



AYELET ISRAELI

FEDOR (TED) LISITSYN

## E-Commerce Analytics for CPG Firms (A): Estimating Sales

The manager of the E-Commerce analytics group stared at the screen, looking through the tasks of the week. The E-Commerce Analytics group was in charge of compiling various online sales reports, as well as making data-driven recommendations for sales and marketing tactics. The young group was established less than a year ago in the over-a-century old CPG firm, but by now the manager was getting requests from many different stake holders in the firm. It was not an easy cultural shift, but finally managers in the traditional firm realized the importance of an analytics group when navigating today's marketplace.

In the last decade or so, hundreds of Digitally Native Vertical Brands (DNVB) and Direct-To-Consumer (DTC) businesses entered virtually every industry and product category, fueled by E-commerce and technology, threatening traditional businesses. One of the major differences between these new types of firms and traditional firms is the distribution model, which allowed the new generation of firms to own the relationship with the consumer, whereas the traditional firms relied heavily on their retail partners. Related, DNVB and DTC brands often had high quality data on each and every end-consumer interaction with the brand, which allowed them to better understand their consumers and cater to their needs. Initially, traditional firms continued working with their original distribution models, adding e-tailers and m-tailers to their mix of channel partners. More recently, and somewhat accelerated by COVID-19, these firms have also started developing their own digital presence, sometimes investing in DTC channels, and have been significantly revamping their data and analytics capabilities.

The first task the manager prioritized was urgent. It was rumored that one of the key channels of distribution, a retailer that sells 100 SKUs of the CPG firm, would soon stop sharing E-Commerce specific information. Instead, the retailer would provide total sales information (combination of their offline and online sales), and likely provide online information only for 10 SKUs. Given the importance of the E-Commerce channel, and the need to perform more research on online trends, the leadership team asked the manager to come up with an idea on how to make an accurate estimation going forward. The manager and the data scientist came up with five different approaches to estimate online sales and wanted to test them using historical data: historic month trends, SKU-level trends, shares of the top 10 SKUs, extrapolating sales using every 10<sup>th</sup> SKU, and average online shares. Which is the preferred methodology? Can the firm rely on these estimates? Is there another methodology that they

---

Professor Ayelet Israeli and Fedor (Ted) Lisitsyn (HBS MBA 2019) prepared this exercise as the basis for class discussion rather than to illustrate either effective or ineffective handling of an administrative situation. The exercise is not based on a single individual or company but is a composite based on the authors' general knowledge and experience.

Copyright © 2021 President and Fellows of Harvard College. To order copies or request permission to reproduce materials, call 1-800-545-7685, write Harvard Business School Publishing, Boston, MA 02163, or go to [www.hbsp.harvard.edu](http://www.hbsp.harvard.edu). This publication may not be digitized, photocopied, or otherwise reproduced, posted, or transmitted, without the permission of Harvard Business School.

are overlooking? If they are able to negotiate with the retailer on the 10 SKUs they will receive data for, which 10 SKUs should they ask for?

## The data

The spreadsheet supplement for this exercise (HBS No. 521-712) contains several tabs. The “Instructions” tab provides detailed instructions and guiding questions to complete the exercise. The “Raw Data” tab contains online sales performance for all 100 SKUs from the retailer for each month, September-November, and an indicator for seasonal items (seasonal=1 indicates seasonal items). The tabs “Trends”, “Top-10”, “Every n-th”, and “Share of Online” can be used to calculate the projections using each of the five methodologies, and the “Output” tab allows a comparison of the different methodologies. Using the performances in September and October, the goal is to compare the different projections for November online sales in order to choose the best methodology.