Author(s): Makkonen, Markus; Frank, Lauri

Title: Identifying the Sales Patterns of Online Stores with Time Series Clustering

Year: 2018

Version: Published version

Please cite the original version:

Makkonen, M., & Frank, L. (2018). Identifying the Sales Patterns of Online Stores with Time Series Clustering. In A. Pucihar, M. Kljajić, P. Ravesteijn, J. Seitz, & R. Bons (Eds.), Bled 2018 : Proceedings of the 31st Bled eConference. Digital Transformation : Meeting the Challenges (pp. 491-506). University of Maribor Press. https://doi.org/10.18690/978-961-286-170-4.34

Identifying the Sales Patterns of Online Stores with Time Series Clustering

MARKUS MAKKONEN & LAURI FRANK

Abstract During the past decades, electronic commerce, especially in the business-to-consumer (B2C) context, has emerged as a popular research topic in information systems (IS). However, this research has traditionally been dominated by the consumer focus instead of the business focus of online stores. In this explorative study, we aim to address this gap in prior research by identifying the most typical sales patterns of online stores operating in the B2C context. By segmenting the monthly sales time series of 399 online stores with time series clustering, we are able to identify four approximately equally sized segments, of which two are characterised by a clear upward or downward trend in the sales and two are characterised by strong seasonal sales variation. We also investigate the potential segment differences in terms of several key business and technical parameters as well as discuss more broadly the applicability of time series clustering to IS.

Keywords: • Electronic Commerce • Business-to-Consumer (B2C) • Online Stores • Sales Patterns • Time Series Clustering • Segmentation •

CORRESPONDENCE ADDRESS: Markus Makkonen, Project Researcher, University of Jyvaskyla, Faculty of Information Technology, PO Box 35, 40014 University of Jyvaskyla, Finland, e-mail: markus.v.makkonen@jyu.fi. Lauri Frank, Ph.D., Research Professor, University of Jyvaskyla, Faculty of Information Technology, PO Box 35, 40014 University of Jyvaskyla, Finland, e-mail: lauri.frank@jyu.fi.

DOI https://doi.org/10.18690/978-961-286-170-4.34 ISBN 978-961-286-170-4

Available at: http://press.um.si.

1 Introduction

During the past decades, electronic commerce, especially in the business-to-consumer (B2C) context, has emerged as a popular research topic in information systems (IS). However, the prior studies on the topic have traditionally been dominated by the consumer focus, such as the factors affecting online consumer behaviour (e.g., Perea y Monsuwé, Dellaert & de Ruyter, 2004; Chang, Cheung & Lai, 2005; Cheung, Chan & Limayem, 2005) and the segmentations of online consumers, based either on self-reported survey data (e.g., Brown, Pope & Voges, 2003; Swinyard & Smith, 2003; Kau, Tang & Ghose, 2003; Bhatnagar & Ghose, 2004a, 2004b; Rohm & Swaminathan, 2004; Brengman et al., 2005; Allred, Smith & Swinyard, 2006; Barnes et al., 2007; Soopramanien & Robertson, 2007) or on click-stream data about the actual visits in online stores (e.g., Moe, 2003; Su & Chen, 2015). In contrast, the business focus of online stores has traditionally gained far less attention. For example, whereas various segmentations exist for online consumers based on their purchase behaviour, no such segmentations have been developed for online stores based on their sales patterns.

In this explorative study, we aim to address this gap in prior research by identifying the most typical sales patterns of online stores operating in the B2C context. With *a sales pattern,* we refer to the temporal variation in the aggregate sales of an online store across all its sold items. Based on our review of prior literature, we find this study to be the first of its kind. Of course, sales patterns (or sales variation) as such have been investigated in several prior studies (e.g., Geurts, 1988). However, these studies have traditionally aimed at forecasting the future sales of an individual item or store instead of investigating the past sales of a wider sample of stores and identifying typicalities in them. Or those studies that done this, have done it only in the context of traditional offline retailing (e.g., Carman & Figueroa, 1986). We see that this kind of an investigation and identification can be highly valuable in both theoretical and practical respects. First, it helps us to better understand the characteristics of online markets and, for example, to anticipate the upcoming seasonal sales peaks, which can be considered important especially for logistic and payment service providers in terms of proper resource allocation. Second, if we are able to find causal explanations for the emergence of specific sales patterns, or at least associate them with specific business and technical parameters of the stores, we may be able to offer the stores competitive advantage by suggesting what kind of strategies help them to move away from sales patterns that are likely to be perceived as mostly negative (e.g., decreasing sales and strong seasonal sales variation) and towards sales patterns that are likely to be perceived as more positive (e.g., increasing and more stable sales).

We conduct the study in two phases by analysing the data collected from 399 online stores operating mainly in the Finnish market. In the first phase, we segment the stores based on their monthly sales time series from January 2016 to December 2017 by using time series clustering. In the second phase, we investigate the potential segment differences in terms of several key business and technical parameters.

