



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

**Centro Nacional de Investigación
y Desarrollo Tecnológico**

Tesis de Maestría

**Caracterización y selección de herramientas de
visualización con orientación a la solución de
problemas en Ciencia de Datos**

presentada por

Ing. Marian González Aguila

como requisito para la obtención del grado de
Maestra en Ciencias de la Computación

Directora de tesis

Dra. María Yasmín Hernández Pérez

Cuernavaca, Morelos, México. Septiembre de 2025.



Cuernavaca, Mor., **19/agosto/2025**

OFICIO No. DCC/235/2025

Asunto: Aceptación de documento de tesis
CENIDET-AC-004-M14-OFICIO

CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de **Marian González Aguila** con número de control **M23CE064**, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado **"Caracterización y selección de herramientas de visualización con orientación a la solución de problemas en Ciencia de Datos"** y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

ATENTAMENTE

Excelencia en Educación Tecnológica®
"Conocimiento y Tecnología al Servicio de México"

Dra. María Yasmín Hernández Pérez
Directora de tesis

Dr. Dante Mújica Vargas
Revisor 1

Dr. Javier Ortiz Hernández
Revisor 2

C.c.p. Depto. Servicios Escolares.
Expediente / Estudiante



2025
Año de
La Mujer
Indígena

Interior Internado Palmira S/N, Col. Palmira,
C. P. 62490, Cuernavaca, Morelos Tel. 01 (777) 3627770, ext. 3201,
e-mail: dcc_cenidet@tecnm.mx | tecnm.mx | cenidet.tecnm.mx





Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO



Centro Nacional de Investigación y Desarrollo tecnológico
Subdirección Académica

Cuernavaca Mor, 20/agosto/2025

Oficio No. SAC/201/2025

Asunto: Autorización de impresión de tesis

MARIAN GONZÁLEZ AGUILA
CANDIDATA AL GRADO DE MAESTRA
EN CIENCIAS DE LA COMPUTACIÓN
P R E S E N T E

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado **"Caracterización y selección de herramientas de visualización con orientación a la solución de problemas en Ciencia de Datos"**, ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

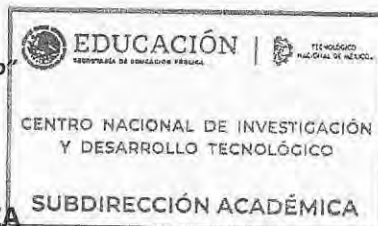
Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

ATENTAMENTE

Excelencia en Educación Tecnológica®

"Conocimiento y Tecnología al Servicio de México"

CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO



c.c.p. Departamento de Ciencias Computacionales
Departamento de Servicios Escolares

CMAZ/lmz



2025
Año de
La Mujer
Indígena

Interior Internado Palmira S/N, Col. Palmira,
C. P. 62490, Cuernavaca, Morelos Tel. 01 (777) 3627770, ext. 4104,
e-mail: acad_cenidet@tecnm.mx | tecnm.mx | cenidet.tecnm.mx

cenidet
Centro Nacional de Investigación
y Desarrollo Tecnológico



Dedicatoria

A mi papá, José†, porque aún en el cielo, lo siento
cerca acompañándome en cada paso del camino.

A mi mamá, Miriam, por su apoyo y su amor en
cada momento.

A Uzi, mi compañero, mi oasis, mi sustento en los
días difíciles. Las palabras se quedan cortas para
expresar lo que significas en mi vida.

Agradecimientos

A Dios, por ser mi fortaleza en cada cosa que hago, en especial en este proyecto.

Al Gobierno de México y a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), por el apoyo económico que hizo posible el desarrollo de esta tesis en calidad de estudiante de tiempo completo.

Al Tecnológico Nacional de México (TecNM) y al Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), por brindarme la oportunidad de cursar la Maestría en Ciencias de la Computación, así como por proporcionar el espacio y los recursos necesarios para llevarla a cabo.

A mi directora de tesis, Dra. María Yasmín Hernández Pérez, por su orientación académica, su valioso apoyo durante este proceso y por los aprendizajes adquiridos bajo su dirección.

Al Dr. Joaquín Pérez Ortega, por haberme aceptado como su estudiante, por su guía en las etapas iniciales de este proyecto y por la calidad humana que siempre me brindó.

A los revisores de este trabajo, el Dr. Dante Mújica Vargas y el Dr. Javier Ortiz Pérez, por compartir sus conocimientos y ofrecerme retroalimentaciones que sin duda contribuyeron a mejorar este trabajo de investigación.

A mi familia, por su amor y por preocuparse siempre por mí.

A mi suegra, Margarita, por sus palabras de aliento en los momentos difíciles, por su cariño y su cercanía.

A mis cuñados, David y Fredy, y sus esposas Lilí y Valeria, por sus ánimos y por quererme como una más de la familia.

A Anita, la asistente del DCC, por todo su apoyo durante los procesos administrativos.

Resumen

En el contexto actual de la Ciencia de Datos, la visualización juega un papel fundamental tanto en la exploración como en la comunicación de los hallazgos derivados del análisis. Sin embargo, el creciente número de herramientas de visualización disponibles, junto con la diversidad de características, precios y niveles de complejidad técnica que presentan, hace que la selección de la herramienta más adecuada para un problema específico resulte una tarea compleja. Esta investigación tuvo como propósito desarrollar un método sistemático que facilite la selección de herramientas de visualización de datos considerando criterios técnicos, contextuales y funcionales. Para ello, se diseñó y propuso el método MATCH-Vis, una metodología de seis pasos basada en la caracterización de herramientas de acuerdo con cinco dimensiones clave: Meta (objetivo de la visualización), Almacenamiento (ubicación de los datos), Tipo de datos, Contexto (perfil del usuario y sus restricciones), y Herramienta (que surge como resultado de las anteriores). Estas dimensiones permiten delimitar claramente los requisitos del escenario de uso y fundamentar la selección con base en criterios relevantes.

A partir de una revisión de la literatura especializada, se identificaron 24 herramientas de visualización ampliamente utilizadas en entornos de Ciencia de Datos, de las cuales se seleccionaron 11 para su análisis detallado. Estas fueron caracterizadas de acuerdo con las dimensiones MATCH y evaluadas empíricamente mediante una encuesta aplicada a 21 especialistas del dominio. Como parte del diseño metodológico, se definieron 11 criterios de evaluación agrupados en cinco categorías (Usabilidad, Complejidad, Integración, Costo y Calidad). Para cuantificar las valoraciones cualitativas, se desarrolló una escala de puntuación de 0 a 5 para evaluar las herramientas de acuerdo con los criterios. La evaluación se concretó sobre seis herramientas de visualización: Matplotlib, Seaborn, Plotly, Ggplot2, Tableau y Power BI. Es decir, se eliminaron de la evaluación aquellas herramientas con más del 50 % de datos faltantes por parte de los participantes en la evaluación. Posteriormente, se construyó una matriz de decisión con las valoraciones obtenidas y se calcularon los pesos objetivos de los criterios usando el método de la entropía, lo cual permitió reducir el sesgo subjetivo en el análisis. Entre los criterios con mayor peso destacaron el Costo, la Asistencia en la selección del tipo de gráfico, y la Exploración interactiva de los datos.

Para comparar las herramientas evaluadas fueron utilizados los siguientes métodos de decisión multicriterio: Ponderación Aditiva Simple (SAW) y la Técnica para el Orden de Preferencia por Similitud a la Solución Ideal (TOPSIS). Los resultados mostraron consistencia entre ambos métodos al aplicar TOPSIS con distancia de Manhattan. Las herramientas Power BI y Plotly se posicionaron como las mejor evaluadas, seguidas por Tableau. Las herramientas más valoradas coincidieron con aquellas que ofrecen mayor grado de interactividad, facilidad de uso y capacidades para integrar fuentes de datos diversas, lo que pone en evidencia la importancia de la experiencia de usuario y la funcionalidad práctica en entornos reales. El método MATCH-Vis se integró como un procedimiento estructurado que guía al usuario desde la definición del escenario de uso hasta la interpretación de los resultados, permitiendo una selección fundamentada, reproducible y adaptable. Esta propuesta responde a una necesidad práctica en el campo de la Ciencia de Datos: contar con una guía clara y objetiva para elegir herramientas de visualización según las características del problema, los datos y el perfil del usuario.

Palabras claves: caracterización, herramientas de visualización de datos, métodos de toma de decisiones multicriterio, MATCH-Vis.

Abstract

In the current context of Data Science, visualization plays a fundamental role in both the exploration and communication of the findings derived from the analysis. However, the growing number of visualization tools available, together with the diversity of features, prices and levels of technical complexity they present, makes the selection of the most appropriate tool for a specific problem a complex task. The purpose of this research was to develop a systematic method to facilitate the selection of data visualization tools considering technical, contextual and functional criteria. To this end, the MATCH-Vis method was designed and proposed, a six-step methodology based on the characterization of tools according to five key dimensions: Goal (objective of the visualization), Storage (location of the data), Type of data, Context (user profile and constraints), and Tool (which emerges as a result of the previous ones). These dimensions make it possible to clearly delimit the requirements of the use scenario and to base the selection on relevant criteria.

Based on a review of the specialized literature, 24 visualization tools widely used in Data Science environments were identified, of which 11 were selected for detailed analysis. These were characterized according to the MATCH dimensions and empirically evaluated by means of a survey applied to 21 domain specialists. As part of the methodological design, 11 evaluation criteria were defined, grouped into five categories (Usability, Complexity, Integration, Cost and Quality). To quantify the qualitative assessments, a rating scale from 0 to 5 was developed to evaluate the tools according to the criteria. The evaluation focused on six visualization tools: Matplotlib, Seaborn, Plotly, Ggplot2, Tableau and Power BI. In other words, those tools with more than 50 % of missing data were eliminated from the evaluation. Subsequently, a decision matrix was constructed with the evaluations obtained and the objective weights of the criteria were calculated using the entropy method, which made it possible to reduce subjective bias in the analysis. Among the criteria with the highest weights were Cost, Assistance in selecting the type of graphic, and Interactive exploration of the data.

The following multi-criteria decision methods were used to compare the tools evaluated: Simple Additive Weighting (SAW) and the Technique for Order of Preference by Similarity to the Ideal Solution (TOPSIS). The results showed consistency between both methods when applying TOPSIS with Manhattan distance. Power BI and Plotly tools were positioned as the best evaluated, followed by Tableau. The top-rated tools coincided with those offering the highest degree of interactivity, ease of use and capabilities to integrate diverse data sources, highlighting the importance of user experience and practical functionality in real-world environments. The MATCH-Vis method was integrated as a structured procedure that guides the user from the definition of the usage scenario to the interpretation of the results, allowing for an informed, reproducible and adaptable selection. This proposal responds to a practical need in the field of Data Science: to have a clear and objective guide to choose visualization tools according to the characteristics of the problem, the data and the user's profile.

Key words: characterization, data visualization tools, multicriteria decision-making methods, MATCH-Vis.

Contenido	Pág.
Resumen	IV
Abstract	V
Lista de Tablas	VIII
Lista de Figuras.....	IX
1 Introducción	10
1.1 Planteamiento del Problema	12
1.2 Justificación	12
1.3 Objetivos.....	13
1.3.1 Objetivo general:	13
1.3.2 Objetivos específicos:.....	13
1.4 Metodología de solución	13
1.4.1 Conocer el dominio de la visualización de datos	13
1.4.2 Definir los criterios	14
1.4.3 Evaluar las herramientas de visualización	14
1.4.4 Integración del método	15
1.5 Alcances y limitaciones.....	15
1.5.1 Alcances.....	15
1.5.2 Limitaciones	16
1.6 Estructura del documento	16
2 Marco Teórico	17
2.1 Ciencia de Datos	18
2.2 Herramientas de visualización de datos	18
2.3 Métodos de toma de decisiones multicriterio	19
2.3.1 Ponderación aditiva simple	20
2.3.2 Técnica de orden de preferencia por su similitud a la solución ideal.....	22
3 Estado del Arte	25
3.1 Comparación de herramientas de visualización	26
3.2 Criterios relevantes en la selección de herramientas de visualización.....	27
3.3 Aplicación de métodos de toma de decisiones multicriterio a la selección de herramientas. .	28
3.4 Conclusiones del análisis del Estado del Arte	29
4 Construcción del método MATCH-Vis	31
4.1 Selección de herramientas de visualización	32
4.2 Determinación de los criterios de evaluación	34
4.2.1 Categoría Usabilidad	35

4.2.2	Categoría Complejidad	35
4.2.3	Categoría Integración	36
4.2.4	Categoría Costo	37
4.2.5	Categoría Calidad	37
4.3	Creación de puntajes de evaluación para los criterios	37
4.4	Construcción de la encuesta	37
4.5	Evaluación de las herramientas y experimentación	38
4.5.1	Análisis descriptivo	38
4.5.2	Tratamiento de los datos faltantes	39
4.5.3	Métodos de toma de decisión multicriterio.....	40
4.5.4	Aplicación del método SAW	41
4.5.5	Aplicación del método TOPSIS	42
4.6	Integración del Método MATCH-Vis	45
4.7	Conclusiones del capítulo	47
5	Resultados y Discusión	48
5.1	Caracterización de las herramientas seleccionadas	49
5.2	Criterios de comparación.....	49
5.3	Resultados obtenidos con los métodos de toma de decisiones multicriterio.....	50
5.4	Comparación con el Estado del Arte.....	52
6	Conclusiones	54
6.1	Conclusiones	55
6.2	Trabajo futuro	55
6.3	Productos Científicos	55
7	Referencias.....	57
8	Anexos	60
8.1	Anexo A: Caracterización de todas las herramientas identificadas en la búsqueda de la literatura 61	
8.2	Anexo B: Encuesta de evaluación	69
8.3	Anexo C: Sistema de puntuación para cuantificar los criterios propuestos para evaluar las herramientas.....	75

Lista de Tablas	Pág.
Tabla 3.1. Trabajos que comparan diferentes herramientas de visualización	26
Tabla 3.2. Trabajos que abordan criterios a tomar en cuenta a la hora de seleccionar una herramienta de visualización	27
Tabla 3.3. Trabajos relacionados con la creación de métodos de selección de herramientas.....	29
Tabla 4.1. Herramientas de visualización seleccionadas para un análisis más profundo	33
Tabla 4.2. Caracterización de herramientas según las dimensiones MATCH	33
Tabla 4.3. Criterios para evaluar las herramientas de visualización.....	34
Tabla 4.4. Porcentaje de datos faltantes por herramientas con respecto al total de participantes	39
Tabla 4.5. Matriz de decisión	40
Tabla 4.6. Matriz normalizada con el método SAW	41
Tabla 4.7. Pesos de los criterios en SAW	41
Tabla 4.8. Ranking de herramientas según el método SAW	42
Tabla 4.9. Matriz normalizada con el método TOPSIS	42
Tabla 4.10. Pesos de los criterios para el método TOPSIS.....	43
Tabla 4.11. Matriz ponderada con método TOPSIS.....	43
Tabla 4.12. Soluciones ideales positivas y negativas con TOPSIS.....	43
Tabla 4.13. Ranking de herramientas según el método TOPSIS.....	44
Tabla 4.14. Ranking de herramientas usando el método TOPSIS con distancia de Manhattan.....	45
Tabla 4.15. Dimensiones MATCH para establecer el escenario de uso de herramientas de visualización	45
Tabla 4.16. Criterios agrupados por las dimensiones del método MATCH-Vis y sus categorías.....	46
Tabla 5.1. Comparación de los ranking de herramientas entre los MCDM.....	50
Tabla 5.2. Pesos asignados a los criterios por el método de la entropía	51
Tabla 5.3. Comparación del método MATCH-Vis con otros trabajos de la literatura.....	53
Tabla 8.1. Características de las 24 herramientas identificadas inicialmente	61
Tabla 8.2. Escala para medir la facilidad de uso.....	75
Tabla 8.3. Escala para medir la flexibilidad en la edición de gráficos.....	75
Tabla 8.4. Escala para medir la asistencia en seleccionar el tipo de gráfico	76
Tabla 8.5. Escala para medir la variedad de tipos de gráficos.....	76
Tabla 8.6. Escala para medir la exploración interactiva de los datos	77
Tabla 8.7. Escala para medir el manejo de grandes conjuntos de datos	77
Tabla 8.8. Escala para medir el criterio fuente de datos	78
Tabla 8.9. Escala para medir el criterio archivos de exportación.....	78
Tabla 8.10. Escala para medir el criterio “Costo”	79
Tabla 8.11. Escala para medir el renderizado de resultados.....	79
Tabla 8.12. Escala para medir la calidad de la visualización.....	79

Lista de Figuras	Pág.
Figura 1.1. Metodología FMDS (Rollins, 2015).....	11
Figura 1.2. Metodología de solución.....	14
Figura 2.1. Marco de trabajo general de los MCDM (Dhiman & Deb, 2020)	20
Figura 2.2. Diagrama de flujo del método SAW (Dhiman & Deb, 2020)	21
Figura 2.3. Diagrama de flujo del método TOPSIS (Dhiman & Deb, 2020)	23
Figura 4.1. Herramientas de visualización identificadas en la investigación.....	32
Figura 4.2. Distribución de participantes según: a) nivel de escolaridad y b) ocupación	38
Figura 4.3. Número de respuestas por herramientas	39

Capítulo 1

Introducción

CAPÍTULO 1: INTRODUCCIÓN

La visualización de datos constituye una parte integral del proceso de Ciencia de Datos, ya que desempeña un papel crucial en la exploración, análisis y comunicación de resultados. Según Prakash et al. (2020), y de acuerdo con la Metodología Fundamental de la Ciencia de Datos (Rollins, 2015) representada en la Figura 1.1, la visualización se utiliza en:

1. La etapa de Comprensión de los datos. En esta etapa, los científicos de datos utilizan visualizaciones para entender la estructura y características de los datos, además de identificar patrones y detectar anomalías.
2. En la etapa de Modelado, las visualizaciones ayudan a interpretar los resultados obtenidos con los modelos estadísticos y algoritmos de aprendizaje automático, facilitando la identificación de relaciones y tendencias.
3. Finalmente, en la etapa de Evaluación, las visualizaciones se vuelven a emplear, donde son esenciales para comunicar los hallazgos a las partes interesadas de manera efectiva, permitiendo una toma de decisiones clara e informada.

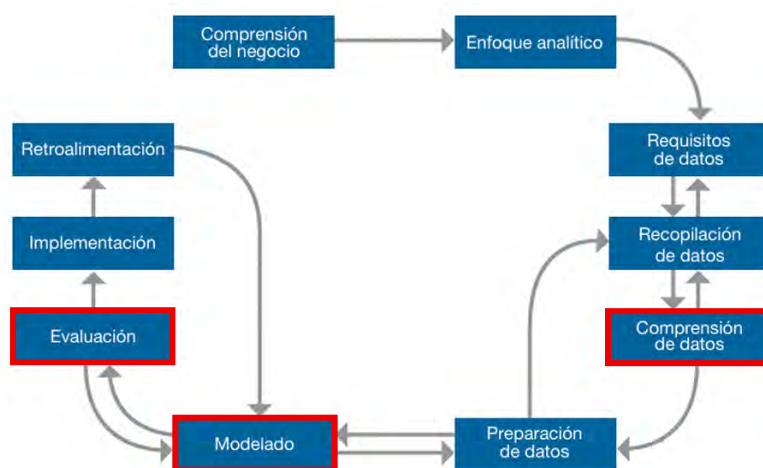


Figura 1.1. Metodología FMDS (Rollins, 2015)

Seleccionar adecuadamente las herramientas de visualización de datos es fundamental para maximizar el valor de los datos analizados. Las herramientas de visualización permiten transformar conjuntos de datos complejos en representaciones gráficas comprensibles, facilitando la identificación de patrones, tendencias y anomalías que podrían pasar desapercibidos en formatos tabulares o textuales. La visualización de datos es esencial para simplificar el análisis de datos, mejorar la comunicación y promover una toma de decisiones más rápida y efectiva en todos los niveles de una organización (Shivaram, 2024).

El uso de herramientas adecuadas de visualización mejora la accesibilidad de la información, permitiendo que tanto expertos como personas sin formación técnica comprendan los hallazgos y participen en el proceso de toma de decisiones. (Hashemi-Pour et al., 2024) señala que las herramientas de visualización

hacen que los datos sean más accesibles y comprensibles, de modo que las personas sin formación técnica puedan interpretarlos y analizarlos.

Es por ello que conocer las herramientas de visualización de datos y sus capacidades es esencial para los científicos de datos y analistas, ya que les permite elegir la herramienta más adecuada para sus necesidades específicas, optimizando así el proceso de análisis y comunicación de datos.

