Case Analysis #3 – Data Warehousing at Smithfield

By Vicky Zhou, Shouyang Wang, and Billy Yu

1.1 Background

Broadly speaking, a Data Warehouse (DW) is a system that handles a large amount of data from various resources. The warehouse functions as a repository that enables business users to collect, centralize, store, perform queries, analyze, report, and visualize data. DW plays a core and irreplaceable role in business intelligence. After all, data-driven decisions are useless in practice without a comprehensive system that can extract insightful information from copious amounts of business data. One crucial piece of advice for business owners is that a DW is not precisely the same as a database designed for transactional purposes. A transactional database is not optimized to perform data analytics tasks. A Data Warehouse, using OnLine Analytical Processing (OLAP), *extracts, transforms, and loads* the said databases from various sources to create and analyze visualizations such as dashboards. DW is used heavily for business intelligence purposes. Data Warehousing is a powerful tool that Smithfield can leverage to reach its strategic goals. BI analytics enables precision agriculture via data collection from the Internet of Things (IoT), AI remote sensors, robots, satellites, customers, etc.

1.2 Strategic Goals Supported by Data Warehousing

Strategic Goals Summary

After understanding the fundamentals of DW, we now illustrate the strategic goals supported by data warehousing. A strategic goal is defined as a "long-term, 'big picture' objective for a business, rather than a short-term tactic that addresses a current problem or challenge." (Milano, 2017) A strategic goal defines the desired result and destination. When establishing those goals, it is integral for business owners to make sure that the plans are measurable, challenging but practical, and realistic at the same time.

Strategic Goal 1 – Increase Revenue by Optimizing Pricing leveraging DW Analytics

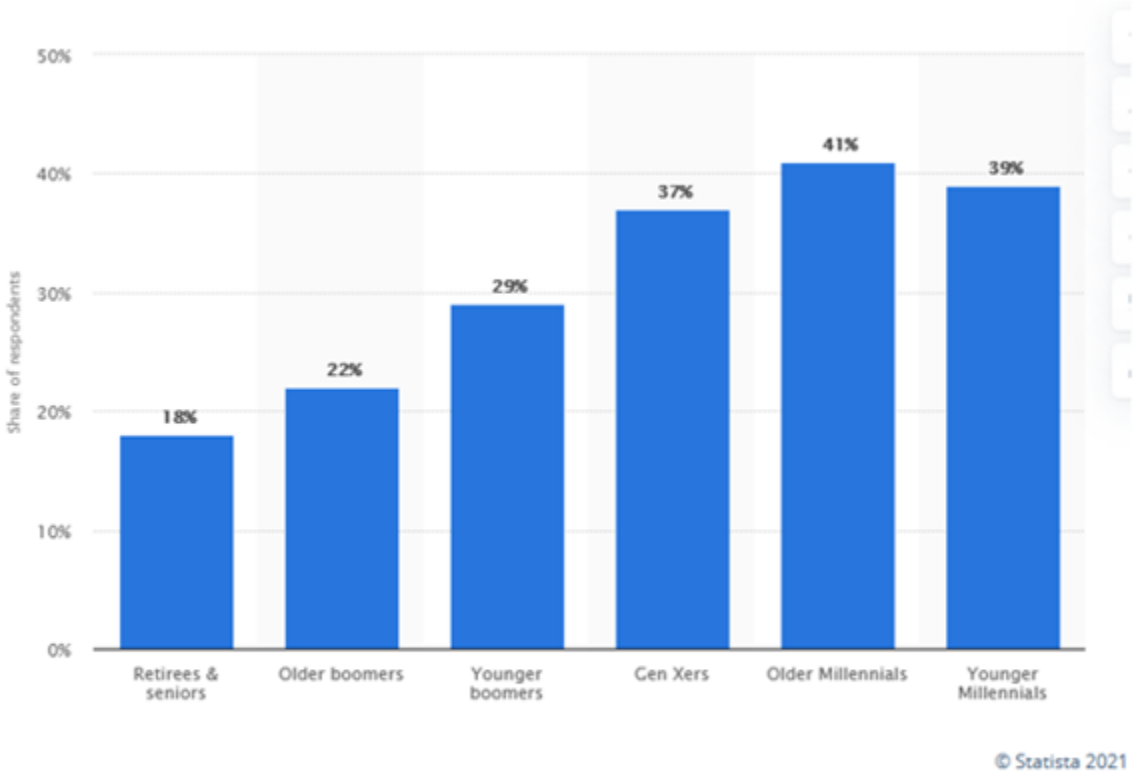
DW Analytics can help Smithfield determine **how much more the average person pays for their organic produce**. Knowing this information would optimize their revenue in the long run.

