**Supplement to  
“Insights into Domain Names and Product Categories of   
e-Commerce Websites: A Case Study of the UAE”**

Rachidatou Ingrid Traoret1, Salama AlDhaheri1, Fatima AlAmeemi1, Gurdal Ertek1

1 College of Business and Economics, United Arab Emirates University, Al Ain, UAE

[700042159@uaeu.ac.ae](mailto:700042159@uaeu.ac.ae), [201801296@uaeu.ac.ae](mailto:201801296@uaeu.ac.ae), [201207945@uaeu.ac.ae](mailto:201207945@uaeu.ac.ae), [gurdal@uaeu.ac.ae](mailto:gurdal@uaeu.ac.ae)

**Abstract.** Understanding e-commerce markets and online consumer behavior is critical for e-commerce retailers in the United Arab Emirates (UAE) and across the world. This study conducted a comprehensive analysis of unique e-commerce data to reveal patterns in domain names, product categories, and associations with other attributes. Using data analytics techniques, the initial dataset was cleaned and augmented to conduct the analysis. Key findings reveal ‘.com’ and ‘.ae’ as dominant domain extensions, numeric patterns in domain names, and ‘apparel’ as the most prevalent product category. An in-depth association mining analysis identifies and lists frequent categories and category combinations. This study provides actionable insights into the naming and category planning of e-commerce retailers in the UAE. Multidimensional analysis establishes an analysis and planning framework for evidence-based e-commerce research that can guide the two critical strategic decisions of domain name and category selection, which are among the important strategic decisions when an e-commerce business is first established. Most importantly, while providing insights for UAE e-commerce markets specifically, the methodology can be applied to any country or region.

**Keywords:** E-commerce, UAE, Online Shopping, Association Mining, Domain Names, Product categories.

1. Introduction

In an era marked by rapid digital transformation and the spread of e-commerce, shopping has overcome the limitations of brick-and-mortar stores. In 2022, the disparity in growth between e-commerce and traditional retail was minimal, with only a marginal difference of 0.2%, highlighting consumers' strong inclination towards online shopping [1]. Online shopping, once a convenience, has evolved to shape the way consumers explore, select, and purchase products and services. This transformation has not only revolutionized the retail landscape, but has also given rise to new consumer behaviors, preferences, and expectations that continue to reshape the virtual marketplace.

The e-commerce industry is rapidly expanding worldwide. By the end of 2023, the global e-commerce market is expected to reach a total value of $6.3 trillion, with a projected growth rate of 11.17% from 2023 to 2027, resulting in a market volume of around $5.56 trillion by 2027 [2]. During the same year, it is anticipated that around 2.64 billion consumers worldwide will have engaged in at least one online purchase, with the global e-commerce market projected to experience an annual growth rate of 8.9% [1, 3]. The peak of global e-commerce growth was observed in 2021 amidst the Covid-19 pandemic, reaching 17.1%. Globally, e-commerce sales are projected to grow by at least 8% per annum in the coming years [3]. The e-commerce retail sector, accounting for 18% of global retail sales in 2020, is expected to represent approximately 22% of global retail sales by 2024 [1].

China has maintained its position as the global leader in e-commerce sales, contributing to more than 50% of total retail sales worldwide, with online sales in China surpassing the $3 trillion mark in 2022 [3]. Meanwhile, the U.S. e-commerce market is predicted to exceed $940.9 billion by 2023, representing slightly over one-third of China's total e-commerce market [1].

Across the Middle East and North Africa (MENA) region, 91% of consumers make online purchases, and 70% prefer digital payment methods [4]. The e-commerce sector in the Middle East is expected to achieve a market volume of US$50 billion by 2025. This growth is supported by factors such as the widespread adoption of advanced technologies, government initiatives, and a young, technology-savvy population with increasing disposable income [4].

The United Arab Emirates (UAE), with its rapid technological advancement and high Internet penetration rate, has emerged as a prominent hub for online shopping in the Middle East. Previous studies have highlighted the growth potential and dynamics of e-commerce in the UAE, emphasizing the importance of understanding the underlying factors that contribute to the success of online platforms [5]. The Business-to-Consumer (B2C) e-commerce market in the country has been projected to grow significantly in recent years [6][7]. UAE is positioned as the 27th largest e-commerce market in the world and is anticipated to generate revenue of approximately US$10,205.5 million by the year 2023 [8]. Recognizing the significance of e-commerce and e-commerce platforms, this research aims to conduct a comprehensive analysis of the top 10,000 e-commerce websites operating in the UAE, examining aspects such as domain-naming patterns and product categories.