The paper proceeds as follows. After this introductory section, Section 2 briefly discusses the basics of time series clustering, which has remained a rarely used data analytics method in IS. After this, we will describe the methodology and results of the study in Sections 3 and 4. The results will be discussed in more detail in Section 5, which also discusses more broadly the applicability of time series clustering to IS. Finally, Section 6 describes the main limitations of the study and proposes some potential paths of future research.

2 Time Series Clustering

As mentioned above, time series clustering has remained a rarely used data analytics method in IS. However, it has been commonly used in other disciplines and domains, such as finance and signal processing. Literature reviews on time series clustering are provided, for example, by Liao (2005), Fu (2011), as well as Aghabozorgi, Shirkhorshidi, and Wah (2015).

Clustering, in general, refers a statistical technique in which objects are grouped together so that they have maximum similarity (or minimum dissimilarity) with the objects in the same group, but minimum similarity (or maximum dissimilarity) with the objects in the other groups. Time series clustering, in turn, refers to a specific case of clustering in which the grouped objects are time series consisting of multiple data points in a chronological order (Liao, 2005; Aghabozorgi, Shirkhorshidi & Wah, 2015). Time series clustering can be divided into three main types (Aghabozorgi, Shirkhorshidi & Wah, 2015): (1) whole time series clustering, (2) subsequence time series clustering, and (3) time point clustering. Of these, whole time series clustering refers to the clustering of a set of separate time series by considering the similarity or dissimilarity of the time series as a whole, whereas subsequence time series clustering and time point clustering refer to the clustering of a set of sub-sequences or time points from a single time series. In this paper, we will concentrate only on whole time series clustering.

The typical process of whole time series clustering can be divided into four main phases (Aghabozorgi, Shirkhorshidi & Wah, 2015): (1) the selection of a representation method, (2) the selection of a distance measure, (3) the selection of a clustering algorithm, and (4) the evaluation of the clustering solution. First, the dimensionality of the raw time series is typically reduced by using a representation method, which can be divided into data adaptive, data non-adaptive, model based, and data dictated approaches (Aghabozorgi, Shirkhorshidi & Wah, 2015). The purpose of this phase is to reduce both the memory and computation requirements of the following phases as well as to filter out outliers and noise.

Second, the similarity or dissimilarity of the time series is calculated by using the selected distance measure. Its selection depends on the characteristics of the time series, the selected representation method, and the overall objective of the clustering, which may be to find similar time series in terms of time, shape, or change (Aghabozorgi, Shirkhorshidi & Wah, 2015). The distance measures can be divided into shape-based, comparison-

based, feature-based, and model-based measures, of which the shape-based measures, especially the Euclidean distance and the dynamic time warping (DTW) distance, are currently the most commonly used ones (Aghabozorgi, Shirkhorshidi & Wah, 2015). The Euclidean distance is suitable for finding similar time series in terms of time, meaning that the potential patterns in the time series must occur synchronously. For example, if there are sales peaks in the sales time series of two online stores, the sales time series are considered similar only if the sales peaks occur at the same time. In contrast, the DTW distance is suitable for finding similar time series in terms of shape, meaning that the potential patterns in the time series may occur asynchronously. For example, in the aforementioned case of two online stores, it would not be necessary for the sales peaks to occur at the same time in order for the sales time series to be considered similar.

Third, after the distances between the clustered objects have been calculated, the selected clustering algorithm can be applied. On a general level, there are six main clustering methods (Aghabozorgi, Shirkhorshidi & Wah, 2015): (1) hierarchical clustering, (2) partitioning clustering, (3) model-based clustering, (4) grid-based clustering, (5) density-based clustering, and (6) multi-step clustering. Of these, hierarchical and partitioning clustering are currently the most commonly used ones (Xu & Wunsch, 2005). On a more specific level, each clustering method is realised by one or more clustering algorithms. For example, hierarchical clustering can be conducted by using either an agglomerative or divisive algorithm. In the agglomerative algorithm, each object first forms its own cluster, which are then merged together in a bottom-up fashion. In contrast, in the divisive algorithm, all the objects first form one large cluster, which is then split into smaller clusters in a top-down fashion. The weakness of both these algorithms is that after conducting a merge or split, the resulting clusters cannot be anymore adjusted. In turn, partitioning clustering first forms k number of clusters (the value of k is pre-assigned before conducting the clustering) and then groups the objects into them so that each cluster finally contains at least one object. The grouping is based on minimising the distance of the objects from the cluster centre. The two most common algorithms for partitioning clustering are k-means and k-medoids, which differ in their definition of the cluster centre. In k-means, the cluster centre is the cluster centroid (i.e., the average of all the clustered objects), whereas in k-medoids, the cluster centre is the cluster medoid (i.e., a prototype object near the cluster centre). Because of this difference, k-medoids is typically more robust against outliers than k-means. The most common realisation of k-medoids is the partitioning around medoids (PAM) algorithm (Kaufman & Rousseeuw, 1990).