Smithfield's organic produce is a competitive advantage that can be leveraged over large factory farms. However, it is also true that organic products are generally more expensive compared to large-quantity produced factory products. The Food and Agriculture Organization (FAO) of the UN notes that *production costs for organic foods are typically higher because of greater labor inputs per unit of output*. (FAO, 2020)

From a business perspective, it does not make sense for Smithfield to set the pricing of its products lower than that of the ones factory farms offer because hormone-injected meat that factories produce has a lower quality compared to organic meat. Because of that, it's a lot cheaper for factories to produce the same quantity. If your organic-food pricing is lower than factory-produced products, it should be evident that you're not profiting as a small farm owner using this pricing strategy. However, if organic farmers set the price too high, they run into the problem that the

business does not have many loyal and returning customers. A significant component of Smithfield's income comes from selling organic farm products. A lack of customers means less revenue.

The concept illustrated here shows that organic farmers have to find a delicate balance between supply and demand when deciding an effective pricing strategy. When it comes to DW analytics, several things are critical for Smithfield to take into consideration. A 2018 study done by Statista shows that less than 50% of consumers are willing to pay a significant extra amount of money for organic products.



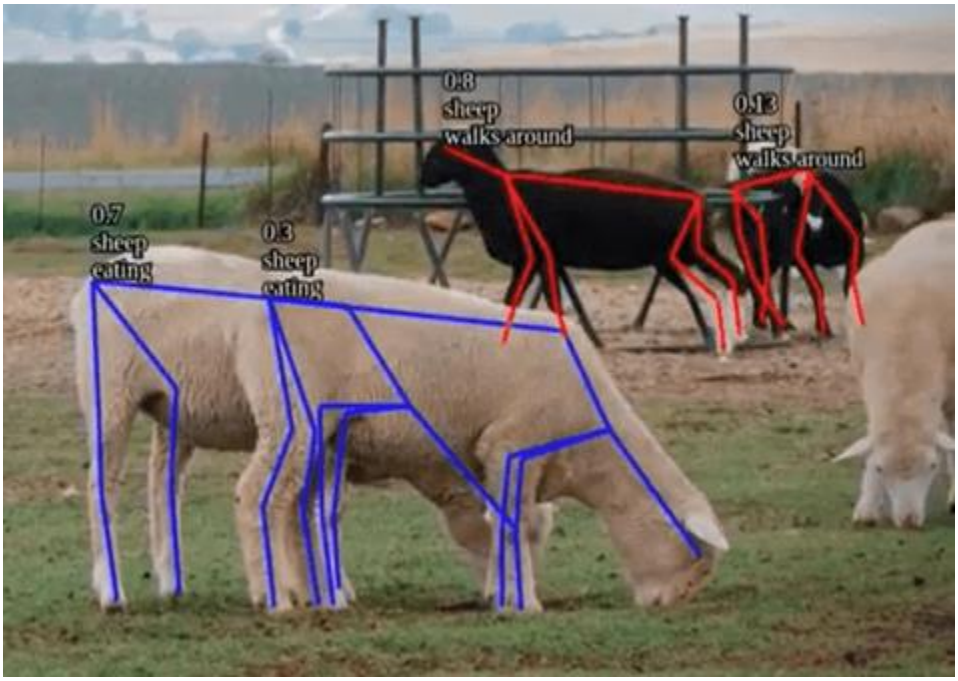
*This statistic depicts the willingness to pay a premium for natural or organic products among consumers in the United States in 2018, broken down by generation.

