Mateusz Bąska

https://orcid.org/0000-0002-4465-7747

Department of Information Technology
Faculty of Management, Computer
Science & Finance
Wroclaw University of Economics,
Wroclaw, Poland
mateusz.baska@ue.wroc.pl

Helena Dudycz

https://orcid.org/0000-0002-4304-6207

Department of Information Technology
Faculty of Management, Computer
Science & Finance
Wroclaw University of Economics,
Wroclaw, Poland
helena.dudycz@ue.wroc.pl

Maciej Pondel

https://orcid.org/0000-0002-1978-6571

Department of Business Intelligence
in Management
Faculty of Management,
Wroclaw University of Economics,
Wroclaw, Poland
maciej.pondel@ue.wroc.pl

Identification of advanced data analysis in marketing: A systematic literature review

Abstract

Aim/purpose – Marketing is an important area of activity for the vast majority of enterprises. Many of them try using marketing data analysis. Both the literature and the practice of many enterprises describe the use of advanced data analysis. However, interpretations of this concept differ. The aim of this paper is to identify the interpretation of advanced data analysis in marketing, in support of decision-making processes applied in the retail trading sector.

Design/methodology/approach – The study was conducted using a systematic literature review, suggested by B. Kitchenham (2004), extended by C. Wohlin & R. Prikladniki (2013). This method was modified and expanded through the division of the whole study into two phases. Each phase is intended to facilitate obtaining answers to different im-

portant research questions. The first phase constitutes an exploratory study, whose results allow the detailed analysis of the literature in the second phase of the study.

Findings – The results of this study of the relevant literature indicate that scholarly publications do not use the phrase 'advanced data analysis', and its context is described with the term 'data analysis'. Another term used broadly within the sphere of data analysis is 'big data'. The concept of 'data analysis' in marketing is focused around the term 'big data analytics' and terms linked to the word 'customer', such as 'customer-centric', 'customer engagement', 'customer experience', 'customer targeting service', and 'customers classification'. The study of the literature undertaken indicates that marketing employs data analysis in such areas as customer needs identification and market segmentation.

Research implications/limitations – The study of the literature review was carried out using selected four databases containing publications, i.e. Web of Science, IEEE, Springer and ACM for the period 2008 to 2018. The research described in the article can be continued in two ways. First, by analysing the literature presented in this paper on advanced data analysis in marketing using the method called snowball sampling. Secondly, the results obtained from the first stage of the study can be used to conduct the study with other databases.

Originality/value/contribution – The main contribution of this work is the proposal of modifying the systematic literature review method, which was expanded through the introduction of two phases. This division of two stages is important for conducting studies of literature when there are no clear, established definitions for the concepts being employed. The result of the study is also a set of ordered terms and their meanings that clearly define advanced data analysis in marketing.

Keywords: advanced data analysis, marketing, systematic review, big data analytics.
JEL Classification: D80, M31, Y9.

1. Introduction

Companies take advantage of advanced data analysis in many fields of activity. Depending on the field, this notion is defined in different manners. What is more, both the performance of data analysis and the tools and methods applied can differ. Advanced data analysis involves analysis carried out using a spreadsheet in MS Excel (functionality of pivot tables), but also Business Intelligence Systems (Sharda, Delen, & Turban, 2014), Decision Support Systems, as well as using methods of machine learning and data mining, deep learning, etc.

One of the primary fields of enterprises' activity is marketing, which can employ advanced data analysis in building customer relationships, thereby contributing to improved financial results. The key question that must be asked is: how can advanced data analysis be used in marketing to enhance a company's growth? In order to examine a range of issues related to advanced data analysis

in marketing, a research method has been applied which facilitates verification of the meaning of the term 'advanced data analysis'.

When analysing the current studies of the systematic review of the literature on data analysis in marketing, it can be concluded that they focus mainly on the Big Data area (Elgendy & Elragal, 2014), and the challenges and directions of data analysis research (Zhao, Fan, & Hu, 2014). In research in the field of marketing, researchers are more and more often paying attention to the use of various data sets for data analysis (Bucklin & Sisemeiro, 2009) and the use of Big Data on the e-commerce market (Akter & Wamba, 2015). The authors of the article have not found publications containing a systematic literature review investigating the use of advanced data analysis in marketing.

The study was conducted using a systematic literature review, suggested by Kitchenham (2004), and a manual search, proposed by Wohlin & Prikladniki (2013). The systematic literature review is a defined, methodical means of identifying, assessing, and analysing published studies in order to explore a specific research question (Kitchenham & Charters, 2007). It is distinct from informal literature reviews because it is planned and carried out following a predefined research procedure. Such an approach can generate better results and determinations than an informal literature review (Niazi, 2015).

The objective of this paper is to identify occurrences of the term 'advanced data analysis' in collections qualifying as 'big data' in marketing, done in support of decision-making processes operating in the retail trading sector. The result of the study is a set of ordered terms and their meanings.

The paper is structured as follows: the next section focuses on issues related to advanced data analysis in marketing. In the succeeding section, the research methodology is discussed. Following that, we present the assumptions of our research and the obtained results. Finally, in the last section, some conclusions are drawn and future directions for advanced data analysis in marketing are considered.

