Prioritizing Retail Digital Transformation Cases by Data Accessibility and Viability

Stanton Man, Kagen Quiballo, and Krishanu Sengupta

MSDS 403: Data Science in Practice

January 31, 2021

As a finalist company competing for a digital transformation project for the Hypothetical Global Organization ("HGT") in Retail, the analytics consultancy ("the Consultancy") shall recommend a data science use case for HGT based on data accessibility and viability. After analyzing and prioritizing six representative digital transformation use cases (each, an "Initiative") based on a Data Source Assessment Matrix (see "Appendix") that weighs different qualities of the data required for the Initiatives (the "Data"), the Consultancy recommends focusing on customer sentiment analysis and targeted marketing Initiatives based on the likelihood of rich, easily available data that can be measured clearly for impact on a business objective.

1. The Data Source Assessment Matrix and the Consultancy's Use Case Prioritization

The Data Source Assessment Matrix (the "Matrix") evaluates the Initiatives based on different properties of the data they require to accomplish a successful digital transformation. It evaluates each Initiative using 6 different criteria, with each criteria scored on a scale of 1-5 (1 being the lowest, 5 being the highest) or 1-10, then 2 summary scores that include important elements from the initial 6 criteria, and finally an overall data accessibility and viability score that averages the 2 summary scores to provide a single score for the Initiative. The 9 criteria in the Matrix are explained in greater detail in the Appendix.

2. Analysis and Results of Scoring Initiatives

Using the Matrix, the Consultancy arrived at an Overall Data Accessibility and Viability Score for each Initiative, and went into more depth with the top 3. The findings that these 3 Initiatives from 2 global companies in Amazon and 7FAM, and a large North American Retailer in Valpak, send a strong signal that similar types of Initiatives should be targeted by the Consultancy for HGT based on their data accessibility and viability.

Customer Sentiment Analysis: Customer sentiment analysis draws insight from customer feedback regarding HGT products and services. Amazon's Initiative leveraged its Comprehend NLP service to process unstructured text such as product reviews and provide reliable and relevant products ("Amazon" 2021). By targeting key terms using NLP, Comprehend eliminates the need to pre-process and clean data, making data source possibilities endless. This Initiative scored highest because of the simple process for finding insights from easily accessible data as well as its direct alignment with the HGT's goals of prioritizing customer's satisfaction and providing top of the line, reliable products. A quarterly customer sentiment analysis that utilizes Amazon Comprehend on HGT's company-accessible data can guide future marketing efforts better allocate resources.

Targeted Marketing: Targeted marketing is a company's segmentation of its customers and planning marketing strategies based on those segment characteristics. Jean company 7FAM used customers' previous purchasing history of similar products to send a targeted email campaign for a new product offering to likely interested customers without spamming their entire email and potentially losing customers ("7" 2018). The relative simplicity of this Initiative stems from the 2 main data sources, an email marketing system and the company's own transaction history. The company focused on what type of previous customer purchasing and browsing history would indicate interest in a new product. The metrics around success, primarily revenue per user, was clear and helped make this Initiative score highly. HGT has similar needs to target its vast customer base intelligently, and has the scale and email and purchasing data systems to accomplish this.

Customer Churn: Customer churn is the percentage of customers that stop using a company's product. Valpak, a large direct marketing company, addressed churn by analyzing touchpoints between account executives and customers, as well as other contacts with customers to predict which customers would churn ("Reducing" 2019). Additionally, their Initiative had a clear metric to judge success, churn. They also had a rich source of customer attributes that could be correlated with churn, including order data. The volume of data and frequency of updates required was considerable, which did cause this Initiative to score lower overall than the top 2. HGT could benefit from learning how Valpak measured success and used their customer data to help reduce churn with its customers.

3. Conclusion and Takeaways

The strongest Initiatives had both the availability and viability of Data and the business experience to interpret that Data and measure success clearly. Amazon's customer sentiment analysis benefited from drawing insight from customer feedback and product reviews, as well as its ability to preprocess unstructured Data. Similarly, 7FAM's targeted marketing Initiative had a tight company focus on the email marketing and customer purchase data which was easily measurable in increase of sales and higher interaction rates to targeted individuals. While Valpak's customer churn Initiative also had highly relevant data its team was familiar with analyzing, the sheer volume of that data and frequency of updates required made its score lower. HGT should focus on its rich history of customer feedback and purchase data to draw insights from successful marketing tactics, like targeted emails, and put customer satisfaction first. These Initiatives provide simple processes to digitize various sectors within the organization by using easily accessible data to create a lasting impact.

Appendix

Data Source Assessment Matrix

The description of the assessment criteria below follow:

- "Types of Input Data Needed" measures the ease of pulling and integrating the Data into the Initiative, with 1 being highly complex and 10 being relatively easy
- "Success Measurement Possible with this data" measures the feasibility of providing key metrics that help judge Initiative success, including an initial baseline, with 1 being very difficult to impossible and 5 extremely straightforward to quantify
- "Volume of data needed" measures the ease of managing and working the Data, with more data meriting a lower score and less a higher score, closer to 5
- "Data Availability" measures the ease of accessing the Data, including any associated cost from acquiring it, with 1 being non-accessible and 5 being more easily available
- "Frequency of data needed" measures both the volume of data and frequency of updates, with 1 being highly complex frequency updates and 5 less frequent and less complex
- "Structured vs Unstructured" measures the amount of preparation the Data needs to be ready for modeling and reporting, with unstructured data such as text, images, and speech requiring more preprocessing and complexity and meriting a 1 score, whereas highly structured data without many if any preprocessing steps would be a 5 score
- "Data Quality" is a summary measure that incorporates volume, frequency, and the structure of the Data to arrive at an overall score, with 5 being high quality
- "Match Keys and Business Knowledge Available around Data" is also a summary measure that focuses on if success measurement was possible with the Data, including if the business was knowledgeable enough about the type of Data to draw a conclusion about the Initiative's success, with 5 being very confident in measuring success given the Data
- "Overall Data Accessibility and Viability Score" is an average of the two summary measures

Use Case	Types of Input Data Needed (1-10)	Success Measurement Possible with this data (1-5)	Volume of Data Needed (1-5)	Data Availability (1-5)	Frequency of Data Needed (1-5)	Data is Structured vs. Unstructured (1-5)	Data Quality (1-5)	Match Keys and Business Knowledge Available around Data (1-5)	Overall Data Accessibility and Viability Score (1-5)
Customer Sentiment Analysis example: Amazon Reviews	8	5	3	4	3	2	4	5	4.5
Targeted Marketing: 7FAM	5	5	4	4	5	5	4	4	4
Customer churn example: Valpak (Through AgileThought)	7	5	1	4	2	4	4	3	3.5
Customer loyalty program: Starbucks	5	2	2	3	1	4	3	3	3
Augmented Reality example: Home Depot Project Color	9	1	5	5	2	1	5	1	3
Customer segmentation : example: MetLife Customer Segmentation	5	2	4	2	4	2	3	2	2.5

References

"Amazon Comprehend." *AWS*. Accessed January 24, 2021. https://aws.amazon.com/comprehend/.

"7 For All Mankind." *Custora, 2018*. https://www.custora.com/customers/7-for-all-mankind.

"Reducing Customer Churn with Predictive Analytics Discovery." *AgileThought, 2019*. https://agilethought.com/client-stories/reducing-customer-churn-predictive-analytics-discovery/.