General Mathematics Vol. 32, No. 1 (2024), 73–86

DOI: 10.2478/gm-2024-0006

sciendo

Evaluating Data Science Tools: A Comparative Analysis of Platforms [1]

Emanuelle I. Perța

Abstract

This study targets organizations in the data science domain seeking optimal platforms for their operations. It employs experimental research using the comparison method and mathematical modeling integrated into the Global Utility Method and TOPSIS Method. The aim is to create hierarchies of three prominent data science platforms: IBM Watson Studio, SAS – Advanced Analytics Software and Google Cloud AI Platform. The analysis was conducted using Microsoft Excel. Various comparison criteria are meticulously applied to ensure mathematical correctness and comprehensiveness. The Data Science Platforms Conclusions, drawn from individual case studies, indicate that different methods yield similar results due to algorithmic variations and data sensitivity. Minor changes were observed in hierarchy positions, attributed to coefficient importance and method sensitivity. There is a necessity for multiple platforms to meet diverse organizational needs. Continuous development and innovation are crucial in this evolving field, where user preferences and technological advancements drive platform suitability.

2010 Mathematics Subject Classification: 68U35.

Key words and phrases: Data Science Platforms, Comparative Analysis, TOPSIS Method, Global Utility Method, Sensitivity Analysis.

[1] *Received 26 March, 2024*
Accepted for publication (in revised form) 22 July, 2024

1 Introduction

Data science involves much more than just data extraction algorithms. Successful individuals in the field of data science must be able to view business problems from a data perspective. There is a fundamental structure for analytical thinking of data and basic principles that should be understood. Data science is a multidisciplinary approach that draws on several "traditional" fields. Much of what has traditionally been studied in the field of statistics is fundamental to data science. Data visualization methods and methodology are vital. There are also certain areas where intuition, creativity, and knowledge of a particular application must be applied. A data science perspective provides practitioners with a framework and principles that give the researcher a systematic way to address issues of extracting useful information from data. [3] In other words, combining elements of mathematics, statistics, and artificial intelligence are used by data science to analyze large amounts of data. Data science is a field that employs interdisciplinary skills and technologies to extract knowledge and insights from data. This data can be structured or unstructured and may come from a variety of sources, including mobile devices, social networks, IoT sensors, historical data, and many others.

1.1 Background

In general, data science platforms are an important tool for organizations looking to extract insights and value from their data. While open-source platforms based on R and Python continue to be popular among data professionals, commercial platforms such as IBM Watson Studio, Microsoft Azure Machine Learning, and Google Cloud AI Platform are also gaining traction due to their ease of use and integrations with other tools. Further research is needed to evaluate the effectiveness of these platforms for different use cases, as well as to identify areas for improvement and innovation.

The October 2012 issue of the Harvard Business Review contained three articles relevant to this editorial: "Big Data: The Management Revolution" (McAfee and Brynjolfsson 2012), "Data Scientist: The Sexiest Job of the 21st Century" (Davenport and Patil 2012), and "Making Advanced Analytics Work for You" (Barton and Court 2012). MIS Quarterly had a special issue on business intelligence, with the main article titled "Business Intelligence and Analytics: From Big Data to Big Impact" (Chen et al. 2012). There are also numerous articles in trade and even lay publications on these topics. There is even a journal, Big Data, which premiered in March 2013. [6]

Mathematics is the foundation of any contemporary scientific discipline. Almost all data science techniques, including machine learning, are based on mathematical algorithms. It goes without saying that all other elements of knowledge are needed - programming skills, a certain insight, and an analytical and curious mindset. Therefore, a solid understanding of the mathematical mechanisms behind algorithms provides an advantage. Knowledge of essential mathematics is particularly important

for other professions as well: hardware engineering, sales, chemical process industry, medicine, business management, etc. Although such fields may require experience with various supporting platforms, the mathematical skills needed in data science can significantly change the understanding of all processes in this field. [5]

1.2 Global Utility Method

The Global Utility Method involves calculating the global utility of each decision option and selecting the optimal option based on the maximum global utility. Decision options are denoted as V_i, and decision criteria as C_j. [1] Although decision criteria can be either maximized or minimized, this method utilizes maximization, following this formula:

$$U_{ij} = \frac{a_{ij} - \min a_{ij}}{\max a_{ij} - \min a_{ij}}, \tag{1}$$

where:

- U_{ij} represents the utility of option V_i according to criterion C_j.
- a_{ij} signifies the consequence of option V_i according to criterion C_j.
- $\max a_{ij}$ is the highest consequence among all, according to criterion j.
- $\min a_{ij}$ is the lowest consequence among all, according to criterion j.