1.1 Planteamiento del Problema

En la actualidad existe gran variedad de herramientas de visualización tanto comerciales como de software libre. Cada herramienta tiene sus fortalezas, limitaciones y enfoques específicos. Muchas veces se pueden utilizar en más de un escenario de aplicación lo que hace difícil el proceso de seleccionar la más adecuada para un problema concreto de Ciencia de Datos. Una problemática a la que se enfrentan los científicos de datos tanto expertos como principiantes radica en que las herramientas deben estar ajustadas no solo al tipo de datos y las características del problema, sino también a las habilidades técnicas del usuario y en el mercado actual podemos encontrar herramientas muy poderosas y difíciles de dominar, así como otras herramientas más accesibles, pero limitadas en funcionalidades y personalización.

En la actualidad no existe un método robusto y universal para evaluar herramientas de visualización, ya que las comparaciones suelen ser subjetivas, y muchas veces basadas en las fortalezas y limitaciones de la herramienta. Sucede también que, en muchos proyectos, la selección de la herramienta se realiza sin un análisis profundo, lo que conlleva a problemas de sobrecarga tecnológica, es decir, se usan herramientas muy complejas para problemas muy simples o hay limitaciones en la resolución del problema al no considerarse las funcionalidades necesarias desde un inicio.

Las herramientas de visualización están en constante evolución y nuevas plataformas emergen rápidamente, lo que plantea la necesidad de caracterizar y comparar estas nuevas herramientas emergentes con las que ya existen. La visualización efectiva no solo ayuda a interpretar la información, sino que influye en la toma de decisiones, por lo que una selección inadecuada puede comprometer la capacidad de los tomadores de decisiones para comprender patrones, valores atípicos y tendencias en los datos.

1.2 Justificación

Con esta investigación se propone el desarrollo de un método que permita caracterizar, evaluar y seleccionar herramientas de visualización adaptadas específicamente a las necesidades reales de los problemas de Ciencia de Datos. Este método ofrece múltiples beneficios: reduce la subjetividad en las evaluaciones de las herramientas, optimiza el proceso de selección tecnológica, garantiza una utilización eficiente de los recursos técnicos y humanos en proyectos reales, lo que trae consigo la mejora de la comunicación visual de los resultados (Munzner, 2015).

Debido a que una visualización efectiva no solo mejora la interpretación de la información, sino que incide directamente en la calidad de las decisiones tomadas a partir de ella, resulta imprescindible proporcionar mecanismos robustos que contribuyan a la correcta elección de herramientas visuales. Esta investigación

contribuye en el proceso de selección de herramientas de visualización dentro del contexto de la Ciencia de Datos aportando claridad y valor práctico.

1.3 Objetivos

1.3.1 Objetivo general:

Construir un método que permita seleccionar la herramienta de visualización de datos adecuada para un problema específico en el enfoque de Ciencia de Datos.

1.3.2 Objetivos específicos:

1. Analizar herramientas de visualización de datos para conocer sobre sus capacidades y los factores que influyen en el proceso de visualización.
2. Establecer criterios para caracterizar y evaluar herramientas de visualización de datos.
3. Evaluar un conjunto de herramientas de acuerdo con los criterios previamente establecidos para aportar evidencia objetiva que respalde el proceso de selección.
4. Integrar los criterios y el conocimiento de la evaluación para estructurar el método de selección de herramientas de visualización de datos y facilitar su aplicación.

1.4 Metodología de solución

En esta sección se presenta la metodología llevada a cabo para construir el método de selección de herramientas de visualización MATCH-Vis. La misma se muestra en la Figura 1.2.

1.4.1 Conocer el dominio de la visualización de datos

La primera fase de la metodología tuvo como propósito generar un entendimiento profundo del dominio de la visualización de datos, reconociendo tanto las herramientas actualmente utilizadas como los aspectos clave que inciden en su selección y aplicación. En primer lugar, se realizó una revisión de literatura científica y técnica, con el objetivo de identificar las herramientas de visualización más utilizadas en el ámbito de la Ciencia de Datos en la actualidad y los factores que inciden en este proceso. Esta búsqueda permitió recopilar información sobre las capacidades, limitaciones y casos de uso asociados a distintas herramientas.

Posteriormente, se llevó a cabo una etapa de experimentación directa con las herramientas, donde se exploraron sus funcionalidades, requisitos técnicos, tipos de visualizaciones soportadas, y experiencia de usuario. Esta aproximación práctica permitió validar la información recolectada en la literatura y generar criterios de evaluación más aterrizados a su comportamiento. Además, se realizaron consultas a expertos con el fin de incorporar una visión empírica sobre las herramientas más relevantes en la práctica profesional. En conjunto, esta fase permitió construir una visión integral del dominio, sirviendo como base para la identificación de las herramientas a incluir en el estudio comparativo.

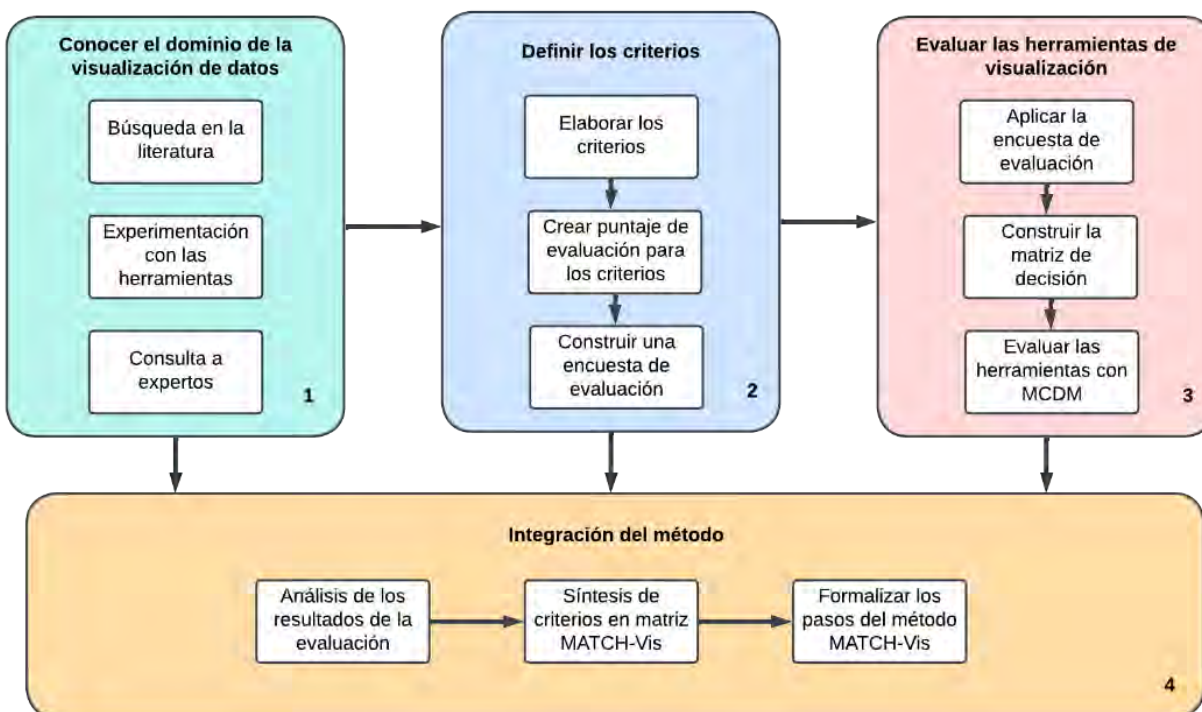


Figura 1.2. Metodología de solución

1.4.2 Definir los criterios

En esta etapa, se elaboraron los criterios para evaluar a las herramientas. Estos criterios fueron contruidos tomando como base una revisión de la literatura especializada, la cual identifica un conjunto de requisitos que toda herramienta de visualización debería satisfacer. Los criterios definidos son de carácter cualitativo, por lo que, se elaboró un sistema que permitiera su cuantificación. Este sistema consistió en descomponer cada criterio en un conjunto de características que lo identifican, asignando una escala de puntuación de 0 a 5, donde un valor 0 representaba que la herramienta no cumple con ninguna de las características asociadas al criterio, mientras que un valor de 5 representa el cumplimiento total de ellas. Las puntuaciones intermedias reflejan niveles parciales de cumplimiento.

Con base en los criterios y el sistema de puntuación diseñado, se elaboró una encuesta para recolectar las valoraciones de los expertos del dominio respecto a las herramientas seleccionadas en la fase anterior. La encuesta solicita información acerca del nivel de estudio y profesión de los participantes, así como sobre su nivel de dominio y experiencia con cada herramienta. Finalmente, se les solicita calificar dichas herramientas con respecto a los criterios, teniendo en cuenta el sistema de puntuación establecido.

1.4.3 Evaluar las herramientas de visualización

En esta tercera etapa se aplicó la encuesta de evaluación a una muestra de 21 expertos del dominio quienes presentaban distintos niveles de formación y experiencia en Ciencia de Datos y visualización. Los resultados de la encuesta fueron procesados, aplicando filtros para excluir las respuestas incompletas de aquellos participantes que indicaron no tener conocimiento sobre alguna herramienta específica tratando de garantizar la calidad de la información. Con los datos válidos se construyó una matriz de decisión que

recoge los puntajes de cada herramienta por cada criterio, promediando las valoraciones mediante una media ponderada que daba mayor peso a las respuestas emitidas por participantes con mayor dominio declarado. Esta matriz se utilizó para aplicar los métodos de toma de decisiones multicriterio SAW y TOPSIS, y con ellos obtener un ranking robusto de las herramientas analizadas. Los pesos de los criterios se asignaron de manera objetiva aplicando el método de la entropía, lo que permitió evitar sesgos subjetivos y reflejar la capacidad discriminativa de cada criterio en función de la variabilidad de las respuestas. Los resultados obtenidos en esta fase proporcionan la base empírica para integrar el método MATCH-Vis y establecer recomendaciones fundamentadas para la selección de herramientas de visualización de datos.

1.4.4 Integración del método

La cuarta y última etapa de la metodología consistió en la integración sistemática de los hallazgos obtenidos en las fases anteriores para construir el método de selección de herramientas de visualización MATCH-Vis. En esta fase se realizó un análisis de los resultados obtenidos tras realizar la evaluación comparativa de las herramientas usando los métodos de toma de decisiones multicriterio. Se realizó una síntesis de los criterios creados agrupando los criterios en cuatro dimensiones fundamentales: **Meta**, **Almacenamiento**, **Tipo de datos** y **Contexto**, las cuáles son la base para formar el acrónimo MATCH. Esta síntesis permitió una representación más clara, coherente y operativa de los factores críticos a considerar en la selección de herramientas.

Finalmente, se procedió a la formalización del método MATCH-Vis, en la que se definieron los pasos a seguir para aplicarlo en escenarios reales de Ciencia de Datos. Esta formalización contempló desde la definición del escenario de uso y la recolección de información técnica o experta, hasta la construcción de la matriz de decisión, la ponderación de los criterios y la aplicación de técnicas de decisión multicriterio. La sistematización de estos pasos dio lugar a una metodología replicable y adaptable que permite recomendar la herramienta de visualización más adecuada de forma objetiva, rápida y fundamentada.

1.5 Alcances y limitaciones

A continuación, se presentan los alcances de este estudio y las limitaciones que se tuvieron en cuenta a la hora de confeccionar el método MATCH-Vis.

1.5.1 Alcances

El alcance de este trabajo se centra en la construcción de un método que permita seleccionar herramientas de visualización de datos en entornos de Ciencia de Datos. Para la elaboración del mismo, el estudio se enfocó, en primer lugar, en hacer un análisis de las principales herramientas de visualización que se emplean en la actualidad en el ámbito de la Ciencia de Datos y posteriormente se hizo una selección de un conjunto con las que se llevó a cabo un análisis más exhaustivo. Las herramientas seleccionadas tienen como principales características que su uso es ampliamente reportado en la literatura y que las mismas son gratuitas o tienen versiones gratuitas. Esto permitió profundizar en sus características para con ellas construir un método sistematizado que permita seleccionar herramientas de visualización de datos teniendo en cuenta factores como la meta o el objetivo del análisis, donde se encuentran almacenados los datos, el tipo de datos y el contexto o las características del usuario.

1.5.2 Limitaciones

Entre las principales limitaciones para el desarrollo de esta investigación se contó con no tener licencia para probar las herramientas comerciales que requieren pago, lo que impidió profundizar en su evaluación comparativa respecto a las herramientas gratuitas seleccionadas. Además, no se consideraron en este estudio, herramientas recientemente lanzadas (últimos 12 meses) a pesar de que cumplieran con los requisitos iniciales, debido a la dificultad práctica de seguir la constante evolución tecnológica y la aparición de nuevo software.

1.6 Estructura del documento

La presente tesis se encuentra estructurada en seis capítulos. El Capítulo 1, como se ha leído, introduce la problemática de la investigación, justificando la necesidad de un método sistemático para seleccionar herramientas de visualización de datos, y plantea los objetivos y la metodología adoptada en esta investigación. En el Capítulo 2 se desarrolla el marco teórico necesario para comprender los principales conceptos y teorías abordados en este trabajo, como son: el campo de la Ciencia de Datos, el papel de la visualización en dicho contexto y los fundamentos de los métodos de toma de decisiones multicriterio utilizados en la evaluación. En el Capítulo 3 se presenta el Estado del Arte, analizando trabajos previos relacionados con la comparación de herramientas de visualización, los criterios propuestos en la literatura a tomar en cuenta a la hora de elegir la herramienta más adecuada y las aplicaciones existentes de técnicas de toma de decisiones multicriterio para seleccionar herramientas, lo que permite fundamentar conceptualmente el enfoque propuesto. El Capítulo 4 describe la construcción del método MATCH-Vis, detallando el proceso de selección y caracterización de herramientas, la definición de criterios de evaluación, el diseño de la encuesta, el tratamiento de los datos recolectados y la aplicación de los métodos SAW y TOPSIS, llevados a cabo en esta investigación. En el Capítulo 5 se discuten los principales hallazgos obtenidos a partir de la comparación de herramientas, analizando los resultados de los rankings generados y su relación con los criterios de evaluación. Finalmente, en el Capítulo 6 se exponen las conclusiones generales del estudio, así como las líneas de trabajo futuro que pueden derivarse del método MATCH-Vis y su aplicación práctica en distintos contextos analíticos.

Capítulo 2

Marco Teórico

En esta sección se definen conceptos claves relacionados con la investigación, entre los que se encuentran la Ciencia de Datos y sus aplicaciones en la era actual. Las herramientas de visualización de datos y su función dentro del flujo de trabajo de Ciencia de Datos, además de los métodos de toma de decisiones multicriterio como instrumento para seleccionar alternativas en escenarios complejos donde inciden muchos factores.

2.1 Ciencia de Datos

La Ciencia de Datos es un campo interdisciplinario que surgió de la intersección de la estadística, la informática y el conocimiento de dominios específicos para extraer conocimientos y valor de los datos. La primera vez que se utilizó este término fue en 1974, cuando Peter Naur, en su libro “Concise Survey of Computer Methods”, lo definió como *“la ciencia de tratar con datos, una vez que se han establecido, mientras que la relación de los datos con lo que representan se delega en otros campos y ciencias”* (Cao, 2017; Schilling et al., 2022). En 2001, William S. Cleveland, formalizó que la Ciencia de datos debía ser reconocida como una nueva disciplina y propuso un plan de acción para expandir las áreas técnicas de la estadística (Schilling et al., 2022).

La Ciencia de Datos ha adoptado y adaptado enfoques de áreas relacionadas, como la minería, aplicando un proceso iterativo que incluye la comprensión del problema de negocio, la preparación de los datos, la exploración de datos, la modelización, la evaluación y la implementación de soluciones. En la actualidad tiene aplicaciones en una amplia gama de sectores, incluyendo: la salud, para la predicción de brotes de enfermedades, personalización de tratamientos médicos y análisis de grandes volúmenes de datos genómicos (Shao, 2024); finanzas, en la detección de fraudes, análisis de riesgos y optimización de carteras de inversión (Chinthamu & Karukuri, 2023; Shao, 2024); marketing, en análisis de comportamiento del consumidor, segmentación de mercado y personalización de campañas publicitarias (Raza & Summair, 2020). En el transporte también se hace uso de la Ciencia de Datos para la optimización de rutas, la predicción de congestión y el análisis de datos de sensores en vehículos autónomos (Rasool & Chaudhary, 2022), al igual que la industria de medios y entretenimiento para hacer recomendación de contenido, análisis de tendencias y personalizar la experiencia del usuario (Shao, 2024).

2.2 Herramientas de visualización de datos

Las herramientas de visualización de datos son componentes esenciales en el proceso de ciencia de datos, lo que permite la transformación de conjuntos de datos complejos en representaciones gráficas intuitivas. Estas herramientas facilitan la identificación de patrones, tendencias y conocimientos, mejorando así la toma de decisiones y la comunicación entre las partes interesadas.

Las herramientas de visualización de datos son aplicaciones de software que crean representaciones gráficas de datos, como tablas, gráficos y mapas. Sirven para simplificar los datos complejos, haciéndolos accesibles para audiencias no técnicas (Swathi Priya et al., 2023), permiten a los usuarios explorar y analizar datos de forma interactiva, identificando patrones y correlaciones (Bikakis, 2022) y apoyan la limpieza y comunicación de los resultados del negocio (Pise & Singampalli, 2023).

En el flujo de trabajo de ciencia de datos, las herramientas de visualización se utilizan para la exploración y descubrimiento (Pise & Singampalli, 2023), el análisis predictivo y la comunicación (Nugroho, 2024).

Varios factores influyen en la efectividad de las visualizaciones de datos, entre ellos se encuentran los principios de diseño (Swathi Priya et al., 2023). Las visualizaciones efectivas requieren un equilibrio entre estética y funcionalidad, asegurando la claridad como principal componente. Otro factor son las características de los datos (Bikakis, 2022). La naturaleza de los datos, como su tamaño, tipo, etc. dicta la elección de las técnicas de visualización. Y finalmente, las necesidades del usuario, ya que comprender los antecedentes y objetivos de la audiencia es crucial para que las visualizaciones se adapten a sus necesidades (Nugroho, 2024).

2.3 Métodos de toma de decisiones multicriterio

Los métodos de toma de decisiones multicriterio (MCDM, por sus siglas en inglés) son un conjunto de metodologías utilizadas para evaluar y priorizar múltiples criterios en los procesos de toma de decisiones (Shaikh, 2025). Este enfoque es esencial en diversos campos, incluidos los negocios, el ejército y la vida cotidiana, donde las decisiones a menudo implican complejas compensaciones entre diferentes factores. MCDM ayuda a los tomadores de decisiones a analizar sistemáticamente las alternativas y seleccionar la opción más adecuada en función de criterios definidos (Majumder & Saha, 2016).

En la actualidad, los MCDM se utilizan en diferentes campos de la ciencia, la ingeniería, la política y hasta las ciencias sociales. Entre sus principales aplicaciones se encuentran la evaluación de riesgo crediticio, la selección y gestión de una cartera de inversión, la adquisición de nuevas empresas, la gestión de recursos naturales, la ubicación de plantas de energía, la planeación de autopistas, la satisfacción al cliente, la evaluación de tecnologías, entre muchas otras (Dhiman & Deb, 2020; Taherdoost & Madanchian, 2023), (Ciardiello & Genovese, 2023). Para determinar cuál la mejor alternativa, los MCDM siguen de manera general un procedimiento estructurado como el que se muestra en la Figura 2.1.

Los MCDM cuentan con diferentes técnicas o enfoques para realizar la evaluación de las alternativas y llegar a la selección de la más apropiada. De acuerdo con (Ciardiello & Genovese, 2023; Dhiman & Deb, 2020; Taherdoost & Madanchian, 2023) algunos de estos enfoques son:

- a) Ponderación aditiva simple o SAW,
- b) Técnica de orden de preferencia por su similitud a la solución ideal o TOPSIS
- c) Evaluación proporcional compleja o COPRAS
- d) Proceso analítico jerárquico o AHP
- e) Optimización multicriterio y solución de compromiso o VIKOR,
- f) Eliminación y selección traducida a la realidad o ELECTRE,
- g) Método de organización de clasificación de preferencias para la evaluación de enriquecimiento o PROMETHEE, entre otros...

En la presente investigación se estarán usando los métodos de toma de decisiones multicriterio SAW (*Simple Additive Weighting*) y TOPSIS (*Technique for Order of Preference by Similarity to Ideal Solution*), debido a que son ampliamente usados para evaluar tecnologías.