2. Research context

The term advanced data analysis, corresponding also to the expressions: 'data exploration' and 'data mining' is defined as a process of automatic detection of non-trivial, unknown, and potentially useful relationships, rules, patterns, similarities, or trends in large datasets (Witten, Frank, Hall, & Pal, 2017). Generally speaking, the task of advanced data analysis is to process data in order to

better understand and use it in decision-making processes. Data exploration or mining is a multi-disciplinary area that integrates a range of research fields such as information systems, databases and data warehouses, statistics, artificial intelligence, parallel computing, operational research, visualisation, and computer graphics. Exploration systems use a broad range of information and communication technologies, Web technologies, information retrieval methods, and geolocation techniques, as well as signal processing and bioinformatics (Pondel & Korczak, 2017).

Marketing (especially digital marketing) is a rich field where advanced data analysis is applied. This fact is caused by the following factors:
1. Variety of marketing data sources (transactions, website visits, social media, reactions to marketing messages, and many more).
2. Ability to trace each move of the customer in e-commerce systems.
3. Many channels of interaction with the customer.

In digital marketing, we can also relatively easily apply results of advanced analytical models and test their efficiency in real marketing communication with customers. Advanced data analysis in marketing can be applied to understand the drivers leading customers to a purchase decision, the preferences of individual customers and their behaviour, but also to forecast future transactions. Such an approach is very useful in the process of targeting individual offers to customers, optimizing the cost of marketing, or defining marketing campaigns that are significantly more efficient than regular ones.

The results of data analysis are used to provide a retailer or a client with a real-time recommendation of a product purchase, discount or marketing activity in order to maximise a selected customer experience factor (purchase probability, customer satisfaction, customer retention risk, product margin).

Advanced data analysis and exploration of large marketing databases are the subjects of many research and application projects (Gordon, Linoff, & Berry, 2011; Han, Pei, & Kamber, 2012; Pawełoszek & Korczak, 2017, Pondel & Korczak, 2017; Setia & Jyoti, 2013). The knowledge acquired can be used automatically in the processes of communication with the client including operations undertaken in order to increase the chance of making a purchase, improve customer satisfaction, reduce the risk of a customer leaving, or optimise the margin on the product.

In recent years, under the influence of IT development, social networks, and artificial intelligence methods, the scope of advanced analysis in marketing has significantly expanded. Today we are able to build recommendation systems that

allow one to determine the rank of a product or preferences that the customer should assign to a given product or group of products (Konstan & Adomavicius, 2013).

In relevant literature (Jannach, Zanker, Felfernig, & Friedrich, 2010; Ricci, Rokach, & Shapira, 2015), application of advanced data analysis in marketing is considered in three main perspectives. From the managerial perspective, we use it to build a decision support system that uses large, heterogeneous data and mechanisms generating recommendations related to the sales strategy and promotion of the products offered. From the customer's perspective, we prepare an advisory system facilitating selection of products in accordance with one's interests, needs, and preferences. From an IT perspective, advanced data analysis is related to a computing platform containing a number of exploration models, integrated with transactional systems of the online and offline store and its environment. This platform must guarantee not only access to various information resources but also scalability of applications operating on a large number of information collections.

The above-mentioned factors proving a high potential of advanced data analysis application in marketing encouraged the authors to conduct literature research in order to identify and organise terms and expressions related to advanced data analysis concentrating on the marketing research field.

3. Research methodology

3.1. A systematic literature review

A systematic literature review is a defined and methodical means of identifying, assessing, and analysing primary studies in order to explore a specific research question (Kitchenham, 2004). It can also help uncover the structure and models of existing studies, and thus identify gaps that could be filled in by scientific research. Systematic reviews are distinct from informal literature reviews owing to their formal planning and methodical execution. This method consists in planning, execution, and description and presentation of results obtained. A good systematic review should be independently repeatable, which thus gives it a different scientific value from that of informal literature reviews. Primary research studies included in the review are examined on the basis of their fulfilment of methodological quality criteria, in order to undertake joint consideration of their possible results.

A systematic literature review is characterised by transparent procedures in order to identify, assess, and synthesise the results of studies associated with the research questions formulated. A research procedure determined in advance ensures transparency and repeatability of the conducted study, while also minimizing bias.

A systematic literature review must include the following:
− clearly formulated criteria for the inclusion and exclusion of particular studies,
− a clearly described strategy for seeking out relevant literature (i.e. research reports),
− systematic coding and analysis of the quality of the included studies,
− synthesis of results of quantitative and/or qualitative studies.

The research procedure of informal literature reviews does not include defining the criteria for inclusion/exclusion of initial results in further analysis. Systematic literature reviews also take into account the assessment of the quality of the results obtained (Niazi, 2015).