$$f(V_i) = \sum_{j=1}^{n} U_{ij} \cdot k_j \tag{2}$$

$$V_{\text{opt}} = \max f(V_i) \tag{3}$$

1.3 Technique for Order Preference by Similarity to Ideal Solution

TOPSIS has as its principle the construction of the positive ideal solution and the negative ideal solution, and the options are ranked based on their distances to these two solutions. For this method, the following algorithm is used [2]:

1. Construction of the normalized matrix:

   $$R = (r_{ij}), i = 1, 2, \ldots, m; j = 1, 2, \ldots, n \tag{4}$$

   through vector normalization:

   $$r_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^{m} a_{ij}^2}}, i = 1, 2, \ldots, m; j = 1, 2, \ldots, n \tag{5}$$

2. Construction of the weighted normalized matrix:

$$V = (v_{ij}), i = 1, 2, \ldots, m; j = 1, 2, \ldots, n; \text{where } v_{ij} = k_j \cdot r_{ij} \tag{6}$$

3. Determination of the positive ideal solution V^{po} and the negative ideal solution V^{ne}:

$$V^{po} = (v_1^{po}, v_2^{po}, \ldots, v_n^{po}) \quad V^{ne} = (v_1^{ne}, v_2^{ne}, \ldots, v_n^{ne}) \tag{7}$$

Criterion C_j is maximized:

$$V_j^{po} = \max_{1 \le i \le m} v_{ij} \quad \text{and} \quad V_j^{ne} = \min_{1 \le i \le m} v_{ij} \tag{8}$$

4. Calculating the distances between options and the two previously calculated solutions:

$$d_i^{po} = \sqrt{\sum_{j=1}^{n} (v_{ij} - V_j^{po})^2} \quad \text{and} \quad d_i^{ne} = \sqrt{\sum_{j=1}^{n} (v_{ij} - V_j^{ne})^2}, i = 1, 2, \ldots, n \tag{9}$$

5. Calculating the relative closeness to the ideal solution:

$$e_i^{po} = 1 - \frac{d_i^{po}}{d_i^{po} + d_i^{ne}} = \frac{d_i^{ne}}{d_i^{po} + d_i^{ne}}, i = 1, 2, \ldots, n \tag{10}$$

$$so, 0 \le e_i^{po} \le 1. \tag{11}$$

6. Ranking based on decreasing values of e_i^{po}.

2 Analysis No.1

Following the Literature Review, a set of 4 criteria for evaluating data science platforms has been established. These will measure the ease of use and feasibility of each criterion. [4]
C_1– Data collection;
C_2– Data inconsistency elimination;
C_3– Data security;
C_4– Efficient collaboration.
The three platforms (Vi) to be analyzed are:
V_1–IBM Watson Studio;
V_2– SAS – Advanced Analytics Software;
V_3– Google Cloud AI Platform.

Table 1: Consequence Matrix and Coefficients

Decision Options / Evaluation Criteria	C1	C2	C3	C4
V_1	4.4	4.3	4.2	4.2
V_2	4.4	4.5	3.8	4.3
V_3	4.4	4.6	3.9	4.4
Importance coefficient of criterion k_j	0.2	0.1	0.4	0.3

The utility matrix illustrated in Table 2 is as follows.

Table 2: Utility Matrix

Decision Options / Evaluation Criteria	C1	C2	C3	C4
V_1	0	0	1	0
V_2	0	0,7	0	0,5
V_3	1	1	0,25	1
Importance coefficient of criterion k_j	0.2	0.1	0.4	0.3

2.1 Analysis No. 1 using the Global Utility Method

For the analysis using the Global Utility Method, the consequence matrix and the importance coefficients of each criterion are constructed in Table 1. All these criteria will be considered as maximization criteria. Utilities are calculated using relation (1).

To determine the global utility of each decision option, relation (2) is used, which is then applied to V_i, as follows:

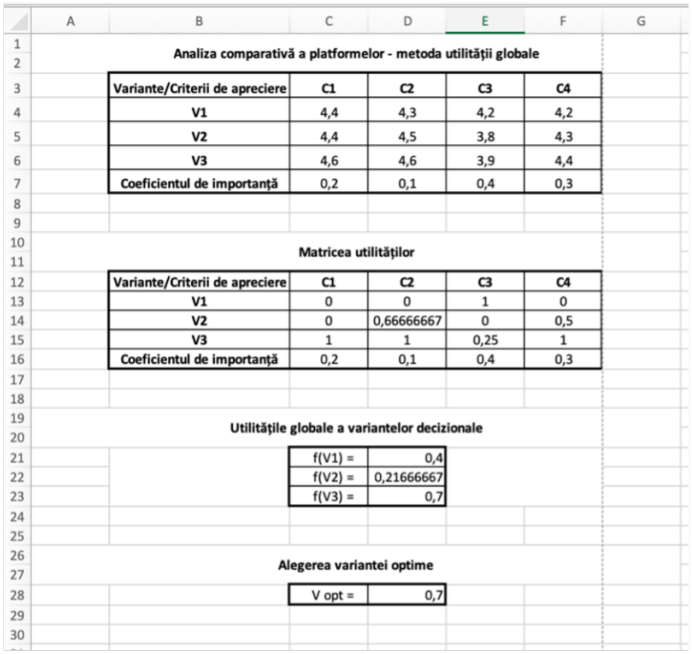
$$f(V_1) = 0 \times 0.2 + 0 \times 0.1 + 1 \times 0.4 + 0 \times 0.3 = 0.4$$
$$f(V_2) = 0 \times 0.2 + 0.7 \times 0.1 + 0 \times 0.4 + 0.5 \times 0.3 = 0.22$$
$$f(V_3) = 1 \times 0.2 + 1 \times 0.1 + 0.25 \times 0.4 + 1 \times 0.3 = 0.7$$

Therefore, since the optimal option is the one with the maximum value, we have that $\max(0.4, 0.22, 0.7) = 0.7$. Thus, option V_3 is the desired one. In Figure 1, the solution is also observed using Excel, and the results are easy to verify.

The conclusions of this analysis emphasize that the optimal option among the considered data science platforms, under the condition that the importance coefficient is maximum for the Data Security criterion, is Google Cloud AI Platform, although it is observed that it does not have the highest utility among the three platforms. The next platform is IBM Watson Studio, and the last one is SAS – Advanced Analytics Software.

Furthermore, conducting a sensitivity analysis aims to analyze a possible change in the hierarchy of platforms. Sensitivity analysis refers to studying how the optimal solution changes when the importance coefficients of the criteria change.

The importance coefficients will be modified in the previous utility matrix. This time, Data Cleaning will be the condition with the highest coefficient.



Analiza comparativă a platformelor - metoda utilității globale

Variante/Criterii de apreciere	C1	C2	C3	C4
V1	4,4	4,3	4,2	4,2
V2	4,4	4,5	3,8	4,3
V3	4,6	4,6	3,9	4,4
Coeficientul de importanță	0,2	0,1	0,4	0,3

Matricea utilităților

Variante/Criterii de apreciere	C1	C2	C3	C4
V1	0	0	1	0
V2	0	0,66666667	0	0,5
V3	1	1	0,25	1
Coeficientul de importanță	0,2	0,1	0,4	0,3

Utilitățile globale a variantelor decizionale

f(V1) =	0,4
f(V2) =	0,21666667
f(V3) =	0,7

Alegerea variantei optime

V opt =	0,7

Figure 1: Global Utility Method in Excel

Table 3: Utility Matrix for the sensitivity analysis

Decision Options / Evaluation Criteria	C1	C2	C3	C4
V_1	0	0	1	0
V_2	0	0,7	0	0,5
V_3	1	1	0,25	1
Importance coefficient of criterion k_j	0.1	0.5	0.3	0.1

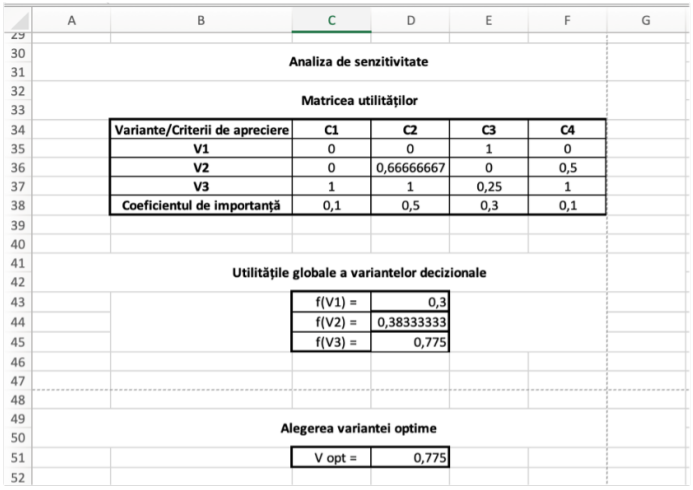
The global utilities of each decision option are as follows:

$$f(V_1) = 0 \times 0.1 + 0 \times 0.5 + 1 \times 0.3 + 0 \times 0.1 = 0.3$$
$$f(V_2) = 0 \times 0.1 + 0.7 \times 0.5 + 0 \times 0.3 + 0.5 \times 0.1 = 0.4$$
$$f(V_3) = 1 \times 0.1 + 1 \times 0.5 + 0.25 \times 0.3 + 1 \times 0.1 = 0.775$$

And once again, it can be observed that the hierarchy of platforms remains the same, as Google Cloud AI Platform has predominantly utility values of 1 (Figure 2). For this reason, the platform ranks first, regardless of changes in coefficients. However, a change can be observed in the other two positions. This demonstrates that as a result of sensitivity analysis, the results can change.