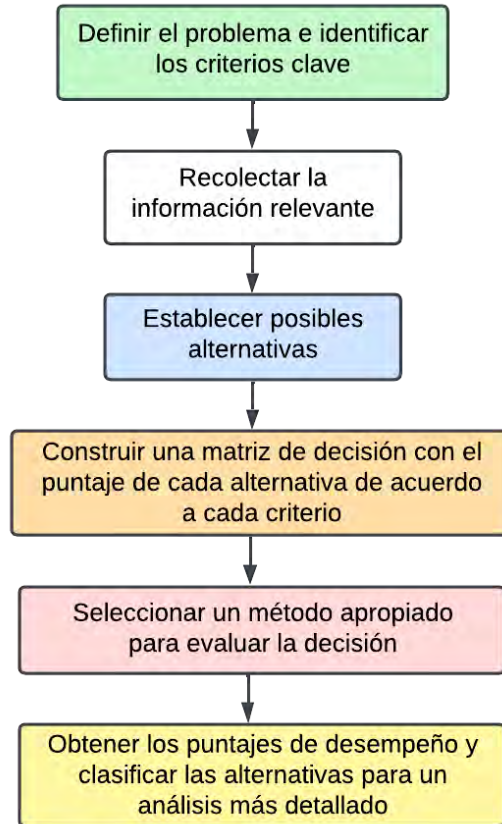


Figura 2.1. Marco de trabajo general de los MCDM (Dhiman & Deb, 2020)

2.3.1 Ponderación aditiva simple

El método de Ponderación Aditiva Simple o SAW, por sus siglas en inglés, es un método presentado por Hwang et al. en 1981 (Dhiman & Deb, 2020) donde el puntaje de evaluación de cada alternativa se calcula haciendo una suma de la multiplicación del peso de cada criterio por el puntaje de desempeño de la alternativa con respecto a dicho criterio. El proceso para determinar el puntaje de evaluación usando el método SAW se describe en la Figura 2.2.

A continuación, se explica paso a paso como llevar a cabo este procedimiento:

Paso 1: El primer paso consiste en determinar el conjunto de alternativas $(A_1, A_2, \dots, A_m)$ a evaluar y los criterios $(C_1, C_2, \dots, C_n)$ que se utilizarán para decidir cuál de ellas es la más adecuada.

Paso 2: Desarrollar una matriz de decisión formada por los puntajes de desempeño de las diferentes alternativas para cada criterio.

$$H = \begin{pmatrix} C_1 & C_2 & \dots & C_n \\ h_{11} & h_{12} & \dots & h_{1j} \\ h_{21} & h_{22} & \dots & h_{2j} \\ \vdots & \vdots & & \vdots \\ h_{i1} & \dots & \dots & h_{ij} \end{pmatrix} \begin{matrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{matrix}$$

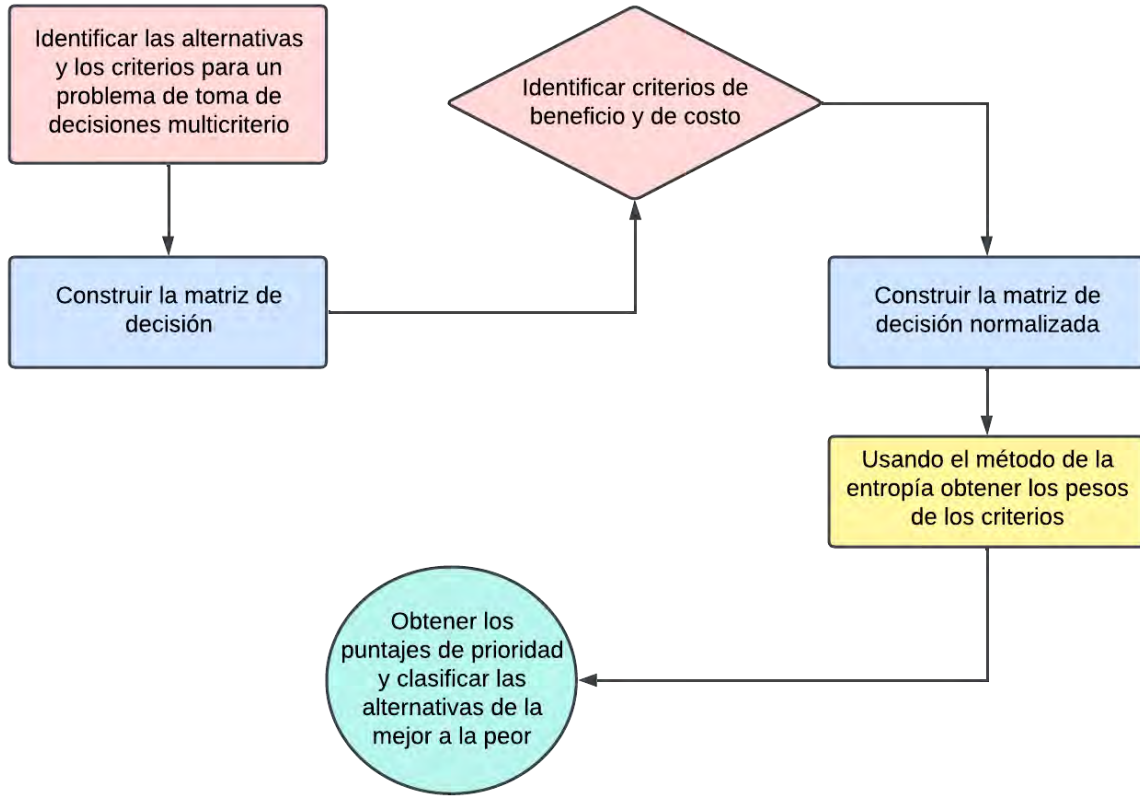


Figura 2.2. Diagrama de flujo del método SAW (Dhiman & Deb, 2020)

Paso 3: Identificar qué criterios son de beneficio y cuáles de costo. Los criterios de beneficio son aquellos maximizan los resultados positivos de las alternativas y se centran en evaluar las ventajas de las mismas como el desempeño, la satisfacción o la calidad. En estos criterios mientras mayor sea el puntaje de desempeño de la alternativa, mejor se califica esta. Mientras que los criterios de costo son aquellos que buscan minimizar los resultados negativos de las alternativas, como los gastos, los riesgos o el consumo de recursos. En estos, mientras menor sea el puntaje de una alternativa mejor se espera que sea dicha alternativa.

Paso 4: Normalizar la matriz de decisión donde cada valor de los criterios debe calcularse teniendo en cuenta si el mismo es de beneficio o de costo. Si el criterio es de beneficio se normaliza de acuerdo con la ecuación (2.1):

$$H_{ij} = \frac{h_{ij}}{\max h_{ij}} \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n, \quad (2.1)$$

si es de costo, según se muestra en la ecuación (2.2):

$$H_{ij} = \frac{\min h_{ij}}{h_{ij}} \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (2.2)$$

Paso 5: Calcular el peso de los criterios usando el método de la entropía. Para ello, primero, se calcula la entropía de cada criterio mediante la fórmula (2.3):

$$e_j = -\frac{1}{\log m} \sum_{i=1}^m H_{ij} \log(H_{ij}) \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n, \quad (2.3)$$

luego, se calcula la dispersión del criterio, de acuerdo con la ecuación (2.4):

$$d_j = |1 - e_j|, j = 1, 2, \dots, n, \quad (2.4)$$

y finalmente con el valor de dispersión se obtiene el peso de cada criterio con la ecuación (2.5):

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j} \quad (2.5)$$

La suma total de todos los pesos debe ser igual a 1. Los pesos de los criterios en problemas de toma de decisiones multicriterio pueden ser asignados por los expertos o decisores, o ser calculados de forma objetiva por el método que se presenta (Taherdoost & Madanchian, 2023). En esta investigación serán calculados por el método de la entropía, tal como se ha descrito anteriormente.

Paso 6: Finalmente, teniendo la matriz normalizada y los pesos de cada criterio se calculan los puntajes de evaluación de cada alternativa siguiendo la fórmula (2.6):

$$S_i = \sum_{j=1}^n w_j H_{ij} \quad (2.6)$$

2.3.2 Técnica de orden de preferencia por su similitud a la solución ideal

El método Técnica de orden de preferencia por su similitud a la solución ideal o TOPSIS se basa en determinar la solución óptima tomando en cuenta la menor distancia Euclidiana con respecto a la solución ideal positiva y la mayor distancia con respecto a la solución ideal negativa (Dhiman & Deb, 2020).

En la Figura 2.3 se muestra el diagrama de flujo para implementar el método TOPSIS, este método comparte los tres primeros pasos con el método SAW.

A continuación, se detallan los mismos:

Paso 1: Determinar el conjunto de alternativas y criterios para determinar cuál es la alternativa más adecuada, al igual que en el método SAW.

Paso 2: Construir la matriz de decisión con los puntajes de cada alternativa para cada criterio.

Paso 3: Identificar cuáles criterios son de beneficio y cuáles de costo.

Paso 4: Normalizar la matriz de decisión siguiendo la fórmula (2.7):

$$H_{ij} = \frac{h_{ij}}{\sqrt{\sum_{j=1}^n h_{ij}^2}} \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (2.7)$$

Paso 5: Calcular el peso de los criterios usando el método de la entropía, al igual que con el método SAW.

Paso 6: Con el peso de los criterios y la matriz normalizada construir la matriz de decisión ponderada, aplicando la siguiente fórmula (2.8):

$$\check{h} = w_j H_{ij} \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (2.8)$$

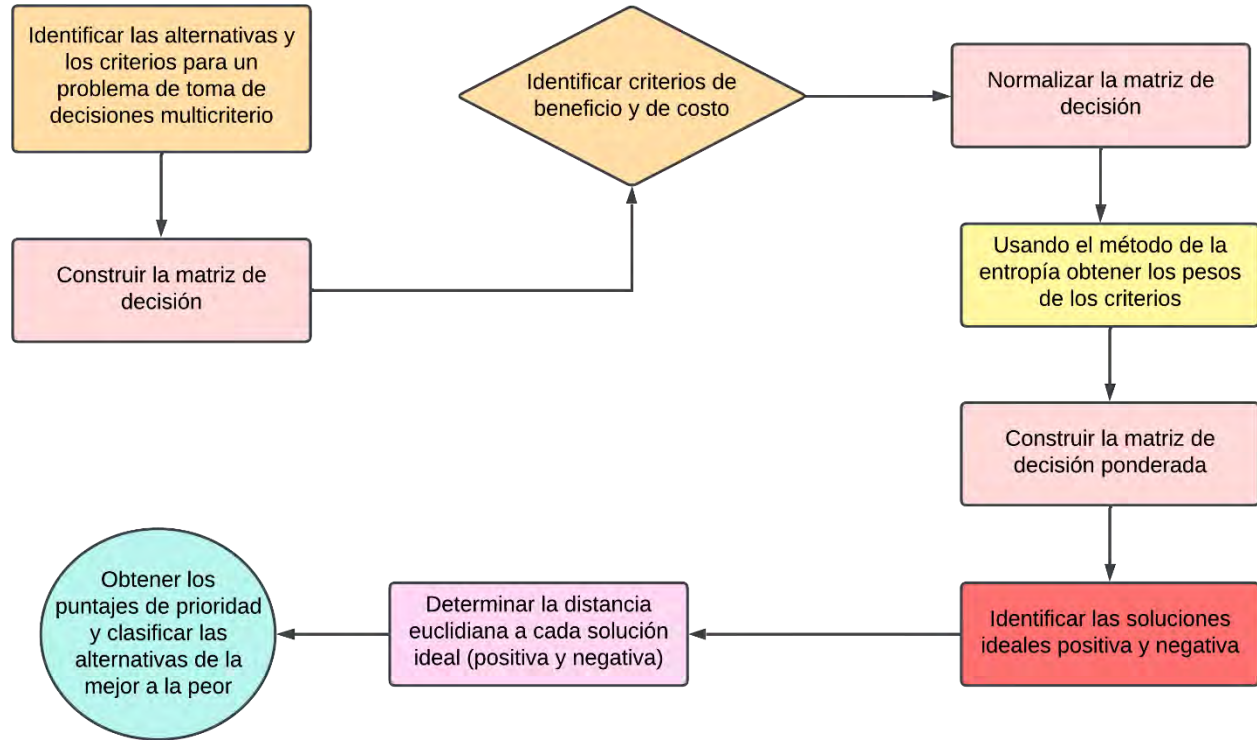


Figura 2.3. Diagrama de flujo del método TOPSIS (Dhiman & Deb, 2020)

Paso 7: Identificar cuál es la solución ideal positiva S^+ y la solución ideal negativa S^- construyendo un vector formado por los valores ponderados de las alternativas de acuerdo con las ecuaciones (2.9) y (2.10) respectivamente:

$$S_j^+ = \{(max h_{ij} | j \in t); (min h_{ij} | j \in n - t | i = 1, 2, \dots, m)\}, \quad (2.9)$$

$$S_j^- = \{(min h_{ij} | j \in t); (max h_{ij} | j \in n - t | i = 1, 2, \dots, m)\}, \quad (2.10)$$

donde t representa los criterios que son de beneficio y $n-t$ los que son de costo.

Paso 8: Determinar la distancia Euclidiana de cada alternativa a la solución ideal positiva S^+ mediante la fórmula (2.11) y a la solución ideal negativa S^- , mediante la fórmula (2.12):

$$D_i^+ = \sqrt{\sum_{j=1}^n (\check{h}_{ij} - S_j^+)^2}, \quad j = 1, 2, \dots, n, \quad (2.11)$$

$$D_i^- = \sqrt{\sum_{j=1}^n (\check{h}_{ij} - S_j^-)^2}, \quad j = 1, 2, \dots, n, \quad (2.12)$$

Paso 9: Finalmente, calcular el coeficiente de proximidad relativa C_i de cada alternativa, siguiendo la ecuación (2.13):

$$C_i = \frac{D_i^-}{D_i^- + D_i^+} \quad i = 1, 2, \dots, m; \quad 0 \leq C_i \leq 1 \quad (2.13)$$

y hacer un ranking con los puntajes de proximidad, dónde la alternativa que tenga el mayor valor de C_i será clasificada como la mejor del conjunto.

Capítulo 3

Estado del Arte

En esta sección se presentan diversos trabajos que abordan tanto el análisis comparativo de herramientas como las metodologías utilizadas para evaluar tecnologías en entornos complejos. Se examinan trabajos que proponen criterios específicos para evaluar herramientas, así como estudios que evidencian la carencia de métodos de selección. Esta revisión permite sustentar la necesidad de diseñar métodos integrales, que articulen dimensiones técnicas y contextuales en la recomendación de herramientas de visualización. Los trabajos analizados se dividieron en tres categorías: i) trabajos relacionados con la comparación de herramientas de visualización, ii) trabajos relacionados con los criterios a tomar en cuenta a la hora de seleccionar una herramienta de visualización, iii) trabajos relacionados con la aplicación de MCDM a la selección de herramientas.

3.1 Comparación de herramientas de visualización

Desimoni & Po (2020) proponen una evaluación empírica de varias herramientas de visualización de software libre. En el artículo se identifican 77 herramientas de visualización de datos conectados, pero solo describen 29 de ellas, las que aún están en funcionamiento. Las herramientas fueron descritas de acuerdo con los siguientes criterios: Tipo de información que muestra, Entrada de datos, Tipo de visualización, Flexibilidad y personalización, Interactividad y Colaboración.

Schmidt (2020) realizó un estudio con las herramientas y aplicaciones más comúnmente utilizadas en visualización de datos, concentrándose en visualización 2D principalmente. Se evaluaron 19 herramientas, 13 de ellas de software libre, entre las que se encuentran las bibliotecas Plotly y Seaborn de Python y Ggplot2 de R, D3.js, Google Charts, Microsoft Power BI, Tableau, entre otras. En este estudio no se menciona ningún criterio de evaluación en específico. Como resultado se creó una matriz de herramientas y técnicas de visualización donde se puede seleccionar la herramienta deseada teniendo en cuenta la técnica de visualización que interesa utilizar.

Shakeel et al. (2022), por su parte, realizan un análisis más profundo del panorama de la visualización de datos, clasificando las herramientas de visualización en dos grupos: lenguajes de programación y plataformas. La investigación realiza una taxonomía de las herramientas de visualización más destacadas en la actualidad y las clasifica según su aparición en artículos científicos. En el trabajo se describen las fortalezas y debilidades de cada herramienta.

En la Tabla 3.1 se muestra un resumen de los trabajos abordados en esta sección y los criterios que se tuvieron en cuenta para realizar la comparación entre las herramientas.

Tabla 3.1. Trabajos que comparan diferentes herramientas de visualización

Artículo	Título	Criterios para la comparación
(Desimoni & Po, 2020)	Empirical evaluation of linked data visualization tools.	Tipo de información que muestra, Entrada de datos, Tipo de visualización, Flexibilidad y personalización, Interactividad, Colaboración.
(Schmidt, 2020)	Usage of visualization techniques in data science workflows	Tipos de técnicas de visualización.

(Shakeel et al., 2022)	Comprehensive State-of-the-Art Survey on data Visualization Tools: research developments, challenges and future domain specific visualization framework.	Popularidad y mención en la literatura, Fortalezas y Debilidades.
------------------------	--	---

3.2 Criterios relevantes en la selección de herramientas de visualización

Lowe & Matthee (2020) realizan una revisión sistemática de la literatura con el objetivo de identificar y clasificar los requisitos esenciales de las herramientas de visualización de datos para manejar datos complejos como *Big Data*. Los autores agrupan estos requisitos en seis categorías: reducción de dimensionalidad, reducción de datos, escalabilidad y legibilidad, interactividad, recuperación rápida de los resultados y asistencia al usuario. El artículo señala que existe la necesidad de gestionar eficazmente las complejidades del *Big Data*, y esto implica mejorar la experiencia del usuario y la flexibilidad de las herramientas para adaptarse a varios tipos de datos y funciones. La investigación también subraya la importancia de la visualización interactiva, que permite a los usuarios explorar los datos de forma más intuitiva.

Nazmuz Sakib (2022) aborda la importancia de la visualización de datos, destacando cómo las representaciones visuales eficaces facilitan una mejor toma de decisiones y comunicación dentro de las organizaciones. En el trabajo se señala la necesidad de contar con herramientas de visualización interactivas que permitan a los usuarios hacer preguntas sobre los datos y recibir respuestas rápidas. La investigación destaca la necesidad de herramientas de visualización fáciles de usar que permitan a los no expertos interactuar con los datos de manera eficaz.

Dogadina & Voronin (2024) realizan un análisis comparativo de cuatro herramientas de visualización de datos muy utilizadas en la actualidad: Tableau, Power BI, D3.js y Matplotlib. Esta evaluación se basa en varios criterios, como la facilidad de uso, la flexibilidad, la escalabilidad, la integración y la calidad de los resultados, que son esenciales para que los usuarios puedan tomar decisiones informadas sobre qué herramienta utilizar para sus necesidades específicas. En el artículo se enfatiza la importancia del contexto a la hora de elegir una herramienta de visualización de datos, señalando que las diferentes herramientas son adecuadas para diferentes aplicaciones.

La Tabla 3.2 muestra los trabajos abordados en esta sección y los criterios propuestos por los autores para seleccionar herramientas de visualización.

Tabla 3.2. Trabajos que abordan criterios a tomar en cuenta a la hora de seleccionar una herramienta de visualización

Artículo	Título	Criterios
(Lowe & Matthee, 2020)	Requirements of Data Visualisation Tools to Analyse Big Data: A Structured Literature Review	Reducción de dimensionalidad, Reducción de datos, Escalabilidad y legibilidad, Interactividad, Recuperación rápida de resultados y Asistencia al usuario.

(Nazmuz Sakib, 2022)	Data Visualization in Data Science	Interactividad, Exploración interactiva de los datos, Renderizado rápido de resultados y Facilidad de uso
(Dogadina & Voronin, 2024)	Comparative Analysis of Data Visualization Tools	Facilidad de uso, Flexibilidad, Escalabilidad, Integración y Calidad de los resultados.

3.3 Aplicación de métodos de toma de decisiones multicriterio a la selección de herramientas.

Hassan & Elragal (2017) proponen un modelo de mejores prácticas para seleccionar herramientas de visualización para Big Data, que sirve como guía para las organizaciones. Este modelo se construyó usando el método Delphi, en el que participaron 14 expertos en Big Data. La investigación identificó 36 criterios como los más relevantes para la selección de este tipo de herramientas. Estos criterios se clasifican en seis tipos principales: técnicos, de visualización, de colaboración y movilidad, operativos, de gobierno de datos y requerimientos gerenciales. Este método de mejores prácticas ayuda a seleccionar y adquirir herramientas de visualización adecuadas, capaces de manejar y analizar Big Data. El estudio aborda una brecha notable en la literatura existente en relación con los criterios para seleccionar herramientas de visualización para Big Data. Si bien estudios previos investigaron características de este tipo de herramientas o las compararon, faltaba un modelo de mejores prácticas para la selección de estas.

Grandhi & Wibowo (2018) presentan un método de toma de decisiones multicriterio diseñado, específicamente, para evaluar y seleccionar herramientas de visualización de Big Data. El proceso de evaluación y selección involucró a un comité de cuatro tomadores de decisiones y consideró cuatro posibles alternativas de herramientas de visualización de Big Data: Aster Discovery, Power BI, Tableau y Looker. Este estudio identifica cuatro criterios cruciales para la evaluación: Costos, Usabilidad, Funcionalidad y Seguridad y sus resultados permitieron observar que los MCDM manejan de manera efectiva la naturaleza multidimensional de este tipo de problemas y aborda los aspectos subjetivos e imprecisos de los procesos de toma de decisiones.