Fourth, the goodness of the resulting clustering solution is evaluated statistically by using either external or internal indices as well as by considering its interpretability in terms of both theory and practice. External indices compare the resulting clustering solution to an ideal clustering solution, which is typically crafted by human experts. In contrast, internal indices evaluate the goodness of the resulting clustering solution based only on the information inherent in the clustered data and the clustering objectives. One of the most commonly used internal indices is the silhouette width (Rousseeuw, 1987), which evaluates both tightness or cohesion (i.e., how similar an object is to the objects in its own

cluster) and separation (how different an object is from the objects in the other clusters). The silhouette width varies from -1 to 1, in which negative values indicate a bad fit to a cluster and positive values indicate a good fit to a cluster. In addition to individual objects, the silhouette width can also be used to evaluate the goodness of individual clusters by calculating their average silhouette widths as well as the goodness of the whole clustering solution by calculating the overall average silhouette width.

3 Methodology

The data for this study was collected from 399 online stores, and it covers both their monthly sales time series from January 2016 to December 2017 as well as their key business and technical parameters at the end of January 2018. The data collection was conducted in co-operation with a Finnish electronic commerce platform provider, which uses a platform-as-a-service (PaaS) model to provide its customers an electronic commerce platform for operating the stores. The participating stores operated mainly in the Finnish market but in several product and service segments, such as clothing, jewellery and accessories, health and beauty, home and garden, sport and hobby, food and beverage, electronics, as well as automotive. All the stores had to fulfil three criteria in order to take part in the study. First, each store had to be first opened before January 2015. Second, each store had to be still open at the end of January 2018. Third, each store had to have at least one sales transaction per month. The first two criteria were intended to ensure that all the stores had been open for the whole two-year period from which the sales data was collected and that no store had opened just before or closed just after the period. In other words, all the stores had more or less an equal status in terms of being relatively mature stores, which had been in business for some time before and were not about to go out of business immediately after the two-year period. The third criterion was intended to ensure that there was an adequate amount of data to be analysed from each store.

In the sales time series, the sales were measured as both sales volume (the number of sales transactions) and sales value (the value of sales transactions). Of these, this study concentrates only on the sales measured as sales volume, but the implications of measuring the sales as sales value are addressed in the limitations. The median monthly sales of the stores ranged from a minimum of three transactions or 108 € to a maximum of 9,545 transactions or 927,476 €, with the median monthly sales of the whole sample being 34 transactions or 2,935 €. The key business and technical parameters of the stores covered their number of items, item variations, item categories, item images, campaigns, campaign items, banners, payment methods, shipping methods, customers, e-mail subscribers, and SMS subscribers. Of these, customers refer to the registered customers of the store, while the e-mail and SMS subscribers refer to the subscribers of the store newsletters via either electronic mail or short message service (SMS).

The time series clustering was conducted by using the *TSclust* version 1.2.4 and *cluster* version 2.0.6 packages of R (Montero & Vilar, 2014). When conducting the time series clustering, no representation method was used, but the sales time series were standardised

so that the sales of each store were measured as standard deviations from its mean sales. This way the differences in the average sales of the stores did not affect the clustering. As a distance metric, we used DTW because we were interested in similarities in terms of shape rather than similarities in terms of time. In turn, as a clustering algorithm, we used PAM. The statistical evaluation of the clustering solutions was based on the silhouette widths, which evaluate both tightness or cohesion and separation. Because not all the assumptions for parametric testing (e.g., normality and homoscedasticity) were met by the data, the statistical significance of the potential cluster differences were investigated by using the non-parametric Kruskal-Wallis (1952) tests and Dunn's (1961, 1964) tests with the Bonferroni correction for multiple testing. These were both conducted by using the *dunn.test* version 1.3.5 package of R.

4 Results

The results are reported in the following two sub-sections, of which the first focuses on finding the proper clustering solution and the second focuses on the potential cluster differences.

4.1 Clustering Solution

As mentioned above, the statistical evaluation of the clustering solutions and the selection of the suitable number of clusters (k) was based on the silhouette widths. The overall average silhouette widths of the clustering solutions with different values of k are reported in the silhouette plots presented in Figure 1. As can be seen, the overall average silhouette widths of the clustering solutions dropped sharply when the value of k exceeded four, suggesting that either a two-cluster, a three-cluster, or a four-cluster solution would be the most suitable one in strictly statistical terms. Of these, after considering also the interpretability of each solution, we decided to select the four-cluster solution because it was seen to give the best insight into the typical sales patterns of online stores in terms of not only the general trend in the sales but also the more specific seasonal sales variation. The silhouette plot of this four-cluster solution, which also reports both the size and the average silhouette width of each cluster, is presented in Figure 2. As can be seen, the sizes of all the four clusters were approximately equal. However, both the average silhouette widths of the clusters and the overall average silhouette width of the whole clustering solution remained relatively low, and all the clusters contained also objects with a negative silhouette width. This suggests that not all the clustered objects really fitted any of the four clusters. This cannot be seen as particularly surprising when considering that we were clustering real-life data in which the sales time series of individual stores were likely to vary considerably.

