Trends and usage pattern of SPSS and Minitab Software in Scientific research

Hilary I. Okagbue¹, Pelumi E. Oguntunde¹, Emmanuela C.M. Obasi², Elvir M. Akhmetshin³

¹Department of Mathematics, Covenant University, Ota, Nigeria
²Department of Computer Science and Informatics, Federal University, Otuoke, Nigeria
³Department of Economics and Management, Kazan Federal University, Elabuga Institute of KFU, Elabuga, Russia

Email: hilary.okagbue@covenantuniversity.edu.ng

Abstract. Most scientific research generates data. Analysis of the data from scientific research helps create new knowledge or a deep understanding of natural phenomena. Statistical software is used mainly in data analysis. SPSS and Minitab appear to be most popular, especially for those that could neither code nor mathematical inclined to handle advanced software such as R, MATLAB, Maple, etc. Trends and usage pattern of SPSS and Minitab Software in Scientific research was studied in this paper with the data obtained from the Scopus database. In their abstracts or keywords, documents that have mentioned SPSS were extracted for the years 2010 to 2019. Frequency analysis showed that the trend of using SPSS and Minitab is steadily increasing, although the use of Minitab is a fraction of SPSS. Minitab is mostly used in engineering, materials science, and computer science, while SPSS is mainly used in medicine, social science, and engineering. Analysis of the document type showed that SPSS and Minitab are mostly stated in abstracts or keywords of research articles, conference papers, review papers, and books indexed in Scopus.

Keywords: Abstract, keywords, Minitab, research, Scopus, software, SPSS, subject classification, statistics,

1. Introduction

Statistical software (SS) are specialized computer programs designed for statistical data analysis. It comes as a standalone package, extension, programming language, or macros/add-ons. Statistical software can be navigated as a graphical user interface (GUI) or scripting. Statistical software is used to analyze data emanating from a carefully organized scientific process of observation and experimentation. The data analysis can then produce a result that can be interpreted to birth scientific knowledge. The data analysis is often done systematically to ensure adequate evaluation of data [1] and hence, describing the data or making statistical inference from the data. The absence of statistical software implies that most scientific or social science data would not graduate into knowledge that can guide in decision making. Also, statistical tools help to precipitate research findings from proper analysis of data, thereby preserving the data's integrity. Again, data analysis can be used to validate research findings in cases where integrity issues are suspected.
Statistical software can be in the form of open-source, public domain, freeware, and proprietary. Examples of open-source SS are R, OpenEpi, JASP, PSPP, ROOT, Scilab, Salstat, SciPy, SOFA statistics, Statistical lab, and others. Examples of public domain SS are X-12 ARIMA, Epi info, and CSPro. Examples of freeware SS are BV4.1, Winpepi, GeoDA, WinBUGS, MINUIT, and MaxStat. Examples of proprietary SS are SPSS, Minitab, Analytica, JMP, LISREL, Maple, Mathematica, MATLAB, OriginPro, SAS, SigmaStat, Statgraphics, Statistica, S-PLUS, and others.

SPSS and Minitab are very common among researchers that do not have competencies in using advanced statistical tools. The two statistical software is easy to use because they are designed mainly to be a graphical user interface, although some coding may be involved for advanced data analysis [2-7]. The two software can handle complex data manipulation with simple predefined steps [8]. SPSS is popularly known as Statistical Package for the Social Sciences or Statistical Product and Service Solutions. SPSS is a widely used statistical software for statistical analysis in social science and market analysis. IBM Inc now owns SPSS. Minitab is a statistics package developed at Pennsylvania State University and distributed by Minitab, LLC, a privately owned company in Pennsylvania. Minitab is often used in process improvement and quality assurance [9-12]. Extensions exist for both software that will enable them to interface with other statistical software or packages in the form of macros or add-ins.

SPSS and Minitab can be used in the following; data transformation, regression analysis, analysis of variance, multivariate analysis of variance, analysis of covariance, t-tests, non-parametric tests, time series, design and analysis of experiments, spatial analysis, survival analysis, dimension reduction, reliability, factor analysis, correspondence analysis, neural network, correlation and others [13-17]. The choice of data analysis will ultimately depend on the suitability of the statistical software [18]. Some selected works that used both statistical software in their data analysis can be assessed in [19-23].