In B. Kitchenham's (2004) opinion, the following characteristics can be distinguished that make a systematic review distinct from a traditional (informal) literature review:

1. A defined and documented research procedure drafted for the systematic literature review prior to commencing the study, with a view to determining research questions and the actions to be performed.
2. Definition and documentation of search strategies within the research procedure in order to locate the largest possible volume of literature relevant to the research area.
3. A description of the criteria for inclusion and exclusion within the research procedure, which should be applied in assessing each result with regards to its further examination.
4. A description of quality assessment mechanisms employed within the procedure, with a view to assessing each study result (i.e. of a given piece of subject literature).

Procedures and systems for systematic literature reviews are well-established in other fields, particularly in medicine (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). Studies with systematic literature reviews have also been conducted in the field of Information and Software Technology, for example, software engineering (Kitchenham, Pretorius, Budgen, Brereton, Turner, Niazi, & Linkman, 2010), software testing (Vahid & Mäntyläc, 2016), and application usability (Quiñones & Rusu, 2017).

3.2. Research questions

The objective of the literature analysis is to identify and order terminology associated with advanced data analysis in the field of marketing with a view to supporting decision-making processes in retail enterprises. The study was divided into two phases, as it was first necessary to define the concept of advanced data analysis. Only after that were we able to proceed to the examination of the possibilities for the use of data analysis in marketing to support decision-making processes in retail enterprises.

Each phase of the study has its own research questions. The first phase of the investigations was focused on identification of the term 'advanced data analysis' through locating answers to the following research questions:

- (RQ 1.1) How is the concept of advanced data analysis defined in the relevant subject literature?
- (RQ 1.2) What keywords are present in the subject literature accompanying the definition of advanced data analysis?

During the second phase of the study, answers were sought to the following research questions:

- (RQ 2.1) What keywords are present in the subject literature accompanying the definition of advanced data analysis within the context of marketing?
- (RQ 2.2) Is advanced data analysis used in the field of marketing in the context of supporting decision-making processes in retail trading enterprises?
- (RQ 2.3) What problems are solved with the use of advanced data analysis in marketing?

The next subsection will present the research procedure applied in the study.

3.3. Study procedure

The research procedure was divided into two phases due to the research questions that in phase 1 were related to the identification of the term advanced analysis of data and keywords occurring together with this term. Phase 1 results will be used in phase 2 for the construction of the search strategy for this phase. The following research procedure was applied in the study:

Phase 1:

1. Defining sources of literature analysis.
2. Defining the strategy and scope of the search.

3. Defining criteria for the selection of scientific papers to be subjected to detailed analysis.
4. Extraction of scientific papers for detailed analysis.
5. Obtaining answers to the formulated research questions (RQ1.1 and RQ1.2).
6. Analysis of the obtained results and formulation of conclusions from the first phase of the study.

Phase 2:

1. Defining sources of literature analysis.
2. Defining a search strategy and scope.
3. Developing criteria for the selection of scientific literature for detailed analysis, including criteria for the inclusion/exclusion of papers and books in the study.
4. Searching for and analysis of scientific literature describing studies aligned with the assumed selection criteria.
5. Acquiring answers to the research questions (RQ2.1, RQ2.2, and RQ2.3).
6. Formulating conclusions from the second research phase.
7. Formulating conclusions from the conducted study.

The study contains two distinct but linked phases, for which a joint systematic review protocol was adopted. The research procedure for each phase is similar; differences have been detailed in particular research questions.

4. Research findings

4.1. Findings of 1st phase of study

The first phase began with defining the databases to be used in the literature review. Initial analyses of available databases were carried out in terms of the number of results obtained for the same search phrase. Based on the results obtained, analysis of the subject matter of database articles, and the possibility of using advanced search, the following databases were selected for the study:

- Web of Science – http://apps.webofknowledge.com,
- IEEE – http://ieeexplore.ieee.org,
- Springer – https://link.springer.com,
- ACM – https://dl.acm.org.

Next, a strategy was developed for searching for scholarly papers and books of relevance to the research questions posed in the first phase. The following strategy was applied in constructing search terms:

1. Determining keywords based on the formulated research questions.
2. Examination of keywords in a given publication referring to the subject matter of the research.
3. Use of the logical operator AND in the search engine to construct a search term that could be applied in databases allowing the use of such a character string with a logical operator.

It was stated on the basis of the formulated research questions that the chain of terms employed in the search would contain the following terms:
− advanced data analysis,
− advanced analysis,
− data analysis,
− advanced analytics.

All of these concepts were entered into the search engine as fully-combined phrases, e.g. 'advanced data analysis', and in the separated form, e.g. 'advanced' AND 'data' AND 'analysis' (Table 1).

It was determined that papers and books written in English, published in the period from 2008 to 2018 would be examined and that the search would be performed in an automated manner.

The study sets out to analyse the latest literature from the last ten years. The resulting set of articles on the subject under study will allow us to continue research and obtain the most important publications from earlier years by using the snowball technique to analyse literature.