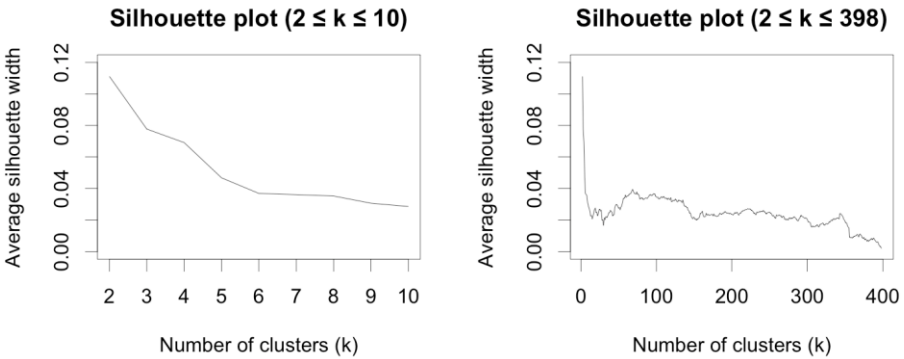

Figure 1: Overall average silhouette widths of the clustering solutions

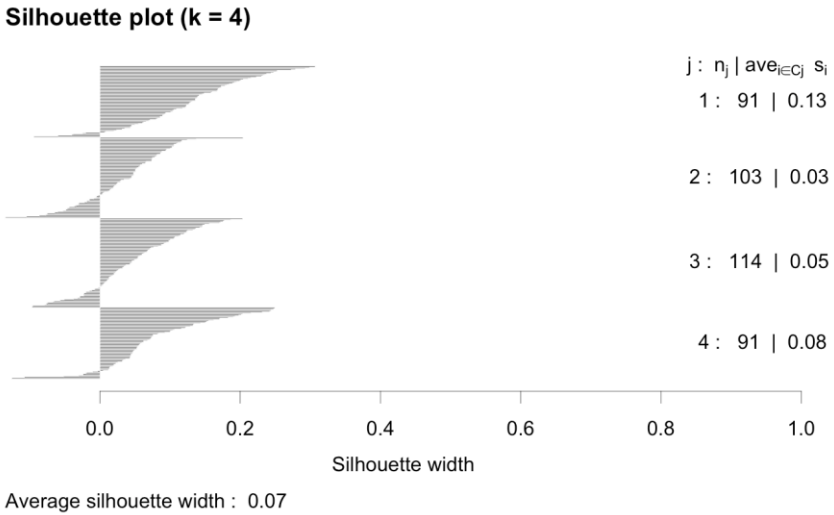

Figure 2: Silhouette widths and cluster sizes of the four-cluster solution

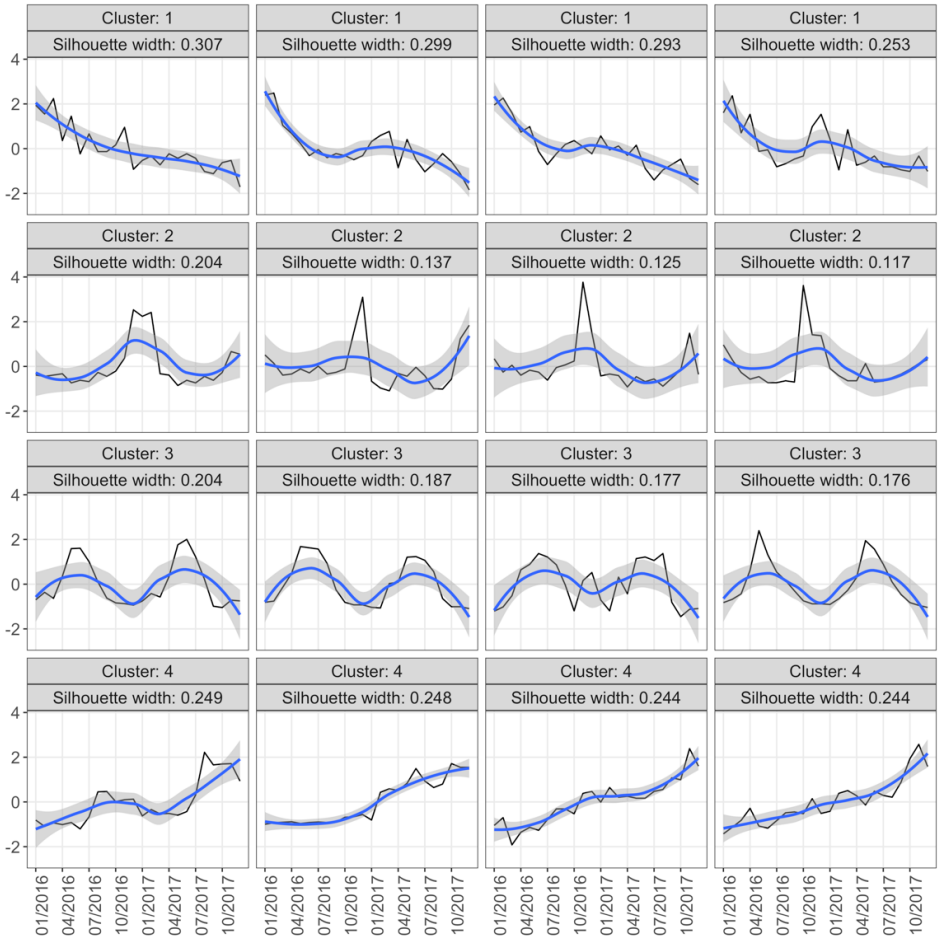

Figure 3: Time series selected as the cluster medoid and with the highest silhouette widths