This paper reports the data analysis of the search done on the Scopus database to determine the frequency of the use of SPSS and Minitab across 27 major subject classification of Scopus. The analysis is further extended to the nature of the documents across the subject classifications. The Scopus subject classifications are; medicine (MED), social sciences (SOC), engineering (ENG), business management and accounting (BMA), biochemistry and molecular biology (BMB), computer science (CSC), nursing (NUR), agricultural and biological sciences (ABS), environmental sciences (ESC), pharmacology, toxicology and pharmaceutical (PTP), arts and humanities (AAH), dentistry (DEN), psychology (PSY), health professions (HPR), physics and astronomy (PAA), decision sciences, immunology and microbiology (IAM), economics, econometrics and finance (EEF), energy (ENE), materials science (MSC), earth and planetary sciences (EPS), multidisciplinary (MUL), neuroscience (NEU), mathematics (MAT), chemical engineering (CEN), chemistry (CHE), veterinary (VET) and undefined (UND).

2. Methods
The summary of the materials used and the adopted methodology are outlined.

Data source: Scopus database.
Scope: The research considered only papers whose abstracts or keywords contain the terms ‘SPSS’ and ‘Minitab’. All document types from the search were analyzed.
Search method: The terms “SPSS” and “Minitab” were searched independently, and the number of times the terms appeared either in the abstracts or keywords or indexed keywords was extracted. Papers in the undefined status were classified as a subject. The search terms were able to bring out results because it is ethical for scholars to report the software they used in their data analysis either in the abstracts or the body of their research works.
Data range: The displayed results from 2010 to 2019 were extracted.
Search period: This query was done on May 30, 2020.
Data extraction: Raw data were retrieved manually from the Scopus database.
Raw data format: Excel.

Data analysis: This was done using Microsoft Excel, SPSS 23.0 and Minitab 17.0.

Statistical analysis: Frequency table, trend plot and Spearman rank correlation. T test was not used because the frequency varies significantly from the frequency analysis.

Data availability: The data is available on request from interested researchers.

3. Result

The search yielded different results for SPSS and Minitab for the years 2010 to 2019, and these are presented in Tables 1 and 2. Medicine (MED), social sciences (SOC), engineering (ENG), business management and accounting (BMA), and biochemistry and molecular biology (BMB) are the top 5 subject areas where SPSS has been used for data analysis in papers indexed in Scopus from 2010 to 2019. On the other hand, mathematics (MAT), chemical engineering (CEN), chemistry (CHE), veterinary (VET) and undefined (UND) are the least five subject areas that make mention of SPSS in their abstracts or keywords.

| SUBJECT | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | Total |
|---------|------|------|------|------|------|------|------|------|------|------|-------|
| MED | 4741 | 4464 | 4006 | 3702 | 3138 | 3046 | 2691 | 2167 | 1784 | 1422 | 31161 |
| SOC | 1254 | 837 | 665 | 567 | 446 | 430 | 309 | 275 | 311 | 238 | 5332 |
| ENG | 893 | 561 | 473 | 378 | 319 | 323 | 278 | 208 | 217 | 161 | 3811 |
| BMA | 761 | 465 | 396 | 351 | 186 | 146 | 115 | 94 | 100 | 82 | 2696 |
| BMB | 656 | 797 | 638 | 713 | 553 | 523 | 498 | 355 | 244 | 153 | 5130 |
| CSC | 649 | 436 | 409 | 279 | 215 | 164 | 139 | 135 | 161 | 153 | 2740 |
| NUR | 567 | 481 | 371 | 351 | 298 | 301 | 255 | 208 | 160 | 168 | 3160 |
| ABS | 418 | 392 | 294 | 272 | 247 | 402 | 249 | 166 | 133 | 130 | 2703 |
| ESC | 393 | 325 | 239 | 169 | 140 | 293 | 149 | 105 | 93 | 61 | 1967 |
| PTP | 385 | 422 | 306 | 442 | 306 | 221 | 157 | 123 | 106 | 71 | 2539 |
| AAH | 328 | 157 | 108 | 91 | 151 | 156 | 83 | 62 | 38 | 25 | 1199 |
| DEN | 324 | 361 | 279 | 215 | 169 | 131 | 116 | 109 | 93 | 70 | 1867 |
| PSY | 223 | 143 | 128 | 100 | 61 | 52 | 65 | 77 | 124 | 83 | 1056 |
| HPR | 195 | 168 | 119 | 116 | 93 | 82 | 74 | 89 | 54 | 54 | 1044 |
| PAA | 182 | 129 | 59 | 68 | 29 | 33 | 21 | 22 | 17 | 30 | 590 |
| DEC | 178 | 131 | 106 | 74 | 49 | 31 | 36 | 19 | 26 | 44 | 694 |
| IAM | 150 | 157 | 102 | 153 | 90 | 108 | 95 | 70 | 44 | 37 | 1006 |
| EEF | 146 | 159 | 230 | 159 | 161 | 138 | 60 | 47 | 28 | 18 | 1146 |
| ENE | 118 | 102 | 124 | 54 | 48 | 34 | 23 | 18 | 15 | 10 | 546 |
| MSC | 116 | 127 | 100 | 62 | 47 | 45 | 28 | 41 | 38 | 33 | 637 |
| EPS | 115 | 109 | 49 | 40 | 28 | 39 | 27 | 34 | 43 | 22 | 506 |
| MUL | 115 | 145 | 72 | 165 | 121 | 82 | 132 | 105 | 93 | 22 | 1052 |
| NEU | 114 | 124 | 110 | 91 | 58 | 55 | 57 | 41 | 32 | 45 | 727 |
| MAT | 112 | 125 | 163 | 86 | 68 | 37 | 47 | 53 | 59 | 47 | 797 |
| CEN | 58 | 94 | 51 | 67 | 60 | 41 | 23 | 23 | 13 | 13 | 443 |
| CHE | 57 | 85 | 38 | 59 | 47 | 33 | 35 | 27 | 15 | 14 | 410 |
| VET | 47 | 36 | 51 | 35 | 31 | 24 | 31 | 31 | 14 | 16 | 316 |
| UND | 0 | 1 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 4 |