Nazliel et al. (2022) presentan una metodología de selección de tecnología para el dominio de la Ciencia de Datos que tiene como objetivo ayudar a los profesionales e investigadores a elegir el stack de tecnología de Ciencia de Datos más adecuado a sus problemas específicos. El enfoque propuesto sintetiza métodos cuantitativos de toma de decisiones multicriterio (MCDM) con enfoques cualitativos de decisiones grupales, utilizando técnicas populares como el Proceso Analítico Jerárquico (AHP) y el Método Delphi para facilitar la toma de decisiones. La metodología propuesta cubre todo el proceso de Ciencia de Datos, de principio a fin, incluidos los niveles de infraestructura, plataforma y aplicación, abordando requerimientos funcionales y no funcionales. Esta metodología permite a los tomadores de decisiones comparar diferentes herramientas y seleccionar las adecuadas basadas en los requerimientos específicos del proyecto.

La Tabla 3.3 muestra los trabajos donde se emplean MCDM para seleccionar herramientas de visualización y los criterios tomados en cuenta durante el proceso.

Tabla 3.3. Trabajos relacionados con la creación de métodos de selección de herramientas

Artículo	Título	No de criterios
(Hassan & Elragal, 2017)	Big Data Visualization Tool: A Best-Practice Selection Model	36 criterios agrupados en seis dimensiones: Técnico, Visualización, Colaboración y movilidad, Operacional, Gobierno de Datos y Gerencial.
(Grandhi & Wibowo, 2018)	A Multi-Criteria Group Decision Making Method for Selecting Big Data Visualization Tools	Costo, Usabilidad, Funcionalidad y Seguridad.
(Nazliel et al., 2022)	Data Science Technology Selection: Development of a Decision-Making Approach	Requisitos funcionales y no funcionales de acuerdo con cada paso del ciclo de trabajo de Ciencia de Datos acordes al problema de negocio.

3.4 Conclusiones del análisis del Estado del Arte

Aunque investigaciones previas han abordado la selección de herramientas de visualización de datos desde diversas perspectivas, el enfoque propuesto en este trabajo introduce aportes diferenciados tanto en alcance como en metodología. Estudios como el de Desimoni & Po (2020), Schmidt (2020) y Shakeel et al. (2022) se centran principalmente en describir y comparar capacidades técnicas de las herramientas, sin ofrecer una metodología que permita una selección sistemática de las mismas en función del contexto y los requerimientos del usuario.

Lowe & Matthee (2020), Nazmuz Sakib (2022) y Dogadina & Voronin (2024) coinciden en la necesidad de contar con criterios bien definidos para seleccionar herramientas de visualización, aunque no ofrecen mecanismos formales para aplicarlos en una evaluación comparativa ni abordan cómo ponderarlos o combinarlos en un método que permita seleccionar herramientas de visualización de forma práctica.

Por otra parte, Hassan & Elragal (2017), Grandhi & Wibowo (2018) y Nazliel et al. (2022) aplicaron métodos de toma de decisiones multicriterio para proponer metodologías que permitan seleccionar herramientas de visualización. En el primer caso, su propuesta se encuentra orientada principalmente a contextos empresariales y no se ofrece información detallada de cómo hacer una asignación de pesos objetivos a los criterios que contribuya al proceso de selección. En Grandhi & Wibowo (2018), por otro lado, el alcance es limitado, ya que simplemente se aplica un MCDM para seleccionar herramientas de visualización, demostrando la capacidad de estos métodos para ser aplicados a este tipo de problemas, pero no se propone una guía replicable en otros contextos y su estudio tiene limitaciones tanto en número de herramientas como en profundidad metodológica. Nazliel et al. (2022), proponen una metodología integral para la selección de tecnologías en Ciencia de Datos, abarcando todo el proceso de la Ciencia de Datos de principio a fin. Su enfoque, al abarcar todo el proceso de Ciencia de Datos, es general en el tratamiento de las herramientas de visualización y no profundiza en las particularidades de estas.

A diferencia de estos trabajos, esta investigación propone un modelo específico denominado MATCH-Vis, enfocado exclusivamente en la selección de herramientas de visualización de datos. Este modelo se

construye a partir de un análisis del dominio, la identificación de criterios derivados de la literatura y su organización en cinco dimensiones clave: Metas, Almacenamiento de Datos, Tipo de datos, Contexto y Herramienta. La evaluación se realiza a través de métodos de toma de decisiones multicriterio como SAW y TOPSIS, con pesos objetivos determinados por el método de la entropía, lo que garantiza rigor y replicabilidad. Además, MATCH-Vis no solo determina la mejor alternativa, sino que ofrece una metodología clara y sistematizada para guiar al usuario en el proceso de selección según el contexto del usuario y los requerimientos específicos del problema de análisis. En este sentido, MATCH-Vis contribuye con una propuesta metodológica original que llena un vacío existente en el abordaje práctico y operativo de la selección de herramientas de visualización dentro del campo de la Ciencia de Datos.

Capítulo 4

Construcción del Método MATCH-Vis

En esta sección se presentan los principales hallazgos obtenidos a lo largo del proceso de investigación. Los mismos serán expuestos siguiendo el orden propuesto en la metodología de solución planteada en la Figura 1.2 y en concordancia con objetivos. Primeramente, se presentan las herramientas de visualización estudiadas, y el proceso llevado a cabo para la selección de un subconjunto de estas para su mayor profundización. Más adelante en la sección, se exponen los criterios extraídos de la literatura para evaluar herramientas de visualización y como se elaboró un sistema de puntajes para cuantificarlos. Finalmente, se describe la elaboración de la encuesta aplicada al conjunto de expertos en Ciencia de Datos que sirvió de base para elaborar la matriz de decisión y aplicar los métodos de toma de decisiones multicriterio SAW y TOPSIS y con esto construir el método de selección de herramientas de visualización MATCH-Vis.

4.1 Selección de herramientas de visualización

Como parte del desarrollo de este trabajo se realizó una búsqueda de información en la literatura especializada para identificar las principales herramientas de visualización de datos utilizadas en los últimos cinco años en entornos de Ciencia de Datos. Dichas herramientas se encuentran representadas en la Figura 4.1.

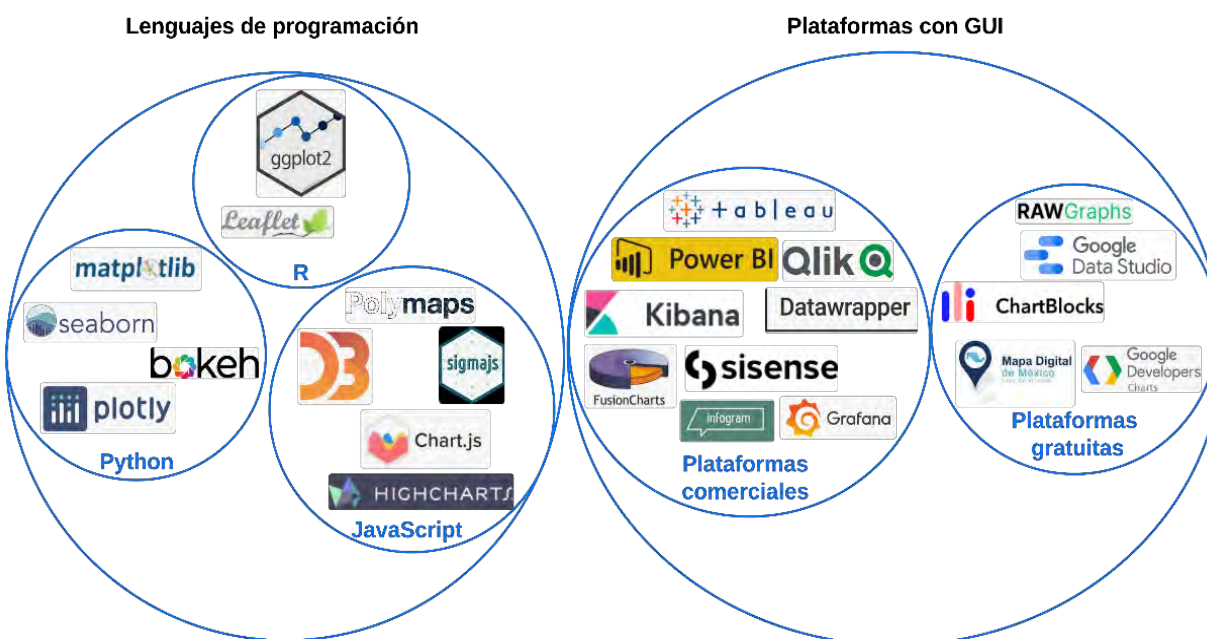


Figura 4.1. Herramientas de visualización identificadas en la investigación

La búsqueda en la literatura permitió caracterizar estas herramientas de acuerdo con su tipo, comunidad de uso, capacidades y limitaciones. Esta caracterización se puede encontrar en el Anexo A.

De este conjunto de 25 herramientas, se escogieron 11 para un análisis más detallado. Las herramientas seleccionadas se muestran en la Tabla 4.1.

Tabla 4.1. Herramientas de visualización seleccionadas para un análisis más profundo

Dominio técnico	Herramienta
Python	Matplotlib Seaborn Bokeh Plotly
R	Ggplot2 Leaflet
JavaScript	D3.js Highcharts Google Charts
Inteligencia de Negocios	Tableau Power BI

Estas herramientas, además de ser consideradas como muy potentes en la literatura reciente se caracterizan principalmente por ser gratuitas o por contar con versiones gratuitas, aunque limitadas, lo que permitió explorarlas en profundidad y evaluar ampliamente sus funcionalidades a través de la experimentación.

Como resultado de la experimentación se pudo construir la Tabla 4.2 que recoge las características principales de estas herramientas de acuerdo con cuatro dimensiones denominadas: Meta, Almacenamiento de datos, Tipo de datos y Contexto. Estas dimensiones forman el acrónimo MATCH, junto a la H de herramienta y servirá para denominar el método que se propone con esta investigación.

La dimensión Meta se refiere a lo que se desea visualizar, al propósito que se tiene con la visualización. La dimensión Almacenamiento, por su parte, hace referencia a dónde se encuentran almacenados los datos. Tipo de datos señala el tipo de datos que se pretende visualizar y la dimensión Contexto las habilidades del usuario, el escenario de uso, las características de la herramienta que se deban considerar, entre otros factores. Finalmente, la dimensión Herramienta corresponde a una descripción de la herramienta en función de las dimensiones anteriores.

Tabla 4.2. Caracterización de herramientas según las dimensiones MATCH

Herramienta	Meta	Almacenamiento	Tipo de datos	Contexto
Matplotlib	Hacer reportes y visualizaciones estáticas. Análisis exploratorio	Archivos locales (CSV, Excel). Datos pequeños y medianos	Numéricos y categóricos	Uso académico, científico. Usuarios con conocimientos de programación en Python. Permite mucha personalización. Gratuita
Seaborn	Análisis estadísticos.	Archivos locales (CSV, Excel). Datos pequeños y medianos	Numéricos y categóricos	Uso académico, científico. Usuarios con conocimientos de programación en Python. Gráficos atractivos con poco código. Gratuita
Bokeh	Insertar visualizaciones en páginas webs. Paneles interactivos	Soporta streaming. Archivos medianos y grandes	Numéricos, series temporales.	Entornos de análisis en tiempo real, y construcción de paneles interactivos en la web. Usuarios con conocimiento de Python. Gratuita

CAPÍTULO 4: CONSTRUCCIÓN DEL MÉTODO MATCH-Vis

Plotly	Paneles interactivos, visualización web	Archivos locales, bases de datos medianas, integración web	Numéricos, categóricos, series temporales	Usuarios que necesitan compartir gráficos en línea. Es compatible con Python, Julia, JavaScript y R. Gratuita
Ggplot2	Visualizaciones estáticas. Análisis exploratorio. Publicaciones científicas	Archivos locales (Excel, CSV). Bases de datos pequeñas y medianas	Numéricos, categóricos, geoespaciales	Investigadores, académicos. Usuarios con conocimiento de R. Gratuita
Leaflet	Mapas interactivos para la web. Análisis geográfico.	Datos geoespaciales, archivos GeoJSON.	Datos georreferenciados	Usuarios con conocimiento de R. Gratuita
D3.js	Paneles interactivos complejos.	Cualquier tipo de fuente. Soporta grandes volúmenes de datos	Numéricos, categóricos, jerárquicos, series temporales	Desarrolladores web avanzados, visualizaciones para portales web. Contiene código preconstruido. Gratuita
Highcharts	Visualizaciones empresariales interactivas.	Bases de datos medianas y grandes. Interacción directa	Numéricos, categóricos, series temporales	Gráficos web responsivos. Es de pago.
Google Charts	Paneles y reportes interactivos sencillos	Conexión sencilla a hojas de cálculo, bases de datos pequeñas/medianas	Numéricos, categóricos, geográficos	Usuarios de Google Suite, pequeños equipos de análisis de datos. Tiene código preconstruido. Gratuita
Power BI	Paneles interactivos empresariales.	Cualquier tipo de fuente de datos: archivos locales, bases de datos, archivos en la nube, streaming	Numéricos, categóricos, geográficos, series temporales	Empresas, analistas de negocio, usuarios sin necesidad de programar. Funciones avanzadas y cooperación son comerciales
Tableau	Paneles interactivos empresariales.	Cualquier tipo de fuente de datos: archivos locales, bases de datos, archivos en la nube, streaming	Numéricos, categóricos, geográficos, series temporales	Empresas, analistas de negocio, usuarios sin necesidad de programar. Es una herramienta comercial

4.2 Determinación de los criterios de evaluación

Luego de haber seleccionado las herramientas para la realización del estudio se observó la necesidad de plantear una serie de criterios que permitan comparar las herramientas entre sí para determinar cuál es la más adecuada a utilizar en un escenario determinado. Estos criterios se construyeron con base en la literatura, ya que se destacan como factores que toda herramienta de visualización debe cumplir y fueron complementados con la experiencia adquirida por parte de la autora durante la experimentación con las herramientas. Los criterios se muestran en la Tabla 4.3.

Tabla 4.3. Criterios para evaluar las herramientas de visualización

Categoría	Criterios
Usabilidad	1. Facilidad de uso (Lowe & Matthee, 2020) (Dogadina & Voronin, 2024) 2. Flexibilidad en la edición de gráficos (Dogadina & Voronin, 2024) 3. Asistencia en seleccionar el tipo de gráfico (Lowe & Matthee, 2020) (Nazmuz Sakib, 2022)
Complejidad	4. Variedad de tipos de gráficos (Nazmuz Sakib, 2022) 5. Exploración interactiva de los datos (Lowe & Matthee, 2020) 6. Manejo de grandes conjuntos de datos (Nazmuz Sakib, 2022)(Dogadina & Voronin, 2024)
Integración	7. Fuentes y tipos de datos (Dogadina & Voronin, 2024)

	8. Archivos de exportación
Costo	9. Costo 10. Renderizado rápido de resultados (Lowe & Matthee, 2020)
Calidad	11. Calidad de la visualización (Dogadina & Voronin, 2024)

Los criterios fueron organizados en cinco categorías diferentes de acuerdo con su similitud y es notorio destacar que tanto esta agrupación como los criterios: Archivos de exportación y Costo fueron propuestos por la autora con base en su experiencia, y no tomados de la literatura. A continuación, se detallan en mayor profundidad los aspectos que evalúa cada criterio:

4.2.1 Categoría Usabilidad

La categoría Usabilidad se refiere a la capacidad que tiene una herramienta para ser utilizada de forma sencilla, eficiente e intuitiva por usuarios con distintos perfiles técnicos y niveles de experiencia. Una alta usabilidad permite a los usuarios aprovechar mejor las funciones del software, disminuyendo la curva de aprendizaje y facilitando su uso. En esta categoría se encuentran agrupados los criterios: facilidad de uso, flexibilidad en la edición de gráfico y asistencia en la selección del tipo de gráfico.

4.2.1.1 Criterio Facilidad de uso

Este criterio evalúa la simplicidad con la que los usuarios pueden interactuar con la herramienta sin requerir conocimientos técnicos avanzados. En (Dogadina & Voronin, 2024) destacan que la facilidad de uso es esencial para garantizar una adopción efectiva de la herramienta por parte de usuarios con diferentes niveles de experiencia. Una interfaz intuitiva y amigable permite a los usuarios centrarse en el análisis de datos en lugar de en la complejidad de la herramienta.

4.2.1.2 Criterio Flexibilidad en la edición de gráficos

La capacidad de personalizar y modificar visualizaciones según las necesidades específicas del usuario es crucial. Una herramienta flexible permite ajustar elementos como colores, etiquetas y tipos de gráficos, facilitando una representación más precisa de los datos y adaptándose a diversos contextos de análisis (Dogadina & Voronin, 2024).

4.2.1.3 Criterio Asistencia en la selección del tipo de gráfico

Este criterio se refiere a la capacidad de la herramienta para sugerir o guiar al usuario en la elección del tipo de visualización más adecuado para sus datos. En (Nazmuz Sakib, 2022) se señala que una asistencia efectiva en este aspecto mejora la comprensión de los datos y evita representaciones inadecuadas que puedan llevar a interpretaciones erróneas.

4.2.2 Categoría Complejidad

La categoría Complejidad hace referencia a la capacidad de la herramienta para gestionar eficazmente visualizaciones que requieren múltiples funcionalidades como representaciones sofisticadas o grandes volúmenes de información. Su importancia reside en que, cuanto mayor sea la capacidad para manejar visualizaciones complejas, más versátil será la herramienta, permitiendo a los usuarios analizar patrones avanzados, explorar datos en profundidad e interactuar dinámicamente con ellos. Esto aumenta significativamente la utilidad del software en contextos analíticos avanzados o científicos. En esta categoría se encuentran los criterios: variedad de tipos de gráficos disponibles, exploración interactiva de los datos y manejo de grandes conjuntos de datos.

4.2.2.1 Criterio Variedad de tipos de gráficos disponibles

Una amplia gama de opciones de visualización permite a los usuarios representar distintos tipos de datos de diferentes maneras, adaptándose a las características específicas de cada conjunto de datos y el mensaje que se quiere transmitir. La disponibilidad de diversos tipos de gráficos es fundamental para abordar distintas necesidades analíticas y comunicativas (Dogadina & Voronin, 2024).

4.2.2.2 Criterio Exploración interactiva de los datos

La capacidad de interactuar con las visualizaciones, como hacer zoom, filtrar o resaltar elementos, facilita una comprensión más profunda de los datos. (Lowe & Matthee, 2020) indican que la interactividad en las herramientas de visualización es esencial para el análisis exploratorio de datos, permitiendo a los usuarios descubrir patrones y relaciones ocultas.

4.2.2.3 Criterio Manejo de grandes conjuntos de datos

Este criterio evalúa la capacidad de la herramienta para procesar y visualizar eficientemente grandes volúmenes de datos sin comprometer el rendimiento. (Dogadina & Voronin, 2024; Nazmuz Sakib, 2022) subrayan que una herramienta eficaz debe mantener un rendimiento óptimo incluso al trabajar con conjuntos de datos de gran tamaño, garantizando tiempos de respuesta adecuados y visualizaciones precisas.

4.2.3 Categoría Integración

La categoría Integración representa la capacidad que tiene la herramienta para interoperar y combinarse de manera eficiente con otras tecnologías, plataformas o sistemas ya existentes. Su impacto está directamente relacionado con la facilidad para conectar diferentes tipos y formatos de fuentes de datos, así como con la posibilidad de exportar visualizaciones resultantes hacia formatos externos o plataformas específicas. La integración efectiva reduce tiempos, mejora la compatibilidad tecnológica y facilita una comunicación más ágil y precisa de los resultados generados. Los criterios comprendidos en esta categoría son: las fuentes y tipos de datos y los archivos de exportación.

4.2.3.1 Criterio Fuentes y tipos de datos

La compatibilidad con múltiples fuentes de datos, como bases de datos, archivos planos o servicios en la nube, es esencial para que una herramienta sea versátil. En (Dogadina & Voronin, 2024) se indica que una amplia capacidad de integración permite a los usuarios consolidar información de diferentes orígenes para un análisis más completo. La integración flexible con diversas fuentes y tipos de datos permite, no solo ampliar la utilidad de la herramienta en diferentes contextos, sino, además, reducir la necesidad de utilizar herramientas adicionales para la transformación o adaptación de los datos antes de su visualización, optimizando así el proceso analítico completo.

4.2.3.2 Criterio Archivos de exportación

La posibilidad de exportar visualizaciones en diversos formatos (PDF, PNG, HTML, etc.) facilita la difusión y presentación de los resultados del análisis. Las herramientas con opciones de exportación variadas permiten a los usuarios compartir sus hallazgos de manera efectiva con diferentes audiencias.

4.2.4 Categoría Costo

La categoría Costo se refiere tanto al aspecto económico directo relacionado con la adquisición o licenciamiento del software, como también a la eficiencia con que se generan y visualizan los resultados en términos de tiempo de procesamiento y renderizado. Esta categoría es especialmente relevante en contextos donde existen restricciones presupuestarias. De igual manera, las herramientas con tiempos de renderizado elevados tienden a ser percibidas como poco funcionales o ineficientes, disminuyendo su aceptabilidad y utilidad práctica en entornos de análisis que demandan inmediatez y agilidad en la generación de resultados visuales. Los criterios asociados a esta categoría son costo y renderizado rápido de resultados.

