

During this early era, even though retail data was limited and data collection was cumbersome, a few innovative retailers took these approaches to heart, gathered data for insights, and acted upon data insights to optimize their operations and improve offerings to customers. One notable example of such a company was the grocery-chain A&P.

Although now defunct, more than 80 years ago, the A&P chain was an innovative and successful retailer that offered consumers a wider range of products than the competition and at lower prices. It was also one of the first retailers to use data to inform its business decisions both internally as well in its offerings to customers. For example, to improve the supply chain between its manufacturing plant and retail stores, "A&P's manufacturing plants earned money because the company learned to use the flow of orders from its [retail] stores to run the plants steadily at full capacity, reducing the waste that comes from expensive factory equipment that is not fully utilized." (Levinson, 2011, pg. 265).

A&P also succeeded because it did what any data-informed retailers (be it brick and mortar or e- commerce) today would do. For example, A&P used customer demand data to meet previously unrecognized regional preferences: "Philadelphians, it found, liked their butter lightly salted, with a light straw color, whereas New Englanders preferred more salt and a deeper yellow coloration." (Levinson, 2011, pg. 105). And the company's "mass of sales data allowed A&P's bakeries to forecast demand with a high degree of accuracy, minimizing returns of stale bread and doughnuts" and thus reducing costs and ultimately retail prices. (Muris and Nuechterlein, 2019).

A&P's understanding of consumer demand and taste also allowed it to create its private-label coffee and was one of the early pioneers of the private-label approach to competing with third-party products. From the company's earliest days, it sold its own leading coffee brand (Eight O'Clock) alongside Maxwell House and Folgers. The same practice of a retailer developing its private label and selling it alongside third-party products is prevalent among retailers of all sorts today, for example, the Kroger chain with its "Kroger" brand, Whole Foods with "365", Walmart with "Sam's Choice", Costco with "Kirkland Signature", Macy's with "Charter Club", CVS with "CVS Health", and Amazon with "Amazon Basics".[1] The results of these approaches were lower prices and delighted customers. During the late 1940s, when the U.S. government was planning to break up A&P because of the concern that it was getting "too big," ordinary consumers wrote letters to the Attorney General. One representative letter read "I am dropping you a line to see if you will try and help us housewives save our A&P. stores. We surely could not make our money go so far in small stores." (Muris and Nuechterlein, 2019).

UPC Scanners and the Big Box Era (1980 – 2000)

The period from the 1980s to the 2000s witnessed two major retail trends: advances in retail technology such as UPS scanners, and the rise of big-box chain stores such as Walmart. These two trends combined to produce the following results:

1. Big-box stores were early adopters of retail technology, which improved operations and insights. This lowered big-box stores' costs, which allowed them to expand their scale and offer more product categories in their stores and at lower prices compared to smaller competitors.

2. Consumers were delighted with lower prices and the convenience of a one-stop shop (e.g., lowered search and transportation costs) and flocked to big-box stores.

3. The increased revenue earned by big-box stores allowed them to further expand in terms of geographical coverage, store size, and product categories. This offered conveniences that further delighted consumers and resulted in even greater market share.

4. The scale and complexity of big box stores' offerings gave them the most reason to make large investments in technology. These new technologies further changed the retail landscape and made retailing generally more efficient.

The efficiency led to big-box stores scaling up and offering more product choices as a one-stop shop for consumers. Just like A&P's innovation of data-based practices improved its business, big-box stores' early adoption of retail technology allowed them to benefit from new data and insights, which improved their businesses and enabled them to scale up.

During this time, big-box chains increased their square footage, employment, sales, and product breadth:

> *"In 1977, 53.7% of general-merchandise stores had total space exceeding 50,000 square feet; by 2007, that fraction was up to 61.5% of general-merchandise stores. Larger stores carry more distinct product lines: stores larger than 50,000 square feet carry 25 - 70% more distinct product lines than smaller stores... The share of general-merchandise stores selling groceries increased from 64% to 88% between 1977 and 2007, and the share selling furniture increased from 42% to 55%. Anecdotally, moreover, many general merchandisers have added groceries to their product offerings; since the late 1980s, Walmart, Target, and Kmart have all added "superstore" formats that include full grocery stores (Basker and Noel, 2009)." (Basker et al., 2012)*

Consumers Benefit From The Increased Product Assortment

The growth of product offerings by big-box chains contributes to increased consumer welfare. As these retailers add new product lines, they offer more benefits to consumers in terms of one-stop shopping. This allows customers to buy multiple items, such as clothing, office supplies, and groceries, in a single location, reducing the purchase costs associated with shopping in multiple stores, including transportation, time, ad mental energy associated with searching (Bonne, 2004; Gogoi, 2008). Walmart itself emphasizes the convenience of one-stop shopping in its annual reports, stating that 22% of consumers identify broad selection and variety as the best thing about the company (Walmart Stores, Inc., 1994, p. 7; Pew, 2005).

Based on the census data from this period and based on an economic model, Basker et al. (2012) show that a retailer's scale economies provide consumers with benefits from one-stop shopping, which in turn generate more incentives for the retailer to increase its product offerings. As a retail chain adds stores and increases sales volume, economies of scale cause the marginal cost of operation to fall. Lower marginal cost induces the chain to increase its range of product offerings, drawing in more customers who take advantage of its one-stop offering. The combination of more products and more customers at each store increases store profit, prompting the retail chain to add even more stores.

The economic repercussions of even a small increase in the number of product lines carried by a general-merchandise

chain can be quite large. In the famous case of Walmart's addition of a full line of groceries to its "Supercenters," Walmart nearly tripled its store count over a twenty-year period while adding grocery lines to most of its stores (Basker and Noel, 2009), and continuing to add other lines, such as pharmacies, a process started even earlier. This process has had far-reaching effects on consumers' access to groceries, on prices, selection, and quality industry-wide. Incumbent supermarkets have reduced their prices by 1-2% on average when a new Walmart Supercenter opens in their city (Basker and Noel, 2009) and reduced their stockouts by up to 24% compared to pre-entry levels (Matsa, 2011).

Big-box chains' expansion increases price competition, which can benefit consumers:

"As large general-merchandise chains add more lines, and more stores, they compete more vigorously with small businesses in an increasing number of product lines and locations. Our model implies that much of the increased competitive pressure on small retailers is due to the fact that growing chains face decreasing marginal cost curves." (Basket et al., 2012, pg. 575)

Similarly, in a review of Walmart's rising influence, Carden (2013) has found that Walmart's effect on prices and consumer access to diverse products has led to a rise in consumers' real wages.