Table 1. Sample queries to databases containing publications for phase 1

Database	Combined phrase	Individual concepts
Springer Link	"advanced+data+analysis"	advanced+AND+data+AND+analysis
Web of Science	TS=("advanced data analysis")	TS=(advanced data analysis)
IEEE Explore	(.QT.advanced data analysis.QT.)	(advanced data analysis)
ACM DL	("advanced data analysis")	(advanced data analysis)

The next step in the study was to develop the following criteria for the selection of scholarly works to undergo detailed analysis:
1. The selected concept is searched for as an entire phrase, or as a connection of the words contained within it.
2. The concept must be present in the title or abstract of the paper/book.
3. The top five results proceeded to the next phase (sorted by the number of citations).

In the next step, we applied criteria for exclusion and inclusion of papers and books in further study. The criteria for the exclusion of publications for detailed analysis were as follows:

1. Papers/books containing proposed solutions for other fields (e.g. life sciences).
2. Studies that do not indicate how a given solution was applied.
3. Incomplete papers, such as summaries and expanded summaries.
4. Papers less than four pages in length published as ShortPapers.
5. Ambiguous and irrelevant papers were also excluded after the search.

Results from the search for defined character strings in selected databases allowed us to compile an initial collection of 11 works (Table 2).

Table 2. Collection of results obtained from an automated search of selected databases with publications in phase 1

Source	Found	Initial selection	Final selection
Springer Link	5,173,608	30	0
Web of Science	358,696	40	5
IEEE Explore	27,694	30	3
ACM DL	41,842	15	3
Total	**5,601,840**	**113**	**11**

Qualitative analysis was then performed to serve as the basis for deciding which publications should be included in further work. With consideration to the purpose of the literature analysis for phase 1, the following criteria were adopted:

1. The publication contains detailed information on the subject of data analysis.
2. The publication includes a detailed description of areas associated with advanced data analysis and contains sufficient information to define and describe that concept.
3. The publication contains studies relevant to the research questions as posed.

The obtained compilation of publications was then analysed, and the following information was extracted:

− authors,
− publication date,
− title,
− keywords.

A total of 11 works to be used in the review of publications, relevant to the purposes of the study and the research questions formulated within it, were identified. After the detailed analysis phase was completed, 4 works were excluded from further analysis owing to the absence of a link with the research field (Table 3).

Table 3. Compilation of publications after the process of extraction in phase 1 was conducted

Author	Date	Title	Keywords
A. Vera-Baquero, R. Colomo-Palacios, & O. Molloy	2013	Business process analytics using a big data approach	business analytics, business process analytics, business process management, big data, information technology
K. Slavakis, G.B. Giannakis, & G, Mateos	2014	Modeling and optimization for big data analytics: (statistical) learning tools for our era of data deluge	big data, compressed sensing, data analysis, learning (artificial intelligence), optimisation, principal component analysis, signal sampling
H. Hu, Y. Wen, T. S. Chua, & X. Li	2014	Toward scalable systems for big data analytics: A technology tutorial	big data, data acquisition, data analysis, data communication, public domain software, storage management
M. Chen, S. Mao, & Y. Liu	2014	Big data: A survey. Mobile networks and application	big data; cloud computing; internet of things; data center; Hadoop; smart grid; data analysis
N. Elgendy & A. Elragal	2014	Big data analytics: A literature review paper	big data; data mining; analytics; decision making
A. Pavlo, E. Paulson, A. Rasin, D. J. Abadi, D. J. DeWitt, S. Madden, & M. Stonebraker	2009	A comparison of approaches to large-scale data analysis	database applications, use cases, database programming
C. P. Chen & C. Y. Zhang	2014	Data-intensive applications, challenges, techniques and technologies: A survey on big data	big data; data-intensive computing; e-science; parallel and distributed computing; cloud computing

The analysis of papers (Table 3) was used to create a list of the frequency with which keywords appear (Table 4).

Table 4. The frequency of keyword appearance – phase 1

Keywords	Frequency in review results	
	Number of appearances	%
1	*2*	*3*
Analytics	1	14
Big data	**5**	**71**
Cloud computing	2	29
Compressed sensing	1	14
Data acquisition	1	14
Data analysis	**3**	**43**
Data centre	1	14
Data communication	1	14
Data mining	1	14
Data-intensive computing	1	14
Database applications	1	14
Database programming	1	14
Decision making	1	14

Table 4 cont.

1	2	3
e-Science	1	14
Hadoop	1	14
Internet of things	1	14
Learning (artificial intelligence)	1	14
Parallel and distributed computing	1	14
Principal component analysis	1	14
Public domain software	1	14
Signal sampling	1	14
Smart grid	1	14
Storage management	1	14
Use cases	1	14

The study led to the determination that the notion of advanced data analysis was not present. No term in this form appeared in the obtained results. The literature does not know a collection of concepts and studies revolving around the term 'advanced data analysis'. As a result, the concept 'data analysis' was applied without the adjective 'advanced' during the second phase of the study.

The results obtained in the first phase of the conducted study also indicate that 'big data' and 'data analysis' were the most frequently occurring keywords in the literature with the adopted criteria for literature analysis. These concepts will be used in the second phase of the study in order to construct search phrases for databases with publications in conjunction with the term of 'marketing'.

4.2. Findings of 2nd phase of study

The second phase of the study was conducted in accordance with the defined research procedure. First, the databases to be used for in-depth literature analysis were defined. We used the same databases as in the first phase.