With this information in mind, Smithfield should leverage customer-based pricing. Customer-based pricing means that the organization collects data and feedback directly from its customers and understands the price consumers are willing to pay for a particular product. Knowing more and quality information about your customers will enable organic farmers to make data-driven decisions to maximize their income.

Strategic Goal 2 – Reduce Workload via Autonomous Livestock Management (DW + AI)

DW Analytics can help Smithfield to monitor the activity of all livestock available. Knowing this information would significantly reduce the time and effort farmers spend on managing livestock. For instance, without a Data Warehouse, a farmer must check on every single livestock in person every day. The traditional way is time-consuming and inefficient. With the help of technologies, such as IoT and AI, the autonomous system leverages remote control of feeding, milking, and

cleaning animals. We can track exactly how much food a cow eats per day and better understand food waste and the livestock's overall health status. This critical information will be presented on a dashboard so farmers can effectively understand all livestock's current situation. On top of that, direct feedback from live cameras can alert farmers about any abnormal activities. It's not possible for one person to monitor all of the livestock 24/7 without a DW and AI system in place. By analyzing the data collected via DW, smart cameras, and AI, farmers can gain some valuable insights. The result can be visualized to support farmers making data-driven decisions, such as when deciding the health of an animal. The picture below demonstrates the concept.



(Source: KEYMAKR Artificial Intelligence, https://keymakr.com/)

Strategic Goal 3 – Making smart and multi-dimensional decisions by analyzing customer behaviors

The most crucial advantage a Data Warehouse has over a regular database is that the user is integrating multiple sources. A DW enables business users to analyze a problem from many different perspectives, which provides a more comprehensive story and improves business intelligence. For instance, a drop in customers could indicate many underlying issues. To understand the root problem, the business owner needs to analyze the historical data from several different databases simultaneously. A drop in meat quality, the pandemic, livestock health, or even political changes could all impact customer behavior. Without an effective DW system in place, a farmer can only assume what happened and make blind decisions without numerical data to back it up. Farmers will invest their time and energy in the wrong areas without an efficient intelligence system.

2. Operational problems data warehousing would solve

Highlight: *Data warehousing improves the speed and efficiency of accessing different data sets and makes it easier for corporate decision-makers to derive insights that will guide the business and marketing strategies that set them apart from their competitors*.("What is Data Warehousing and Why is it Important?," 2017)

- Data warehousing helps Smithfield to store information and make wise business decisions by avoiding operational mistakes, predicting the uncertainties, improving human resource efficiency, and maintaining a better relationship with business partners.

An effective data warehouse shall help the users extract meaningful information and support the decision making process. Like most farms. Smithfield is not using an efficient data warehouse, or they barely have any data storage methods. With the increasing demands for meat and cereal production, farms shall implement a new strategy, such as precision farming, to meet the massive demands. To achieve this, a data warehouse plays an important role. Data warehousing helps Smithfield store confidential farming information, such as information about arable land, seeding, fertilization, insecticide, harvesting, storage, breeding, etc, which can impact the operation significantly. For example, a new employee can feed the wrong food for the cows and cause significant health issues, and there is no way for Smithfield to find out that things go wrong until they notice one of the cows is sick. However, by using the data warehouse, the farm manager can easily access the feeding data/record onsite or remotely, and crosscheck things to ensure safety. More importantly, the new employee can also access the data warehouse under employee view, he/she can easily find out the previous eating habits and feeding schedule for the cows to avoid these human mistakes.

Moreover, by implementing data warehousing, Smithfield can also obtain information such as climate, soil, and others in an easier way, and use the information to provide the farm with the most scientific suggestions. For example, Smithfield can use the climate data stored for the past days, months, and years to predict the climate changes for the future. During the dry season, Smithfield shall use human intervention, such as irrigating more often and keeping all the livestock hydrated to reduce the losses caused by bad weather. Last but not least, Smithfield can also use data warehousing to store confidential information such as their employee data and client information. This way, the HR does not have to store all the employee information, such as a home address, SSN, and bank account in a dusty locker, instead somewhere secure enough and with easy access. So it is payroll time, it will greatly improve the work efficiency by saving up all the time to pull files and update information. On the other hand, a data warehouse helps Smithfield store and access client information, such as their order date, quantity, category, delivery date, special requirement, client's private information, etc. Thus, the sales team can utilize this information to build and maintain a long-standing relationship by knowing deeply about who the clients are, what they want, and how we can better serve them.