It is worth noting that state laws aimed to protect small retailers from having the same brands sold at lower prices by big stores also accelerated big stores' private label strategies – "'Competition was reduced, pricewise; but it was increased in an unexpected way. The attraction of "sure profits" on fixed prices lured thousands of new merchants into business. As they scrambled for customers, price-cutting returned in another form. Now, in order to undersell the "protected" independents, most large chain stores put out their own brands. R. S. Macy & Co. has over 1,400 such items. Gimbels offers its own brand of bonded liquor.'" (Time, 1948).

Case Study – Walmart's Use Of Retail Technology And Data For Growth

Walmart is an example of a big-box chain retailer that used technology and data to thrive in this era. Below is a brief timeline of how Walmart has pioneered the adoption of retail information technology to fuel its expansion from the 1970s to the early 2000s (CIO, 2007; Wang, 2006).

As early as 1975, Walmart leased an IBM 370/135 computer system to maintain inventory control for all merchandise in the warehouse and distribution centers and to prepare income statements for each store. Walmart also had electronic cash registers in more than 100 of its stores to record point-of-sale (POS) data to improve inventory control. Then in 1977, it built a companywide computer network and deployed a system for ordering merchandise from suppliers.

In 1983, the company began to use bar codes for scanning POS data, and in 1984, store associates started using Texlon handheld terminals when reordering merchandise. Upon scanning a shelf label, the Texlon handheld terminals provided a description of the merchandise, information on prior quantities ordered, and other data.

In the early 1990s, Walmart created a prototype data warehouse of historical store sales data and deployed the Retail Link system to strengthen supplier partnerships. The system provided vendors with information on sales trends and inventory levels, so suppliers like P&G could tap into Retail Link to see how well its products are selling.

Then, starting in 1996, in the early days of the Internet, Walmart made Retail Link an electronic data interchange available via the Internet and began using the Internet as an application platform and for data exchange with thousands of its global suppliers. In 2005, Walmart deployed radio frequency identification (RFID) technology.

Over this period, Walmart grew from 125 stores and $340 million in sales in 1975 to $349 billion in sales, nearly 2 million employees, and 6,775 stores worldwide by the end of 2006. (CIO, 2007).

To manage this remarkable growth, Walmart used data provided by technology to manage its complex network of global suppliers and inventory levels across all stores, greatly improving the efficiency of the supply chain. With stores across a large geography, millions of customers, and many product categories, Walmart also had access to information about consumer product preferences. These data were gathered item by item at checkout and then recorded, mapped, and updated by store, state, and region. These data insights on "what products sell where and when" allowed for better demand prediction and more timely inventory allocation across different regions.

Thus, these various data allowed Walmart to improve efficiency at all levels of its operations, from the front of the store in terms of demand prediction and understanding store operations (such as how many cashiers are needed during certain hours at a particular store), to the back of the store, where details about manufacturers' punctuality in filling inventory are recorded to use for future negotiations.

A 2004 New York Times article provided a vivid account of Walmart's use of data to guide business decisions.

When Hurricane Frances was about to make landfall on Florida's Atlantic coast, Walmart used data on what had happened when Hurricane Charley struck several weeks earlier: "Backed by the trillions of bytes' worth of shopper history that is stored in Walmart's computer network, the company could 'start predicting what's going to happen, instead of waiting for it to happen', as put by Walmart's CIO Linda M. Dillman.

The experts mined the data and found that the stores would indeed need certain products - and not just the usual flashlights. 'We didn't know in the past that strawberry Pop-Tarts increase in sales, like seven times their normal sales rate, ahead of a hurricane,' Ms. Dillman said in a recent interview. 'And the pre- hurricane top-selling item was beer.' Thanks to those insights, trucks filled with toaster pastries and six- packs were soon speeding down Interstate 95 toward Walmarts in the path of Frances. Most of the products that were stocked for the storm sold quickly, the company said.

Such knowledge, Walmart has learned, is not only power. It is profit, too." (New York Times, 2004). Data are also used for insight into the planning of new stores and the expectation of consumer demand and taste in the new region: "…achieving sleeker operations is not the whole story. In many ways, data are used to forecast and drive Walmart's business. 'we use it in real estate decisions, understanding what the draw is like and what the customers will be like' Ms. Dillman said." (New York Times, 2004). This approach is in the same spirit as Kelley (1958)'s original treatise on the need for retailers to gather regional demand data to inform where to open new stores and how large stores should be.

As can be seen above, Walmart and other big box retailers fundamentally use data to follow exactly what "Marketing Principles" preach – carry the right products (and enough inventory, but not too much) for the right customers, and do so profitably. Furthermore, a benefit of POS and UPC data also has been to correct inaccuracies in what are deemed the "bestsellers". Genres that had long been deemed culturally unimportant – ranging from romance novels to country music – were often undercounted in sales because the New York Times, Billboard, and other list-makers relied on information from specialist books and music retailers. When Walmart's barcode data became available, bestseller lists could better reflect purchases by consumers who shopped at Walmart, thereby providing Walmart with a more accurate sense of cultural reality.

Modern Omnichannel Era (I.E., 2010 – Present)

The modern era in retail is characterized by "big data". Retailers during this era face a rapidly changing market landscape, due to the need to adapt to the ways that modern customers shop: in-store, online through mobile channels, and via various other connected devices. These data have enabled retailers to introduce new business practices, such as personalized recommendations and omnichannel retail, further enhancing customer convenience and shopping experiences.

As the cost of data storage and processing continues to decline, retailers are collecting more data, including purchase data from enterprise systems (e.g., quantity purchased, price and cost of each item, size of discounts applied, composition of the shopping basket, and time and date of purchase) as well as social media and demographic information about customers. Big data allow retailers to create massive data warehouses that combine multiple data sets and thus uncover unique insights. For example, retailers can combine customer purchase and loyalty data, demographic information (e.g., age, gender), and geographic data (e.g., store locations and weather forecasts) to build better demand models than those in the previous era and provide timely personalized product recommendations. This leads to higher customer satisfaction and better firm performance.

For example, in 2013, Australia's largest grocery retailer Woolworths acquired 50% of Quantium, a data brokering firm to enhance its customer information strategy and data analytics. The goal was to design customized promotions to increase sales (Mitchell, 2016). Quantium partnered with Facebook to collect and analyze customer profile data including lifestyle characteristics, and Woolworths uses this information to design tailor-made direct marketing promotions. The result was also beneficial for consumers and increased consumer satisfaction - the retailer generated a 24 percent increase in customers' reported satisfaction with the marketing communications they receive (Pascoe, 2017).