2.2 Analysis No. 1 using the TOPSIS Method

For the TOPSIS method, starting from the Consequence Matrix and Coefficients considered in Table 1, the normalized matrix is calculated in Figure 3.



Analiza de senzitivitate

Matricea utilităților

Variante/Criterii de apreciere	C1	C2	C3	C4
V1	0	0	1	0
V2	0	0,66666667	0	0,5
V3	1	1	0,25	1
Coeficientul de importanță	0,1	0,5	0,3	0,1

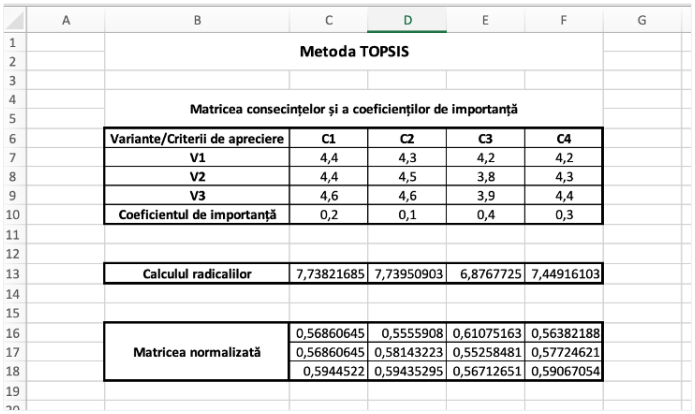
Utilitățile globale a variantelor decizionale

f(V1) =	0,3
f(V2) =	0,38333333
f(V3) =	0,775

Alegerea variantei optime

V opt =	0,775

Figure 2: Sensitivity Analysis in Excel



Metoda TOPSIS

Matricea consecințelor și a coeficienților de importanță

Variante/Criterii de apreciere	C1	C2	C3	C4
V1	4,4	4,3	4,2	4,2
V2	4,4	4,5	3,8	4,3
V3	4,6	4,6	3,9	4,4
Coeficientul de importanță	0,2	0,1	0,4	0,3

Calculul radicalilor	7,73821685	7,73950903	6,8767725	7,44916103

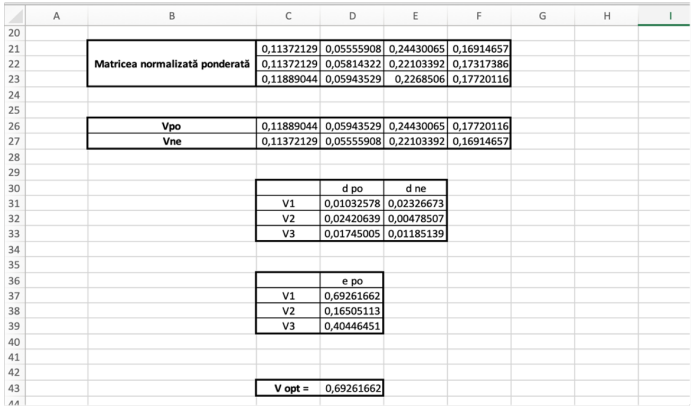
Matricea normalizată	0,56860645	0,5555908	0,61075163	0,56382188
	0,56860645	0,58143223	0,55258481	0,57724621
	0,5944522	0,59435295	0,56712651	0,59067054

Figure 3: Calculation of the normalized matrix in Excel

For the following steps, the algorithm implementation has been done in Excel due to their complexity. The optimal option will be the one with the maximum value, meaning the platforms will be ranked in descending order. Thus, the optimal option is IBM Watson Platform, followed by Google Cloud AI Platform, and lastly SAS – Advanced Analytics Software. The results can be seen in Figure 4.