Total 8733 7947 6719 6437 5211 5114 4459 3554 3018 2323 53515
Table 2: Frequency of the use of Minitab from 2010 to 2019

| SUBJECT | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | Total |
|---------|------|------|------|------|------|------|------|------|------|------|-------|
| MED     | 15   | 18   | 18   | 18   | 15   | 18   | 18   | 18   | 22   | 20   | 169   |
| SOC     | 4    | 5    | 2    | 0    | 4    | 4    | 5    | 9    | 3    | 8    | 47    |
| ENG     | 112  | 112  | 88   | 77   | 51   | 62   | 59   | 40   | 37   | 36   | 674   |
| BMA     | 17   | 17   | 14   | 12   | 7    | 8    | 5    | 2    | 3    | 2    | 87    |
| BMB     | 11   | 14   | 14   | 14   | 10   | 6    | 6    | 7    | 2    | 9    | 98    |
| CSC     | 35   | 30   | 25   | 10   | 11   | 12   | 8    | 8    | 8    | 8    | 158   |
| NUR     | 5    | 3    | 0    | 1    | 2    | 1    | 2    | 1    | 0    | 15   | 15    |
| ABS     | 29   | 20   | 14   | 12   | 10   | 10   | 6    | 7    | 2    | 9    | 148   |
| ESC     | 22   | 30   | 16   | 15   | 13   | 19   | 15   | 6    | 7    | 9    | 152   |
| PTP     | 5    | 6    | 10   | 15   | 13   | 8    | 3    | 4    | 5    | 2    | 71    |
| AAH     | 0    | 1    | 1    | 1    | 1    | 1    | 2    | 0    | 1    | 9    | 9     |
| DEN     | 1    | 3    | 1    | 0    | 0    | 3    | 1    | 2    | 2    | 0    | 13    |
| PSY     | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 1    | 1     |
| HPR     | 0    | 1    | 1    | 0    | 1    | 0    | 1    | 0    | 1    | 5    | 5     |
| PAA     | 28   | 24   | 13   | 14   | 9    | 8    | 10   | 10   | 9    | 8    | 133   |
| DEC     | 6    | 10   | 12   | 7    | 1    | 4    | 1    | 1    | 0    | 2    | 44    |
| IAM     | 2    | 4    | 4    | 6    | 4    | 1    | 2    | 5    | 2    | 36   | 36    |
| EEF     | 0    | 2    | 2    | 1    | 3    | 1    | 1    | 0    | 0    | 0    | 11    |
| ENE     | 18   | 13   | 14   | 8    | 12   | 7    | 5    | 3    | 8    | 5    | 93    |
| MSC     | 55   | 75   | 41   | 34   | 29   | 14   | 15   | 24   | 15   | 14   | 316   |
| EPS     | 6    | 12   | 6    | 2    | 3    | 3    | 3    | 4    | 4    | 4    | 46    |
| MUL     | 7    | 5    | 1    | 7    | 4    | 1    | 2    | 4    | 5    | 3    | 39    |
| NEU     | 1    | 0    | 1    | 0    | 0    | 0    | 1    | 0    | 0    | 1    | 4     |
| MAT     | 8    | 12   | 8    | 8    | 9    | 10   | 6    | 10   | 9    | 5    | 85    |
| CEN     | 28   | 27   | 16   | 13   | 14   | 13   | 11   | 10   | 10   | 10   | 158   |
| CHE     | 16   | 23   | 18   | 15   | 16   | 11   | 9    | 11   | 5    | 8    | 132   |
| VET     | 3    | 5    | 4    | 3    | 3    | 2    | 1    | 1    | 2    | 1    | 25    |
| UND     | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0     |
| Total   | 252  | 267  | 183  | 176  | 147  | 149  | 136  | 116  | 99   | 83   | 1608  |