New Data Sources Available To Retailers

While the retail principle of meeting customer needs profitably remains timeless, specific practices are constantly changing. New academic research and business practices often arise from new data and new attendant methods of analysis. In contrast to the product-level data that guided big box retailers from the previous era, consider the wide variety of data sources available to retailers today. These data sources can be categorized into three large groups: (1) Traditional Enterprise Data Capture; (2) Customer Characteristics and Data Capture From Social Media; and (3) Customer Location-Based Data. The intersection of these groups provides insights and possibilities that arise from capturing and modeling diverse and better data.

Traditional Enterprise Data Capture

The first group of data sources comes from traditional sales data obtained through UPC scanners and inventory data obtained from ERP or SCM software. These sources provide insights into product performance and the four elements of marketing (product, price, promotion, and place) at the store, aisle, shelf, and other levels. By utilizing these data, retailers can study the relationships between products by analyzing market baskets and cross-category dependence (e.g., Blattberg, Kim, and Neslin, 2008). They can also monitor aggregate sales and inventory patterns for each SKU, determine the sensitivity of prices and shelf space to changes in demand at different levels of aggregation (i.e., at the category, brand, and SKU levels), and evaluate the impact of prices, promotions, and product attributes on sales.

Customer Characteristics and Data Capture From Social Media

The second group of data sources is focused on identifying consumers and obtaining information about them, such as demographics, purchase history, preferences, response to promotions, product return history, and basic contacts like email addresses. This type of data collection enables a household- purchase-occasion to become the unit of analysis

A common way of identifying customers is through the use of loyalty or bonus cards, where customers sign up in exchange for discounts and promotional offers. The advent of household-specific panel data has made it possible to estimate household-specific preferences and to design household-specific promotions and marketing campaigns (Rossi et al., 1996). To link all relevant information about a customer from multiple sources, a single customer ID is used.

Retailers can also gather information from the customers' online behavior, including online browsing recorded via cookies inquiries, complaints, email click-throughs, and data purchased from syndicated sources about their online searches and products viewed but not necessarily purchased. This information can be used to better understand consumers' purchase contexts and responsiveness to promotions.

With the rise of social media usage, retailers can also gather information from customers' social media accounts, obtained through either syndicated means or by customers voluntarily providing their social media information. This information can provide insight into all stages of the customer decision-making process, from awareness to action (Dubois et al., 2016), and enable behavior-based targeting, preference identification, and word-of-mouth recommendation (Wang et al., 2013; Zhang, 2019). These behavioral- based targeting practices assure retailers that their advertisements are more relevant to the consumer. Instead of being bombarded with irrelevant ads, consumers are more likely to receive ads for products they actually want or need, and this can help consumers make more informed purchasing decisions, leading to greater satisfaction with their purchases. Further, behavioral-based price promotions can help consumers save money by offering promotions and discounts on products they are more likely to need.

Customer Location–Based Data

The third group of data sources involves using customers' locations to understand their preferences and purchasing tendencies, and to design targeted marketing efforts. Smartphones have greatly impacted the way location-based data is collected and used, with data being gathered through mobile shopping apps installed on the devices (Ghose and Han, 2014). The real-time location data of customers within or near retail stores can provide valuable context that can be utilized to make marketing messages and offers more relevant and impactful to their behavior and attention (Luo et al., 2014).

Additionally, data is also gathered from habits and unconscious behaviors of customers, such as their eye movement patterns when viewing a product or web page (Wedel and Pieters, 2000) or the path they take in a physical store which can be tracked using RFID chips in shopping carts (Larson et al., 2005). These behaviors are ones that customers are not aware of or unable to articulate in traditional surveys.

New Practices

Mature retailers can continue to prosper, without adding new stores, if they focus on growing sales in their existing stores using data to optimize the drivers of sales, including assortment, inventory, price, and store staffing levels. Examples of retailers discussed in the article that are doing this well include Home Depot and McDonald's (Fisher et al., 2017). New data can enhance retailers' productivity and efficiency through better execution of existing operations and in developing business innovations. Here we provide some examples in both areas.

Better Execution Of Existing Operations

Big data and analytics can help retailers execute their current business models better. The retail business is such that even small improvements— such as increases in sales (or reductions in lost sales), reductions in bad inventory, or reductions in labor costs—could improve profits substantially. In addition to enhancing the retailer's market capitalization, superior execution can also provide the cash flow needed to invest in innovation projects and modernization.

Below, we provide some areas where retailers can employ data to better execute:

Optimizing Store Assortments

Periodically, retailers modify the assortment of goods available in their stores by eliminating some products and incorporating new ones in response to shifting consumer demands and newly introduced products. Determining if a potential new product will sell and the extent to which its sales will complement or negatively impact existing products is a

challenge. Accordingly, Fisher and Vaidyanathan (2014) have proposed a method to address this challenge. This method involves determining the characteristics of products in a particular category, estimating the demand for those characteristics based on sales data of existing products, and finally, estimating the demand for a potential new product by evaluating its constituent characteristics. The application of this approach has led to a 3-6% increase in revenue.

Optimizing Order Fulfillment Speed

To enhance customer satisfaction, many omnichannel retailers are making significant investments to reduce the time it takes to fulfill customer orders. While it is expected that customers will be pleased with quicker fulfillment, retailers want to know - does the increased revenue from faster fulfillment time justify the cost? Fisher et al. (2016) illustrate a case of a U.S. clothing retailer that added a second distribution center, reducing the order fulfillment time for a portion of their customers from 5 to 9 days down to 2 to 4 days. Through a comparison of revenue before and after this change, the retailer discovered that this resulted in a 4% increase in revenue, which more than made up for the cost of the second distribution center. The same approach was used to determine that a third distribution center would not be cost-effective, given the retailer's current volume. By utilizing data, retailers can make informed decisions on their investments in order fulfillment speed.

Maximizing Revenue Recovery During Store Closing and Liquidation

Retailers periodically need to close underperforming stores or liquidate entire retail chains, with the goal of recovering as much money as possible through the sale of the inventory. The use of inventory as collateral for loans by lenders such as banks has become more common as retailers have become more efficient at liquidation. To make the liquidation process even more efficient, Craig and Raman (2016) explored ways to optimize it by tracking store performance during liquidation and adjusting prices, inventory, and liquidation duration as necessary. They found that data-based forecasting and optimization could have improved the profit obtained during a liquidation by 2 to 5 percentage points, as determined by the net recovery on cost (the profit obtained during a liquidation expressed as a percentage of the cost value of the liquidated inventory). Professional liquidation firms understand the

importance of using data to make informed decisions during the liquidation process.