Next, a strategy was developed for searching for scientific papers and books of relevance to the posed research questions within the second phase of the study. The obtained results from the first phase allowed us to determine that the search strings to be employed would include the terms 'data analysis' and 'big data'. The authors of the study decided to add the term 'decision support systems', which specifies the field of support for decision-making processes within the conducted study. All three terms were paired with the concepts 'marketing' and 'retailing', which are relevant to this study in its second phase. The paired concepts are as follows:

− marketing AND data analysis,
− marketing AND big data,

‒ marketing AND decision support systems,
‒ retailing AND data analysis,
‒ retailing AND big data,
‒ retailing AND decision support systems.

The next step in the study was to develop criteria for the selection of scientific works for detailed analysis, including criteria for the inclusion and exclusion of papers and books for the remainder of the study. The following criteria for publication selection were established:

1. The selected concept is searched for as a complete phrase or as a connection of the words contained within it.
2. The concept must be present in either the title or abstract of the paper/book.
3. Publications were sorted by the number of citations.
4. Only results in the form of papers and books were sought.

The following criteria were applied in the study for the inclusion of publications for detailed analysis:

1. Papers and books containing a description of advanced data analysis in marketing.
2. Papers and books containing a description of problems and difficulties arising with the use of data analysis in marketing.
3. The first 5 publications listed in search results, if their subject matter did not diverge from the narrowed research area.

The following criteria were applied in the exclusion of publications from detailed analysis:

1. Papers and books containing a description of an application of data analysis that did not concern an enterprise, particularly the field of marketing.
2. Papers and books that do not provide details on the scope and/or manner of application of data analysis in marketing.
3. Incomplete papers, such as summaries and expanded summaries.
4. Papers presenting opinions without empirical examples.

After establishing assumptions and search criteria, publications relevant to the research questions were acquired in accordance with the strategy that had been developed. First, analysis of a paper or book was based on the title and an analysis of the abstract, rejecting those which were clearly of insignificance for the literature review being conducted. If these elements gave rise to doubts as to the inclusion/exclusion of a work (with regard to the formulated criteria), the paper/book proceeded to the next step of the analysis. Next, selected works were reviewed by way of reading the introduction, the section detailing the research

method and procedure, and conclusions. Criteria for inclusion/exclusion in the next part of the study were applied to the information contained in those parts of the publication. If it was not sufficient to read those parts of a given work in order to make such a determination, the work was read in full to determine whether it was or was not relevant to the objective of the study. In this manner, a list of scholarly works containing studies of importance to the posed research questions was obtained.

After the application of the search string (i.e. the developed search criteria) in selected databases and a detailed analysis, we obtained the number of publications relevant to the posed research questions (Table 5).

Table 5. Compilation of results obtained following an automatic search of selected databases with publications within phase 2

Source	Found	Initial selection	Final selection
Springer Link	4,777,448	30	2
Web of Science	621,096	30	6
IEEE Explore	67,694	30	3
ACM DL	21,392	20	1
Total	**5,487,630**	**110**	**12**

The collection of papers and books was then analysed, leading to the same information being compiled (Table 6) as in the first phase (Table 3).

Table 6. Compilation of publications after the extraction process within phase 2 was conducted

Author	Date	Title	Keywords
1	*2*	*3*	*4*
P. S. H. Leeflang, P. C. Verhoef, P. Dahlström, & T. Freundt	2014	Challenges and solutions for marketing in a digital era	online marketing; big data; analytics; customers; firm strategies; social media
S. Akter & S. F. Wamba	2016	Big data analytics in e-commerce: A systematic review and agenda for future research	big data analytics; e-commerce; business value
S. K. Fan, R. Y. K. Lau, & J. L Zhao	2015	Demystifying big data analytics for Business Intelligence through the lens of the marketing mix	big data analytics; Business Intelligence; marketing intelligence; marketing mix; survey versus log data
N. R.Sanders	2016	How to use big data to drive your supply chain	decision making; supply chain; technology
D. Kridel & D. Dolk	2013	Automated self-service modelling: Predictive analytics as a service	customer targeting service; predictive analytics; service-oriented architecture; model management; automated modelling
D. Zhao	2013	Frontiers of big data business analytics: Patterns and cases in online marketing	

Table 6 cont.

1	2	3	4
B. J. Calder, E. C. Malthouse, & E. Maslowska	2016	Brand marketing, big data and social innovation as future research directions for engagement	customer engagement; consumer
B. Grandhi, N. Patwa, & K. Saleem	2017	Data-driven marketing for growth and profitability	data-driven marketing; customer-centric; value proposition; digital data; marketing metrics; marketing analytics; marketing audit; marketing dashboard; digital marketing and marketing investment; ROMI (Return On Marketing Investment); diffusion of innovation
R. Yeung & W. Yee	2012	Application of cluster analysis and discriminant analysis in market segmentation and prediction	customers classification; group membership; cross-border shopping; motivation; attitudes
H. Aguinis, L. E. Forcum, & H. Joo	2012	Using market basket analysis in management research	research design, methodology, data analysis
A. Pabedinskaitė, V. Davidavičienė, & P. Milišauskas	2014	Big data driven e-commerce marketing	e-commerce; marketing; big data;
B. Pierański & S. Strykowski	2017	Towards a personalized virtual customer experience	personalization; virtual retailing; virtual reality; customer experience

A total of 12 papers were obtained for use in the detailed analysis of publications relevant to the objective of the study and the formulated research questions. After the detailed analysis stage of the extracted literature was completed, 6 works were rejected owing to a lack of relevance to the research area. These works are presented in Table 6.