Figure 3 illustrates the four-cluster solution by presenting the for each cluster the sales time series selected as the cluster medoid and three additional sales time series with the highest silhouette widths. These represent the most typical sales patterns of each cluster. In the case of the first cluster, the cluster medoid is the fourth sales time series with the lowest silhouette width, whereas in the case of the other three clusters, the cluster medoid is the first sales time series with the highest silhouette width. The black curve represents the standardised monthly sales. In order to smooth local variations, also a blue LOESS curve and its 95 % confidence interval are presented (Cleveland, 1979; Cleveland & Devlin, 1988). As can be seen, there are considerable differences between the four clusters. The typical sales time series in the first cluster have a clear downward trend, whereas the typical sales time series in the fourth cluster have a clear upward trend. In contrast, the typical sales time series in the remaining two clusters are characterised by strong seasonal variation. In the case of the second cluster, there seems to be a sharp sales

peak timed at around Christmas, whereas in the case of the third cluster, the sales peaks seem to be less sharp and timed at the summer months. However, most importantly, in both these latter cases, the sales time series lack a clear upward or downward trend.

4.2 Cluster Differences

Table 1 reports the medians of 14 store parameters for the whole sample and for each of the four clusters as well as the differences between the sample median and the cluster medians in percentages. The store parameters include the monthly sales measured as both sales volume (N) and sales value (€) as well as the 12 key business and technical parameters. Respectively, Table 2 reports the results of the Kruskal-Wallis tests and the Dunn's tests that were used to investigate the statistical significance of the differences in the cluster medians.

Table 1: Sample and cluster medians of the store parameters

Store parameter	All	Cluster 1		Cluster 2		Cluster 3		Cluster 4	
	Median	Median	Difference	Median	Difference	Median	Difference	Median	Difference
Monthly sales (N)	34	27	-21 %	30	-12 %	31	-9 %	68	+100 %
Monthly sales (€)	2,935	2,008	-32 %	1,937	-34 %	3,537	+21 %	6,230	+112 %
Items	329	225	-32 %	241	-27 %	425	+29 %	770	+134 %
Item variations	167	109	-35 %	152	-9 %	141	-16 %	214	+28 %
Item categories	43	33	-23 %	32	-26 %	47	+9 %	61	+42 %
Item images	679	539	-21 %	629	-7 %	614	-10 %	1,629	+140 %
Campaigns	6	5	-17 %	5	-17 %	5	-17 %	8	+33 %
Campaign items	101	62	-39 %	94	-7 %	94	-7 %	187	+85 %
Banners	10	7	-30 %	10	0 %	8	-20 %	23	+130 %
Payment methods	3	2	-33 %	3	0 %	3	0 %	3	0 %
Shipping methods	5	4	-20 %	6	+20 %	5	0 %	6	+20 %
Customers	494	485	-2 %	383	-22 %	429	-13 %	710	+44 %
E-mail subscribers	412	436	+6 %	358	-13 %	355	-14 %	646	+57 %
SMS subscribers	207	206	0 %	180	-13 %	189	-9 %	275	+33 %

Table 2: Results of the Kruskal-Wallis tests and the Dunn's tests (* = p < 0.001, ** = p < 0.01, * = p < 0.05, n.s. = statistically not significant)**

Store parameter	Kruskal-Wallis test			Dunn's tests					
	H	df	p	1 vs. 2	1 vs. 3	1 vs. 4	2 vs. 3	2 vs. 4	3 vs. 4
Monthly sales (N)	493.009	3	***	n.s.	n.s.	***	n.s.	***	***
Monthly sales (€)	656.341	3	***	n.s.	***	***	***	***	***
Items	18.921	3	***	n.s.	n.s.	***	n.s.	**	n.s.
Item variations	3.749	3	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Item categories	15.740	3	***	n.s.	n.s.	**	n.s.	**	n.s.
Item images	18.145	3	***	n.s.	n.s.	***	n.s.	*	*
Campaigns	7.839	3	*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Campaign items	9.828	3	*	n.s.	n.s.	*	n.s.	n.s.	n.s.
Banners	24.578	3	***	n.s.	n.s.	***	n.s.	**	**
Payment methods	7.226	3	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Shipping methods	13.509	3	***	**	n.s.	**	n.s.	n.s.	n.s.
Customers	9.031	3	*	n.s.	n.s.	n.s.	n.s.	n.s.	*
E-mail subscribers	9.843	3	*	n.s.	n.s.	n.s.	n.s.	n.s.	*
SMS subscribers	5.755	3	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

As can be seen, statistically significant differences between the clusters were found in terms of all the store parameters except for the number of item variations, payment methods, and SMS subscribers. Most of the differences concerned the fourth cluster, which was found to have higher monthly sales measured as both sales volume and sales value as well as a higher number of item images and banners in comparison to all the other clusters. The fourth cluster was also found to have a higher number of items and item categories in comparison to the first and second cluster, a higher number of campaign items and shipping methods in comparison to the first cluster, and a higher number of customers and e-mail subscribers in comparison to the third cluster. In addition, a few differences were found between the other three clusters, of which the second cluster was found to have a higher number of shipping methods in comparison to the first cluster, whereas the third cluster was found to have higher monthly sales measured as sales value in comparison to the first and second cluster.