Improving Customer Experience Via Business Innovation

Retailers that provide the best shopping experiences optimize value for customers by reducing the friction and inconvenience associated with the shopping process. Examples of such retailers include Lululemon, Apple, Nike, Sephora, Target, Walmart, Kroger, and Warby Parker. For instance, Walmart's Express Delivery service enables shoppers to order items online and have them picked by a personal shopper, who then delivers the items in the preferred manner. Target also provides similar delivery options through Shipt, one of its acquisitions. These delivery options allow traditional brick-and-mortar retailers to compete with Amazon, and in many cases provide brick-and-mortar a competitive advantage with consumers who want products immediately.

However, according to Jindal et al. (2021), the fundamentals of retailing such as a large assortment of products, competitive prices, and convenient purchasing options remain crucial for shoppers. This highlights the enduring nature of marketing principles and retailers' core goals despite advancements in technology and data.

Tracking Consumer Movements

Historical analyses in the retail industry have generally used data that was aggregated on a monthly or weekly basis. However, with the advent of time-stamped data, it is now possible to continually monitor various aspects of customer behavior, product offerings, stock levels, in-store displays, and store environments. The assumption that these factors are static can only be considered an approximation. For instance, consider a retailer who wants to understand the impact of offering discounts or altering product placement on the flow of customers in the store, including how long customers spend at a specific location, the items they choose to put in their shopping baskets, and the order in which they do so. With the addition of a time dimension, a database containing consumer in-store movements linked to their purchases (as discussed in Hui, Bradlow, and Fader, 2009) can now provide answers to these questions.

The following discusses retailers' use of in-store video recording: "The source and quality of data has changed. Grocers are relying on their own proprietary research to decide how and where and at what price to place products—their own brands as well as those from CPG. Kroger and Walmart are using increasingly sophisticated software to decide where to place items and which products to shelve next to one another —factors that can move sales up or down several percentage points. The software—which can incorporate video surveillance of shoppers and how they react to displays and the time they take in selecting, or not, a product and was widely exhibited at the National Retail Federation in January—helps them create much more sophisticated plan-o-grams." (Lempert, 2020).

Geo-Targeting Based on Customer's Real-Time Location

The capability to track a customer's spatial location at any given moment has created new opportunities for retailers, where a customer's geospatial location may impact the efficacy of marketing efforts (Dhar and Varshney, 2011). When a customer's geospatial location is linked to a firm's CRM database, retailers can leverage tremendous value as a customer's purchase history can then be connected to the products they are physically near, enabling the most precise level of geo-targeting. Researchers have shown that shopping mall operators can monitor customers' movements by logging into Wi-Fi to document their shopping patterns and gain insight into their shopping context, thus providing more accurate product recommendations in bundle form (as per the findings of Ghose et al., 2019).

As discussed in the below 2013 New York Times article, various brick-and-mortar retailers are experimenting with these approaches:

> *Like dozens of other brick-and-mortar retailers, Nordstrom wanted to learn more about its customers — how many came through the doors, how many were repeat visitors — the kind of information that e- commerce sites like Amazon have in spades. So last fall the company started testing new technology that allowed it to track customers' movements by following the Wi-Fi signals from their smartphones.*
>
> *Nordstrom's experiment is part of a movement by retailers to gather data about in-store shoppers' behavior and moods,*

> *using video surveillance and signals from their cellphones and apps to learn information as varied as their sex, how many minutes they spend in the candy aisle and how long they look at merchandise before buying it.*
>
> *All sorts of retailers — including national chains, like Family Dollar, Cabela's and Mothercare, a British company, and specialty stores like Benetton and Warby Parker — are testing these technologies and using them to decide on matters like changing store layouts and offering customized coupons. (New York Times, 2013)*

These data allow brick-and-mortar retailers the same visibility on customers' shopping journey beyond the final transaction as e-commerce retailers who have long had access to customers' browsing data aided by browser cookies:

> *'Brick-and-mortar stores have been disadvantaged compared with online retailers, which get people's digital crumbs,' said Guido Jouret, the head of Cisco's emerging technologies group, which supplies tracking cameras to stores. Why, Mr. Jouret asked, should physical stores not 'be able to tell if someone who didn't buy was put off by prices, or was just coming in from the cold?' The companies that provide this technology offer a wide range of services.*
>
> *RetailNext, based in San Jose, Calif., adds data from shoppers' smartphones to deduce even more specific patterns. If a shopper's phone is set to look for Wi-Fi networks, a store that offers Wi-Fi can pinpoint where the shopper is in the store, within a 10-foot radius, even if the shopper does not connect to the network, said Tim Callan, RetailNext's chief marketing officer.*
>
> *The store can also recognize returning shoppers, because mobile devices send unique identification codes when they search for networks. That means stores can now tell how repeat customers behave and the average time between visits. (New York Times, 2013)*

Omnichannel Retail and Showrooming

The past 20 years have witnessed an increase in the number of channels through which consumers obtain product, experience, purchase, and post-purchase information.

Consequently, consumers are displaying a tendency to engage in "research shopping" - accessing information from one channel while purchasing from another (Verhoef, Neslin, and Vroomen, 2007). This has led to efforts to collect data from multiple channels). The collection, integration, and analysis of such omnichannel data help retailers in several ways: 1) understanding, tracking, and mapping the customer journey across channels; 2) evaluating profit impact; and 3) better-allocating marketing budgets to channels. Realizing that information gathering and actual purchase may happen at different points of time and that consumers often require assistance in making purchase decisions, retailers have started experimenting with relatively new ideas like showrooming—wherein the customer searches in the offline channels and buys online (Rapp et al., 2015), and webrooming — where the customer's behavior is the opposite.

Our recent research in the omnichannel domain (Zhang et al., 2022; Zhang and Neslin, 2022) provides a good case study of this showrooming effect and how it can benefit multiple stakeholders - enhanced customer experience and welfare via more accurate product fit and reduced search cost, improved retailer profit and operations, and better-informed real-estate decisions for retailers and city planners. Using large-scale customer-level transaction data provided by a large national outdoor product retail chain, we find that most new consumers prefer to buy products that they want to closely inspect in brick-and-mortar stores, then, having learned about the quality and fit, they would later buy the same products online for convenience. This important insight into customer behaviors allows retailers to better manage the customer experience by better-allocating resources across physical and online shopping channels. Consumers get greater convenience for the purchase and delivery of known products while also getting better service and information when buying products that they are unfamiliar with, thereby promoting a better match of consumers' preferences with services. The result is higher customer satisfaction from getting the right products, reduced wasted time, frustration, and wasted resources in buying the wrong products and returning them. This research has implications for the future of retail store designs and the efficient use of physical real estate.