Analysis of the 12 papers led to the creation of a list detailing the frequency with which keywords appeared (Table 7).

Table 7. The frequency of keywords occurrence – phase 2

Keywords	Frequency of appearance	
	Number of appearances	%
1	2	3
Big data	**2**	**5.0**
Big data analytics	**2**	**5.0**
Business Intelligence	1	2.5
Business value	1	2.5
Consumer	1	2.5
Customers	1	2.5
Customer centric	1	2.5
Customers classification	1	2.5
Customer engagement	1	2.5

Table 7 cont.

1	2	3
Customer experience	1	2.5
Customer targeting service	1	2.5
Data analysis	1	2.5
Data-driven marketing	1	2.5
Decision making	1	2.5
Digital data	1	2.5
Digital marketing and marketing investment	1	2.5
e-commerce	**2**	**5.0**
Marketing	1	2.5
Marketing analytics	1	2.5
Marketing audit	1	2.5
Marketing dashboard	1	2.5
Marketing intelligence	1	2.5
Marketing metrics	1	2.5
Marketing mix	1	2.5
Methodology	**2**	**5.0**
Model management	1	2.5
Motivation	1	2.5
Online marketing	1	2.5
Personalization	1	2.5
Predictive analytics	1	2.5
Research design	1	2.5
ROMI (Return On Marketing Investment)	1	2.5
Service-oriented architecture	1	2.5
Social media	1	2.5
Supply Chain	1	2.5
Survey versus log data	1	2.5
Technology	1	2.5
Value proposition	1	2.5
Virtual reality	1	2.5
Virtual retailing	1	2.5

The results obtained from the study indicate the following links of concepts with one another: 'big data analytics' with the entire collection of concepts associated with a customer ('customer-centric,' 'customer engagement', 'customer experience', 'customer targeting service', 'customers classification').

Analysis of publications based on the adopted research procedure allowed for the determination of keywords associated with the term 'advanced data analysis' (RQ 1.2), whose context is described through the following concepts: 'data analysis' and 'big data'. Concepts identified as consistent with the meaning of 'advanced data analysis' were used in the second phase of the study.

In the second phase, we identified keywords appearing in the literature jointly with a definition of data analysis in the context of marketing. Additionally, we examined whether and in what manner data analysis supports decision-making processes in retail enterprises, as well as what problems can be solved using data analysis. In the first phase of the study, keywords were defined as

necessary for constructing character strings in search engines. Apart from the terms 'data analysis' and 'big data', generated by the first phase, the term 'decision support systems' was added, which specifies the sphere for support of decision-making processes within the framework of the study. All three terms were paired with the concepts 'marketing' and "retailing'. The second phase of the study allowed us to obtain a group of concepts connecting the fields of data analysis and marketing (RQ 2.1). The term which appeared with the greatest frequency in the analysed publications was 'big data analytics.' An important word recurring very frequently in results was 'customer', but more important was the distinction of a set of key terms like 'customer-centric', 'customer engagement', 'customer experience', 'customer targeting service', and 'customers classification'.

The conducted study confirmed that data analysis is used in the decision-making processes of enterprises engaged in retail sales (RQ 2.2). Data analysis techniques such as predictive analysis, referring to data exploration procedures (employing statistical techniques like multiple regression), are used in order to create forecasts supporting managerial decision-making (Kridel & Dolk, 2013; Zhao, 2013).

Data analysis allows companies in the retail sales industry to make more efficient use of data, increase conversion rates, enhance decision-making processes, and improve their customers' position (Calder, Malthouse, & Maslowska, 2016). The most technologically developed manner of employing data analysis is online data analysis, which facilitates the process of making decisions in real-time for companies involved in retail trading (Aguinis, Forcum, & Joo, 2013; Leeflang, Verhoef, Dahlström, & Freundt, 2014). The issue of market and customer segmentation is an important element in the marketing strategy of every enterprise (Fan, Lau, & Zhao, 2015). Data analysis supplies methods and techniques (RQ 2.3) that assist in solving such problems (Akter & Wamba, 2016).

The greatest benefits derived from the use of data analysis are visible in such areas as customer personalisation, which allows for the identification of customers who demonstrate the highest likelihood of returning to visit the company again, thereby increasing the chances they will again purchase products or services (Yeung & Yee, 2015). The creation of a dynamic pricing policy using data analysis facilitates determining the price that will maximise revenue and/or profit (Akter & Wamba, 2016).