4.2.4.1 Criterio Costo

Este criterio considera el precio de adquisición y mantenimiento de la herramienta, así como la disponibilidad de versiones gratuitas o de código abierto. El costo es un factor determinante en la elección de herramientas de visualización, especialmente para organizaciones o usuarios con recursos limitados.

4.2.4.2 Criterio Renderizado rápido de resultados

La velocidad con la que la herramienta genera visualizaciones influye en la eficiencia del análisis. (Lowe & Matthee, 2020) señala que un renderizado rápido mejora la experiencia del usuario y permite una exploración más fluida de los datos.

4.2.5 Categoría Calidad

La categoría Calidad hace referencia al nivel de claridad, precisión visual y efectividad comunicativa de las representaciones gráficas generadas por la herramienta. Las visualizaciones claras y atractivas permiten una interpretación más sencilla, rápida y exacta de los datos, facilitando decisiones mejor informadas. El impacto de contar con alta calidad visual es crucial para garantizar que los mensajes extraídos de los análisis sean correctamente entendidos, minimizando malentendidos y mejorando la percepción global de los resultados obtenidos. Esta categoría contiene al criterio calidad de la visualización.

4.2.5.1 Criterio Calidad de la visualización

Este criterio evalúa la claridad, estética y efectividad de las visualizaciones producidas por la herramienta. (Dogadina & Voronin, 2024) destaca que una alta calidad en las visualizaciones facilita la comprensión de los datos y mejora la comunicación de los hallazgos.

4.3 Creación de puntajes de evaluación para los criterios

Como se explicó previamente en la sección de Metodología, los criterios definidos son de carácter cualitativo por lo que se elaboró un sistema de puntuación con una escala de 0 a 5 para cuantificar cada criterio. Los detalles de dicha escala se presentan en el Anexo C.

4.4 Construcción de la encuesta

Tomando como base estos criterios y las herramientas previamente seleccionadas se conformó una encuesta para recabar valoraciones de expertos en Ciencia de Datos de las herramientas con respecto a los criterios. Esta encuesta se presenta en el Anexo B.

La encuesta lanzada solicita información personal a los participantes, como su nivel de estudios y profesión, además de su nivel de dominio y experiencia con las herramientas seleccionadas. A los participantes les fue requerido evaluar las herramientas, tomando en cuenta el sistema de puntuación previamente creado para cuantificar los criterios y se les invitó a sugerir otras herramientas que no se hubiesen considerado en la investigación, solicitando información sobre el nivel de dominio y experiencia que tenían con las mismas. Finalmente, se les solicitó evaluarlas también, buscando con este enfoque enriquecer los resultados del estudio.

4.5 Evaluación de las herramientas y experimentación

Durante la investigación se realizó una evaluación para comparar las diferentes herramientas usando los métodos de toma de decisiones multicriterio SAW y TOPSIS. Como procedimiento inicial para aplicar estos métodos matemáticos se utiliza una matriz de decisión, creada a partir de los resultados de la aplicación de la encuesta.

4.5.1 Análisis descriptivo

La encuesta fue contestada por un total de 21 participantes, los cuales fueron seleccionados debido a que pertenecen a diversos sectores profesionales que en su labor diaria hacen uso de herramientas de visualización de datos. La muestra se encuentra distribuida de la siguiente forma, tal como se muestra en la Figura 4.2.

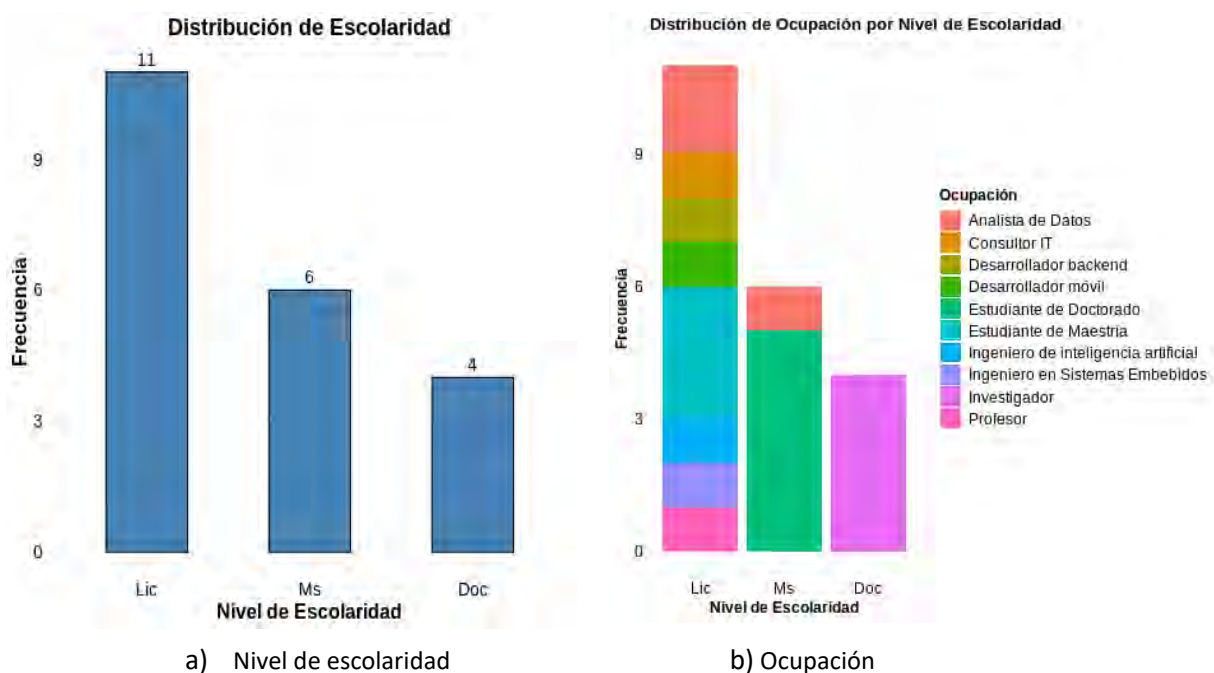


Figura 4.2. Distribución de participantes según: a) nivel de escolaridad y b) ocupación

Como resultado de la encuesta se observó que no todos los participantes dominaban todas las herramientas obteniéndose un número de respuestas por herramientas tal como se muestra en la Figura

4.3. La distribución del nivel de dominio de los diferentes niveles de escolaridad por herramienta se puede observar en el Anexo E.

4.5.2 Tratamiento de los datos faltantes

Es apreciable a simple vista que el número de datos faltantes en la muestra es representativo. Esto provocó que se hiciera necesario entender el patrón de datos faltantes, determinándose que el mismo es MAR (*Missing At Random*), por sus siglas en inglés, es decir “perdidos al azar”. Este tipo de patrón se explica en que la ausencia de los datos faltantes depende de una o varias variables observadas (Jakobsen et al., 2017). En este caso de estudio estuvo determinado por el nivel de conocimiento que tenían los participantes de cada herramienta. Cuando un encuestado contestaba “no tener conocimiento” sobre la herramienta, no podía por ende realizar una evaluación de esta de acuerdo con los criterios propuestos. En la Figura 4.3 se muestra el porcentaje de datos faltantes por herramienta.

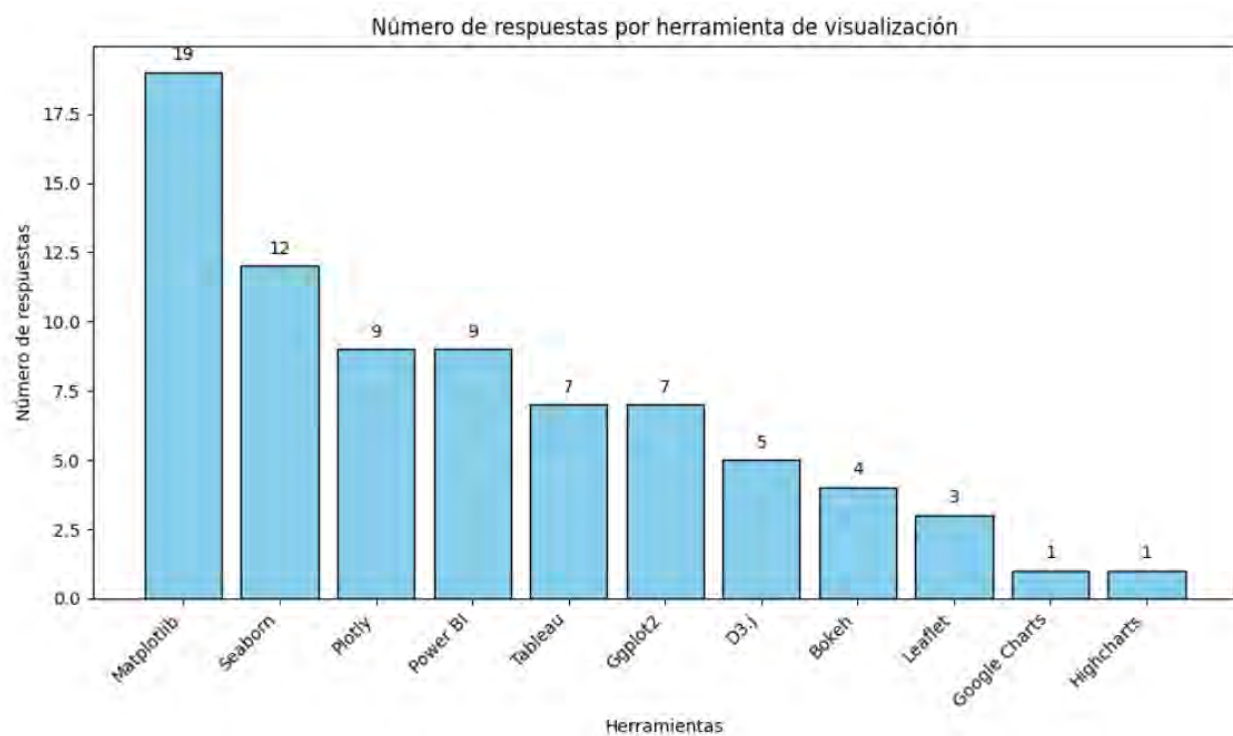


Figura 4.3. Número de respuestas por herramientas

Tabla 4.4. Porcentaje de datos faltantes por herramientas con respecto al total de participantes

Herramienta	Datos faltantes (% del total)
Highcharts	95.23
Google Charts	95.23
Bokeh	80.95
Leaflet	85.71
D3.js	76.19
Ggplot2	66.66

Tableau	66.66
Plotly	57.14
Power BI	57.14
Seaborn	42.86
Matplotlib	0.095

De acuerdo con (Azur et al., 2011; Stef van Buuren, 2018) una tasa de valores faltantes superior al 50% introduce errores en el análisis, ya que conduce a especulaciones sobre los datos. Estos casos se pueden tratar eliminando las variables con muy baja tasas de respuesta o haciendo un análisis de sensibilidad. Es por ello que se tomó la decisión de eliminar del estudio las variables con menos de 7 respuestas por participantes, es decir, las herramientas: Bokeh, Leaflet, D3.js, Highcharts y Google Charts debido a su baja tasa de respuesta. Los métodos de toma de decisiones multicriterio fueron aplicados a las herramientas restantes: Matplotlib, Seaborn, Plotly, Ggplot2, Tableau y Power BI.

4.5.3 Métodos de toma de decisión multicriterio

Una vez definidas las herramientas para llevar a cabo la evaluación se procedió a construir la matriz de decisión con las respuesta de los expertos en cada uno de los criterios definidos para cada herramienta. Para ello, en primer lugar, se procedió a eliminar los valores faltantes y luego se calculó por cada herramienta para cada criterio la media ponderada de las evaluaciones, bajo la fórmula (4.1):

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \quad (4.1)$$

Finalmente, la matriz de decisión quedó tal como se representa en la Tabla 4.5.

Tabla 4.5. Matriz de decisión

Criterios	Matplotlib	Seaborn	Plotly	Ggplot2	Tableau	Power BI
Facilidad de uso	2.88	2.57	3.23	3	4.44	4.42
Flexibilidad en la edición de gráficos	3.5	3.29	3.77	2.7	4.22	4.5
Asistencia en la selección del tipo de gráfico	2.28	2.29	3.23	2.7	4.33	4.25
Variedad de tipo de gráficos disponibles	3.94	3.36	3.92	3.3	4.22	4.5
Exploración interactiva de los datos	2.75	2.36	3.92	2.7	4.67	4.5
Manejo de grandes conjuntos de datos	3.6	3.07	3.46	3.8	4.44	4.08
Fuentes y tipos de datos	3.38	3.14	4	3.3	4.67	4.33
Archivos de exportación	3.62	3.57	4.15	3.7	4.44	4.33
Costo	4.19	4.14	4.08	4.4	1.44	2.67
Renderizado rápido de resultados	3.12	3.14	3.69	3.4	4.11	4.67
Calidad de la visualización	3.59	3.71	4	3.9	4.89	4.58

4.5.4 Aplicación del método SAW

Siguiendo el flujo descrito en la Figura 2.2, el siguiente paso para aplicar SAW luego de tener la matriz de decisión es definir si los criterios son de costo o beneficio. En nuestro caso de estudio todos los criterios son de beneficio, debido a que los criterios de costo: “Costo” y “Renderizado rápido de resultados” fueron descritos en su escala de evaluación de forma que obtener un puntaje alto en los mismos significaba que la herramienta era gratuita o prácticamente gratuita y que renderizaba sus gráficos de forma muy eficiente, respectivamente. Por lo tanto, los criterios se normalizaron siguiendo la ecuación (2.1).

La matriz de decisión normalizada quedó, tal como se muestra en la Tabla 4.6.

Tabla 4.6. Matriz normalizada con el método SAW

Criterios	Matplotlib	Seaborn	Plotly	Ggplot2	Tableau	Power BI
Facilidad de uso	0.647	0.579	0.727	0.675	1.000	0.994
Flexibilidad en la edición de gráficos	0.778	0.730	0.838	0.600	0.938	1.000
Asistencia en la selección del tipo de gráfico	0.526	0.527	0.746	0.623	1.000	0.981
Variedad de tipo de gráficos disponibles	0.926	0.790	0.923	0.776	0.993	1.000
Exploración interactiva de los datos	0.589	0.505	0.841	0.579	1.000	0.964
Manejo de grandes conjuntos de datos	0.689	0.691	0.779	0.855	1.000	0.919
Fuentes y tipos de datos	0.723	0.673	0.857	0.707	1.000	0.929
Archivos de exportación	0.816	0.804	0.935	0.833	1.000	0.975
Costo	0.952	0.942	0.927	1.000	0.328	0.606
Renderizado rápido de resultados	0.670	0.673	0.791	0.729	0.881	1.000
Calidad de la visualización	0.735	0.760	0.818	0.798	1.000	0.937

Con la matriz de decisión normalizada, se calculó el peso de los criterios siguiendo los siguientes pasos. Primero, se calculó la entropía de cada criterio mediante la fórmula (2.3). Luego, se obtuvo la dispersión para cada criterio mediante la fórmula (2.4) y finalmente con el valor de dispersión se obtiene el peso de cada criterio aplicando la ecuación (2.5). Los pesos de los criterios quedaron tal como se muestra en la Tabla 4.7.

Tabla 4.7. Pesos de los criterios en SAW

Criterios	Pesos
Facilidad de uso	0.1091
Flexibilidad en la edición de gráficos	0.0644
Asistencia en seleccionar el tipo de gráfico	0.1685
Variedad de tipos de gráficos	0.0237
Exploración interactiva de los datos	0.1674
Manejo de grandes conjuntos de datos	0.0465
Fuentes y tipos de datos	0.0532

Archivos de exportación	0.0188
Costo	0.2655
Renderizado rápido de resultados	0.0526
Calidad de la visualización	0.0304

Luego, con la matriz normalizada y los pesos de cada criterio se calcularon los puntajes de evaluación de cada alternativa siguiendo la fórmula (2.6). Los puntajes de evaluación y el orden del ranking de herramientas se muestran en la Tabla 4.8.

Tabla 4.8. Ranking de herramientas según el método SAW

Ranking	Herramienta	Puntaje
1	Power BI	0.88
2	Plotly	0.83
3	Tableau	0.81
4	Ggplot2	0.75
5	Matplotlib	0.73
6	Seaborn	0.69

4.5.5 Aplicación del método TOPSIS

Al igual que con el método SAW, el siguiente paso en el método TOPSIS es identificar si los criterios son de costo o beneficio, y esta clasificación quedó establecida desde el método anterior. Cada elemento de la matriz de decisión se normalizó siguiendo la fórmula (2.7) y dejando como resultado la matriz que se representa en la Tabla 4.9.

Tabla 4.9. Matriz normalizada con el método TOPSIS

Criterios	Matplotlib	Seaborn	Plotly	Ggplot2	Tableau	Power BI
Facilidad de uso	0.335	0.300	0.377	0.350	0.518	0.515
Flexibilidad en la edición de gráficos	0.385	0.361	0.415	0.297	0.465	0.495
Asistencia en la selección del tipo de gráfico	0.283	0.284	0.401	0.335	0.537	0.527
Variedad de tipo de gráficos disponibles	0.418	0.356	0.416	0.350	0.448	0.451
Exploración interactiva de los datos	0.312	0.267	0.445	0.306	0.529	0.510
Manejo de grandes conjuntos de datos	0.339	0.340	0.383	0.421	0.492	0.452
Fuentes y tipos de datos	0.358	0.334	0.425	0.350	0.495	0.460
Archivos de exportación	0.371	0.366	0.425	0.379	0.455	0.444
Costo	0.469	0.464	0.456	0.492	0.162	0.298
Renderizado rápido de resultados	0.342	0.344	0.404	0.372	0.450	0.511
Calidad de la visualización	0.354	0.366	0.394	0.385	0.482	0.452

CAPÍTULO 4: CONSTRUCCIÓN DEL MÉTODO MATCH-Vis

Posteriormente, se calculó el peso de los criterios siguiendo la fórmula (2.5) planteadas en el método SAW y los resultados fueron los que se muestran en la Tabla 4.10.

Tabla 4.10. Pesos de los criterios para el método TOPSIS

Criterios	Pesos
Facilidad de uso	0.1091
Flexibilidad en la edición de gráficos	0.0644
Asistencia en seleccionar el tipo de gráfico	0.1685
Variedad de tipos de gráficos	0.0237
Exploración interactiva de los datos	0.1674
Manejo de grandes conjuntos de datos	0.0465
Fuentes y tipos de datos	0.0532
Archivos de exportación	0.0188
Costo	0.2655
Renderizado rápido de resultados	0.0526
Calidad de la visualización	0.0304

Con estos pesos se construyó la matriz de decisión ponderada, expuesta en la Tabla 4.11 y luego se determinaron las soluciones ideales positiva y negativas, tal como indica la Tabla 4.12.

Tabla 4.11. Matriz ponderada con método TOPSIS

Criterios	Matplotlib	Seaborn	Plotly	Ggplot2	Tableau	Power BI
Facilidad de uso	0.037	0.033	0.041	0.038	0.057	0.056
Flexibilidad en la edición de gráficos	0.025	0.023	0.027	0.019	0.030	0.032
Asistencia en la selección del tipo de gráfico	0.048	0.048	0.068	0.056	0.091	0.089
Variedad de tipo de gráficos disponibles	0.010	0.008	0.010	0.008	0.011	0.011
Exploración interactiva de los datos	0.052	0.045	0.074	0.051	0.089	0.085
Manejo de grandes conjuntos de datos	0.016	0.016	0.018	0.020	0.023	0.021
Fuentes y tipos de datos	0.019	0.018	0.023	0.019	0.026	0.024
Archivos de exportación	0.007	0.007	0.008	0.007	0.009	0.008
Costo	0.124	0.123	0.121	0.131	0.043	0.079
Renderizado rápido de resultados	0.018	0.018	0.021	0.020	0.024	0.027
Calidad de la visualización	0.011	0.011	0.012	0.012	0.015	0.014

Tabla 4.12. Soluciones ideales positivas y negativas con TOPSIS

Criterios	S_j^+	S_j^-
Facilidad de uso	0.057	0.033
Flexibilidad en la edición de gráficos	0.032	0.019

Asistencia en seleccionar el tipo de gráfico	0.091	0.048
Variedad de tipos de gráficos	0.011	0.008
Exploración interactiva de los datos	0.089	0.045
Manejo de grandes conjuntos de datos	0.023	0.016
Fuentes y tipos de datos	0.026	0.018
Archivos de exportación	0.009	0.007
Costo	0.131	0.043
Renderizado rápido de resultados	0.027	0.018
Calidad de la visualización	0.015	0.011

Finalmente, se calcularon las distancias euclidianas de cada herramienta a las soluciones ideales positiva y negativa y se calculó el coeficiente de proximidad relativa C_i de cada alternativa mediante la fórmula (2.13).