2.3 Conclusions for Analysis No. 1

According to the first analysis conducted using two methods, one of which underwent sensitivity analysis, a change in the platform order can be observed after altering the calculation method. Both methods result in different but still close rankings. It is noticeable that SAS – Advanced Analytics Software ranks last in both methods, but according to the sensitivity analysis applied to the first method, the platform



Matricea normalizată ponderată	0,11372129	0,05555908	0,24430065	0,16914657
	0,11372129	0,05814322	0,22103392	0,17317386
	0,11889044	0,05943529	0,2268506	0,17720116
Vpo	0,11889044	0,05943529	0,24430065	0,17720116
Vne	0,11372129	0,05555908	0,22103392	0,16914657

	d po	d ne
V1	0,01032578	0,02326673
V2	0,02420639	0,00478507
V3	0,01745005	0,01185139

	e po
V1	0,69261662
V2	0,16505113
V3	0,40446451

V opt =	0,69261662

Figure 4: Resolution through the TOPSIS method in Excel

reaches the second position. The sensitivity analysis caused a change in the last two positions but left Google Cloud AI Platform in the first position. It can be confirmed that there is no single platform that is the optimal solution, regardless of the method. In Analysis No. 1, we considered four criteria that are commonly encountered issues in specialized organizations. These criteria were chosen initially to establish an overview of the common needs of organizations, with other criteria considered in subsequent analyses.

3 Analysis No.2

The next criteria will be chosen in such a way that there is a connection between them. This aspect is crucial when it comes to comparison, ensuring that the analysis results are clear and accurate. For the second analysis of platforms, different criteria will be chosen for evaluation. These criteria will measure the ease of use and the feasibility of the criterion, focusing on artificial intelligence-specific aspects. Once again, we will aim for maximization, so the problems will be approached using the same formulas.

C_1– Predictive Analysis: Utilizes data extraction methods, statistics, machine learning, mathematical modeling, and artificial intelligence to make future predictions about unknown events.

C_2– Forecasting: Predicts a sequence of values by estimating based on previous data.

C_3– Simulation: Uses large quantities of data and usage to simulate scenarios or conditions from the real world.

C_4– Automation: Artificially increases the amount of data by generating new data points from existing data.

C_5– Machine Learning: Automatically analyzes large amounts of data, automating the data analysis process and making real-time data-based predictions without any human intervention, using machine learning algorithms.

Table 4: Consequence Matrix and Importance Coefficients

Decision Options / Evaluation Criteria	C_1	C_2	C_3	C_4	C_5	C_6
V_1	4,3	4,1	3,9	4,2	4,2	4,3
V_2	4,2	4,1	3,9	4,3	4,4	4,5
V_3	4,5	4,3	4,3	4,5	4,5	4,4
Importance coefficient of criterion k_j	0,20	0,45	0,15	0,10	0,05	0,05

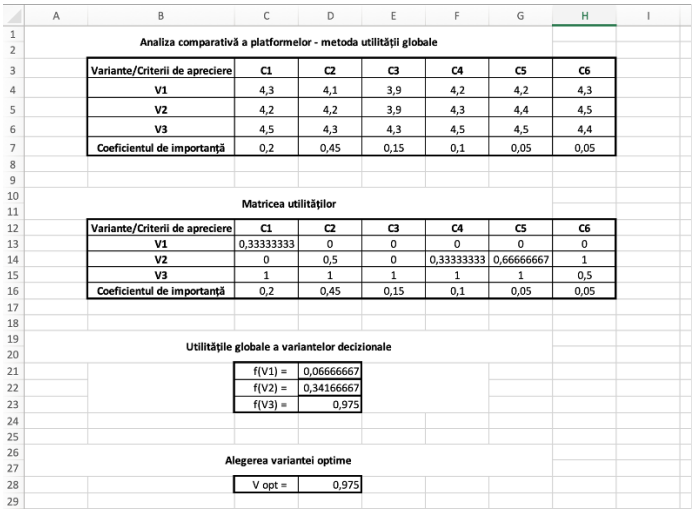
C_6– Other Advanced Analyses: Data analysis methodology that uses deep learning, cognitive analysis, and other statistical methods to analyze information from a variety of data sources.

3.1 Analysis No. 2 using the Global Utility Method

The specific calculations have been performed in Excel, resulting in obtaining the results quickly and with high accuracy. This way, the results can be achieved in a very short time, and with much greater precision. These can be seen in Figure 5, where the resolution of the problem according to the method used is clear and complete.

After the calculations, $f(V_3) > f(V_2) > f(V_1)$, the ranking of the analyzed platforms will be as follows:

1. Google Cloud AI Platform;
2. SAS – Advanced Analytics Software;
3. IBM Watson Studio.



Analiza comparativă a platformelor - metoda utilității globale

Variante/Criterii de apreciere	C1	C2	C3	C4	C5	C6
V1	4,3	4,1	3,9	4,2	4,2	4,3
V2	4,2	4,2	3,9	4,3	4,4	4,5
V3	4,5	4,3	4,3	4,5	4,5	4,4
Coeficientul de importanță	0,2	0,45	0,15	0,1	0,05	0,05