Engineering (ENG), materials science (MSC), computer science (CSC), agricultural and biological sciences (ABS), physics and astronomy (PAA) and chemical engineering (CEN) are the top 6 subject areas where Minitab has been used for data analysis in papers indexed in Scopus from 2010 to 2019. On the other hand, arts and humanities (AAH), psychology (PSY), health professions (HPR), economics, econometrics and finance (EEF) and undefined (UND) are the least five subject areas that make mention of Minitab in their abstracts or keywords.

A further look at Tables 1 and 2 shows that Minitab's usage pattern is a fraction of SPSS, which implies that researchers used SPSS more than Minitab.

The frequency analysis is depicted on a trend plot to show the two statistical software trends and can be seen in Figures 1 and 2.

It could be seen from Figure 1, that the trend of the use of SPSS is increasing steadily The trend of the use of Minitab is increasing steadily, although, there is a downward trend from 2018 to 2019.
Figure 1: The trend of SPSS usage in scientific research

Figure 2: The trend of Minitab usage in scientific research

Table 3: Document type that have mentioned SPSS in their abstracts or keywords

| Document type      | Frequency |
|--------------------|-----------|
| Article            | 49104     |
| Conference Paper   | 3141      |
| Review             | 741       |
| Book Chapter       | 229       |
| Undefined          | 84        |
| Data Paper         | 68        |
| Book               | 57        |
| Note               | 35        |
| Erratum            | 31        |
| Retracted          | 20        |
| Letter             | 15        |
| Editorial          | 11        |
| Short Survey       | 9         |
| Conference Review  | 6         |
| Abstract Report    | 1         |

Table 4: Document type that have mentioned Minitab in their abstracts or keywords

| Document type      | Frequency |
|--------------------|-----------|
| Article            | 1131      |
| Conference Paper   | 420       |
| Book               | 21        |
| Book Chapter       | 15        |
| Review             | 9         |
| Data Paper         | 4         |
| Undefined          | 3         |
| Conference Review  | 2         |
| Short Survey       | 1         |
| Note               | 0         |
| Erratum            | 0         |
| Retracted          | 0         |
| Letter             | 0         |
| Editorial          | 0         |
| Abstract Report    | 0         |

The breakdown of the document type from Table 3 showed that article, conference paper, review, and book chapter are the ones with the most frequency in decreasing order for SPSS. According to Minitab, the most top document type is article, conference paper, book, and book chapter (Table 4).

Spearman rank correlation was performed to show the relationship between the SPSS and Minitab data from 2010 to 2019. The result presented in Table 5 showed no significant correlation between the pair for years, as mentioned above. The pattern of Tables 1 and 2 are different for SPSS and Minitab.
Table 5: Correlation between the documents that have stated their use of SPSS and Minitab in their papers.

| Year | Correlation |
|------|-------------|
| 2019 | 0.223       |
| 2018 | 0.169       |
| 2017 | 0.180       |
| 2016 | 0.255       |
| 2015 | 0.257       |
| 2014 | 0.351       |
| 2013 | 0.207       |
| 2012 | 0.221       |
| 2011 | 0.225       |
| 2010 | 0.044       |

4. Conclusion
SPSS and Minitab will continue to be most sought statistical data analysis tools, especially for those with low statistical or mathematical competencies. The trend analysis showed that SPSS is utilized more than Minitab, and no correlation exists between the two. Hence, further research works are needed to show if competency in one software affects the other. The extent to which the two statistical software is to be used will greatly depend on the nature of the analysis. SPSS and Minitab will continue to be patronized by researchers, the areas of application notwithstanding. Adequate awareness is needed to educate researchers on the use of Minitab since the present work has shown that it is far less utilized than SPSS.