Product Innovation And Diversity Increase Consumer Welfare

Retailer's analyses of customers' shopping data help launch private labels successfully and hence improve retailer's productivity, but it can also create a positive feedback loop – customers' shopping data from private labels inform the retailers about emerging consumer trends, further providing insights for product innovation that fit these emerging trends. This better match between products and consumers' taste lead to enhanced consumer welfare.

Neiman and Vavra (2019) investigated the phenomenon over the past 15 years termed "niche consumption", where, in contrast to previous eras where a few "superstar" products dominated grocery categories, households are increasingly buying different goods from one another.

Using the Nielsen Homescan dataset that covers a large fraction of spending on groceries and other household nondurables to study the shopping behavior of thousands of households from 2004-2016. They build a model to show increasing product variety drives these divergent trends. When more products are available, households select products better matched to their tastes, which leads to large consumer welfare gains. They attribute this trend in product innovation to "recent progress in supply chain integration, big-data marketing research, targeted advertising, and the growing importance of online sales" (Neiman and Vavra, 2019).

Consumers' welfare enhancements, in product innovation through retailers' private label initiatives, are also illustrated in the following statistics:

> "Consumers are increasingly purchasing P.L.s for value, not simply to save money. A 2016 Nielsen study found that 70% of U.S. households believe that P.L.s are good alternatives to national brands. An IRI study finds that P.L.s also create value by helping consumers save time in categories with vast selections of products. A more recent study finds that 90% of U.S. adults who switched to a P.L. did so due to "quality for price." Finally, U.S. consumers save over $40 billion annually by choosing P.L.s over national brands" (Dube, 2022)

Case Study – Retail Analytics For Private Label

Private labels have had a long tradition in retailing. In the 1920s, the vertically integrated A&P used data to tailor its own private-label offerings to customers' preferences. Because merchandise decisions can make or break a retailer, retailers continually experiment with the merchandise assortment to identify what is most relevant to customers. Although private labels and store brands have been available for decades, they account for an increasing share of retailers' assortments. They have evolved from being only low-priced, low-quality options to spanning a multi-tiered assortment, in which the premium private-label brands offer great quality and good prices—as well as greater profit margins for retailers. Retailers' private-label practices are now improved with retailers' access to data. Industry reports have proclaimed that retailing is "one of the hottest markets for big data analytics" (Ingram Micro, 2018), that "big data is especially promising and differentiating for retailers" (IBM Analytics, 2018), or that big data will be "a complete game changer in the retail sector" (Dekimp, 2020).

Gielens et al. (2021) discuss how advances in data and analytics enable retailers to gain a broader understanding of various customer journey stages (from browsing to buying to consumption).

These advances support greater targeting of niche segments and more precise personalization. Retailers and their design teams thus gain new insights into the needs of different segments, which has prompted them to offer a wider variety of brands (national and private label). Some of the biggest brick-and-mortar retailers (e.g., Walmart, Carrefour, Aldi, Lidl, Trader Joe's) and start-ups (e.g., Flamingo, Brandless.com) already offer their customers smart private-label brands. Today, the use of customer data and analytics to manage retailer private label programs has become ubiquitous amongst traditional brick-and-mortar chains:

> Retailers are grounding their private label strategies in deep shopper insight and powerful data mining, including loyalty card analysis...Through analysis of loyalty card and POS scan data, these retailers have refined their assortments, pricing, and consumer targeting efforts with significant success. Today, both Kroger and Tesco offer multi- tiered

> private label products, managed as part of a comprehensive portfolio." (Houlihan Lokey, 2017).

> "Analysts say that Tesco's big advantage over major international rivals, which also include Germany's Aldi and Lidl, is its unrivaled ability to manage vast reams of data and translate that knowledge into sales. While data crunching may sound dull, it has given Tesco two major advantages: an unmatched ability to operate multiple retail formats ranging in size from convenience stores to hypermarkets and the market knowledge to offer what many analysts say is the best and broadest range of house brands from any retailer. (CSP News, 2008)

Kroger's partnership with Dunnhumby has allowed it to use loyalty card data for various purposes, including the determination of successful new private label launches such as Private Selection meats and cheeses:

> "Through its joint venture with London-based Dunnhumby — which specializes in data management, customer analysis and insight-led planning — Kroger crunches customer loyalty card data to better understand its customers and identify opportunities to enhance their shopping experience in its stores. The retailer is able to target promotions based on individual shopper preferences — to a level that's far above what most other retailers are able to do today.

> Hertel says the joint venture has facilitated a great deal of shopper insights development, at both a key category level and a total store level.

> It also can be of tremendous help on the product development and assortment sides of the equation.

> 'The era of product customization has come to the grocery industry,' Meloche adds. 'This relationship will allow Kroger to keep its finger on the consumer's pulse and adapt their product offerings for various locales as needed. This can really only be done with private label.' (Canning, 2010)

Safeway has also established private label brand management teams that "...increasingly focus on consumers and consumer analytics in developing new lines and new products," relying in particular on "...data and insight from proprietary data systems

to get granular information on what's working." (Dairy Foods, 2012).

Albertson's and Rite Aid are relying on data analytics to integrate their private label offerings in one another's stores:

> *"Combining with Rite Aid will enable Albertsons to better cater to today's health and wellness- conscious consumer, as the supermarket retailer will more than double its pharmacy counters and get access to the drug chain's EnvisionRx pharmacy benefits management network and*

services. The merged company also will have extensive scale on both the East and West Coasts.

'That is an operator's dream come true in realizing synergies from both a cost perspective and revenue perspective,' Donald [Albertsons Cos. President and Chief Operating Officer] said. The companies also will use data analytics to integrate their loyalty programs and bring Albertsons' own-brand food and fresh items to Rite Aid's front end and, in turn, Rite Aid's health and beauty brands to Albertsons stores." (Redman, 2018)

Case Study – How Brands Benefit From Retailers' Access To Consumer Data

As large retailers become increasingly omnichannel and adopt similar business models, the line between traditional brick-and-mortar and e-commerce retailers has blurred. For example, the large retailers' understanding of customers' behavior enabled them to launch advertising services to help brands improve advertising efficiency.