La Tabla 4.13 muestra un ranking con los puntajes de proximidad, dónde las alternativas fueron ordenadas de acuerdo con su valor de C_i .

Tabla 4.13. Ranking de herramientas según el método TOPSIS

Ranking	Herramienta	CI
1	Plotly	0.72
2	Ggplot2	0.61
3	Power BI	0.59
4	Matplotlib	0.56
5	Seaborn	0.54
6	Tableau	0.44

La literatura sugiere que usar otras medidas de distancia en TOPSIS puede ofrecer una mayor robustez frente a datos faltantes y menor sensibilidad a valores extremos (Ciardiello & Genovese, 2023). En este contexto, se decidió seguir experimentando con TOPSIS, pero esta vez usando la distancia de Manhattan, la cual se calcula mediante las fórmulas siguientes Para calcular la distancia de la alternativa a la solución ideal positiva se utiliza la ecuación (4.2):

$$S_j^+ = \sum_{j=1}^n |v_{ij} - v_j^+|, \quad (4.2)$$

mientras que para calcular la distancia de la alternativa a la solución ideal negativa se emplea la ecuación (4.3):

$$S_j^- = \sum_{j=1}^n |v_{ij} - v_j^-| \quad (4.3)$$

Donde, v_{ij} es el valor ponderado del criterio j para la alternativa i , v_j^+ es el valor ideal positivo (el mejor valor para el criterio j) y v_j^- es el valor ideal negativo (el peor valor para el criterio j) respectivamente.

CAPÍTULO 4: CONSTRUCCIÓN DEL MÉTODO MATCH-Vis

La distancia de Manhattan ha sido recomendada a utilizar en MCDM donde existen datos escasos, incompletos o con variabilidad alta, tal como sucede en nuestro caso de estudio. La distancia euclidiana eleva las diferencias al cuadrado y esto puede amplificar la influencia de datos atípicos, mientras que, la distancia de Manhattan permite una interpretación más directa y estable de las diferencias.

Aplicando TOPSIS con distancia de Manhattan se consiguió que los puntajes de proximidad de cada herramienta fuesen los mostrados en la Tabla 4.14.

Tabla 4.14. Ranking de herramientas usando el método TOPSIS con distancia de Manhattan

Ranking	Herramienta	CI
1	Power BI	0.75
2	Plotly	0.65
3	Tableau	0.62
4	Ggplot2	0.48
5	Matplotlib	0.42
6	Seaborn	0.35

4.6 Integración del Método MATCH-Vis

Posteriormente, los resultados obtenidos y el conocimiento adquirido durante la realización de la investigación se integraron en un método que permite la selección de herramientas de visualización de una forma sistemática y estructurada. Este método se denominó MATCH-Vis y los pasos para su aplicación se exponen a continuación:

Paso 1: Definir el escenario de uso según las dimensiones MATCH.

Definir el escenario de uso según las dimensiones MATCH: meta, almacenamiento de datos, tipo de datos, contexto y herramienta, tal como se muestra en la Tabla 4.15.

Tabla 4.15. Dimensiones MATCH para establecer el escenario de uso de herramientas de visualización

Dimensión	Descripción	Preguntas para la guía
M (Meta)	Se refiere al propósito principal de la visualización de datos.	¿Cuál es el objetivo de la visualización? (exploración de los datos, visualización estática, visualización web...)
A (Almacenamiento)	Hace referencia al lugar donde están ubicados los datos que se desean visualizar.	¿Dónde se encuentran almacenados los datos? (archivos locales, bases de datos distribuidas, la nube...)
T (Tipo de datos)	Describe las características estructurales de los datos a visualizar.	¿Qué tipo de datos se van a visualizar? (numéricos, categóricos, georreferenciados, jerárquicos, conectados, series de tiempo...)
C (Contexto)	Considera el perfil del usuario final, su nivel de experiencia técnica y los recursos disponibles.	¿Quién es el usuario final? (científico de datos, usuario de negocio, estudiante...) ¿Cuál es el nivel de experiencia del usuario? (es principiante, sabe programar o no)

CAPÍTULO 4: CONSTRUCCIÓN DEL MÉTODO MATCH-Vis

		¿Cuánto está dispuesto a invertir en la herramienta?
H (Herramienta)	Corresponde a la herramienta de visualización seleccionada según las características definidas en las dimensiones anteriores.	Herramienta que corresponde a las características anteriores

Paso 2: Definir criterios de evaluación

Seleccionar los criterios relevantes para comparar las herramientas. Puede usar todos, solo algunos o integrar nuevos:

1. Facilidad de uso
2. Flexibilidad en la edición de gráficos
3. Asistencia en la selección del gráfico
4. Variedad de tipos de gráficos
5. Exploración interactiva de los datos
6. Manejo de grandes conjuntos de datos
7. Fuentes y tipos de datos
8. Archivos de exportación
9. Costo
10. Renderizado rápido de resultados
11. Calidad de la visualización

En la Tabla 4.16 se muestran los criterios agrupados de acuerdo con las categorías a las que pertenecen y a las dimensiones MATCH a las que se encuentran asociados.

Tabla 4.16. Criterios agrupados por las dimensiones del método MATCH-Vis y sus categorías

	Meta	Almacenamiento	Tipo de datos	Contexto
Usabilidad				Facilidad de uso Flexibilidad en la edición de gráficos Asistencia en la selección del gráfico
Complejidad	Variedad de tipos de gráficos Exploración interactiva de los datos		Manejo de grandes conjuntos de datos	
Integración	Archivos de exportación	Fuentes y tipos de datos	Fuentes y tipos de datos	
Costo				Costo
Calidad	Calidad de la visualización			

Paso 3: Construir la matriz de decisión

- 3.1. Identificar herramientas candidatas.
- 3.2. Recolectar evaluaciones por herramienta y por criterio.
 - 3.2.1. Usar encuestas, entrevistas o revisión técnica. La encuesta utilizada en esta investigación se puede encontrar en el Anexo B.
 - 3.2.2. Para medir cuantitativamente las evaluaciones usar los puntajes propuestos por cada criterio. Ver Anexo C.

Paso 4: Ponderar los criterios

- 4.1. Utilizar Método de entropía para calcular los pesos (evita sesgos subjetivos).
- 4.2. Usar asignación por expertos si se prefiere.

Paso 5: Aplicar el método de evaluación

- 5.1. Elegir el método multicriterio:
 - SAW: simple y efectivo cuando los criterios son de beneficio.
 - TOPSIS: permite identificar alternativas cercanas al ideal positivo y alejadas del negativo.Para concordancia en los resultados de ambos métodos utilizar TOPSIS con distancia de Manhattan.
- 5.2. Calcular puntuaciones finales

Paso 6: Interpretar y seleccionar

- 6.1. Analizar los resultados del ranking.
- 6.2. Validar si la herramienta mejor posicionada cumple con las restricciones del contexto.

4.7 Conclusiones del capítulo

En este capítulo se desarrolló paso a paso la construcción del método MATCH-Vis mediante la integración coherente del estudio del dominio, la caracterización de las herramientas, la evaluación empírica y los métodos de toma de decisiones multicriterio. Se identificó un conjunto inicial de 25 herramientas de visualización destacadas en la literatura reciente, de ahí se seleccionaron 11 con base en su disponibilidad gratuita y amplia adopción. Estas herramientas fueron caracterizadas de acuerdo con las dimensiones MATCH (Meta, Almacenamiento, Tipo de datos, Contexto y Herramienta), lo que permitió contextualizar su aplicabilidad en distintos escenarios de Ciencia de Datos. Se establecieron, además, 11 criterios de evaluación, agrupados en cinco categorías clave (Usabilidad, Complejidad, Integración, Costo y Calidad), incluyendo aportaciones originales como los criterios Archivos de exportación y Costo y la categorización propuesta. La recolección de datos permitió obtener valoraciones empíricas a través de la aplicación de una encuesta a especialistas, aunque debido a la presencia de datos faltantes fue necesario excluir cinco herramientas del estudio posterior. El análisis de pesos por entropía en los métodos de toma de decisiones multicriterio SAW y TOPSIS reveló que el criterio Costo tuvo la mayor influencia en el proceso de decisión, seguido por aspectos vinculados a la asistencia gráfica e interactividad. La aplicación de los métodos SAW y TOPSIS permitió generar un ranking comparativo entre las seis herramientas con las que se realizó el estudio finalmente, destacando a Power BI y Plotly como las opciones más robustas, aunque con ligeras variaciones según la métrica de distancia utilizada. Finalmente, los resultados obtenidos fueron integrados en un procedimiento estructurado de seis pasos que da forma al método MATCH-Vis, proporcionando una guía sistemática, reproducible y adaptable para la selección de herramientas de visualización de datos, con potencial de aplicación tanto en entornos académicos como profesionales.

Capítulo 5

Resultados y Discusión

En este capítulo se presenta una discusión de los principales resultados obtenidos con la realización de esta investigación.

5.1 Caracterización de las herramientas seleccionadas

La aplicación del método MATCH-Vis permitió evaluar de manera estructurada un conjunto representativo de herramientas de visualización. Las 24 herramientas seleccionadas inicialmente para el estudio se categorizaron en dos grupos principales: bibliotecas de lenguajes de programación (Python, R y JavaScript) y plataformas GUI, tanto gratuitas como comerciales. Estas herramientas fueron caracterizadas de acuerdo con su costo, comunidad y ventajas y desventajas de su uso. Del total de herramientas 11 se clasificaron como bibliotecas (Python: Matplotlib, Seaborn, Plotly, Bokeh; R: ggplot2, Leaflet; JavaScript: D3.js, Highcharts.js, Chart.js, Polymaps.js, Sigma.js) y 13 plataformas GUI: 5 no comerciales: Google Charts, RAWGraph, ChartBlocks, Google Data Studio, Kibana y 8 comerciales: Tableau, Power BI, Datawrapper, Infogram, Qlik, FusionCharts, Grafana, Sisense.

El análisis mostró que las bibliotecas ofrecen mayor personalización y capacidad de integración a costa de una curva de aprendizaje pronunciada y la necesidad de saber programación. Las plataformas GUI, por su parte, priorizan la facilidad de uso, la interactividad, la productividad y la colaboración con más restricciones de personalización y/o licenciamiento con respecto a las bibliotecas de programación. Se observó, además, que coexisten soluciones especializadas por dominio, como por ejemplo Leaflet y Polymaps para cartografía y Sigma.js para redes, junto con herramientas generales orientadas a dashboards como Tableau, Power BI, Grafana y Kibana.

Esta caracterización permitió tener una visión comparativa que sirvió de insumo directo para construir la matriz de decisión del método propuesto. Permitió además realizar un mapeo preliminar entre perfil de usuario y tipo de herramienta, por ejemplo, para los usuarios sin conocimientos de programación se hace más conveniente el uso de plataformas GUI, mientras que los usuarios técnicos prefieren las bibliotecas. Esta asociación es útil para la dimensión Contexto del método MATCH-Vis. La identificación de elementos diferenciadores entre las herramientas como la interactividad, la variedad de gráficos, la conectividad a fuentes y opciones de exportación, el costo, entre otros fueron usados como criterios posteriormente para ser tomados en cuenta a la hora de seleccionar la herramienta adecuada y ponderados por entropía para su evaluación con los métodos SAW y TOPSIS. Esta base permitió sustentar la selección final tras la validación contextual en el Paso 6 del método MATCH-Vis.

5.2 Criterios de comparación

La investigación permitió establecer 11 criterios como los necesarios para llevar a cabo el proceso de selección de una herramienta de visualización. Estos criterios fueron agrupados en cinco categorías: Usabilidad, Complejidad, Integración, Costo y Calidad de la Visualización. Dichas categorías encierran los aspectos fundamentales a tomar en cuenta para seleccionar una herramienta que cumpla con los requisitos necesarios para llevar a cabo el proceso de visualización de datos. Los criterios se distribuyeron entre estas categorías evitando que existiera un solapamiento en los mismos, por ejemplo, el criterio “Flexibilidad en la edición de gráficos” se refiere a la capacidad que tiene la herramienta para permitir la personalización de los mismos y cambios en el color, los ejes, tipos de letras, etc., mientras que el criterio “Variedad de tipos de gráficos” se refiere a la diversidad de gráficos o técnicas de visualización que ofrece

la herramienta. La Tabla 5.3 muestra una comparación de los criterios propuestos en este estudio con respecto a los abordados en otros trabajos reportados en la literatura.

5.3 Resultados obtenidos con los métodos de toma de decisiones multicriterio

Los métodos de toma de decisiones multicriterio SAW Y TOPSIS fueron aplicados a la matriz de decisión obtenida tras aplicar la encuesta de evaluación a los expertos del dominio. Este paso tuvo como objetivo realizar un ranking de las mejores herramientas de acuerdo con las valoraciones de los expertos por cada criterio.

La Tabla 5.1 muestra la comparación entre los rankings de herramientas alcanzados con los métodos SAW y TOPSIS en sus dos variantes, tanto con distancia Euclidiana, como con distancia de Manhattan.

Tabla 5.1. Comparación de los ranking de herramientas entre los MCDM

Herramienta	SAW	TOPSIS-E	TOPSIS-M
Power BI	1	3	1
Plotly	2	1	2
Tableau	3	6	3
Ggplot2	4	2	4
Matplotlib	5	4	5
Seaborn	6	5	6

Los resultados obtenidos con los métodos SAW y TOPSIS usando la distancia de Manhattan coinciden, esto se debe a que se cumplen ciertas propiedades de aditividad, sensibilidad a los pesos, normalización lineal y escala comparable (Ciardiello & Genovese, 2023). Ambos métodos se basan en una función de valor aditiva, de modo que el valor de cada alternativa se obtiene calculando la suma ponderada de las evaluaciones de los criterios. En el experimento realizado los pesos asignados a cada criterio son los mismos para cada uno de los métodos utilizados y los métodos SAW y TOPSIS con distancia de Manhattan cumplen ciertas propiedades formales similares, tales como la normalización lineal y escala comparable. En ambos métodos los valores de la matriz de decisión obtenidos con la encuesta fueron normalizados de forma lineal y llevados a una misma escala donde los criterios quedaron expresados en la misma unidad adimensional y con la misma dirección de preferencia, es decir todos los criterios fueron tratados como criterios de maximización o beneficio.

Por otro lado, los resultados entre el método SAW y TOPSIS con distancia Euclidiana no coinciden debido a las diferencias en la forma de ponderar las alternativas. Mientras que SAW suma directamente los valores ponderados, TOPSIS con distancia Euclidiana calcula las distancias a las soluciones ideales positivas y negativas siguiendo las ecuaciones (2.11) y (2.12), lo que eleva al cuadrado cada diferencia y los criterios que presentan valores más alejados del ideal influyen de manera más desproporcionada en el índice de cercanía. Esto quiere decir, que la distancia Euclidiana tiende a hacer menos notables las pequeñas diferencias entre las alternativas y a enfatizar las grandes, lo que puede invertir el orden de las alternativas que son similares en la mayoría de los criterios, pero difieren en uno solo (Ciardiello & Genovese, 2023). Con la distancia de Manhattan, ver ecuaciones (4.2) y (4.3), la suma de los valores absolutos mantiene una

relación más lineal con la agregación SAW por lo que los rankings entre estos dos métodos de toma de decisiones multicriterio coinciden más. Aunque ambos métodos empleen la misma normalización y los mismos pesos, la diferencia principal radica en que la distancia Euclidiana introduce una sensibilidad cuadrática a las discrepancias entre alternativas e ideales, mientras que SAW utiliza una agregación lineal; esa diferencia estructural produce rankings distintos.

La similitud observada entre los rankings generados sugiere que las herramientas con mayor reconocimiento y experiencia entre los usuarios encuestados tienden a presentar mejor desempeño en criterios como facilidad de uso, exploración interactiva, flexibilidad y rendimiento gráfico. Los resultados nos permiten observar que las herramientas mejor calificadas fueron las herramientas que incorporan componentes interactivos. Esta preferencia puede atribuirse a la capacidad de dichas herramientas para facilitar el análisis de información de manera más intuitiva y eficiente, al permitir una exploración más profunda de los datos.

En contraste, las herramientas basadas en visualizaciones estáticas, aunque ampliamente reconocidas y utilizadas, no alcanzaron los mismos niveles de valoración. Este resultado puede explicarse por la mayor barrera de entrada que representan para usuarios principiantes, ya que su uso suele requerir conocimientos técnicos especializados del lenguajes de programación. Además, las visualizaciones estáticas presentan limitaciones en términos de dimensionalidad, al restringirse generalmente a representar variables en solo dos dimensiones. Esto puede convertirse en una dificultad a la hora de interpretar la visualización, especialmente para usuarios que no están familiarizados con las técnicas de visualización ni con las particularidades de los datos analizados.

La utilización del enfoque de entropía para determinar los pesos objetivos aportó rigor al análisis multicriterio, al permitir que los datos determinaran el nivel de discriminación de cada criterio sin intervención subjetiva. En la Tabla 5.2 se puede observar que los mayores pesos estuvieron asociados a los criterios Costo, Asistencia en la selección del tipo de gráfico, Exploración Interactiva de los datos y Facilidad de uso, lo cual sugiere que estos factores son los que más contribuyen a diferenciar el desempeño de las herramientas de visualización evaluadas y, por tanto, tienen mayor influencia en la toma de decisiones en este contexto. Esto evidencia que los usuarios valoran no solo la funcionalidad técnica, sino también la accesibilidad económica y la facilidad de interacción con la herramienta como aspectos clave para su adopción.

Tabla 5.2. Pesos asignados a los criterios por el método de la entropía

Criterios	Pesos
Facilidad de uso	0.1091
Flexibilidad en la edición de gráficos	0.0644
Asistencia en seleccionar el tipo de gráfico	0.1685
Variedad de tipos de gráficos	0.0237
Exploración interactiva de los datos	0.1674
Manejo de grandes conjuntos de datos	0.0465
Fuentes y tipos de datos	0.0532
Archivos de exportación	0.0188

Costo	0.2655
Renderizado rápido de resultados	0.0526
Calidad de la visualización	0.0304

La investigación evidencia que, cuando los datos están parcialmente incompletos, pero siguen un patrón MAR (*Missing At Random*), el análisis multicriterio es todavía viable y confiable, siempre que se adopten estrategias de ponderación adecuadas. Los resultados invitan a futuras investigaciones que contrasten estas metodologías en escenarios con distintos patrones de ausencias y con poblaciones de usuarios más especializadas.

5.4 Comparación con el Estado del Arte

A continuación, la Tabla 5.3, muestra una comparación entre el método MATCH-Vis para selección de herramientas de visualización y otros trabajos reportados en la literatura con un enfoque similar. Es notable destacar que el método MATCH-Vis toma en cuenta el escenario de uso y las características del usuario a diferencia de otros métodos propuestos en la literatura como los de Hassan & Elragal (2017) y Nazliel et al. (2022). También propone un enfoque sistemático para seleccionar la herramienta de visualización adecuada, basado en métodos de toma de decisiones bien sustentados en el campo de Investigación de Operaciones como son los métodos SAW y TOPSIS, a diferencia de trabajos como los de Schmidt (2020) y Majeed et al. (2020), que solo hacen comparaciones descriptivas sin ofrecerle al usuario una guía para la selección de la herramienta, especialmente para usuarios novatos en el campo de la Ciencia de Datos.

CAPÍTULO 5: RESULTADOS Y DISCUSIÓN

Tabla 5.3. Comparación del método MATCH-Vis con otros trabajos de la literatura

Artículo	Herramientas	Enfoque	Criterios	Dimensiones
(Hassan & Elragal, 2017)	No se mencionan	MCDM (Método Delphi)	36 criterios *	Técnica, Visualización, Colaboración y movilidad, Operación, Gobernanza de Datos y Requerimientos gerenciales
(Grandhi & Wibowo, 2018)	Aster Discovery, Power BI, Tableau, Looker	MCDM (TOPSIS difuso intuicionista)	Costo, Usabilidad, Funcionalidad, Seguridad	N/A
(Schmidt, 2020)	Plotly, Seaborn, Ggplot2, VegaLite, D3.js, Google Charts, Chart.js, Apexcharts, dygraphs, Bokeh, RAWGraphs, .Net LiveCharts, Qt Charts, Power BI, Tableau, SAS Visual Analytics, Highcharts, Quadrigram, MatLab	Comparación	Tipo de visualización	N/A
(Majeed et al., 2020)	NodeXL, GEPHI, UCINet, PAJEK, MuxViz,	Comparación	Funcionalidad principal, Cobertura, Código abierto/Software propietario, Formato de entrada, Archivo de exportación, Tipos de visualización, Disponibilidad de código fuente, Standalone/Dependiente, Filtros, Fortalezas, Debilidades.	Características generales, Características de entrada/salida, Visualización, Funcionalidades y soporte, Fortalezas y debilidades
(Nazliel et al., 2022)	No se mencionan	MCDM (AHP + Método Delphi)	Requisitos funcionales y no funcionales	Estrategia tecnológica, Restricciones legislativas, Costo, Capacidades Organizacionales
MATCH-Vis	Matplotlib, Seaborn, Plotly, Bokeh. Ggplot2, Leaflet, D3.js, Highcharts, Sigma.js, Polymaps.js, charts.js, Google Charts, RAWGraph, ChartBlocks, Google Data Studio, Tableau, Power BI, Qlik, Datawrapper, Infogram, Grafana, FusionCharts, Kibana, Sisense	MCDM (SAW/TOPSIS + contexto)	Facilidad de uso, Flexibilidad, Asistencia en seleccionar el tipo de gráfico, Variedad de tipos de gráficos, Exploración interactiva de los datos, Manejo de grandes conjuntos de datos, Fuentes y tipos de datos, Archivos de exportación, Costo, Renderizado rápido de resultados, Calidad de la visualización	Meta, Almacenamiento de datos, Tipo de datos, Contexto + Usabilidad, Complejidad, Integración, Costo, Calidad de la visualización

Los 36 criterios planteados en el trabajo de Hassan & Elragal (2017), no fueron descritos en esta tabla comparativa debido a la extensión de estos. Se sugiere ir a la referencia para más información.