Matricea utilităților

Variante/Criterii de apreciere	C1	C2	C3	C4	C5	C6
V1	0,33333333	0	0	0	0	0
V2	0	0,5	0	0,33333333	0,66666667	1
V3	1	1	1	1	1	0,5
Coeficientul de importanță	0,2	0,45	0,15	0,1	0,05	0,05

Utilitățile globale a variantelor decizionale

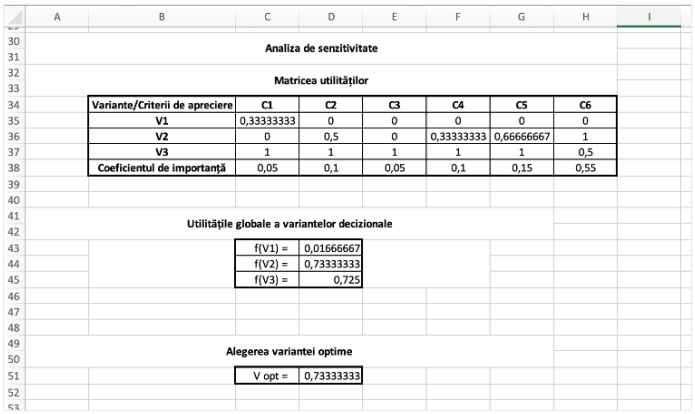
f(V1) =	0,06666667
f(V2) =	0,34166667
f(V3) =	0,975

Alegerea variantei optime

V opt =	0,975

Figure 5: Problem Resolution Using the Global Utility Method in Excel

By applying a sensitivity analysis, I am looking for a certain change that could occur in that hierarchy. Thus, we will have: In Figure 6, it can be observed that $f(V_2) > f(V_3) > f(V_1)$. This means that the hierarchy of platforms will be reversed for the first two positions, but IBM Watson Studio remains in the third position.



Analiza de senzitivitate

Matricea utilităților

Variante/Criterii de apreciere	C1	C2	C3	C4	C5	C6
V1	0,33333333	0	0	0	0	0
V2	0	0,5	0	0,33333333	0,66666667	1
V3	1	1	1	1	1	0,5
Coeficientul de importanță	0,05	0,1	0,05	0,1	0,15	0,55

Utilitățile globale a variantelor decizionale

f(V1) =	0,01666667
f(V2) =	0,73333333
f(V3) =	0,725

Alegerea variantei optime

V opt =	0,73333333

Figure 6: Problem Resolution Using the Global Utility Method in Excel

3.2 Analysis No. 2 using the TOPSIS Method

After inputting the initial data and performing the relevant calculations in MS Excel (Figure 7), the optimal result was obtained, which is also in this case still option 3, the Google Cloud AI Platform. Additionally, the hierarchy of platforms is the same as in the global utility method, which strengthens these results.

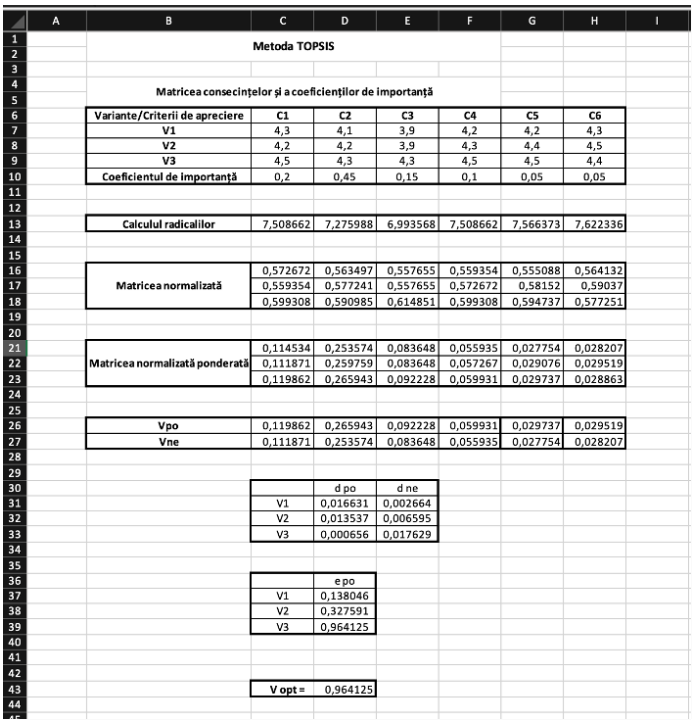
In Figure 8, a sensitivity analysis was applied to the TOPSIS method to see if the changes that occurred in the previous method are also present here.