3. Overview of the importance of real-time data access in a data warehouse

Real-time data is information that is delivered immediately to the collection. Real-time data access means that users can access information directly from the centralized places that help organizations process information and make decisions based on requests.

Firstly, real-time data access helps Smithfield supervise and monitor farms more efficiently. By implementing a real-time data monitoring system, farms can monitor the activities, such as agricultural planting, aquaculture, livestock and poultry breeding, vegetable greenhouses, and surrounding areas with no time delay. For example, Smithfield uses the system to monitor all the livestock in real-time to avoid anything abnormal and dangerous to ensure farm safety. More importantly, the real-time data access assists in achieving the automatic tracking function, that is the system can track any moving object, zoom in on the target, and track for the target to record the movement and send out alarms to the owner. These actions don't require any manpower and save up the associated costs.

Secondly, real-time data access also helps Smithfield increase revenue. That is, the farm can improve customer service, operational efficiency, and employee engagement. With the help of a real-time data dashboard, the farm sales representative can pull clients' information and provide real-time pricing based on the current market trends immediately without digging through piles of files when they call. Thus, the customers feel they are well treated and secure a better deal compared with other farms. Moreover, the owner knows immediately what is happening on the farm by looking at the dashboard, and what kind of improvements that the farm needs to be done. (Barnett, 2017) Also, it is always good for employees to see how they are doing at work and what is the monthly goal/expectation of them, so they can always reflect and make immediate adjustments to reach their goals. Real-time data access provides this opportunity to all employees and helps them to understand their up to date performance to improve their engagement and efficiency.

In conclusion, as *most executives (77%) responding to a recent survey conducted by IDC in partnership with InterSystems agree that lack of timely access to data is inhibiting their businesses.* (McKendrick, 2017) Real-time data access can help businesses capture opportunities and respond to threats promptly that greatly improves productivity, efficiency, and revenue eventually.

4.1 Applications of a data warehouse (functional uses)

Precision Livestock Farming

As we mentioned in previous sections, Smithfield can develop its livestock management by implementing a data warehouse with the integration of advanced technologies such as the Internet of Things. Precision Livestock Farming is a commonly applied method among meat farms, which applies technological advances to the monitoring of and data collection from individual animals within large herds with the hope of optimizing the welfare and contribution of each animal (Benjamin, 2019). With the assistance of smart devices such as Electronic Ear Tags, Smithfield can generate an electronic ID for each livestock, upload associated data to the warehouse, and get insightful analytics that can precisely assist the management of livestock.

Lean and Agile Farm Production and Operation

The operation of a farm can always be uncertain since farm businesses assess product cycles in terms of seasons or weather events (Australian Farm Institute, 2017). By implementing a data

warehouse, Smithfield can effectively store valuable historical information that is relevant to the daily operation of the farm and use data analytics to predict future events that may affect the normal production. With this kind of system, Smithfield will be ahead of influencing factors such as weather, traffic, etc., and operate in turbulent environments.

Customer Relationship Management (CRM)

Data Warehouse plays a fundamental role in CRM since it can serve as a repository for all customer-related information: operational data, interaction data, customer-profile data, and demographic and behavioral data (Khan et al., 2012). Therefore, for Smithfield's Bed & Breakfast services, the data warehouse can significantly contribute to decisions such as customers' entertainment needs, accommodation preference, special need for food allergies and children activities, etc.

4.2 Which data warehouse applications will be the most vital and why?

In our opinion, the implementation of ***Precision Livestock Farming*** is the most vital part as Smithfield develops its data warehouse system. We will discuss the significance of Precision Livestock Farming with the assistance of a data warehouse in three aspects: Product Quality, Customer Satisfaction, and Competitive Pricing.