5 Discussion and Conclusions

In this explorative study, we aimed at identifying the most typical sales patterns of online stores operating in the B2C context. By segmenting the monthly sales time series of 399 online stores with time series clustering, we were able to identify four approximately equally sized segments, each with its characteristic sales pattern. Two of the segments were characterised by a clear upward or downward trend in the sales, whereas the other two segments were characterised by strong seasonal sales variation. In one of the segments, the seasonal sales peaks lasted for about three or four months and were timed at the summer months, whereas in the other segment, they only lasted for about one or two months and were timed at around Christmas. However, most importantly, the two latter segments lacked a similar clear upward or downward trend in the sales that was found in the two former segments, meaning that the sales of the stores stayed about the same each year. Therefore, if the stores are aiming at increasing their sales in the long-

term, it seems that they are likely to benefit from strategies that decrease their short-term sales variation. The explanation for this finding is most likely linked to the differences between the segments in terms of risk perceptions and investment propensity, which are traditionally expected to be inversely related (Caballero, 1991). The stores with low seasonal sales variation obviously operate in a more certain business environment, which can be expected to have a positive effect on their investment propensity and result in increased sales if the return on investment is positive (or decreased sales if the return on investment is negative). In contrast, the stores with high seasonal sales variation obviously operate in a more uncertain business environment, which can be expected to have a negative effect on their investment propensity and hinder their growth. For example, whereas the stores with steady sales have the chance to both identify the potential changes in their economic environment and assess their return on investment throughout the year, the stores with seasonal sales can typically do this only during the on-peak months. During the off-peak months, they can often only speculate with the future while waiting and preparing for the next sales peak.

In addition, we investigated the potential differences between the online stores in each segment in terms of several store parameters. Here, especially the fourth segment was found to differ considerably from the other three segments. For example, the stores in the fourth segment were found to have higher monthly sales measured as both sales volume and sales value. Of course, this cannot be seen as particularly surprising when considering the clear upward trend in their sales. In addition, the stores in the fourth segment were found to be forerunners in terms of several other key business and technical parameters. These latter differences can be considered from two perspectives.

On one hand, the differences between the stores in the fourth segment as well as the stores in the second and third segment can be seen as providing insights on the business and technical parameters that are most closely associated with the reduction of seasonal sales variation. For example, the stores in the fourth segment were found to have a higher number of items and item categories in their selection, especially in comparison to the stores in the second segment. This can be expected to offer them more opportunities for diversification in terms of having in their selection also items with less seasonal sales patterns or at least items with different kinds of seasonal sales patterns. The outcomes of this diversification are likely to be seen also in the fact that the stores in the fourth segment were found to have more banners and item images on their site as well as more registered customers and e-mail subscribers, especially in comparison to the stores in the third segment. After all, a more diverse item selection is likely to attract more customers to register to the store and subscribe to its newsletters. Respectively, a more diverse item selection not only offers the store more opportunities but also requires it to have more item images on its site and use more banner advertisement in order to gain attention for the individual items in its vast selection.

On the other hand, the differences between the stores in the fourth segment and the stores in the first segment can be seen as providing insights on the business and technical parameters that are most closely associated with having an upward rather than a

downward trend in the sales. Many of these differences concerned the same parameters that were already discussed above. For example, the stores in the fourth segment were found to have more items and item categories in their selection as well as more item images and banners on their site also in comparison to the stores in the first segment. This would seem to suggest that these parameters are closely associated with not only reducing seasonal sales variation but also promoting the overall sales of the stores. This suggestion is supported by several prior studies. For example, the width and depth of the item selection have been found important especially from the perspective of capitalising the so-called "long tail" phenomenon (Anderson, 2006), which states that in many online stores, a considerable share of the sales comes not only from a few top items sold in high quantities but also from numerous niche items sold in low quantities individually, but in high quantities collectively. In turn, item images have been found important particularly in terms of promoting the flow of online shopping, the perceived usefulness of the stores, and the return intention to the stores through improved perceived diagnosticity (Jiang & Benbasat, 2004, 2007), whereas banner advertising has been found to positively affect purchase probabilities especially in the case of current customers (Manchanda et al., 2006). In addition, differences were found in the number of campaign items and shipping methods, which were also found to be higher in the stores belonging to the fourth segment in comparison to the stores belonging to the first segment. These can be seen to highlight especially the importance of active marketing as well as offering consumers a wide range of alternatives in terms of delivering the purchased items. Interestingly, the range of alternatives in terms of paying for the purchased items seemed not to be as important.