Similar to the sensitivity analysis applied to the global utility method, the results of the analysis applied to the TOPSIS method have changed. It is observed that the hierarchy of platforms is the same, namely:

1. SAS – Advanced Analytics Software;
2. Google Cloud AI Platform;
3. IBM Watson Studio.

3.3 Conclusions for Analysis No. 2

It can be observed that the ranking of platforms changes after changing the calculation method. The two methods lead to different but still close rankings. IBM Watson Studio ranks last in both methods, and according to the sensitivity analysis applied to the methods, the platform remains in the same position. Only the sensitivity analysis applied to the first method caused a change in the top two positions, leaving IBM Watson Studio in the last position. The second sensitivity analysis



Metoda TOPSIS

Matricea consecințelor și a coeficienților de importanță

Variante/Criterii de apreciere	C1	C2	C3	C4	C5	C6
V1	4,3	4,1	3,9	4,2	4,2	4,3
V2	4,2	4,2	3,9	4,3	4,4	4,5
V3	4,5	4,3	4,3	4,5	4,5	4,4
Coeficientul de importanță	0,2	0,45	0,15	0,1	0,05	0,05

Calculul radicalilor	7,508662	7,275988	6,993568	7,508662	7,566373	7,622336

Matricea normalizată	0,572672	0,563497	0,557655	0,559354	0,555088	0,564132
	0,559354	0,577241	0,557655	0,572672	0,58152	0,59037
	0,599308	0,590985	0,614851	0,599308	0,594737	0,577251

Matricea normalizată ponderată	0,114534	0,253574	0,083648	0,055935	0,027754	0,028207
	0,111871	0,259759	0,083648	0,057267	0,029076	0,029519
	0,119862	0,265943	0,092228	0,059931	0,029737	0,028863

Vpo	0,119862	0,265943	0,092228	0,059931	0,029737	0,029519
Vne	0,111871	0,253574	0,083648	0,055935	0,027754	0,028207

	d po	d ne
V1	0,016631	0,002664
V2	0,013537	0,006595
V3	0,000656	0,017629

	e po
V1	0,138046
V2	0,327591
V3	0,964125

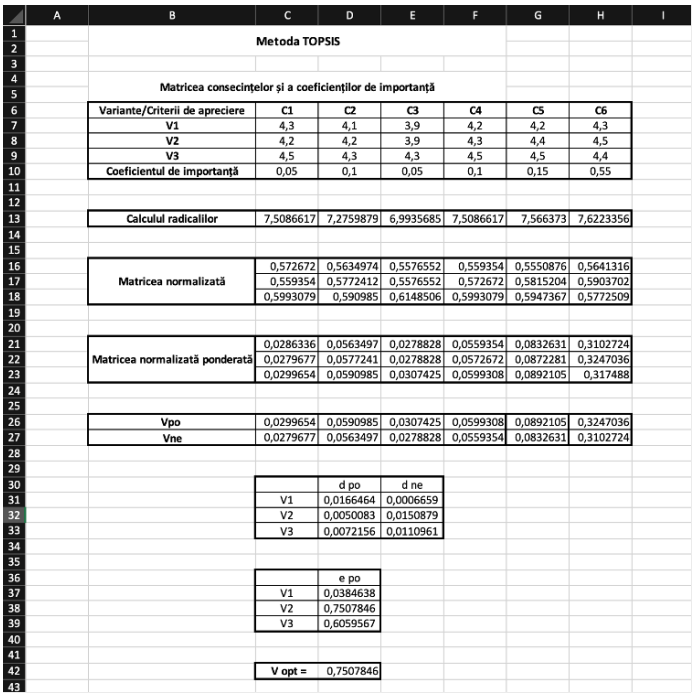
V opt =	0,964125

Figure 7: Problem Resolution Using the TOPSIS Method in Excel

applied to the TOPSIS method left the hierarchy unchanged, which refutes the hypothesis. It can be confirmed that there is no single platform that represents the optimal solution, regardless of the method. However, it can be observed that in 3 out of the 4 cases analyzed, the hierarchy was the same. Although there is no single platform that represents the optimal solution for all studied cases, an important conclusion can be drawn by analyzing the obtained results.

4 Conclusions

Using different methods can indeed lead to identical results. One reason for this is the variation in algorithms and techniques, resulting in different operation modes and applicability, making it difficult for results to remain the same. Additionally, some methods may assume a certain data distribution, while others may have fewer restrictions in this regard. Sensitivity to input data can also affect the obtained results. It was normal and expected for this to happen, so it is important to understand the specific context of the analysis and the algorithms involved in each method. In both cases where sensitivity analyses were applied, the obtained hierarchies were not significantly influenced. There were some changes between positions, but there were no major changes that would create entirely different hierarchies. There are several factors for which these analyses did not lead to major changes, such as the