To the best of our knowledge, the present study is the first of its kind, by presenting a methodology and sample country-specific analysis that can guide two critical strategic planning problems that every e-commerce start-up across the globe faces:

* What should be our domain name?
* Which product categories should we offer?

The contributions of the current study become more pronounced when related literature is reviewed, as presented next.

1. Literature Review
   1. E-commerce Research

In the e-commerce sector, the quality and relevance of data plays a pivotal role and has witnessed exponential growth globally, with emerging markets demonstrating high potential. While there is a rich body of research on e-commerce, we cite only a few studies that provide a taste of types of research streams.

Farhangmehr *et al*. [9] compared consumer perceptions of hypermarkets and traditional stores to highlight the factors that influence retail format selection. Their findings revealed that product assortment, quality, and flexibility in payment options are key drivers in Portugal. Although insightful, their study was limited to store formats rather than online channels. Yadav and Pathak [10] stressed that understanding these emerging trends is critical for retailers aiming to cater to diverse consumer segments through targeted assortments, recommendations, and promotions.

One can find detailed literature review papers on various streams of e-commerce research, including social commerce [11], information disclosure [12], cross-border e-commerce [13], consumer-to-consumer e-commerce [14], after-sales attributes [15], e-commerce and globalization [16], besides many other reviews.

Furthermore, one can obtain holistic perspective of different aspects of e-commerce and e-commerce research through taxonomy papers, including taxonomies for privacy concerns [17], strategic grouping [18], queries and search behavior [19], products [20], security challenges [21], content creator-based business models [22], and e-commerce concept labels [23], besides many other taxonomies.

* 1. Research for UAE

Studies have highlighted the increasing adoption of online shopping among consumers in the Middle East, attributing it to increased internet penetration, smartphone usage, and changing lifestyles [5].

Several studies have focused on the e-commerce sector in the UAE: According to Alzaabi *et al*. [5], several factors have contributed to the UAE’s e-commerce boom, including technological readiness, logistics infrastructure, increased selection, and competitive pricing, compared to brick-and-mortar stores. However, gaps persist in understanding the specific categories and products driving online sales. Dhanapal *et al*. [24] attempted to segment online shoppers in the UAE based on their demographic attributes and concluded that different age groups demonstrate distinct motivations and preferences. However, their study did not examine the market composition.

* 1. Analysis of Domain Names

Research on the analysis of domain names typically focuses on cyber security aspects [25][26] or the detection of certain types of domains, such as gambling sites [27]. [25] uses data from Internet protocol Domain Name Systems (DNS) to develop a system that automatically detects malicious activity as early warning signals for potential cyber-attacks. Using a large dataset, [26] showed that the current privacy-preserving enhancements currently under standardization, such as DoH and eSNI, are not very effective and call for further actions to avoid retrieving the names of servers. According to [27], most of the training datasets for gambling domain names (GDN) in previous work used GND blacklists from publicly available sources; however, their accuracy and authenticity have not been verified, and they did not provide satisfactory classification results. To improve these results, the authors propose a certified and textual analysis-based classification method (CT-GDNC) to classify GDN training data, which proves to be significant and results in better accuracy compared to previous methods.

To assess the role of e-commerce websites in metro and small business resilience during the Covid-19 pandemic, [28] used a new source of population data on the domain name host. The findings imply that metros with high concentrations of small businesses with online presence experienced more positive economic perceptions and outcomes from April to December 2020, suggesting that digitally enabled small businesses played an important role in local economic resilience. For [29], the web domain name value depended on its ability to attract web traffic (visitors) and transform them into cash-generating customers. This confirms the results in [30] on the importance of incorporating search engine optimization (SEO) techniques into websites, as it contributes to improving brand visibility and, therefore, increases the odds of being noticed by future potential customers.

* 1. Product Category and Portfolio Management

Most research on product category and portfolio management focuses on the application of machine learning, natural language processing, deep learning, and neural networks to classify products into different categories. The rapid increase in e-commerce has led to a significant challenge in navigating an abundance of products. To overcome this challenge, [31] used a natural language processing algorithm (Term Frequency-Inverse Document Frequency) and supervised machine learning techniques (support vector machines) to classify product description data into four categories: electronics, household items, books, and clothing. The results of the research have proven to be significant and could contribute to improving customer experience and overall operational efficiency. [32] developed a generalized interactive real-time system that provides category recommendations for different e-commerce scenarios. The system then uses a convolutional sequence-to-sequence approach for product category recognition, which is effective and efficient for real-world data.