The customer-level data that retailers such as Walmart, Kroger, and Target, collect on "which customers bought what and when" can be highly useful for brands aiming to better connect with consumers. These omnichannel retailers are offering advertising services supplemented by their data to allow these brands to advertise more efficiently and accurately.

For example, in a partnership with 84.51°, a data science company that helps its clients understand consumer preferences and develop strategies to improve customer experiences, Kroger developed Kroger Precision Marketing, an advertising service combining Kroger's customer data with 84.51° targeting and personalization science. The service helps advertisers better understand their ad campaign performance between in-store and online campaigns with data on transactions that take place both online and in-store.

In a public presentation, Walmart claims to be "leading the omnichannel evolution, touting that its "scale is truly massive" and that "our proprietary data encompasses 95% of U.S. house-holds", such that "Walmart.com is a daily destination for searching and browsing...16 million searches per day, 717 page views per second, and 12 million visits per day." (Walmart Media Group, 2020). Walmart also created its own advertising service, Walmart Advertising Partners. Like Kroger, Walmart offers brands the ability to better connect with Walmart customers. Walmart also recently launched Walmart+, a membership service designed to compete with Amazon Prime, that offered its members unlimited free delivery, early access to deals, and various discounts and convenience features.

Similarly, Target also created its own advertising service Roudel, which uses insights from the data it collects to help create advertising campaigns and connect brands with consumers.

These brick-and-mortar retailers' advertising businesses are very similar to that of Amazon which leveraged the data it collects to create its own advertising service, Amazon Advertising. These large retailers, with their rich data and customer insights, are not only using their data internally but creating services that utilize that data beyond their consumer business and share it with external brands to ensure greater success for all parties. These retailers compete with each other not just for consumers but also for brands to utilize their digital advertising services.

Analysis Of Public Companies Using Customer Data To Improve The Customer Experience

To understand the prevalence of how contemporary retailers use customer data to enhance the customer experience, we reviewed the annual 10-K reports that publicly-traded companies must file with the U.S. Securities and Exchange Commission. These reports provide investors with a detailed perspective on everything that is material to the company's financials, operations, and risks of success.

To identify retail companies, we employed a data source that provides a listing of the largest retailers by market cap and then excluded all companies that were not traded on a U.S. stock exchange.[2] That search identified 113 companies with 10-Ks in the U.S. Securities and Exchange Commission's EDGAR database.

It is important to note that those 113 companies provide a very comprehensive picture of retailers in the U.S. Although average employment across all 113 companies was 80,700 employees, the individual firms included in that list varied from 35 to 2.3 million employees. They also spanned a large number of retail industries that included automotive dealers (e.g., America's Car-Mart), bookstores (e.g., Barnes and Noble), clothing (e.g., Abercrombie and Fitch), electronics (e.g., Best Buy), furniture (e.g., Haverty's), grocery (e.g., Kroger), home improvement (e.g., Home Depot), pharmacies (e.g., CVS), and recreational vehicles (e.g., Camping World),. In addition, they also included a mix of traditional brick-and-mortar stores (e.g., Macy's) and strictly e-commerce companies (e.g., Wayfair).

Our analysis showed that the annual reports for 92 (81 percent) of the 113 companies made specific reference to using customer data to enhance the customer experience. For U.S. retail companies with a market cap of $1 billion or more, 84 percent of them mentioned using customer data in their most recent 10-K filing.

Some examples of those references include:

"We also use our just for U loyalty program, including leveraging customer and transaction information with data driven analytics to provide both personalized deals and digital coupons, as well as gas and grocery rewards, to target promotional activity and improve our customers' experience." Albertsons.[3]

"Clients increasingly engage with us through digital methods including our website and social media. To capitalize on these trends and continue increasing our client base, we are investing in data analytics to improve the client journey from the moment clients begin browsing online or enter our Showrooms. This will allow us to target clients with personalized digital offerings to increase online conversion and client lifetime value." Arhaus.[4]

"We are migrating to an industry-leading customer database and customer marketing platform to personalize the customer journey and manage all customer communication in an effort to increase traffic across all channels and increase conversion." Chico's.[5]

"As our customers expect a more personalized experience, our ability to collect, use and protect relevant customer data is important to our ability to effectively meet their expectations." Home Depot.[6]

"We also need to collect, use and share relevant customer data to effectively meet customer expectations of a more personalized experience." Lowes.[7]

"In addition, our ability to create a personalized guest experience through the collection and use of accurate and relevant guest data is important to our ability to differentiate from other retailers." Target.[8]

"We are also using the data and information provided by loyalty members to personalize the experience to the user and improve the communication and offering. We plan to

continue to further enhance this program during fiscal 2022." Tillys.[9]

"We are uniquely positioned with a deep understanding of our customers and their preferences which enables us to personalize experiences and target communications and promotions through our Customer Relationship Management (CRM) platform." ULTA.[10]

Many of the companies that do not refer to the use of customer data in their 10-Ks may still be doing so even if they do not explicitly mention it in their report. For example, Buckle (one of the companies that did not mention the use of customer data in its 10-K) has a customer loyalty program that rewards customers when they provide data to the company on their purchases.

This review of companies' annual reports shows that the use of customer data to enhance the customer experience is extremely widespread. It includes retailers of all sizes across a very wide spectrum of retailers, including both traditional brick-and-mortar retailers as well as those that specialize in e-commerce.

Additionally, beyond the retail industry, other customer-facing industries collect and use data on customer behavior. For example, in the tourism, travel, and hospitality industries there has long been widespread use of customer-loyalty programs to target promotions, anticipate future purchases, and reward repeat customers.

---

[2] Largest retail companies by market cap (companiesmarketcap.com) Accessed November 3, 2022.

[3] A grocery and drug chain retailer that operated 2,276 stores in 2022 under 24 different trade names including Albertsons, Safeway, Vons, Randalls, and Balducci's Food Lovers Market. Albertsons Companies, Inc. Form 10-K for the fiscal year ended February 26, 2022, accessed December 12, 2022.

[4] A provider of home furnishings that operated a network of 71 traditional showrooms, 5 Design Studios, and 3 Outlets, as well as 58 showrooms with in-home interior designers in 2021. Data source: stockanalysis.com, accessed December 12, 2022.

[5] A specialty retailer of woman's clothing that operated 1,266 stores in 2022. Data source: stockanalysis.com, accessed December 12, 2022.

