

Analytics Driving Supply Chain Segmentation for Lenovo

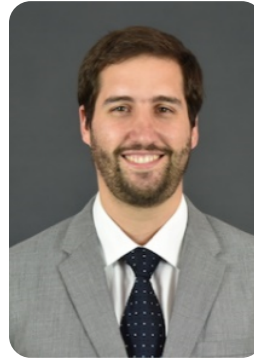
By Luiz Gosling and Javier Urrutia
Advisor: Maria Jesus Saenz, PhD

Topic Areas: Supply Chain Strategy,
Segmentation, K-Means Clustering

Summary: This research provided an approach for a Machine Learning application for customer-oriented supply chain segmentation at Lenovo. A k-Means model on Importance and Complexity dimensions identified four major segments among their Hyperscale Business Unit. A supply chain policy analysis suggested the design of three distinct policies and strategies for the identified clusters. Lenovo will expand the model for further business units, and integrate more segment characteristics and performance metrics, particularly capturing Customer Experience and Cost Efficiency.



Prior to MIT, Luiz Gosling was an Associate Partner at Infra Partners, coordinating M&A assessments and operational due diligences for clients in the Logistics Industry. Prior to M&A, worked for a Fortune 500 mining company and studied Industrial Engineering (BEng) and Logistics (MSc) at PUC-Rio, in Brazil.



Prior to MIT, Javier Urrutia worked as Project Controller at Moly-Cop Chile, a global leader in grinding media and mining chemicals, where he managed large scale capital projects. Studied Financial Engineering (BEng) and Operations Research (MEng) at Universidad de Concepcion, in Chile.

KEY INSIGHTS

- 1. Lenovo DCG's client-product portfolio has four supply chain segments.**
- 2. High service levels are observed across Lenovo DCG's portfolio yet cost efficiency differs.**
- 3. Our analysis accounts for two factors (Complexity and Importance) and recommends three distinct supply chains for Lenovo DCG Hyperscale business unit.**
- 4. A framework for the implementation of a Machine Learning approach to customer segmentation is provided and embraced by Lenovo.**

Introduction

The choice of segmented supply chain is a strategic-level decision with long-term impact on a firm's performance. Although being common in the academic literature, two realities contrast in supply chain management: While a vast number of

companies still perceive supply chain as cost-centers, only a few have adopted profit-driver approaches. The second group embraced segmentation to deploy end-to-end supply chain strategies that match the segments' requirements to the company's capabilities, adding sustainable value in the process.

Business Context

As cloud, search engine and social media providers ("Cloud providers" or "Hyperscale") become increasingly focus on optimizing CAPEX, they also have to balance fast-paced innovation and superior customer experience. As a result, players such as Lenovo Data Center Group (DCG) act as solution providers, offering (i) Hardware Design, (ii) Procurement, and (iii) System Integrator options for their clients. System Integrators (SI) are middle-men between clients and their suppliers, while also being vendors themselves. As such, SIs provide physical logistics, and design and model clients' solutions (i.e. data centers' infrastructure and software capabilities), including the selection and management of vendors globally.

Lenovo understands that their current "one size fits all" supply chain approach fails to address individual

needs in a precise manner. All things considered, Lenovo is potentially limiting its ability to deliver superior perceived value for bespoke clients and higher efficiency for cost-sensitive ones.

Lenovo proposed an assessment of their Hyperscale supply chain to better serve their portfolio. Thus, this capstone had three goals: (i) review frameworks in the supply chain segmentation literature; (ii) identify key **clients** and **products** portfolios through **quantitative methods**; and (iii) propose supply chain **policy guidelines** for each identified group, creating baselines for change.

Methodology

This project was conducted in five phases: (i) **Literature Review** covering supply chain strategies and segmentation; (ii) **Company and Sector Immersion**, interviews with executives and key-people from Lenovo, technical visits, and developed our research questions and hypotheses; (iii) **Data Gathering**, covering three years of records on sales, inventory, client, channel, product, and delivery data from Lenovo. (iv) the **Analysis** phase consisted of a Dimension Reduction on the descriptive variables and the application of clustering algorithms on the data collected. Lastly, the (v) **Research Insights** phase provided quantitative insights on the results, as well as proposed a framework that revealed targets of future work through a workshop.