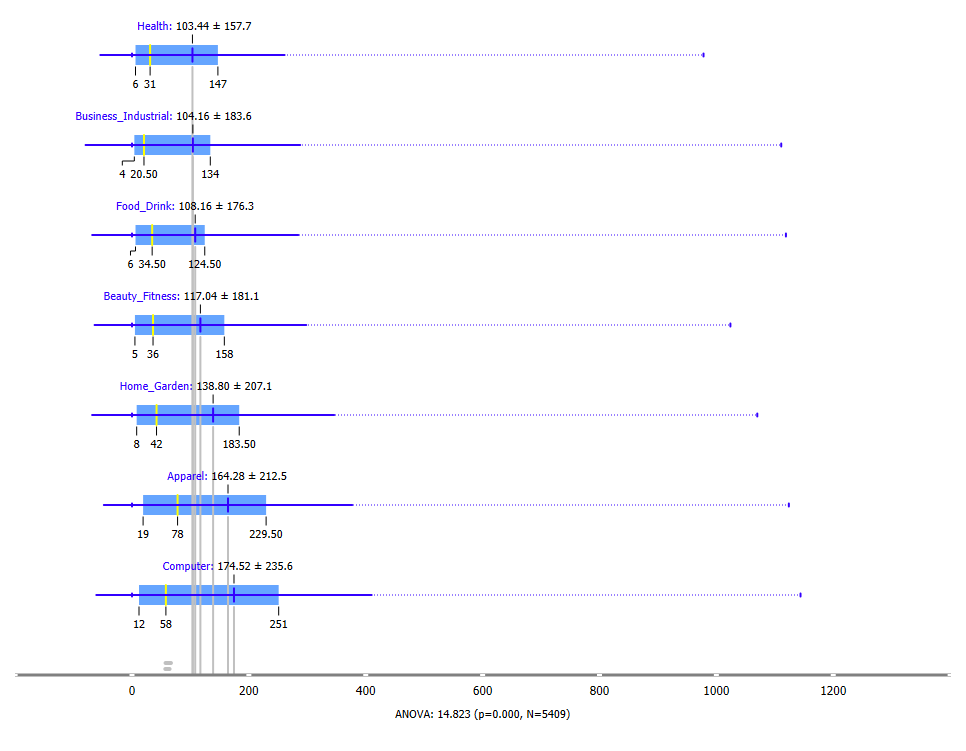
According to [33], the top five EU-27 product categories with their respective percentages are: any physical good, 97%; clothes, shoes, or accessories, 68%; physical multimedia (music, DVDs, books, etc.), 34%; any physical goods from private persons, 34%; and deliveries from restaurants, fast food chains, and catering services, 30%. Research on e-commerce category analysis has mostly been conducted using market basket analysis (MBA), which applies the Apriori algorithm or other association rule mining algorithms [34] to unravel the hidden relations between item categories and item sets. In this context, [35] used K-means clustering-based MBA on UK-domiciled online retailer data to identify distinct patterns for bundling sales recommendations to promote sales and sales. The results show that the categories with the most dominant associative purchase patterns are Bags and Accessories, Decoratives, Homeware and Stationaries. Moreover, Bags and Accessories have a large number of outward directing linkages. [36] applied the Apriori algorithm to analyze the Instacart data in the US to generate association rules for a recommender system. The results show that banana is the most preferred item in the data that can easily match strawberries, apples, cucumbers, avocados, and lemons.

1. Methodology
   1. Data

Table 1 in this supplement, which was referred to in the paper, lists the names of the attributes of interest (that were considered in this particular study) in the augmented Dataset E, along with their data type (strings and integer attributes), sample values for a selected observation, and description.