[6] A chain of home improvement stores that operated 2,317 stores in 2021. Data source: stockanalysis.com, accessed December 12, 2022.

[7] A chain of home improvement stores that operated 1,971 stores in 2022. Data source: stockanalysis.com, accessed December 12, 2022.

[8] A general merchandiser that operated approximately 2,000 stores in 2022. Data source: stockanalysis.com, accessed December 12, 2022.

[9] A specialty retailer of casual clothing, footwear, and accessories that operated approximately 241 stores in 2022. Data source: stockanalysis.com, accessed December 12, 2022. 10 A beauty products retailer that operated approximately 1,308 stores in 2022. Data source: stockanalysis.com, accessed December 12, 2022.

[10] A beauty products retailer that operated approximately 1,308 stores in 2022. Data source: stockanalysis.com, accessed December 12, 2022.

Conclusion

This report demonstrates that retailers have always used as much consumer data as they have been able to obtain in order to improve their business practices, their customer service, and the variety of products that they offer to shoppers. As shown by the guidance of academics and the examples of data use throughout the past century, while technologies and data types have become more sophisticated, the fundamental reasons why retailers use data have not changed: to generate insights that increase sales and reduce costs. For consumers, when retailers use data about their shopping behavior, it provides them with more convenience, an improved shopping experience, and better products that better fit their needs and preferences.

Jonathan Zhang, PH.D.

Fletcher Mangum, PH.D

David Zorn, PH.D

NetChoice

College of Business, Rockwell Hall, Room 9

Ft. Collins, Colorado 80523

(970) 491 - 2238

BIZ.COLOSTATE.EDU

MANGUM economics

4198 Cox Road, Suite 104

Glen Allen, Virginia 23060

(804) 322 - 7777

MANGUMECONOMICS.COM

References

Abernathy, F. H., Dunlop, J. T., Hammond, J. H., & Weil, D. (1999). A stitch in time: Lean retailing and the transformation of manufacturing–lessons from the apparel and textile industries. Oxford University Press.

Applebaum, W. (1951). Studying customer behavior in retail stores. Journal of Marketing, 16(2), 172- 178.

Banks, S. (1950). The relationships between preference and purchase of brands. Journal of Marketing, 15(2), 145-157.

Bartels, R. D. (1944). Marketing principles. Journal of Marketing, 9(2), 151-157.

Basker, E., & Noel, M. (2009). The evolving food chain: competitive effects of Wal- Mart's entry into the supermarket industry. Journal of Economics & Management Strategy, 18(4), 977-1009.

Basker, E., Klimek, S., & Hoang Van, P. (2012). Supersize it: The growth of retail chains and the rise of the "Big-Box" store. Journal of Economics & Management Strategy, 21(3), 541-582.

Berthiaume, D. (2020). Marks & Spencer enhances private label product development. Chain Store Age, https://chainstoreage.com/marks-spencer-enhances-private-label- product-development Accessed on Dec 1, 2022.

Blattberg, R. C., Kim, B. D., & Neslin, S. A. (2008). Market basket analysis. In Database Marketing (pp. 339-351). Springer, New York, NY.

Bonne, J. (2004). Big Store or Small? Why Consumers Care, MSNBC, February 4, http://www.msnbc.msn.com/id/4001039/4001039. Accessed on Nov 20, 2022.

Canning, K. (2010). Store Brand Superhero. Store Brands, July 1, 2010 https://storebrands.com/store- brand-superhero. Accessed on Dec 2, 2022.

Carden, A. (2013). Retail innovations in American economic history: The rise of mass- market merchandisers. In Routledge handbook of major events in economic history (pp. 419-431). Routledge.

CIO (2007), "45 Years of Walmart History: A Technology Time Line", https://www.cio.com/article/274537/infrastructure-45-years-of-walmart-history-a- technology-time- line.html. Accessed Nov 1, 2022.

CSP News (2008). Tesco: "Walmart's Worst Nightmare". December 31, 2008, https://www.cspdailynews.com/company-news/tesco-walmarts-worst-nightmare. Accessed on Dec 4, 2022.

Dairy Foods (2012). How Safeway is building its own brands. Dairy Foods, October 12, 2012, https://www.dairyfoods.com/articles/88918-how-safeway-is-building- its-own-brands. Accessed on Dec 1, 2022.

Das, N., Falaris, E. M., & Mulligan, J. G. (2009). Vintage effects and the diffusion of time-saving technological innovations. The B.E. Journal of Economic Analysis & Policy, 9(1).

Dekimpe, M. G. (2020). Retailing and retailing research in the age of big data analytics. International Journal of Research in Marketing, 37(1), 3-14.

Doms, M. E., Jarmin, R. S., & Klimek, S. D. (2004). Information technology investment and firm performance in U.S. retail trade. Economics of Innovation and new technology, 13(7), 595-613.

Dubois, D., Bonezzi, A., & De Angelis, M. (2016). Sharing with friends versus strangers: How interpersonal closeness influences word-of-mouth valence. Journal of Marketing Research, 53(5), 712-727.

Dubé, J. P. (2022). Amazon Private Brands: Self-Preferencing vs Traditional Retailing. Working Paper. Available at SSRN 4205988.

Fisher, M., & Vaidyanathan, R. (2014). An algorithm and demand estimation procedure for retail assortment optimization. Management Science, 60(10): 2401−2415.

Fisher, M., Gallino, S., & Xu, J. (2016). The value of rapid delivery in online retailing. Wharton School Working paper, Available at SSRN 2573069.

Fisher, M., V. Gaur, H. Kleinberger. 2017. Curing the Addiction to Growth. Harvard Business Review, January−February 2017.

Foley, C. F., Raman, A., & Craig, N. C. (2012). Inventory-based lending industry note. Harvard Business School Case, (612-057).

Foster, L., Haltiwanger, J., & Krizan, C. J. (2006). Market selection, reallocation, and restructuring in the U.S. retail trade sector in the 1990s. The Review of Economics and Statistics, 88(4), 748-758.

Gielens, K., Ma, Y., Namin, A., Sethuraman, R., Smith, R. J., Bachtel, R. C., & Jervis, S. (2021). The future of private labels: towards a smart private label strategy. Journal of Retailing, 97(1), 99-115.

Ghose, A., & Han, S. P. (2014). Estimating demand for mobile applications in the new economy. Management Science, 60(6), 1470-1488.

Ghose, A., Li, B., & Liu, S. (2019). Mobile targeting using customer trajectory patterns. Management Science, 65(11), 5027-5049.