Product Quality

Precision Livestock Farming with an efficient data warehouse can significantly increase Smithfield Meat Farm's product quality in many ways. If we look at dairy farming, there can be tons of data generated from milking and feeding robots, temperature sensors, animal tracking systems, dairy herd management systems, and databases maintained by dairy herd improvement associations (Schuetz, Schausberger & Michael Schrefl, 2018). While data analysis can be time-consuming, data warehousing can hugely speed up the process so that farmers can get instant feedback on tasks to complete and make sure products are always of a high quality.

Customer Satisfaction

Customer satisfaction is closely related to Smithfield's product quality since customers do not want their produce from the organic farm to have the same quality as those from the mass meat production industry. By implementing Precision Livestock Farming, Smithfield can keep answering customer's needs with great qualities so that it can constantly increase customer satisfaction rate. In addition, as we mentioned in the previous section, the data warehouse serves as a repository for customer information. High customer satisfaction can also bring Smithfield more customers so that it can seek chances to expand its business by introducing its various services such as the Bed & Breakfast and other farming recreational activities to new customers.

Competitive Pricing

As high-quality products and good customer satisfaction rate being the fundamentals of the brand, Smithfield can further take advantage of those and engage in valuable and flexible pricing

strategies to be competitive in the market. In addition, with real-time monitoring of animals, Smithfield can use Precision Livestock Farming to accurately calculate the revenue of each animal and the related product so that the farm can plan for the pricing and budget more efficiently and systematically.

5. Data Warehouse (DW) Initiative Challenges

Understanding the basics of DW and BI software (Difficulty level: ★)

Most small business owners have no technical background in Business Intelligence Software. It is particularly true for traditional farmers. Farmers often rely on primitive methods, such as writing on paper, to record data. Analyze a large amount of data is not possible using primitive tools. Traditional farmers make experience-based decisions rather than relying on accurate quantitative data and modern technology.

As mentioned previously, a data warehouse should be used for data analysis purposes only instead of everyday transactional processing. The most initial challenge for business owners is understanding the difference between different databases (e.g., Database, DW, Data Lake, etc.) and their roles served in business functions. Abusing a transactional database can be problematic in the long run. Without understanding the basics, it is very likely for most non-technical people to abuse databases.

Picking the right DW software (Difficulty level: ★★)

Farmers need to understand that it is not merely a matter of buying the most expensive DW solution out there. There are many DW software available to business owners, such as Apache Hive, Amazon Redshift, Oracle Autonomous Data Warehouse, and Treasure Data. Depending on your business needs and the amount of data you have, it is imperative for farmers to do sufficient research before picking the right DW software that fits their farm's business functions. For instance, Apache Hive is relatively easy to set up and cheap as well. However, it doesn't have constant updates or support, and the processing speed is slow compared to other DW software.

Finance and Cost (Difficulty level: ★★★★)

Finance is a challenge that farmers have to face. It's not that most farmers are too poor to afford a DW system. Some farmers might not feel the need to use a DW when they can't directly profit from the DW. There are four main contributions to this:

1. There are too many DW hidden costs, and maintenance is costly. (eg. Why should I pay Amazon $100 for this nonsense?)

2. A farmer might only gain low-quality and useless business information from the DW due to poor data collection and analytics. (eg. I know this nonsense without using a DW)

3. A farmer might hire an expert to manage the system, which costs money. (eg. Now I am paying an extra worker $50 a day for this nonsense)

4. A farmer might not see a short-term profit. (eg. I am not making any money using this nonsense)

When a farmer sees no profit from leveraging a DW system, especially at the initial stage, it is not likely they will continue using the system.

Technological complexity (Difficulty level: ★★★★★★★★)

The DW software is as good as nothing if business users are not leveraging it correctly. The learning curve for most DW software is relatively flat. However, most farmers are not experts in computers, and many need training in computer skills to perform basic tasks. Performing a standard query would be a challenge farmers have to conquer to gain business insights.