**Table 1.** Data attributes in Dataset E that were the focus of the presented research.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Data Type (Role) | Sample Value | Description |
| ID | integer (key) | 13779 | Rank of the website/platform in the world |
| domain | string | track24.net | Domain name |
| NumberStringInDomainName | integer | 24 | If any, the number string in the domain name |
| NumberOfDigits | integer | 2 | If the domain name has numbers in it, the number of digits of that number in the domain name |
| NumberOfComponents | integer | 2 | Number of components in the domain name, excluding www |
| DomainType | string | net | Domain extension (type) |
| categories | string list | Business\_Industrial Mail\_Package\_ Delivery | Categories, separated by space |
| city | string | Dubai | City where the e-commerce company is registered |
| state | string | Dubai | State (the UAE Emirate) where the e-commerce company is registered |
| estimated\_monthly\_sales | integer | 1,553,273.71 | Monthly sales estimated |
| products\_sold | integer | 262 | Number of products sold in the platform |
| Apparel | binary | 0 | Whether the website includes the mentioned category (1) or not (0 or empty cell) |
| Home\_Garden | binary | 0 |  |
| Food\_Drink | binary | 0 |  |
| Beauty\_Fitness | binary | 0 |  |
| Business\_Industrial | binary | 1 |  |
| Face\_Body\_Care | binary | 0 |  |
| Food | binary | 0 |  |
| Health | binary | 0 |  |
| Computers | binary | 0 |  |



**Fig. 5.** Distribution of number of products sold, in relation to product categories.

1. Results and Discussion

The last analysis concerned the number of products sold on e-commerce websites in relation to product categories. Due to space limitations, this analysis is provided in this supplemental document. This analysis was conducted by querying Dataset G from Dataset E. Dataset G denotes the number of products sold on the e-commerce website for each website that had a category (e.g., Apparel). The box plot visualization and ANOVA results shown in Fig. 5 provide several insights:

1. E-commerce websites are ranked according to the average number of products sold (vertical gray lines in Fig. 5), based on the following categories: computer, apparel, home garden, beauty fitness, food drink, business industry, and health.
2. The distribution of values is right skewed for all categories, as deduced from the median values (vertical white lines within the boxplots), which are well below the mean values (vertical gray lines).
3. Based on both the median and mean values, the number of products needed to establish Health and Business\_Industrial e-commerce businesses is the lowest.
4. Websites that sell Apparel have the highest median value of 78.
5. 90% of the websites have less than around 400 products (the right-hand value of the horizontal blue lines do not go beyond 400).
6. 75% of the websites have less than around 250 products (251, to be exact).
7. Parametric ANOVA test suggests very strongly, with a p value very close to zero, that the differences in means (as estimated by averages) across categories is statistically significant. As can be read from the individual box plots, the data for at least some of the categories are most likely not normally distributed, thus, applying a non-parametric comparison test, such as the Kruskal-Wallis H Test, followed by a non-parametric pairwise comparison test, such as the Dunn test is more appropriate [37]. Yet, the ANOVA results are still valuable as they signal the high likelihood that the mean number of products across categories are not equal.

Acknowledgement

The authors thank Mr. Sumit Ghugharwal, Founder of the Dubai-based My Lead Fox SaaS software, for allowing the research team to use the dataset for research and to publish the analysis results of the filtered dataset publicly.