5. Conclusions and future works

This paper presents a study concerning the use of advanced data analysis in marketing by companies operating in the retail trading sector. The study consisted in a systematic literature review. The foundation of the applied research procedure was the approach proposed by B. Kitchenham (2004) and extended by C. Wohlin & R. Prikladniki (2013). The main contribution of this work is the proposal of modifying the systematic literature review method, which was expanded through the introduction of two phases. Although there are many similarities in the manner in which they were conducted, this distinction is important for conducting studies of literature if there are no clear, established definitions for the concepts being employed. Each phase is intended to facilitate obtaining answers to different important research questions. The first phase constitutes an exploratory study, whose results allow for the second phase of the study to be conducted. With respect to the present study, key information was obtained concerning the application of advanced data analysis in marketing, particularly in companies within the retail trading sector.

The results of this study of the relevant literature indicate that scholarly publications do not use the phrase 'advanced data analysis', and its context is described with the term 'data analysis'. Another term used broadly within the sphere of data analysis is 'big data'. In the subject-relevant literature, the concepts of data analysis in marketing are nested within the context of enterprises in the retail trading sector, focused around the concept of 'big data analytics' and terms linked to the word 'customer', such as 'customer-centric', 'customer engagement', 'customer experience', 'customer targeting service', and 'customers classification'. The study undertaken of the literature indicates that marketing employs data analysis in such areas as customer needs identification and market segmentation, which solves many problems that would be time-consuming, difficult, and even impossible to manage without the use of techniques and concepts borrowed from data analysis.

An additional result of the conducted research, important in the perspective of their continuation, is a list of articles and works that exactly relate to the analysed area. Using the method called snowball sampling, an analysis of the literature in these publications will be carried out, which will allow identifying works of importance to this area published before 2008. These studies are important for conducting in-depth analyses of the following issues: which marketing activities relate to data analysis, what data analysis process is used, etc.

When conducting literature-oriented research into advanced data analysis in marketing, one needs to concentrate only on searching for publications related to the phrase 'big data'. Research performed so far indicates that the term 'big data' should be interpreted as a synonym of the expression 'advanced data analysis'.

To summarise, we may say that enterprises in the retail trading sector employ data analysis to support decision-making processes. One example of such activity is the dynamic determination of prices for products and services using methods of exploration and machine learning, based on data gathered from customers. Further studies will explore the issue of data analysis in retail trading enterprises oriented at the customer, as well as support for decision-making processes using real-time data analysis.

References

Aguinis, H., Forcum, L. E., & Joo, H. (2013). Using market basket analysis in management research. *Journal of Management, 39*(7), 1799-1824. doi: 10.1177/014920 6312466147

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: A systematic review and agenda for future research. *Electronic Markets, 26*(2), 173-194. doi: 10.1007/s12525-016-0219-0

Bucklin, R. E., & Sismeiro, C. (2009). Click here for Internet insight: Advances in clickstream data analysis in marketing. *Journal of Interactive Marketing, 23*(1), 35-48.

Calder, B. J., Malthouse, E. C., & Maslowska, E. (2016). Brand marketing, big data and social innovation as future research directions for engagement. *Journal of Marketing Management, 32*(5-6), 579-585. doi: 10.1080/0267257X.2016.1144326

Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big data. *Information Sciences*, *275*, 314-347. doi: 10.1016/j.ins.2014.01.015

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19(2)*, 171-209. doi: 10.1007/s11036-013-0489-0

Elgendy, N., & Elragal, A. (2014). Big data analytics: A literature review paper. *Industrial Conference on Data Mining* (pp. 214-227). Berlin: Springer-Cham, doi: 10.1007/978-3-319-08976-8_16

Fan, S., Lau, R. Y., & Zhao, J. L. (2015). Demystifying big data analytics for Business Intelligence through the lens of marketing mix. *Big Data Research, 2*(1), 28-32. doi: 10.1016/j.bdr.2015.02.006

Gordon, S. Linoff, M., & Berry, J.A. (2011). *Data Mining Techniques: For marketing, sales, and customer relationship*. New York: John Wiley & Sons.

Grandhi, B., Patwa, N., & Saleem, K. (2017). Data-driven marketing for growth and profitability. *10th Annual Conference of the EuroMed Academy of Business*, 675-694.

Han, J., Pei, J., & Kamber, M. (2012). *Data Mining: Concepts and techniques*. Amsterdam: Elsevier.

Hu, H., Wen, Y., Chua, T. S., & Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE Access, 2*, 652-687. doi: 10.1109/ACCESS. 2014.2332453.

Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). *Recommender systems: An introduction*. Cambridge: Cambridge University Press.

Kitchenham, B. (2004). *Procedures for performing systematic reviews*. Joint Technical Report, Keele: Keele University TR/SE-0401 and NICTA 0400011T.1, July.

Kitchenham, B., & Charters, C. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Keele University and Durham University Joint Report–EBSE 2007-001.

Kitchenham, B., Pretorius, R., Budgen, D., Brereton, O. P., Turner, M., Niazi, M., & Linkman, S. (2010). Systematic literature reviews in software engineering – a tertiary study. *Journal Information and Software Technology, 52*, 792-805. doi: 10.1016/j.infsof.2008.09.009

Konstan, J.A., & Adomavicius, G. (2013). Toward identification and adoption of best practices in algorithmic recommender systems research [in] *Proceedings of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation* (pp. 23-28). New York: ACM.