In addition to the aforementioned theoretical and practical insights concerning electronic commerce and online stores in particular, this study also contributes to IS research in general by providing an illustrative case example of a data analytics method for investigating the temporal dimension of various IS phenomena, which has often been overlooked in prior IS studies. For example, although time unarguably plays an important part in phenomena like technology acceptance and use, most prior studies, such as the numerous applications and extensions of the technology acceptance model (TAM) by Davis (1989) and the unified theory of acceptance and use of technology (UTAUT) by Venkatesh et al. (2003), have examined them only by using cross-sectional data collected at a single time point instead of longitudinal data collected at multiple time points. By analysing this kind of temporal data with time series clustering or other corresponding data analytics methods, more rigorous, relevant, and richer arguments can typically be made on the causalities concerning the phenomenon in question. For example, instead of treating IS usage as a simple dichotomous construct by dividing individuals into users and non-users, also the more complex differences in their temporal usage patterns can be considered. Therefore, backed by our positive experiences from this study, we strongly encourage their usage in future IS studies.

6 Limitations and Future Research

We consider this study to have three main limitations. First, the analysed sales time series covered only a two-year period and were collected only from online stores operating mainly in the Finnish market. Therefore, it is impossible to say anything about the sales patters that exceed this two-year period or how generalisable the sales patterns found in this study are to other markets. Even the external validity of our sample to the Finnish market remains difficult to assess because although the 399 online stores included both very small and very large stores in the Finnish scale, there does not exist any reference statistics on the total number of online stores operating in the Finnish market or on their distributions in terms of sales volume and sales value. Second, we were not able to obtain from the partnering electronic commerce platform provider longitudinal data on the key business and technical parameters from the two-year period but only cross-sectional data from January 2018. Therefore, we can only make associative rather than causal claims on the relationships between the store parameters and the store sales. Third, we were also not able to obtain from the partnering electronic commerce platform provider data on whether all the stores had been open for the whole two-year period or whether some of them had been temporarily closed, for example, due to maintenance. This data could have obviously been used as a valuable control variable in our study. However, although such temporary suspensions were possible, we do not believe that any of them were particularly long-lasting because all the stores participating in the study had to have at least one sales transaction per month.

Finally, it should be noted that in addition to clustering the sales time series in which the sales were measured as sales volume, we also replicated the clustering with the sales time series in which the sales were measured as sales value. This resulted in a three-cluster solution, in which the first three clusters were more or less identical to the ones of the aforementioned four-cluster solution, but the fourth cluster with the clear upward sales trend was missing. This would seem to suggest that even though some stores were able to grow their sales in terms of sales volume, this did not necessarily result in a similar growth in sales in terms of sales value, potentially due to the price erosion associated with the growing sales volumes. Whether such price erosion actually exists and how severe a problem it is for online stores, would be an interesting and important topic of future studies.

Another potential path of future research would be to concentrate more precisely on the causal explanations for the sales patterns found in this study through a more thorough control of the store parameters. These parameters do not have to be strictly quantitative, as they were in this study, but they can also be more qualitative, such as assessments of the usability and aesthetics of the stores as well as information on the primary product or service segment in which the stores operate and their level of diversification. In addition, instead of investigating the sales patterns within one common time period, one could also investigate them within a certain time period since the opening of each store (e.g., the sales of the first year).

References

Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering – A decade review. *Information Systems*, *53*, 16–38. doi: 10.1016/j.is.2015.04.007

Allred, C. R., Smith, S. M., & Swinyard, W. R. (2006). E-shopping lovers and fearful conservatives: a market segmentation analysis. *International Journal of Retail & Distribution Management*, *34*(4/5), 308–333. doi: 10.1108/09590550610660251

Anderson, C. (2006). *The Long Tail: Why the Future of Business is Selling Less of More*. New York, NY: Hyperion.

Barnes, S. J., Bauer, H. H., Neumann, M. M., & Huber, F. (2007). Segmenting cyberspace: a customer typology for the internet. *European Journal of Marketing*, *41*(1/2), 71–93. doi: 10.1108/
03090560710718120

Bhatnagar, A., & Ghose, S. (2004a). A latent class segmentation analysis of e-shoppers. *Journal of Business Research*, *57*(7), 758–767. doi: 10.1016/S0148-2963(02)00357-0

Bhatnagar, A., & Ghose, S. (2004b). Segmenting consumers based on the benefits and risks of Internet shopping. *Journal of Business Research*, *57*(12), 1352–1360. doi: 10.1016/S0148-2963(03)
00067-5

Brengman, M., Geuens, M., Weijters, B., Smith, S. M., & Swinyard, W. R. (2005). Segmenting Internet shoppers based on their Web-usage-related lifestyle: a cross-cultural validation. *Journal of Business Research*, *58*(1), 79–88. doi: 10.1016/S0148-2963(02)00476-9

Brown, M., Pope, N., & Voges, K. (2003). Buying or browsing? An exploration of shopping orientations and online purchase intention. *European Journal of Marketing*, *37*(11/12), 1666–1684. doi: 10.1108/03090560310495401

Caballero, R. J. (1991). On the Sign of the Investment–Uncertainty Relationship. *American Economic Review*, *81*(1), 279–288.