Clustering was a way of classifying Lenovo's client-product portfolio, which translates into analytically grouping sets of data, using measurable characteristics that ensure: (i) similarity among points in the same group; and (ii) differences between points in different groups. We chose *k-Means* as the preferred algorithm.

Regarding data collection, a few areas of information were not adequate for the intended application: (i) inventory, given its granularity, not fitting the client-product level; (ii) manufacturing and shipping capacity due to granularity. Moreover, the data gathered covered only North American. That said, the data collected was paramount in providing factual quantitative information on the customer-segmentation exercise, including sales trends, geographical footprint, channels performance, and service performance.

6775 Sales records were consolidated and grouped onto 142 unique client-product pairs. 13 variables

were identified as features for describing client-products. Quantitative data was normalized before application of EFA and subsequent clustering.

Due to statistical significance, not all features were included in the cluster analysis: The original list of 13 was reduced to 10: Sales in Dollars, Sales in Units, Demand Uncertainty, Order Frequency, Average Order Size, Distribution Channel, Total Lead Time, Manufacturing Lead Time, Destinations, and Delivery on Time (%). Total Lead Time and Delivery on Time (%) were kept as target metrics.

The resulting factors were named **Importance** and **Complexity**, per their composition. Importance included Sales Volume and Revenue features, while Complexity was composed of Demand Uncertainty and Number of Destinations (countries).

Results and Discussions

The *k-Means* elbow curve suggested 3 to 4 clusters as ideal. Figure 1 shows the graphical representation of the four clusters.

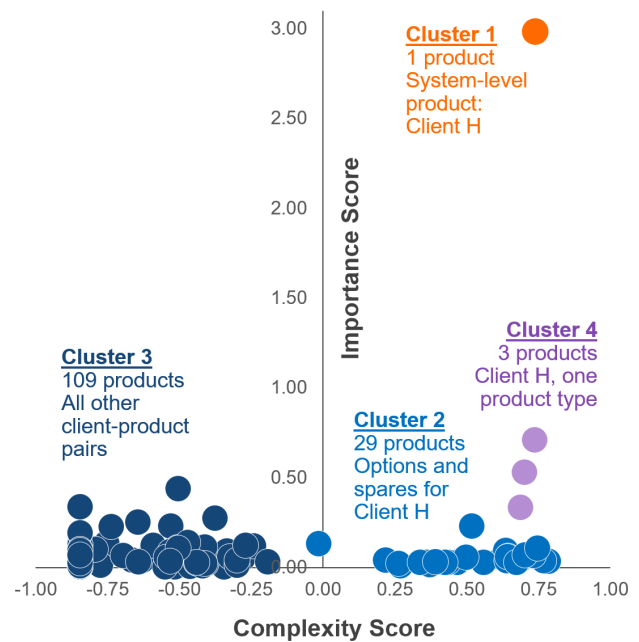


Figure 1: clusters on the Complexity-Importance space.

C1 is an outlier composed by one single product, supplied to one single client (H). C2 has 29 components (tier-2, tier-3 Bill of Material goods) that are also supplied exclusively to Client H. C3 has 109 products and could be further segmented to define accurate policies for more segregated groups; a few large clients in the cluster, for example, have a clear share-of-wallet expansion target for the foreseeable future, which should be considered, among other

factors. C4 represents items supplied to Client H with demand and supply particularities that differentiate them from C2. **Figure 2** summarizes results from the ANOVA and Tukey analysis, for inter-cluster heterogeneity.

Policies	Cluster1	Cluster2	Cluster3	Cluster4
Cluster Description	System-level Product	Options and Spares for Client H	All other Clients	Client H, one product type
Cluster Size (members)	1	29	109	3
Avg Importance Score	2.979	0.051	0.058	0.523
Importance Position	High	Low	Low	Median
Avg Complexity Score	0.740	0.515	-0.716	0.710
Complexity Position	High	Median	Low	High
Target Metric 1 Total Lead Time	47.437	30.943	26.606	46.787
Target Metric 2 Manufacturing Lead Time	23.373	11.455	9.955	22.784
Target Metric 3 Delivery on Time %	0.884	0.748	0.763	0.890

Figure 2: Complexity and Importance among clusters.

Removing C1 from the analysis, the means comparison reveals that, for Importance, C2, C3 and C4 are all significantly different. Regarding Complexity, we found that C2 and C4 show a non-significant means difference at a 0.05 level. As similar levels of performance are noticed between target metrics, the need for additional targets becomes evident (e.g. Cost Efficiency), which would enable further differentiation between groups and more accurate segmentation. This is important due to the fact that the four clusters are fundamentally different, yet the available metrics do not capture such differences in execution.

