**Size distribution for seasonal products**

**Quentin Meeus, Koen Graat, Wesley Fitters, Lennard Albarda,**

**Simone Latisa, Vincent Smeets**

Jheronimus Academy of Data Science (JADS)

**Abstract**

**1. Introduction**

A well-known problem for retailers is estimating stock sizes for seasonal products. The fast development of technology and science causes an increase in demand for fashionable or seasonal products. These products are associated with high levels of uncertainty and short cycles of sales. That is why the management for purchasing certain products is essential and complicated at the same time. Capital is a crucial part of purchasing but especially small and medium-sized enterprises often experience a scarcity in working capital. For that reason, the focus of a retailer should be on efficient management of logistic processes (Shi, Guo and Fung, 2017).

A way to define the management of logistic processes is: “The process of planning, implementing and controlling the efficient, cost-effective flow and storage of raw materials, in-process inventory, finished goods and related information flow from point-of-origin to point-of-consumption for the purpose of conforming to customer requirements” (Makepeace, Tatham, & Wu, 2017). Part of the logistic management is inventory control, which focuses on the quantity of orders, lead-time and safety stock. Seasonal demand affects the control of inventory. This influence on inventory can be explained from a shorter sales period, unstable demand for seasonal products and longer lead-times related to production. Over the last couple of years, an increasing demand for seasonal products has been detected compared to the basic demand products. The goal for retailers is to manage this growing demand of seasonal products efficiently in order to avoid problems with inventory control.

The changing market demand after order placements by retailers possibly are the cause for under-or overstocking (Bernhardsson and Johansson, 2017).

This increasing demand for seasonal products comes from a growing trend towards diversity and individualization. Other reasons for this trend in demand are global competition, rapid product development and growing flexibility in the manufacturing process. In the last decades, the average department store has expanded up to five times his original size. Results of such an increase are difficulties in predicting short-term demands for individual items. Dependable forecasting models for seasonal products are difficult to construct due to their smaller life cycles and rapidly changing collections. For that reason, many companies experience difficulties in calculating stocks and the beginning of the season (Bernhardsson and Johansson, 2017). On the other hand, companies that do not account for seasonality patters in demand will experience a systematic mismatch in supply and demand for items at store level. This mismatch results in structural higher costs for companies than necessary. Although retailers are aware of this problem, many of these retailers lack the ordering systems that have the technical capabilities to include seasonal trends (Ehrenthal, Honhon and Van Woensel, 2014).

Scientific research focuses to a large extend on predicting sales and stock levels for these products. Difficulties are experienced in the prediction models due to demand uncertainty, the product variability and a lack of historical sales data. Complex models, like ARIMA, regression, neural networks or data mining models were applied. In order to advance the predictive capacities, hybrid models are introduced that combine the advantages of various models (Liu et al, 2013). Eventually, the idiosyncrasies within the fashion markets could only be explained through highly advances models which could hardly be adopted by retail companies (Beheshti-Kashi et al, 2015). For that reason, this research will not focus on a complex predictive model but on advanced insights into the available company data.

A retail company that experiences difficulties incorporating seasonal demand into their inventories is Shoeby. Shoeby is a company that started in 1981 with a single store in Den-Bosch but has become a chain with more than 230 stores in the Netherlands in less than forty years. This chain not only distributes its products through their stores but also offers a substantial online sales platform, including a Shoeby app. For this retailer, over-or understocking is a recurring problem and their inventory systems do not account for seasonality. Leftover stocks at the end of the season lead to unnecessary costs for Shoeby and for that reason it is essential to focus on these trends in the future. The available sales and stock data of Shoeby ranges between August 2015 and December 2017. As discussed, constructing a predictive model is unrealistic due to lack of sufficient historical data and complexity of the available datasets. Providing Shoeby with methods and tools to generate highly advanced exploratory results related to their current data should lay the groundwork for possible predictive models in the future. This results into the following research question:

*To what extend can a size distribution analysis give insights into the seasonal products of Shoeby?*

In order to answer this question a series of steps will be executed throughout the paper. Firstly, the method is described. Then, the used dataset is described including the treatments and manipulations applied to the data. The results shall be presented thereafter, followed by general discussion and some concluding remarks.