Acknowledgement
Covenant University is acknowledged for creating an enabling environment that is suitable for research.

Reference
[1] Begum, K. J., & Ahmed, A. (2015). The importance of statistical tools in research work. International Journal of Scientific and Innovative Mathematical Research, 3(12), 50-58.
[2] Cronk, B. C. (2019). How to use SPSS®: A step-by-step guide to analysis and interpretation. Routledge.
[3] Wagner III, W. E. (2019). Using IBM® SPSS® statistics for research methods and social science statistics. Sage Publications.
[4] Amos, I. S. (2019). IBM SPSS Amos Overview.
[5] Baarda, B., & van Dijkum, C. (2019). Introduction to Statistics with SPSS. Routledge.
[6] Rajathi, A., & Chandran, P. (2019). SPSS for You. MJP Publisher.
[7] Trihendradi, C. (2019). Step by step IBM SPSS 21: Analisis Data Statistik.
[8] Capilla, C. (2015). The Role of Statistical Software in Teaching Data Analysis. INTED2015 Proceedings, 1137-1144.
[9] De Muth, J. E. (2019). Practical Statistics for Pharmaceutical Analysis: With Minitab Applications (Vol. 40). Springer Nature.
[10] Allen, T. T. (2019). Software overview and methods review: Minitab. In Introduction to Engineering Statistics and Lean Six Sigma (pp. 575-600). Springer, London.
[11] Fardillah, F., Ruhimat, A., & Priatna, N. (2020). Self regulated Learning Student Through Teaching Materials Statistik Based on Minitab Software. Journal of Physics: Conference
Parthasarathy, R. (2019). Applied Statistics Manual: A Guide to Improving and Sustaining Quality with Minitab. *Quality Progress*, 52(8), 62-62.

Argyrous, G. (2005-11-23). Statistics for Research: With a Guide to SPSS. London: SAGE. ISBN 978-1-4129-1948-7.

Bryman, A. & Cramer, D. (2011). Quantitative Data Analysis with IBM SPSS 17, 18 and 19: A Guide for Social Scientists. New York: Routledge. ISBN 978-0-415-57918-6.

Levesque, R. (2007). SPSS Programming and Data Management: A Guide for SPSS and SAS Users (4th ed.). Chicago, Illinois: SPSS Inc. ISBN 978-1-56827-390-7.

SPSS 15.0 (2006). Command Syntax Reference. Chicago, Illinois: SPSS Inc.

Wellman, B. (1998). "Doing It Ourselves: The SPSS Manual as Sociology's Most Influential Recent Book". In Clawson, Dan (ed.). Required Reading: Sociology's Most Influential Books. Amherst: University of Massachusetts Press. pp. 71–78. ISBN 978-1-55849-153-3.

Abatan, S. M. & Olayemi, M. (2014). The Role of Statistical Software in Data Analysis. *International Journal of Applied Research and Studies*, 3(8). Available at SSRN: [https://ssrn.com/abstract=2532326](https://ssrn.com/abstract=2532326).

Okagbue, H. I., Adamu, M. O., Oguntunde, P. E., Opanuga, A. A., Owoloko, E. A., & Bishop, S. A. (2017). Datasets on the statistical and algebraic properties of primitive Pythagorean triples. *Data in Brief*, 14, 686-694.

Okagbue, H. I., & Teixeira da Silva, J. A. (2020). Correlation between the CiteScore and Journal Impact Factor of top-ranked library and information science journals. *Scientometrics*, 124, 797-801.

Korableva, O. N., Kalimullina, O. V., & Mityakova, V. N. (2019). Designing a system for integration of macroeconomic and statistical data based on ontology. Advances in Intelligent Systems and Computing, 998, 157-165. doi:10.1007/978-3-030-22868-2_12

Klochko, E., & Brizhak, O. (2019). Prospects of using virtual technologies in modern corporate business systems. Advances in Intelligent Systems and Computing, 726, 308-319. doi:10.1007/978-3-319-90835-9_36

Rosa, A. T. R., Pustokhina, I. V., Lydia, E. L., Shankar, K., & Huda, M. (2019). Concept of electronic document management system (EDMS) as an efficient tool for storing document. *Journal of Critical Reviews*, 6(5), 85-90. doi:10.22159/jcr.06.05.14