References

1. Keenan, M.: Global Ecommerce Statistics: Trends to Guide Your Store in 2024. (2023). Shopify Plus. https://www.shopify.com/enterprise/global-ecommerce-statistics. Accessed: 25 December 2023.
2. Statista. eCommerce - Worldwide. https://www.statista.com/outlook/dmo/ecommerce/worldwide. Accessed: 25 December 2023.
3. Adobe Experience Cloud Team. Top ecommerce statistics for 2023 (2023). https://business.adobe.com/blog/basics/2023-ecommerce-statistics. Accessed: 25 De-cember 2023.
4. Lau, L.: Ecommerce in the Middle East: Statistics and Trends in 2023. Orienteed e-commerce Solutions. (2023). https://orienteed.com/en/ecommerce-middle-east-statistics-trends-2023/. Accessed: 25 December 2023.
5. Alzaabi, K., Hamad, M.O., Omar, R., Romle, A.R.: Organizational factors and e-commerce adoption in SMEs of United Arab Emirates: mediating role of perceived stra-tegic value. International Journal of Entrepreneurship 25(1), 1-18 (2021).
6. Market Research. (2023). United Arab Emirates B2C Ecommerce Market Opportuni-ties Databook – 100+ KPIs on Ecommerce Verticals (Shopping, Travel, Food Service, Media & Entertainment, Technology), Market Share by Key Players, Sales Channel Anal-ysis, Payment Instrument, Consumer Demographics - Q1 2023 Update. https://www.marketresearch.com/PayNXT360-v4075/United-Arab-Emirates-B2C-Ecommerce-33669535/. Accessed: 25 December 2023.
7. Dubai Chamber. UAE e-commerce market forecast to reach $9.2 billion by 2026. (2022). https://www.dubaichamber.com/en/media-center/news/uae-e-commerce-market-forecast-to-reach-9-2-billion-by-2026/. Accessed: 26 Oct. 2023.
8. ECBD. eCommerce Market UAE - Data, Trends, Top Stores (2022). ECDB. https://ecommercedb.com/markets/ae/all. Accessed: 25 December 2023.
9. Farhangmehr, M., Marques, S., Silva, J.: Consumer and retailer perceptions of hy-permarkets and traditional retail stores in Portugal. Journal of Retailing and Consumer services 7(4) 197-206 (2000). https://doi.org/10.1016/s0969-6989(00)00019-9
10. Yadav, R., Pathak, G.S.: Young consumers’ intention towards buying green prod-ucts in a developing nation: Extending the theory of planned behavior. Journal of Cleaner Production 135, 732-739 (2016). https://doi.org/10.1016/j.jclepro.2016.06.120
11. Attar, R.W., Almusharraf, A., Alfawaz, A., Hajli, N.: New Trends in E-Commerce Research: Linking Social Commerce and Sharing Commerce: A Systematic Literature Review. Sustainability, 14(23), 16024 (2022). https://doi.org/10.3390/su142316024
12. Kolotylo-Kulkarni, M., Xia, W., Dhillon, G.: Information disclosure in e-commerce: A systematic review and agenda for future research. Journal of Business Research, 126, 221-238 (2021). https://doi.org/10.1016/j.jbusres.2020.12.006
13. Hazarika, B.B., Mousavi, R.: Review of cross-border E-commerce and direc-tions for future research. Journal of Global Information Management (JGIM), 30(2), 1-18 (2021). https://doi.org/10.4018/jgim.20220701oa02
14. Yuan, C.H., Wu, C.H., Wang, D., Yao, S., Feng, Y.: Review of consumer-to-consumer E-commerce research collaboration. Journal of Organizational and End User Computing (JOEUC), 33(4), 167-184 (2021). https://doi.org/10.4018/joeuc.20210701.oa8
15. Ferraz, R.M., da Veiga, C.P., da Veiga, C.R.P., Furquim, T.S.G., da Silva, W.V.: After-Sales Attributes in E-Commerce: A Systematic Literature Review and Future Research Agenda. Journal of Theoretical and Applied Electronic Commerce Research, 18(1), 475-500 (2023). https://doi.org/10.3390/jtaer18010025
16. Penu, O.K.A., Boateng, R., Babatope, L.O.: E-Commerce and Globaliza-tion: A Bibliometric Review of 24 Years of Research. Digital Innovations, Business and Society in Africa: New Frontiers and a Shared Strategic Vision, 3-29 (2022). https://doi.org/10.1007/978-3-030-77987-0\_1
17. Bandara, R., Fernando, M., Akter, S.: Privacy concerns in E-commerce: A taxonomy and a future research agenda. Electronic Markets, 30(3), 629-647 (2020). https://doi.org/10.1007/s12525-019-00375-6
18. Holland, C.P., Gutiérrez-Leefmans, M.: A taxonomy of SME e-commerce platforms derived from a market-level analysis. International Journal of Electronic Commerce, 22(2), 161-201 (2018). https://doi.org/10.1080/10864415.2017.1364114
19. Sondhi, P., Sharma, M., Kolari, P., Zhai, C.: A taxonomy of queries for e-commerce search. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 1245-1248 (2018, June). https://doi.org/10.1145/3209978.3210152