Metoda TOPSIS

Matricea consecințelor și a coeficienților de importanță

Variante/Criterii de apreciere	C1	C2	C3	C4	C5	C6
V1	4,3	4,1	3,9	4,2	4,2	4,3
V2	4,2	4,2	3,9	4,3	4,4	4,5
V3	4,5	4,3	4,3	4,5	4,5	4,4
Coeficientul de importanță	0,05	0,1	0,05	0,1	0,15	0,55

Calculul radicalilor	7,5086617	7,2759879	6,9935685	7,5086617	7,566373	7,6223356

Matricea normalizată	0,572672	0,5634974	0,5576552	0,559354	0,5550876	0,5641316
	0,559354	0,5772412	0,5576552	0,572672	0,5815204	0,5903702
	0,5993079	0,590985	0,6148506	0,5993079	0,5947367	0,5772509

Matricea normalizată ponderată	0,0286336	0,0563497	0,0278828	0,0559354	0,0832631	0,3102724
	0,0279677	0,0577241	0,0278828	0,0572672	0,0872281	0,3247036
	0,0299654	0,0590985	0,0307425	0,0599308	0,0892105	0,317488

Vpo	0,0299654	0,0590985	0,0307425	0,0599308	0,0892105	0,3247036
Vne	0,0279677	0,0563497	0,0278828	0,0559354	0,0832631	0,3102724

	d po	d ne
V1	0,0166464	0,0006659
V2	0,0050083	0,0150879
V3	0,0072156	0,0110961

	e po
V1	0,0384638
V2	0,7507846
V3	0,6059567

V opt =	0,7507846

Figure 8: Sensitivity Analysis of the TOPSIS Method in Excel

existence of importance coefficients that strongly influence the results. Additionally, this depends on the method used, as some methods may be less sensitive to coefficient variations, but sensitivity analyses may also have certain limitations. In both analyses, a change in the optimal variant was observed depending on the parameters and data of the problem. It would be unrealistic to assume that there is a single platform that meets all requirements and maximizes them. Additionally, depending on the requirements of each organization, their importance, and the quality of qualified personnel working with the data, the hierarchy of platforms may differ. As observed in the analyses, some users may need strong data analysis capabilities, while others may be interested in visualization, team collaboration, etc. It is difficult for a single platform to meet all these needs optimally. Moreover, being a continuously developing field where requirements, data quantities, competition, are increasing, each platform must be in continuous development and innovation to develop functionalities, features, and even new algorithms. The various needs and preferences of users can make one platform more suitable than others based on technology and infrastructure.

5 Limitations of the research

Given the complexity of the scientific endeavor, the following limitations of this research have been identified:

- In making a decision at the organizational level, there are multiple factors that can intervene, and the analysis that should be conducted needs to meet all the organization's specific requirements.

- The criteria chosen in this work, although of major importance in this field, vary from one organization to another, as do the importance coefficients.

Furthermore, the identified limitations open up new research directions for the implementation of calculation methods that are as specific and easy to use as possible within any organization, to assist in making decisions regarding the optimal variant for each. In the field of data science, all information changes at a very rapid pace, which is why the hierarchy of platforms can change at any time. Therefore, there is a need for studies that are as clear, accurate, and up-to-date as possible, keeping pace with the changes that occur in technology.

References

[1] S. Albu, M.-L. Marian, *The applicability of the maximum global utility method in the activity of the real estate consulting service*, Open Journal of Applied Sciences, vol. 12, no. 11, 2022, 1719-1729, https://doi.org/10.4236/ojapps.2022.1211117

[2] D. Lixandroiu, *Modelarea deciziei economice*, Editura ECONOMICA, 2014

[3] F. Provost, T. Fawcett, *Data science and its relationship to big data and data-driven decision making*, Big Data, vol. 1, no. 1, 2013, 51-59, https://doi.org/10.1089/big.2013.1508

[4] R. Sharma, *Common data science challenges of 2024 [with solution]*, KnowledgeHut, 2014, https://www.knowledgehut.com/blog/data-science/data-science-challenges

[5] T. Sarkar, *Essential math for data science - towards data science*, Medium, 2018, https://towardsdatascience.com/essential-math-for-data-science-why-and-how-e88271367fbd

[6] M. A. Waller, S. E. Fawcett, *Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management*, Journal of Business Logistics, vol. 34, no. 2, 2013, 77-84, https://doi.org/10.1111/jbl.12010

Emanuelle-Ioana Perța
[1]Nicolae Bălcescu Land Forces Academy
Faculty of Military Management
Engineering and Management in military communications systems
550170, Sibiu, Romania
[2]Bucharest University of Economic Studies
Faculty of Cybernetics, Statistics and Informatics
Cybernetics and Statistics
010374, Bucharest, Romania
e-mail: muntean.emanuelle@yahoo.com