Capítulo 6

Conclusiones y trabajo futuro

CONCLUSIONES Y TRABAJO FUTURO

En esta sección se presentan las conclusiones que abordan los temas fundamentales abordados en el desarrollo de esta tesis.

6.1 Conclusiones

La caracterización de las herramientas permitió evidenciar las elecciones por parte del usuario de acuerdo con su nivel de dominio técnico. Mientras las bibliotecas se distinguen por su capacidad de personalización e integración, presentan como limitación la necesidad de conocimientos de código avanzados; en contraste, las plataformas GUI priorizan la facilidad de uso, la intuitividad y la colaboración, aunque restringen la flexibilidad y suelen implicar costos de licenciamiento.

La investigación permitió evidenciar que el dominio de herramientas de visualización por parte de los participantes no garantiza su adecuación en todos los contextos analíticos. Las herramientas con mayor valoración según los métodos de toma de decisión multicriterio no necesariamente fueron las más conocidas por los encuestados, sino aquellas que presentaron componentes interactivos y mejor capacidad para facilitar la exploración de datos.

El uso del método de entropía para la ponderación de criterios permitió una asignación objetiva y libre de sesgos, al permitir que los datos determinaran el nivel de discriminación de cada criterio sin intervención subjetiva.

La comparación entre los métodos multicriterio SAW y TOPSIS con la métrica de Manhattan, reveló concordancia en las herramientas mejor evaluadas, resultado que coincide con lo reportado en la literatura.

El método MATCH-Vis propuesto demostró ser efectivo para sistematizar la selección de herramientas de visualización, permitiendo a los analistas tomar decisiones informadas con base en criterios relevantes al contexto.

6.2 Trabajo futuro

Aplicar el método MATCH-Vis a más casos de uso y en contextos reales de análisis de datos dentro de organizaciones o instituciones educativas, para validar su generalización, utilidad práctica y adaptabilidad a distintos dominios.

Realizar pruebas combinando un sistema de ponderación híbrido mediante técnicas como AHP para comparar resultados.

Ampliar la base de participantes para observar si los patrones de evaluación y uso de herramientas cambian en función del tipo de experiencia.

Desarrollar una aplicación digital que implemente MATCH-Vis como asistente para la selección de herramientas de visualización, facilitando su uso y adopción por parte de los usuarios.

6.3 Productos Científicos

Como parte de esta investigación se obtuvieron los siguientes productos científicos:

CONCLUSIONES Y TRABAJO FUTURO

- Presentación del artículo titulado “Visualización de Datos: Principios, Técnicas y Herramientas”, presentado en el 4to Congreso Internacional de Ciencia y Tecnología Aplicada (CITCA), celebrado del 21 al 23 de mayo de 2025 en el TecNM/CENIDET.
- Presentación del proyecto de tesis denominado “Caracterización y selección de herramientas de visualización y georreferenciación con orientación a la solución de problemas de Ciencia de Datos” en el Concurso “3+3 Minutos sobre mi tesis” llevado a cabo el día 31 de marzo de 2025, en el CENIDET/TecNM, en el marco de la Divulgación de la Ciencia y Tecnología a Jóvenes.

Referencias

- Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*, 20(1), 40–49. <https://doi.org/10.1002/mpr.329>
- Bikakis, N. (2022). Big Data Visualization Tools. En *Encyclopedia of Big Data Technologies* (pp. 1–8). Springer International Publishing. https://doi.org/10.1007/978-3-319-63962-8_109-2
- Canales Luna, J. (2023, octubre 20). *Exploring 12 of the Best Data Visualization Tools in 2023 With Examples*. <https://www.datacamp.com/blog/12-of-the-best-data-visualizations-tools>.
- Cao, L. (2017). Data science: A comprehensive overview. En *ACM Computing Surveys* (Vol. 50, Número 3). Association for Computing Machinery. <https://doi.org/10.1145/3076253>
- Chapman, C. (2023, diciembre 1). *A Complete Overview of the Best Data Visualization Tools*. <https://www.toptal.com/designers/data-visualization/data-visualization-tools>.
- Chinthamu, N., & Karukuri, M. (2023). Data Science and Applications. *Journal of Data Science and Intelligent Systems*, 1(2), 83–91. <https://doi.org/10.47852/bonviewjdsis3202837>
- Ciardiello, F., & Genovese, A. (2023). A comparison between TOPSIS and SAW methods. *Annals of Operations Research*, 325(2), 967–994. <https://doi.org/10.1007/s10479-023-05339-w>
- Desimoni, F., & Po, L. (2020). Empirical evaluation of Linked Data visualization tools. *Future Generation Computer Systems*, 112, 258–282. <https://doi.org/10.1016/j.future.2020.05.038>
- Dhiman, H. S., & Deb, D. (2020). Multi-criteria Decision-Making: An Overview. En *Studies in Systems, Decision and Control* (Vol. 253, pp. 19–36). Springer International Publishing. https://doi.org/10.1007/978-981-15-0275-0_2
- Dogadina, V., & Voronin, A. (2024). Comparative Analysis of Data Visualization Tools. *International Journal of Computing, Programming and Database Management*, 5(1), 49–51. <https://doi.org/10.33545/27076636.2024.v5.i1a.93>
- Grandhi, S., & Wibowo, S. (2018). A Multi-Criteria Group Decision Making Method for Selecting Big Data Visualization Tools. *Journal of Telecommunication, Electronic and Computer Engineering*, 10, 67–72.
- Guerrero, N. (2021, enero 7). *Librerías de JavaScript para visualización de información*. Programa en Línea. <https://www.programaenlinea.net/las-5-mejores-librerias-javascript-la-visualizacion-informacion/>
- Hashemi-Pour, C., Brush, K., & Burn, E. (2024, agosto 16). *What is data visualization?* TechTarget. <https://www.techtarget.com/searchbusinessanalytics/definition/data-visualization>
- Hassan, A. M., & Elragal, A. (2017). BIG DATA VISUALIZATION TOOL: A BEST-PRACTICE SELECTION MODEL. En Powell P., Rodrigues L., Nunes M.B., & Isaias P. (Eds.), *10th IADIS International*

REFERENCIAS

- Conference Information Systems* (pp. 59–68). Institute of Electrical and Electronics Engineers (IEEE).
- Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials - A practical guide with flowcharts. *BMC Medical Research Methodology*, 17(1). <https://doi.org/10.1186/s12874-017-0442-1>
- Lowe, J., & Matthee, M. (2020). Requirements of Data Visualisation Tools to Analyse Big Data: A Structured Literature Review. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12066 LNCS, 469–480. https://doi.org/10.1007/978-3-030-44999-5_39
- Majeed, S., Uzair, M., Qamar, U., & Farooq, A. (2020, noviembre 5). Social Network Analysis Visualization Tools: A Comparative Review. *Proceedings - 2020 23rd IEEE International Multi-Topic Conference, INMIC 2020*. <https://doi.org/10.1109/INMIC50486.2020.9318162>
- Majumder, M., & Saha, A. K. (2016). Multi Criteria Decision Making. En M. Majumder & A. K. Saha (Eds.), *Feasibility Model of Solar Energy Plants by ANN and MCDM Techniques* (pp. 9–12). Springer Singapore. https://doi.org/10.1007/978-981-287-308-8_3
- Munzner, T. (2015). *Visualization analysis & design* (1a ed.) [Book]. CRC Press Taylor & Francis Group. <https://doi.org/https://doi.org/10.1201/b17511>
- Nazliel, K., Kayabay, K., Gokalp, M. O., Gokalp, E., & Eren, E. (2022). Data Science Technology Selection: Development of a Decision-Making Approach. *2022 IEEE Technology and Engineering Management Conference: Societal Challenges: Technology, Transitions and Resilience Virtual Conference, TEMSCON EUROPE 2022*, 172–178. <https://doi.org/10.1109/TEMSCONEUROPE54743.2022.9802054>
- Nazmuz Sakib, S. M. (2022). *Data Visualization in Data Science*. <https://doi.org/10.33774/coe-2022-c3kk5-v2>
- Nugroho, A. Y. (2024). Data science: Building visual understanding. *Jurnal Teknologi Cerdas*, 1(1), 1–6. <https://doi.org/10.70310/7y9mnh17>
- Pise, A. A., & Singampalli, D. T. (2023). An Introduction to Data Visualisation: An Overview of Visualizing Data. En K. K. Hiran, K. Hemachandran, A. Pise, & B. J. Rabi (Eds.), *Handbook of Research on AI and Knowledge Engineering for Real-Time Business Intelligence* (pp. 34–53). IGI Global. <https://doi.org/10.4018/978-1-6684-6519-6.ch003>
- Rasool, M., & Chaudhary, V. K. (2022). Applications of Data Science in Respective Engineering Domains. *International Journal of Scientific Research in Science and Technology*, 71–75. <https://doi.org/10.32628/ijrst22958>
- Raza, Q. U., & Summair, M. (2020). Applications of Data Science. En *Data Science Concepts and Techniques with Applications* (pp. 25–40). Springer Singapore. https://doi.org/10.1007/978-981-15-6133-7_2
- Rollins, J. B. (2015). *Metodología Fundamental para la Ciencia de Datos*. <https://www.ibm.com/downloads/cas/WKK9DX51>

REFERENCIAS

- Schilling, F.-P., Flumini, D., Fuchslin, R. M., Gavagnin, E., Geller, A., Quarteroni, S., Stadelmann, T., & Scensei, A. G. (2022). Foundations of Data Science: A Comprehensive Overview Formed at the 1st International Symposium on the Science of Data Science. *KIT Scientific Publishing*, 2, 1–21. <https://doi.org/10.5445/IR/1000146422>
- Schmidt, J. (2020). Usage of visualization techniques in data science workflows. *VISIGRAPP 2020 - Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 3, 309–316. <https://doi.org/10.5220/0009181903090316>
- Shaikh, Z. P. (2025). Multi-Criteria Decision Making (MCDM) in Organizations: Conceptualization of Theory and Methods. En S. Aouadni & I. Aouadni (Eds.), *Recent Theories and Applications for Multi-Criteria Decision-Making* (pp. 151–180). IGI Global. <https://doi.org/10.4018/979-8-3693-6502-1.ch006>
- Shakeel, H. M., Iram, S., Al-Aqrabi, H., Alsboui, T., & Hill, R. (2022). A Comprehensive State-of-the-Art Survey on Data Visualization Tools: Research Developments, Challenges and Future Domain Specific Visualization Framework. *IEEE Access*, 10, 96581–96601. <https://doi.org/10.1109/ACCESS.2022.3205115>
- Shao, Q. (2024). The Frontier Exploration and Application Prospect of Data Science. *Applied and Computational Engineering*, 116(1), 73–78. <https://doi.org/10.54254/2755-2721/116/20251743>
- Shivaram, P. R. (2024, noviembre 12). *Benefits of Data Visualization: Transforming Numbers into Narratives*. acceldata. <https://www.acceldata.io/blog/benefits-of-data-visualization-transforming-numbers-into-narratives>
- Stef van Buuren. (2018). *Flexible Imputation of Missing Data* (2a ed.). Chapman & Hall/CRC. <https://stefvanbuuren.name/fimd/>
- Survey Point Team. (2023, octubre 7). *Top 10 Data Visualization Tools to Transform Your Data*. Survey Point. <https://surveypoint.ai/knowledge-center/top-10-data-visualization-tools/>
- Swathi Priya, C., Potnuru, S. S., Jha, I., K., H., & Dudekula, C. S. (2023). Data Insights by Using Data Visualization and Exploration. En *Artificial Intelligence and Knowledge Processing: Methods and Applications* (pp. 72–83). BENTHAM SCIENCE PUBLISHERS. <https://doi.org/10.2174/9789815165739123010008>
- Taherdoost, H., & Madanchian, M. (2023). Multi-Criteria Decision Making (MCDM) Methods and Concepts. *Encyclopedia*, 3(1), 77–87. <https://doi.org/10.3390/encyclopedia3010006>

Anexos

8.1 Anexo A: Caracterización de todas las herramientas identificadas en la búsqueda de la literatura

En este anexo se exponen las características principales de las 24 herramientas de visualización de datos identificadas en la literatura, estas se muestran en la Tabla 8.1. Estas herramientas fueron elegidas, ya que en los artículos revisados figuraban como las mejores herramientas de visualización en los últimos años.

Tabla 8.1. Características de las 24 herramientas identificadas inicialmente

Herramienta	Tipo	Característica	Costo	Comunidad	Ventajas	Desventajas
Matplotlib (Canales Luna, 2023)	Biblioteca de Python.	Es la biblioteca por excelencia de Python para visualizaciones.	Gratuita.	Tiene gran respaldo en la comunidad de Ciencia de Datos.	Permite exportar las visualizaciones en múltiples formatos de archivo, como .PDF, .PNG, y .SVG.	Elevada curva de aprendizaje, ya que se necesita saber Python para usarla.
Seaborn (Canales Luna, 2023)	Biblioteca de Python.	Seaborn es una biblioteca creada para superar las dificultades que trae crear una visualización utilizando Matplotlib.	Gratuita.	Cuenta con gran respaldo en la comunidad de Ciencia de Datos	Permite crear gráficos atractivos de forma sencilla y proporciona una interfaz de alto nivel para gráficos estadísticos. Está construida encima de Matplotlib y cuenta con una buena documentación y numerosos ejemplos.	Sus opciones de personalización son limitadas en comparación con Matplotlib. No proporciona gráficos interactivos
Plotly (Canales Luna, 2023; Survey Point Team, 2023)	Biblioteca de Python.	Es un paquete de gráficos compatible con varios lenguajes de programación como R, Python, JavaScript, Julia, MatLab y Pearl.	Gratuita.	Debido a su simplicidad y excelentes resultados, Plotly es la biblioteca estándar para visualizaciones de datos interactivas en muchas grandes empresas de tecnología y tiene una amplia comunidad en GitHub.	Su sintaxis es muy simple e intuitiva y puede ser usado por personas que no tengan conocimientos técnicos para crear gráficos interactivos mediante el uso de la GUI.	La cantidad de paletas de colores y otros aspectos estéticos son limitados en comparación con otras herramientas. Las visualizaciones realizadas con una versión comunitaria de Plotly son siempre públicas.

ANEXOS

Bokeh (Canales Luna, 2023)	Biblioteca de Python.	Su fortaleza son las visualizaciones interactivas. Sus visualizaciones están diseñadas fundamentalmente para ser insertadas en navegadores web.	Gratuita.	Es la biblioteca de visualización de Python con mejor ranking en GitHub.	Proporciona gráficos elegantes que van desde gráficos sencillos hasta dashboards complejos con conjuntos de datos en tiempo real. Bokeh puede ser personalizado de manera adicional a través de JavaScript.	Tiene una flexibilidad limitada a la hora de crear los gráficos y la combinación entre Python y JavaScript puede ser un poco complicada.
Ggplot2 (Canales Luna, 2023)	Biblioteca de R.	Ggplot2 es un paquete de R que permite modelar gráficos visualmente atractivos. Está inspirado en la metodología de visualización de datos llamada "la gramática de gráficos", cuya idea es especificar de forma independiente los componentes del gráfico y luego combinarlos.	Gratuita.	Ggplot2 es la biblioteca más popular para crear visualizaciones en R. Cuenta con una gran comunidad de usuarios y desarrolladores que contribuyen con tutoriales, ejemplos, paquetes adicionales y soporte.	Tiene una sintaxis simple e intuitiva.	Este paquete puede llegar a ser computacionalmente más lento que otros paquetes de R.
Leaflet (Canales Luna, 2023)	Biblioteca de R.	Es una biblioteca de R que sirve para crear mapas atractivos e interactivos para la web.	Gratuita.	Tiene una gran comunidad de usuarios y desarrolladores.	Tiene una sintaxis simple e intuitiva y soporta múltiples objetos para mapas como líneas, polígonos, mosaicos y archivos GeoJSON. Leaflet, además, permite incrustar los mapas creados en documentos RMarkdown y aplicaciones Shiny.	Los mapas estáticos creados con este paquete no son tan bonitos como los formados con ggplot2.
D3.js (Chapman, 2023; Guerrero, 2021; Survey	Biblioteca de JavaScript.	Tiene una gran cantidad de plantillas para gráficos disponibles.	Gratuita.	Cuenta una comunidad de usuarios muy grande y activa y cuenta con	Poderosa y personalizable.	Requiere tener conocimientos de JavaScript.

ANEXOS

Point Team, 2023)				numerosos recursos como tutoriales, foros y libros.		
Highcharts.js (Guerrero, 2021; Survey Point Team, 2023)	Biblioteca de JavaScript.	Es una de las bibliotecas de JavaScript más completa y popular entre los programadores web.	Gratuita.	Gran comunidad en Stack Overflow. Hasta desarrollaron una API.	Proporciona gráficos interactivos optimizados para dispositivos móviles. Las gráficas se pueden descargar en .PDF, .PNG, .JPG o .SVG. Admite una gama de diversos tipos de gráficos.	Necesita de una licencia para uso. Es comercial
chart.js (Chapman, 2023; Guerrero, 2021)	Biblioteca de JavaScript.	Es una biblioteca de JavaScript que permite crear gráficos animados e interactivos.	Gratuita.	Tiene una comunidad grande y activa, lo que no tan destacable como D3.js	Representa bien los gráficos en todos los navegadores modernos. Los gráficos creados también son responsivos.	Tipos de gráficos muy limitados en comparación con otras herramientas.
Polymaps.js (Chapman, 2023)	Biblioteca de JavaScript.	Biblioteca de JavaScript dedicada a la cartografía.	Gratuita.	Su comunidad es pequeña en comparación con otras bibliotecas de JavaScript.	Los resultados son mapas dinámicos y responsivos en una variedad de estilos, desde superposiciones de imágenes hasta mapas de símbolos y mapas de densidad.	Solo crea un tipo de visualización: mapas.
Sigma.js (Chapman, 2023)	Biblioteca de JavaScript.	Su único propósito es crear gráficos de red.	Gratuita.	La comunidad de Sigma.js es moderada en tamaño y actividad, con recursos y contribuciones disponibles, pero no tan extensas como D3.js o Highcharts.js.	Es altamente personalizable.	Solo crea un tipo de visualización: gráficos de red.
Google Charts (Canales Luna, 2023; Chapman, 2023)	Plataforma GUI no comercial.	Google Charts, es una poderosa herramienta para construir gráficos	Gratuita.	Tiene una comunidad robusta, debido a la accesibilidad y	Tiene una gran variedad de tipos de gráficos. Los gráficos de salida	Más allá de los tutoriales y el foro disponibles, el soporte es limitado. No es

ANEXOS

2023; Guerrero, 2021)		interactivos y fácil de personalizar.		popularidad de la herramienta.	funcionan en navegadores sin el uso de complementos adicionales.	tan flexible como otras herramientas comerciales.
RAWGraph (Canales Luna, 2023)	Plataforma GUI no comercial.	Es una plataforma de visualización de datos de código abierto gratis e intuitiva.	Gratuita.	Cuenta con una comunidad muy pequeña de usuarios.	Permite a los usuarios representar los datos rápida y cómodamente sin necesidad de código e incluye la posibilidad de insertar la visualización creada en otro sitio web. Es una herramienta de “arrastre y suelte”, adecuada para entornos de aprendizaje gracias a su interfaz amigable. No hay necesidad de crear un usuario o registrar una dirección de correo para trabajar con RAWGraph.	No tiene mucha flexibilidad, ya que solo se pueden personalizar algunos parámetros y su colección de gráficos es limitada.
ChartsBlocks (Chapman, 2023; Survey Point Team, 2023)	Plataforma GUI no comercial.	ChartBlocks es una herramienta de visualización interactiva, de arrastre y suelte, muy fácil de usar.	Gratuita.	Cuenta con una comunidad pequeña.	El asistente de creación de gráficos ayuda a los usuarios a elegir exactamente los datos correctos para sus gráficos antes de importarlos. La herramienta permite una personalización extensa y el asistente es fácil de usar para importar los datos necesarios.	No tiene capacidad para crear mapas.
Google Data Studio (Survey	Plataforma GUI no comercial.	Herramienta de visualización de arrastre y suelte para hacer reportes	Gratuita.	Tiene una comunidad grande con presencia	Permite conectarse a una gran cantidad de bases de datos.	Está estrechamente integrado con otros productos de Google, lo

ANEXOS

Point Team, 2023)		dinámicos y cuadros de mando interactivos.		de muchos tutoriales en la web.		que limita su capacidad de integración fuera del ecosistema de Google.
Tableau (Canales Luna, 2023; Chapman, 2023; Survey Point Team, 2023)	Herramientas de Inteligencia de Negocios	Potente herramienta de visualización de datos para inteligencia de negocios. Tiene una amplia comunidad de soporte.	De pago, pero con versión gratuita.	Cuenta con cientos de videos tutoriales para aprender a manejar la herramienta, así como blogs y forúms donde se puede interactuar con la comunidad.	La plataforma ofrece una amplia gama de plantillas para gráficos y es posible importar los datos desde múltiples fuentes. Tableau es capaz de manejar un gran volumen de datos y puede crear gráficos y paneles de visualización en unos pocos minutos. Tiene una versión gratuita con algunas funcionalidades limitadas, entre ellas está que los cuadros o gráficos que ahí se realicen deben ser compartidos públicamente en la web.	La versión Premium es bastante costosa y la versión gratuita no permite hacer análisis privados.
Power BI (Canales Luna, 2023; Survey Point Team, 2023)	Herramientas de Inteligencia de Negocios	Herramienta de análisis de datos basada en la nube, que permite importar datos de diferentes fuentes, analizarlos y presentar dicho análisis a través de visualizaciones, reportes y paneles complejos.	De pago, pero tiene una versión gratuita.	Power BI es la herramienta de Inteligencia Empresarial más usada en la actualidad.	Está diseñado para ser usado de manera simultánea entre usuarios y departamentos de una compañía, lo que permite tener vistas de la información en tiempo real.	Entre los principales contras se encuentra que tiene capacidades limitadas de intercambio de datos, ya que al exportar a CSV solo se pueden exportar 30 000 filas, mientras que a Excel solo 150 000 y su interfaz de usuario está muy cargada.