Granbois, D. H. (1968). Improving the study of customer in-store behavior. Journal of Marketing, 32(4_part_1), 28-33.

Harmon, C. K., & Adams, R. (1984). Reading between the Lines: An Introduction to Bar Code Technology. North American Technology. Inc., Peterborough, NH.

Holmes, T. J. (2001). Bar codes lead to frequent deliveries and superstores. RAND Journal of Economics, 708-725.

Houlihan L. (2017). The Re-Emergence of Private Label Strategy. White Paper, https://www.hl.com/uploadedFiles/11_Blogs/Strategic-Consulting/The-Re- Emergence- of-Private- Label-WP.pdf. Accessed on Nov 20, 2022.

Jarmin, R. S., Klimek, S. D., & Miranda, J. (2009). The role of retail chains: National, regional and industry results. In Producer dynamics: New evidence from micro data (pp. 237-262). University of Chicago Press.

Jia, P. (2008). What happens when Walmart comes to town: An empirical analysis of the discount retailing industry. Econometrica, 76(6), 1263-1316.

Jindal, R. P., Gauri, D. K., Li, W., & Ma, Y. (2021). Omnichannel battle between Amazon and Walmart: Is the focus on delivery the best strategy?. Journal of business research, 122, 270-280.

Kelley, E. J. (1958). The importance of convenience in consumer purchasing. Journal of Marketing, 23(1), 32-38.

Lempert, P. (2020). How In-Store Brands Are Taking Over In-Store. Winsight Grocery Business, April 10, 2020, https://www.winsightgrocerybusiness.com/retailers/how- store-brands-are-taking-over-store. Accessed Nov 30, 2022.

Levinson, Marc (2011). The Great A&P and the Struggle for Small Business in America. Hill and Wang.

Matsa, D. A. (2011). Competition and product quality in the supermarket industry. The Quarterly Journal of Economics, 126(3), 1539-1591.

Mitchell, S. (2016). Woolworths Sitting on Big Data Goldmine. Financial Review,, https://www.afr.com/companies/retail/woolworths-sitting-on-big-data-goldmine- 20161013-gs1cgw Accessed Nov 30, 2022.

Muris, T. J., & Nuechterlein, J. E. (2019). Antitrust in the internet era: The legacy of United States v. A&P. Review of Industrial Organization, 54(4), 651-681.

Levin, S. G., Levin, S. L., & Meisel, J. B. (1987). A dynamic analysis of the adoption of a new technology: the case of optical scanners. The Review of Economics and Statistics, 12-17.

Luo, X., Andrews, M., Fang, Z., & Phang, C. W. (2014). Mobile targeting. Management Science, 60(7), 1738-1756.

Neiman, B., & Vavra, J. S. (2019). The rise of niche consumption (No. w26134). National Bureau of Economic Research.

New York Times (2004). What Walmart Knows About Customers' Habits. November 14, 2004. https://www.nytimes.com/2004/11/14/business/yourmoney/what-walmart-knows-about-customers-habits.html. Accessed Nov 17, 2022.

New York Times (2013). Attention, Shoppers, Stores is Tracking Your Cell. July 14, 2013. https://www.nytimes.com/2013/07/15/business/attention-shopper-stores-are- tracking-your-cell.

Pessemier, E. A. (1959). A new way to determine buying decisions. Journal of Marketing, 24(2), 41-46.

Pascoe, M. (2017). Woolies Is Playing Smart. Coles Is Not, andIt Shows. Sydney Morning Herald, https://www.smh.com.au/business/companies/why-coles-is-in-the-dog- house-it-thinks-i-have-a-cat- 20171102-gzd8w3.html. Accessed Nov 20, 2022.

Pew Research Center (2005), Walmart: A Good Place to Shop but Some Critics Too. Pew Research Center for the People and the Press, December 15, 2005, http://people- press.org/reports/display.php3?ReportID=265. Accessed Nov 2, 2022.

Progressive Grocer (2006). Kroger Winning with Shopper Data, Says Vice Chairman. May 4, 2006, https://progressivegrocer.com/kroger-winning-shopper-data-says-vice-chairman. Accessed on Dec 1, 2022.

Rapp, A., Baker, T. L., Bachrach, D. G., Ogilvie, J., & Beitelspacher, L. S. (2015). Perceived customer showrooming behavior and the effect on retail salesperson self- efficacy and performance. Journal of Retailing, 91(2), 358-369.

Redman, R. (2018). Albertsons' Jim Donald sees "improving momentum". Supermarket News, July 17, 2018 https://www.supermarketnews.com/retail-financial/albertsons-jim-donald-sees-improving- momentum Accessed on Dec 3, 2022.

Rossi, P. E., McCulloch, R. E., & Allenby, G. M. (1996). The value of purchase history data in target marketing. Marketing Science, 15(4), 321-340.

Time (1948). Fair Trade? Nov, 15, 1948. https://content.time.com/time/subscriber/article/0,33009,853450,00.html. Accessed on Feb 1, 2023.

Trivedi, M., Gauri, D. K., & Ma, Y. (2017). Measuring the efficiency of category-level sales response to promotions. Management Science, 63(10), 3473-3488.

Verhoef, P. C., Neslin, S. A., & Vroomen, B. (2007). Multichannel customer management: Understanding the research-shopper phenomenon. International journal of research in marketing, 24(2), 129-148.

Walmart Stores, Inc. (1994) Walmart Annual Report. Bentonville, AR.

Walmart Media Group (2020). Spark growth in your business. https://s3-prod.adage.com/s3fs-public/2020-10/WMG%20Intro%20Deck.pdf. Accessed Nov 1, 2022.

Wang, J., Aribarg, A., & Atchadé, Y. F. (2013). Modeling choice interdependence in a social network. Marketing Science, 32(6), 977-997.

Wang, J. (2006). Economies of I.T. systems at Walmart-An historical perspective. Journal of Management Information and Decision Sciences, 9(1), 45.

Wedel, M., & Pieters, R. (2000). Eye fixations on advertisements and memory for brands: A model and findings. Marketing science, 19(4), 297-312.

Zhang, J. Z. (2019). Dynamic customer interdependence. Journal of the Academy of Marketing Science, 47(4), 723-746.

Zhang, J. Z., Chang, C. W., & Neslin, S. A. (2022). How physical stores enhance customer value: The importance of product inspection depth. Journal of Marketing, 86(2), 166-185.

Zhang, J. Z., & Neslin, S. A. (2022). Leveraging the Physical Store to Boost Customer Value - Providing the Right Engagement at the Right Time. California Management Review, June.