20. Aanen, S.S., Vandic, D., Frasincar, F.: Automated product taxonomy mapping in an e-commerce environment. Expert Systems with Applications, 42(3), 1298-1313 (2015). https://doi.org/10.1016/j.eswa.2014.09.032
21. ALadan, M.I.: E-commerce security challenges: a taxonomy. Journal of Eco-nomics, Business and Management, 4(10), 589-593 (2016). Available at: http://www.joebm.com/vol4/457-MH0011.pdf. Accessed: 25 December 2023. https://doi.org/10.18178/joebm.2016.4.10.457
22. Mileros, M.D., Lakemond, N., Forchheimer, R.: Towards a Taxonomy of E-Commerce: Characterizing Content Creator-based Business Models. Technology Inno-vation Management Review 9, 10: 62–77 (2019). Available at: https://www.timreview.ca/article/1276. Accessed: 25 December 2023.
23. Kejriwal, M., Shen, K.: A Lightweight Global Taxonomy Induction System for E-Commerce Concept Labels. In Cyber Security and Social Media Applications, pp. 99-114 (2023). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-33065-0\_5
24. Dhanapal, S., Vashu, D., Subramaniam, T.: Perceptions on the challenges of online purchasing: a study from “baby boomers”, generation “X” and generation “Y” point of views. Contaduría y Administración 60, 107-132 (2015). https://doi.org/10.1016/j.cya.2015.08.003
25. Vinayakumar, R., Poornachandran, P., Soman, K.P.: Scalable Framework for Cyber Threat Situational Awareness Based on Domain Name Systems Data Analysis. In: Roy, S., Samui, P., Deo, R., Ntalampiras, S. (eds) Big Data in Engineering Applica-tions. Studies in Big Data, vol 44. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-8476-8\_6. Available under https://link.springer.com/chapter/10.1007/978-981-10-8476-8\_6.
26. Trevisan, M., Soro, F., Mellia, M., Drago, I., Morla, R.: Does domain name en-cryption increase users' privacy?. ACM SIGCOMM Computer Communication Review, 50(3), 16-22 (2020). https://doi.org/10.1145/3411740.3411743
27. Sun, G., Ye, F., Chai, T., Zhang, Z., Tong, X., Prasad, S.: Gambling Domain Name Recognition via Certificate and Textual Analysis. The Computer Journal, 66(8), 1829-1839 (2023). https://doi.org/10.1093/comjnl/bxac043. Available under https://academic.oup.com/comjnl/article/66/8/1829/6570684.
28. Mossberger, K., Martini, N.F., McCullough, M., Tolbert, C.J.: Digital economic activity and resilience for metros and small businesses during Covid-19. Small Business Economics, 60(4), 1699-1717 (2023). https://doi.org/10.1007/s11187-022-00674-x
29. Moro-Visconti, R.: Domain Name and Website Valuation. In The Valuation of Digital Intangibles: Technology, Marketing, and the Metaverse. pp. 349-381 (2022). Springer International Publishing. https://doi.org/10.1007/978-3-031-09237-4\_11
30. Ologunebi, J., Taiwo, E.O.: The Importance of SEO and SEM in improving brand visibility in E-commerce industry; A study of Decathlon, Amazon and ASOS. (November 20, 2023). https://doi.org/10.2139/ssrn.4638890
31. Pakpahan, D., Siallagan, V., Siregar, S.: Classification of E-Commerce Product Descriptions with The Tf-Idf and Svm Methods. Sinkron: Jurnal dan Penelitian Teknik Informatika, 8(4), 2130-2137 (2023). https://doi.org/10.33395/sinkron.v8i4.12779
32. Hasson, I., Novgorodov, S., Fuchs, G., Acriche, Y.: Category recognition in e-commerce using sequence-to-sequence hierarchical classification. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining, pp. 902-905 (2021). https://doi.org/10.1145/3437963.3441717
33. Lone, S., Weltevreden, J.W.J.: 2023 European E-commerce Report. Amster-dam/Brussels: Amsterdam University of Applied Sciences & Ecommerce Europe (2023). https://www.upu.int/UPU/media/wwwUpuIntUniversalPostalUnionAboutUpuBodiesConsultativeCommittee/2023EuropeanEcommerceReportEn.pdf
34. Ertek, G., Demiriz, A.: A framework for visualizing association mining re-sults. In International Symposium on Computer and Information Sciences, pp. 593-602 (2006). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/11902140\_63
35. Lim, T.: K-Means Clustering-Based Market Basket Analysis: UK Online E-Commerce Retailer. In ICIT. pp. 126-131 (2021). https://doi.org/10.1109/icit52682.2021.9491703
36. Gurudath, S.: Market Basket Analysis & Recommendation System Using Associa-tion Rules. Master of Science in Big Data Management and Analytics, Griffith College, Dublin (2020).
37. Ertek, G., Tokdemir, G., and Hammoudi, M. M.: Graph-Based Visuali-zation of Stochastic Dominance in Statistical Comparisons. 2019 IEEE/ACS 16th Inter-national Conference on Computer Systems and Applications (AICCSA), Abu Dhabi, United Arab Emirates, 2019, pp. 1-7 (2019). https://doi.org/10.1109/AICCSA47632.2019.9035354