On top of that, knowing the basics is not enough to retrieve valuable business information. A data warehouse is merely a place to analyze the data you collected and organized. The quality of the data, the accuracy, and data collection methods are equally important. For instance, the data collected from remote sensors and AI need to be transferred via different Operating Systems and interfaces. A DW is just one part of the equation. The technology gets exponentially more complex as big data, cloud technology, AI, and the Internet of Things get involved.

6. Primary benefits of a data warehouse implementation

Lower operational cost and higher revenue
With the implementation of a data warehouse, Smithfield can make data-driven decisions such as precisely allocating resources to each livestock with the assistance of Precision Livestock Farming. As a result, farmers can effectively manage the budget and minimize any waste due to misallocation or operational error. Besides, data warehouse analytics can increase the overall production of the farm so that Smithfield can expect higher revenue in the long run.

More efficient use and management of labor and other resources
Using data warehouse analytics, Smithfield can accurately identify influencing factors both inside and outside the farm. For example, by analyzing and predicting times of high demands, Smithfield can make decisions about when to hire more employees or purchase necessary tools and machinery. This is especially valuable during current pandemic conditions since the operation of the farm and the Bed & Breakfast services can be turbulent. As a result, Smithfield can efficiently and systematically use and manage all its capital and operate sustainably.

Increased customer satisfaction
As mentioned in previous sections, the implementation of a data warehouse can help Smithfield understand customer's behavior and preference and enhance its customer relationship management. By keeping track of customer's interactions and analyzing rises and drops of customers using the data warehouse, Smithfield can constantly refine its products and services and thus increase the customer satisfaction rate.

References

FAO. “FAQ.” Organic Agriculture: Why is organic food more expensive than conventional food? Accessed January 30, 2021. http://www.fao.org/organicag/oa-faq/oa-faq5/en/.

Milano, Steve. “What Is a Strategic Goal?” Small Business - Chron.com. Chron.com, November 21, 2017. https://smallbusiness.chron.com/strategic-goal-74875.html.

Wunsch, Nils-Gerrit. “Willingness to Pay More for Natural/Organic Products U.S. 2018.” Statista, November 26, 2020. https://www.statista.com/statistics/912204/willingness-pay-premium-natural-organic-generational-us/.

What is Data Warehousing and Why is it Important? (2017, May 30). Retrieved from HerzingUniversity website: https://www.herzing.edu/blog/what-data-warehousing-and-why-it-important#:~:text=Data%20warehousing%20improves%20the%20speed

Barnett, M. (2017, August 1). The Importance of Real Time Data & Five Reasons Why You Need it. Retrieved from Fourth Source website: https://www.fourthsource.com/data/importance-real-time-data-five-reasons-need-22014

Australian Farm Institute. (2017, June 5). Farmers have always been Lean and Agile. https://www.farminstitute.org.au/farmers-have-always-been-lean-and-agile/

Benjamin, M. (2019, September 11). Precision Livestock Farming Technology for Remote Monitoring of Pig Welfare and Health: Part 1. Michigan State University. https://www.canr.msu.edu/news/precision-livestock-farming-technology-for-remote-monitoring-of-pig-welfare-and-health

McKendrick, J. (2017, November 1). Rise of the Real-Time Enterprise: It’s All About the Data. Retrieved January 30, 2021, from Database Trends and Applications website: https://www.dbta.com/Editorial/Trends-and-Applications/Rise-of-the-Real-Time-Enterprise-Its-All-About-the-Data-121160.aspx

Christoph G. Schuetz, Simon Schausberger & Michael Schrefl (2018) Building an active semantic data warehouse for precision dairy farming, Journal of Organizational Computing and Electronic Commerce, 28:2, 122-141, DOI: 10.1080/10919392.2018.1444344

Khan, Abeer & Ehsan, N. & Mirza, Ebtisam & Sheikh, Zahoor. (2012). Integration between Customer Relationship Management (CRM) and Data Warehousing. Procedia Technology. 1. 239–249. 10.1016/j.protcy.2012.02.050.