The establishment of the additional targets would allow Lenovo to periodically assess segments' performance. Machine Learning approach, consist of periodic reruns with comparisons between outputs in order to identify levers (supply chain policies) that improve target metrics (e.g. maximization of Customer Experience and Cost Efficiency, or minimization of Lead Time). Changing policies and remeasuring – through data collection or simulation – creates a formal, robust process that reveals the individual changes' impact. **Figure 3** illustrates a methodological approach.

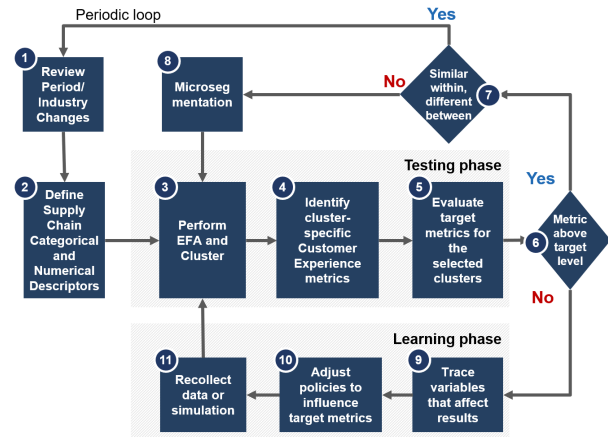


Figure 3: Machine Learning implementation for Customer Driven Supply Chain Policies with an Evaluation Loop.

The cluster analysis tests confirmed four clusters as an appropriate number among the firm's hundreds of client-products. Among those clusters, one is highlighted as an absolute outlier, and analyses were re-run without it, with no significant change in results.

Moreover, such discussion should build a guideline for future work, indicating focus areas for Lenovo and clarifying how proposed clusters differ not only in performance, but also in policies and strategies. To achieve that, a workshop was conducted with a team of experts from Lenovo. A matrix-shaped framework was presented to the team as a way to differentiate policies, given that – in the current practice – supply chain policies are not defined for each cluster suggested by this study. Policies were grouped into five categories: Sourcing, Inventory, Production, Fulfillment, and Customers. **Figure 4** illustrates the framework and provides a skimmed synthesis of the populated matrix after the workshop.

Policy	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster Description	System-Level Product, Client H	Options and Spares, Client H only	All other Clients	Client H, one type of product
Cluster size	1	29	109	3
Sourcing	Very important, defines the cluster	Somewhat important, supports the cluster	Very important, defines the cluster	Similar to C1
Inventory	Somewhat important, supports the cluster	Very important, defines the cluster	Somewhat important, supports the cluster	
Production	Not as important, no direct effect on cluster	Not as important, components not manufactured by DCG	Heterogeneous	
Fulfillment	Somewhat important, supports the cluster	Not as important, no direct effect on cluster	Includes various clients. Further segmentation will help better understand this cluster.	
Customers	Very important, defines the cluster	Very important, defines the cluster		

Figure 4: Policy-cluster matrix with insights from the workshop.

Each cluster had its guiding-policies identified and discussed and insights on what could be an area of focus for future adjustment or analytical work. While

C1 is strongly defined and affected by Sourcing policies, C2 is mostly driven and regulated by Inventory policies, given the nature of its products. These and other examples of quanti- and qualitative similarities and differences among clusters evidences the potential cluster-based segmentation of Lenovo's supply chain execution, likely improving adherence to policies that are often conflicting between clusters, improving service, and preventing inefficiencies.

Conclusions

As our clustering exercise revealed four different segments among Lenovo Hyperscale's portfolio, independent strategies are being internally assessed with policies based on the client-product nature. Through a policy-cluster framework, we defined current state of policies for each cluster, which now can guide internal discussions to adapt part of Lenovo's operations to a segmented reality, in which clusters with specific needs are handled with capabilities that adequately match requirements. Lastly, the Machine Learning Loop embraced by Lenovo will ensure the continuous use of resources to improve upon these results.

We hope that further studies will build upon these results to produce relevant knowledge on this increasingly important field and that Lenovo works towards continuously improving their supply chain to better serve their clients and provide sustainable value to their shareholders.