Carman, H. F., & Figueroa, E. E. (1986). An Analysis of Factors Associated with Weekly Food Store Sales Variation. *Agribusiness*, *2*(3), 375–390. doi: 10.1002/1520-6297(198623)2:3<375::AID-AGR2720020308>3.0.CO;2-I

Chang, M. K., Cheung, W., & Lai, V. S. (2005). Literature derived reference models for the adoption of online shopping. *Information & Management*, *42*(4), 543–559. doi: 10.1016/j.im.2004.02.006

Cheung, C. M. K., Chan, G. W. W., & Limayem, M. (2005). A Critical Review of Online Consumer Behavior: Empirical Research. *Journal of Electronic Commerce in Organizations*, *3*(4), 1–19. doi: 10.4018/jeco.2005100101

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836. doi: 10.2307/2286407

Cleveland, W. S., & Devlin, S. J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, *83*(403), 596–610. doi: 10.2307/2289282

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. doi: 10.2307/249008

Dunn, O. J. (1961). Multiple Comparisons Among Means. *Journal of the American Statistical Association*, *56*(293), 52–64. doi: 10.2307/2282330

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics*, *6*(3), 241–252. doi: 10.2307/1266041

Fu, T.-c. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence, 24*(1), 164–181. doi: 10.1016/j.engappai.2010.09.007

Geurts, M. D. (1988). The Impact of Misrepresentative Data Patterns on Sales Forecasting Accuracy. *Journal of the Academy of Marketing Science*, *16*(3–4), 88–94. doi: 10.1007/BF02723364

Jiang, Z., & Benbasat, I. (2004). Virtual Product Experience: Effects of Visual and Functional Control of Products on Perceived Diagnosticity and Flow in Electronic Shopping. *Journal of Management Information Systems*, *21*(3), 111–147. doi: 10.1080/07421222.2004.11045817

Jiang, Z., & Benbasat, I. (2007). Investigating the Influence of the Functional Mechanisms of Online Product Presentations. *Information Systems Research*, *18*(4), 454–470. doi: 10.1287/isre.1070.
0124

Kau, A. K., Tang, Y. E., & Ghose, S. (2003). Typology of online shoppers. *Journal of Consumer Marketing*, *20*(2), 139–156. doi: 10.1108/07363760310464604

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ: Wiley.

Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. doi: 10.2307/2280779

Liao, T. W. (2005). Clustering of time series data – a survey. *Pattern Recognition*, *38*(11), 1857–1874. doi: 10.1016/j.patcog.2005.01.025

Manchanda, P., Dubé, J.-P., Goh, K. Y., & Chintagunta, P. K. (2006). The Effect of Banner Advertising on Internet Purchasing. *Journal of Marketing Research*, *43*(1), 98–108. doi: 10.1509/jmkr.43.1.
98

Moe, W. W. (2003). Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream. *Journal of Consumer Psychology*, *13*(1–2), 29–39. doi: 10.1207/S15327663JCP13-1&2_03

Montero, P., & Vilar, J. A. (2014). TSclust: An R Package for Time Series Clustering. *Journal of Statistical Software*, *62*(1), 1–43. doi: 10.18637/jss.v062.i01

Perea y Monsuwé, T., Dellaert, B. G. C., & de Ruyter, K. (2004). What drives consumers to shop online? A literature review. *International Journal of Service Industry Management*, *15*(1), 102–121. doi: 10.1108/09564230410523358

Rohm, A. J., & Swaminathan, V. (2004). A typology of online shoppers based on shopping motivations. *Journal of Business Research*, *57*(7), 748–757. doi: 10.1016/S0148-2963(02)00351-X

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. doi: 10.1016/0377-0427(87)90125-7

Soopramanien, D. G. R., & Robertson, A. (2007). Adoption and usage of online shopping: An empirical analysis of the characteristics of "buyers", "browsers" and "non-internet shoppers". *Journal of Retailing and Consumer Services*, *14*(1), 73–82. doi: 10.1016/j.jretconser.2006.04.002

Su, Q., & Chen, L. (2015). A method for discovering clusters of e-commerce interest patterns using click-stream data. *Electronic Commerce Research and Applications*, *14*(1), 1–13. doi: 10.1016/j.elerap.
2014.10.002

Swinyard, W. R., & Smith, S. M. (2003). Why People (Don't) Shop Online: A Lifestyle Study of the Internet Consumer. *Psychology & Marketing*, *20*(7), 567–597. doi: 10.1002/mar.10087

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478. doi: 10.2307/30036540

Xu, R., & Wunsch, D. (2005). Survey of Clustering Algorithms. *IEEE Transactions on Neural Networks*, *16*(3), 645–678. doi: 10.1109/TNN.2005.845141