**2. The method**

The method describes how insights within the shoeby data are generated. Multiple steps have been applied to the problem to create a complete overview. Firstly, a size distribution tool is implemented, secondly measuring cure show and calculates the differences between a normal distribution and the actual distribution and finally a dashboard displays all the results in one simple tool.

**2.1. Size distribution tool**

The method related to this research is based on analyzing the sales and stock data between 2015 and 2017. These datasets contained a large amount of unnecessary features which have been dropped in order to use the data more efficiently.

**3. The Dataset**

Netscalers is a data-driven company that works together with Shoeby in order to tackle company problems by data anlysis. In cooperation with Netscalers, datasets related to Shoeby’s sales history were distributed. These datasets contained information about the sales, the Shoeby inventory and

**3.1. The sales dataset**

The sales dataset has been extensively used throughout the research. This dataset contains more than a million rows and 148 columns with specific information related to sales between august 2015 and September 2017. First of all, a selection of useful columns is made and implemented within the raw datasets. This selection contains the 15 most relevant columns, such as the order time, the brand, the season, the sizes and the quantities. A new column containing the right date and time format is made and the old time and date columns were dropped. Also, all the items that showed “Kosten verzending” (cost shipping) or “AFTERPAY (WEB) Nederland” as a product description are dropped from the sales dataset. This means that the original dataset is reduced by more than 250.000 rows. Furthermore, the order date has been introduced as the new index for the dataset and the data is grouped per day. This grouped data is combined with specific filters, namely Web Shop Code, Merchandise Code, Brand and Season, which form the new columns. The quantities sold and returned are aggregated for each product and size. These two quantities have been represented in the last two columns of the new dataset. The horizontal component code, which contains the sizes of the items, is renamed to sizes. Finally, a filter is applied to include only the items that possessed a size between XXS and XXXL. A substantial amount of items did not have a workable size assigned to it so these shall not be used for further analysis.

The reshaped sales dataset now contains around 230 thousand rows and seven columns. These columns display useful information for each item with a workable size per day. It shows the Web Shop and Merchandise code, from which brand and season the product originates and the quantity sold and returned.

**3.2. The inventory dataset**

A large dataset concerning inventory data was provided by a PDEng student called Valentine Tuyishime. This student has been working on Shoeby stock data in cooperation with Netscalers for extensive period of time. Edwin van Dongen, an employee at Netscalers, addressed the possible that Valentine could have important insights concerning this research. The inventory datasets came with specific guidelines for cleaning the data in order to keep the useful information, which resulted in the merged stocks dataset.

The merged stock dataset contains more than a million lines and 20 columns. This set provides information about the types of products, their opening and closing inventories, product descriptions and possible discounts between February 2016 and September 2017. For some products the opening inventories or other important information is unknown, so these products will not be incorporated within the inventory dataset.

References

Bernhardsson, R., & Johansson, H. (2017). A CASE STUDY ON HOW TO IMPROVE ORDER QUANTITIES WITH SEASONAL DEMAND. *E–Proceeding*, *2017*, 180.

Beheshti-Kashi, S., Karimi, H. R., Thoben, K. D., Lütjen, M., & Teucke, M. (2015). A survey on retail sales forecasting and prediction in fashion markets. Systems Science & Control Engineering, 3(1), 154-161.

Ehrenthal, J. C. F., Honhon, D., & Van Woensel, T. (2014). Demand seasonality in retail inventory management. European Journal of Operational Research, 238(2), 527-539.

Liu, N., Ren, S., Choi, T. M., Hui, C. L., & Ng, S. F. (2013). Sales forecasting for fashion retailing service industry: a review. Mathematical Problems in Engineering, 2013.

Makepeace, D., Tatham, P., & Wu, Y. (2017). Internal integration in humanitarian supply chain management: Perspectives at the logistics-programmes interface. Journal of Humanitarian Logistics and Supply Chain Management, 7(1), 26-56.

Shi, J., Guo, J. E., & Fung, R. Y. (2017). Decision support system for purchasing management of seasonal products: A capital-constrained retailer perspective. Expert Systems with Applications, 80, 171-182.

<http://www.thaiejournal.com/wp-content/uploads/ICIBE-2017.pdf#page=189>

<https://onlinelibrary.wiley.com/doi/epdf/10.1002/cplx.21729>