Kridel, D., & Dolk, D. (2013). Automated self-service modelling: Predictive analytics as a service. *Information Systems and e-Business Management, 11(1)*, 119-140. doi: 10.1007/s10257-011-0185-1

Leeflang, P. S., Verhoef, P. C., Dahlström, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. *European Management Journal, 32*(1), 1-12. doi: 10.1016/j.emj.2013.12.001

Niazi, M. (2015). Do systematic literature reviews outperform informal literature reviews in the software engineering domain? An initial case study. *Arabian Journal for Science and Engineering, 40*(3), 845-855. doi: 10.1007/s13369-015-1586-0

Pabedinskaitė, A., Davidavičienė, V., & Milišauskas, P. (2014). Big data driven e-commerce marketing. *In Proceedings of 8th International Scientific Conference Business and Management-Spausdinta* (pp. 645-654). Vilnus, Lietuva. doi: 10. 3846/bm.2014.079

Pavlo, A., Paulson, E., Rasin, A., Abadi, D. J., DeWitt, D. J., Madden, S., & Stonebraker, M. (2009). A comparison of approaches to large-scale data analysis. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data* (pp. 165-178). Providence, RI: ACM. doi: 10.1145/1559845.1559865

Pawełoszek, I., & Korczak, J. (2017). From data exploration to semantic model of customer. In *Intelligent Systems Conference (IntelliSys)* (pp. 382-388). London, 7-8 September. doi: 10.1109/IntelliSys.2017.8324322

Pierański, B., & Strykowski, S. (2017). Towards a personalized virtual customer experience. In *Asian Conference on Intelligent Information and Database Systems* (pp. 185-195). Berlin: Springer. doi: 10.1007/978-3-319-56660-3_17

Pondel, M., & Korczak, J. (2017). A view on the methodology of analysis and exploration of marketing data. In *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems. Annals of Computer Science and Information Systems* (Vol. 9, pp. 1135-1143). Prague, September 3-6. doi: 10.15439/2017F442

Quiñones, D., & Rusu, C. (2017). How to develop usability heuristics: A systematic literature review. *Computer Standards & Interfaces, 53*, 89-122. doi: 10.1016/j.csi.2017.03.009

Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: Introduction and challenges. In *Recommender systems handbook* (pp. 1-34). Boston, MA: Springer, doi: 10.1007/978-0-387-85820-3

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-based medicine: How to practice and teach EBM*. 2nd ed., Edinburgh: Churchill Livingstone. doi: 10.1177/088506660101600307

Sanders, N. R. (2016). How to use big data to drive your supply chain. *California Management Review, 58*(3), 26-48.doi: 10.1525/cmr.2016.58.3.26

Setia, S., & Jyoti, D. (2013). Multi-level association rule mining: A review. *International Journal of Computer Trends and Technology (IJCTT), 6*(3), 166-170.

Sharda, R., Delen, D., & Turban, E. (2014). *Business intelligence and analytics: Systems for decision support*. Harlow: Pearson Educational.

Slavakis, K., Giannakis, G. B., & Mateos, G. (2014). Modelling and optimization for big data analytics: (statistical) learning tools for our era of data deluge. *IEEE Signal Processing Magazine*, *31(5)*, 18-31. doi: 10.1109/MSP.2014.2327238

Vahid, G., & Mäntyläc, M. V. (2016). A systematic literature review of literature reviews in software testing. *Journal Information and Software Technology, 80*, 195-216. doi: 10.1016/j.infsof.2008.09.009

Vera-Baquero, A., Colomo-Palacios, R., & Molloy, O. (2013). Business process analytics using a big data approach. *IT Professional*, *15*(6), 29-35. doi: 10.1109/MITP.2013.60

Witten, I., Frank, E., Hall, M., & Pal, C. (2017). *Data mining: Practical machine learning tools and techniques*. Burlington, MA: Morgan Kaufman. doi: 10.1016/C2015-0-02071-8

Wohlin, C., & Prikladniki, R. (2013). Systematic literature reviews in software engineering. *Information and Software Technology, 55*(6), 919-920. doi: 10.1016/j.infsof.2008.09.009

Yeung, R., & Yee, W. (2015). Application of cluster analysis and discriminant analysis in market segmentation and prediction. In J. Mendy & S. G. Geringer (Eds.), *Leading Issues in Business Research Methods* (Vol. 2, pp. 63-79). Sonning Common, RG: ASPI.

Zhao, D. (2013). Frontiers of big data business analytics: Patterns and cases in online marketing. In J. Liebowitz (Ed.), *Big data and business analytics* (pp. 43-67). Boston, MA: Auerbach Publications.

Zhao, J. L., Fan, S., & Hu, D. (2014). Business challenges and research directions of management analytics in the big data era. *Journal of Management Analytics, 1*(3), 169-174. doi: 10.1080/23270012.2014.968643