ANEXOS

Datawrapper (Canales Luna, 2023; Survey Point Team, 2023)	Plataforma GUI comercial.	Es una herramienta de visualización creada para agregar gráficos y mapas a las historias de las noticias.	De pago, pero tiene una versión gratuita que permite hacer visualizaciones ilimitadas.	Es ampliamente usada por los periodistas y profesionales de los medios de comunicación que necesitan hacer visualizaciones para sus artículos.	No es necesario tener conocimientos de código y su curva de aprendizaje es baja. La herramienta trae incorporado un comprobador de daltonismo, que hace que sus visualizaciones sean amigables para las personas con esta enfermedad. DataWrapper cuenta con una versión gratuita con la que se pueden crear vistas ilimitadas, aunque con capacidades más limitadas.	Su versión Premium es mucho más cara que la de otras herramientas competitivas y solo es posible importar datos de unas pocas fuentes.
Infogram (Survey Point Team, 2023)	Plataforma GUI comercial.	Es una herramienta de visualización de arrastrar y soltar.	De pago, pero tiene una versión gratuita muy limitada.	Infogram tiene una base de usuarios considerable, especialmente entre diseñadores gráficos, profesionales de marketing y comunicadores.	Las visualizaciones pueden ser exportadas en un gran número de formatos como .PNG, .JPG, .GIF, .PDF, and .HTML. Incluye más de 35 tipos de gráficos y más de 550 tipos de mapas.	Tiene acceso a menos fuentes de datos que otras aplicaciones y su versión gratuita solo permite realizar hasta 10 proyectos y tener acceso a solo 13 tipos de mapas.
Qlik (Canales Luna, 2023; Survey Point Team, 2023)	Plataforma GUI comercial.	Herramienta de visualización de datos diseñada para utilizar análisis aumentados para entender los datos.	De pago y cuenta con un plan gratuito de 30 días.	Aunque grande, su comunidad es algo más pequeña en comparación con la de Tableau y Power BI.	Cuenta con una interfaz intuitiva de “arrastre y suelte” y está dirigida principalmente a la confección de cuadros interactivos. Qlik permite realizar la preparación de los datos de una manera visual y avanzada, ya que	Capacidades limitadas de extracción de datos, faltan algunos gráficos básicos y precio elevado en comparación con la competencia.

ANEXOS

					puede ordenar y limpiar los datos automáticamente.	
FusionCharts (Chapman, 2023)	Plataforma GUI comercial.	Opción basada en JavaScript para crear gráficos cuadros de mando web e interactivos.	De pago con una versión de prueba gratis durante 14 días.	FusionCharts cuenta con una comunidad moderada en comparación con otras herramientas de visualización más populares.	Proporciona código listo para usar para todas las variaciones de gráficos y mapas, lo que facilita la integración en sitios web incluso para aquellos diseñadores con conocimientos de programación limitados. Incluye más de 150 tipos de gráficos y 1000 tipos de mapas.	Muy cara, comienza con casi 500 USD para una licencia de desarrollador.
Grafana (Chapman, 2023)	Plataforma GUI comercial.	Es un software de visualización de datos que permite a los usuarios crear cuadros de mando dinámicos y otros tipos de visualizaciones.	Gratis, pero tiene algunas funcionalidades que son de pago.	Tiene una comunidad muy grande.	Admite fuentes de datos mixtas, anotaciones y funciones de alerta personalizables, y se puede ampliar mediante cientos de complementos disponibles. Las funciones de exportación permiten a los diseñadores compartir instantáneas de paneles e invitar a otros usuarios a colaborar. Grafana admite más de 50 fuentes de datos mediante complementos.	No ofrece tantas opciones de personalización visual como otras herramientas.
Kibana (Survey Point Team, 2023)	Plataforma GUI no comercial.	Es una herramienta de código abierto muy popular que permite analizar datos.	Gratis.	Cuenta con una comunidad activa, especialmente en el contexto de ELK	Es una gran herramienta para el estudio de datos no estructurados basado en búsquedas. Cuenta con	Kibana está diseñado específicamente para trabajar con datos

ANEXOS

				(Elasticsearch, Logstash, Kibana).	paneles de control interactivos y gráficos como mapas de calor, histogramas y elementos visuales geoespaciales.	almacenados en Elasticsearch.
Sisense (Survey Point Team, 2023)	Herramientas de Inteligencia de Negocios	Herramienta de BI que facilita a las personas sin conocimientos técnicos preparar y visualizar datos y crear paneles.	De pago, con una versión de prueba de 15 días.	Sisense tiene una presencia significativa en el ámbito empresarial y entre los desarrolladores de BI.	Cuenta con presencia de inteligencia artificial para guiar las visualizaciones. Tiene una gran cantidad de conectores de datos y filtrado interactivos. Los grandes conjuntos de datos se manejan centrándose en el rendimiento.	Tiene una curva de aprendizaje más empinada que Tableau y Power BI para los nuevos usuarios, especialmente aquellos sin experiencia previa en BI o análisis de datos.

8.2 Anexo B: Encuesta de evaluación

En el presente anexo se muestra una parte de la encuesta lanzada a los expertos para recopilar información relevante relacionada con las herramientas de visualización y construir la matriz de decisión. La encuesta completa se puede encontrar en: <https://forms.gle/Qn3bw9JvdDvRAsnRA>

Encuesta para Expertos en Herramientas de Visualización de Datos

Esta encuesta tiene como objetivo recopilar información de contacto de profesionales y expertos en herramientas de visualización de datos. Nos interesa conocer su experiencia y nivel de conocimiento en herramientas específicas. La información recabada será utilizada exclusivamente para fines de investigación y colaboración en proyectos relacionados con visualización de datos. Agradecemos de antemano su colaboración y apoyo.

* Indica que la pregunta es obligatoria

1. Correo electrónico *

2. Grado máximo de estudios*

(Marca solo un óvalo)

- ☐ Licenciatura o Ingeniería
- ☐ Maestría
- ☐ Doctorado
- ☐ Otros: _____

3. Ocupación*

(Marca solo un óvalo)

- ☐ Profesor
- ☐ Investigador
- ☐ Científico de Datos
- ☐ Analista de Datos
- ☐ Estudiante de Doctorado
- ☐ Estudiante de Maestría
- ☐ Otros: _____

4. Institución de procedencia*

5. Años de experiencia en el sector*

(Marca solo un óvalo)

- ☐ Menos de año
- ☐ 1 a 3 años
- ☐ más de 3 años

Biblioteca Matplotlib de Python

6. ¿Cuál es el nivel de conocimiento que tiene sobre la herramienta Biblioteca Matplotlib de Python? *

(Marca solo un óvalo)

ANEXOS

- ☐ No tengo conocimiento sobre ella (*Ir a la pregunta 19*)
- ☐ Básico (*Ir a la pregunta 7*)
- ☐ Intermedio (*Ir a la pregunta 7*)
- ☐ Avanzado (*Ir a la pregunta 7*)

Años de Experiencia con Matplotlib

7. ¿Cuánto tiempo de experiencia tiene con la biblioteca Matplotlib de Python? *

(*Marca solo un óvalo*)

- ☐ menos de 1 año
- ☐ 1 a 3 años
- ☐ más de 3 años

Evaluación de Matplotlib como herramienta de visualización de Datos

La evaluación de la herramienta se realizará de acuerdo con varios criterios, los cuales se desglosan en diferentes características.

8. Facilidad de Uso *

La herramienta:

1. No requiere conocimientos de programación.
2. Proporciona una interfaz gráfica fácil de entender.
3. Cuenta con documentación, tutoriales o soporte en línea para resolver dudas.
4. Incorpora asistencia que le indique al usuario que hacer.
5. Proporciona una interfaz de "arrastrar y soltar".
6. No demanda períodos extensos de aprendizaje para las funciones avanzadas.

(*Marca solo un óvalo*)

No cuenta con ninguna de las características anteriores.

- ☐ 0 (*No cuenta con ninguna de las características anteriores*)
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 (*Tiene cinco o más características anteriores*)

9. Flexibilidad en la edición de gráficos *

La herramienta:

1. Permite que los gráficos puedan ser personalizados en aspectos básicos como color, tipo de letra, y tamaño.
2. Acepta personalización avanzada, como modificar ejes, etiquetas y leyendas.
3. Permite que se incluyan elementos interactivos como zoom, descripciones emergentes y filtros dinámicos.
4. Posibilita seleccionar entre un conjunto de plantillas predefinidas.
5. Admite la edición del código fuente y/o plantillas para personalizaciones más avanzadas y la creación de gráficos completamente personalizados.

(*Marca solo un óvalo*)

ANEXOS

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 *(Tiene las cinco características anteriores)*

10. Ayuda en seleccionar el tipo de gráfico *

La herramienta:

1. Incluye un asistente que oriente al usuario en seleccionar el gráfico adecuado de acuerdo con las características de los datos y el objetivo del análisis.
2. Sugiere tipos de gráficos básicos como barras, líneas, pastel, dispersión, área, entre otros según el tipo de datos.
3. Sugiere gráficos más complejos como mapas, diagramas de árbol, diagramas de cajas y bigotes, entre otros.
4. Inhabilita los gráficos que no sean adecuados para los datos.
5. Proporciona ejemplos, tutoriales o documentación que explique cómo elegir el gráfico más adecuado en diferentes escenarios de análisis.

(Marca solo un óvalo)

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 *(Tiene las cinco características anteriores)*

11. Variedad del tipo de gráficos *

La herramienta:

1. Incluye gráficos básicos como barras, líneas o pastel.
2. Proporciona gráficos más elaborados tales como: dispersión, histogramas, diagramas de cajas y bigotes y diagramas de área.
3. Ofrece gráficos especializados como mapas de calor, gráficos de burbuja, gráficos de radar, gráficos de red, entre otros.
4. Permite combinar gráficos para generar otros más avanzados.
5. Brinda la posibilidad de diseñar nuevos tipos de visualizaciones mediante scripts o configuraciones avanzadas.

(Marca solo un óvalo)

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 *(Tiene las cinco características anteriores)*

12. Exploración interactiva de los datos *

La herramienta:

ANEXOS

1. Permite al usuario acercarse o alejarse de una visualización y desplazarse dentro de ella.
2. Proporciona filtros dinámicos para seleccionar los datos directamente en la visualización.
3. Muestra información adicional como etiquetas, valores exactos o descripciones contextuales, al pasar el cursor sobre un elemento del gráfico o hacer clic en él.
4. Permite que el usuario interactúe con el gráfico para seleccionar partes específicas de los datos
5. Posibilita la conexión de visualizaciones de modo que, al interactuar con una, las demás se actualicen.

(Marca solo un óvalo)

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 *(Tiene las cinco características anteriores)*

13. Manejo de grandes conjuntos de datos

La herramienta:

1. Admite conjuntos de datos que superen los millones de registros
2. Manipula los conjuntos de datos de gran tamaño de manera que no se degrade el rendimiento de la misma
3. Soporta el manejo de datos tipo LOB (large objects)
4. Adapta las visualizaciones a pantallas de diferentes tamaños y resoluciones sin comprometer la legibilidad
5. Permite que los usuarios exploren los gráficos a escala sin afectar el rendimiento ni la calidad visual

(Marca solo un óvalo)

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 *(Tiene las cinco características anteriores)*

14. Fuentes y tipos de datos *

La herramienta:

1. Soporta datos desde archivos locales
2. Conecta con servidores de base de datos
3. Permite la integración con flujos de datos en tiempo real
4. Tiene la capacidad de representar datos geoespaciales
5. Visualiza datos jerárquicos y/o conectados

(Marca solo un óvalo)

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 *(Tiene las cinco características anteriores)*

ANEXOS

15. Archivos de exportación *

La herramienta:

1. Ofrece la opción de exportar las visualizaciones en PDF
2. permite la exportación de los gráficos en formatos estáticos como PNG o JPEG
3. Posibilita la exportación a formatos vectoriales como SVG o EPS
4. Admite la exportación en formatos interactivos como HTML
5. Exporta los datos a formatos de hoja de cálculo como Excel o CSV

(Marca solo un óvalo)

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 *(Tiene las cinco características anteriores)*

16. Costo *

La herramienta:

1. No necesita pago de una licencia inicial para su uso
2. No tiene costos de suscripción mensuales o anuales
3. No cobra cargos extras por añadir nuevos usuarios y colaboradores al sistema
4. No cobra cargos para instalar actualizaciones o acceder a características más avanzadas
5. No tiene costos por el volumen de datos que puede manejar

(Marca solo un óvalo)

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 *(Tiene las cinco características anteriores)*

17. Renderizado rápido de resultados *

La herramienta:

1. Genera gráficos con pequeños conjuntos de datos de manera eficiente
2. Representa gráficos con múltiples capas, detalles o niveles de personalización sin afectar el tiempo de respuesta
3. Gráfica grandes volúmenes de datos sin que el proceso de visualización se vuelva lento
4. Actualiza los gráficos rápidamente cuando se realizan cambios en los datos o se aplican filtros
5. Garantiza tiempos de renderizado rápidos en diferentes entornos como navegadores web, dispositivos móviles o software de escritorio.

(Marca solo un óvalo)

- ☐ 0 *(No cuenta con ninguna de las características anteriores)*
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4

ANEXOS

- 5 (*Tiene las cinco características anteriores*)

18. Calidad de la visualización

La herramienta:

1. Construye gráficos legibles y fáciles de interpretar
2. Permite ajustar diversos aspectos de los gráficos, como color, fuentes, grosor de línea y escalas
3. Genera gráficos atractivos, estructurados y que cumplen con los principios de diseño visual
4. Ofrece una variedad de tipos de gráficos avanzados, como gráficos 3D, mapas geoespaciales, gráficos interactivos y otros
5. Mantiene la calidad visual sin importar el dispositivo o el tamaño de pantalla

(Marca solo un óvalo)

- 0 (*No cuenta con ninguna de las características anteriores*)
- 1
- 2
- 3
- 4
- 5 (*Tiene las cinco características anteriores*)

*Continuar la encuesta repitiendo el proceso desde la **pregunta 6 a la 18** para cada una de las herramientas a evaluar.*

8.3 Anexo C: Sistema de puntuación para cuantificar los criterios propuestos para evaluar las herramientas

A continuación, se muestra como quedan desglosados cada uno de los criterios creados para evaluar las herramientas de visualización y el sistema de puntuación creado para cuantificarlos.

Características de criterio “Facilidad de uso”

La herramienta:

1. No requiere conocimientos de programación.
2. Proporciona una interfaz gráfica fácil de entender.
3. Cuenta con documentación, tutoriales o soporte en línea para resolver dudas.
4. Incorpora asistencia que le indique al usuario que hacer.
5. Proporciona una interfaz de "arrastrar y soltar".
6. No demanda períodos extensos de aprendizaje para las funciones avanzadas.

Tabla 8.2. Escala para medir la facilidad de uso

Escala	Descripción
0	No cuenta con ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene cinco o más características anteriores.

Características del criterio “Flexibilidad en la edición de gráficos”

La herramienta:

1. Permite que los gráficos puedan ser personalizados en aspectos básicos como color, tipo de letra, y tamaño.
2. Acepta personalización avanzada, como modificar ejes, etiquetas y leyendas.
3. Permite que se incluyan elementos interactivos como zoom, descripciones emergentes y filtros dinámicos.
4. Posibilita seleccionar entre un conjunto de plantillas predefinidas.
5. Admite la edición del código fuente y/o plantillas para personalizaciones más avanzadas y la creación de gráficos completamente personalizados.

Tabla 8.3. Escala para medir la flexibilidad en la edición de gráficos

Escala	Descripción
0	No cuenta con ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.

5	Tiene las cinco características anteriores.
---	---

Características del criterio “Asistencia en seleccionar el tipo de gráfico”

La herramienta:

1. Incluye un asistente que oriente al usuario en seleccionar el gráfico adecuado de acuerdo con las características de los datos y el objetivo del análisis.
2. Sugiere tipos de gráficos básicos como barras, líneas, pastel, dispersión, área, entre otros según el tipo de datos.
3. Sugiere gráficos más complejos como mapas, diagramas de árbol, diagramas de cajas y bigotes, entre otros.
4. Inhabilita los gráficos que no sean adecuados para los datos.
5. Proporciona ejemplos, tutoriales o documentación que explique cómo elegir el gráfico más adecuado en diferentes escenarios de análisis.

Tabla 8.4. Escala para medir la asistencia en seleccionar el tipo de gráfico

Escala	Descripción
0	No tiene ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene las cinco características anteriores.

Características del criterio “Variedad de tipo de gráficos”

La herramienta:

1. Incluye gráficos básicos como barras, líneas o pastel.
2. Proporciona gráficos más elaborados tales como: dispersión, histogramas, diagramas de cajas y bigotes y diagramas de área.
3. Ofrece gráficos especializados como mapas de calor, gráficos de burbuja, gráficos de radar, gráficos de red, entre otros.
4. Permite combinar gráficos para generar otros más avanzados.
5. Brinda la posibilidad de diseñar nuevos tipos de visualizaciones mediante scripts o configuraciones avanzadas.

Tabla 8.5. Escala para medir la variedad de tipos de gráficos

Escala	Descripción
0	No tiene ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.

5	Tiene las cinco características anteriores.
---	---

Características del criterio “Exploración interactiva de los datos”

La herramienta:

1. Permite al usuario acercarse o alejarse de una visualización y desplazarse dentro de ella.
2. Proporciona filtros dinámicos para seleccionar los datos directamente en la visualización.
3. Muestra información adicional como etiquetas, valores exactos o descripciones contextuales, al pasar el cursor sobre un elemento del gráfico o hacer clic en él.
4. Permite que el usuario interactúe con el gráfico para seleccionar partes específicas de los datos
5. Posibilita la conexión de visualizaciones de modo que, al interactuar con una, las demás se actualicen.

Tabla 8.6. Escala para medir la exploración interactiva de los datos

Escala	Descripción
0	No tiene ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene las cinco características anteriores.

Características del criterio “Manejo de grandes conjuntos de datos”

La herramienta:

1. Admite conjuntos de datos que superen millones de registros.
2. Manipula los conjuntos de datos de gran tamaño de manera que no se degrade el rendimiento de la misma.
3. Soporta el manejo de datos tipo LOB (objetos grandes).
4. Adapta las visualizaciones a pantallas de diferentes tamaños y resoluciones sin comprometer la legibilidad
5. Permite que los usuarios exploren los gráficos a escala sin afectar el rendimiento ni la calidad visual

Tabla 8.7. Escala para medir el manejo de grandes conjuntos de datos

Escala	Descripción
0	No tiene ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene las cinco características anteriores.

Características del criterio “Fuentes y tipos de datos”

La herramienta:

1. Soporta datos desde archivos locales.
2. Conecta con bases de datos.
3. Permite la integración con flujos de datos en tiempo real.
4. Tiene la capacidad de representar datos geoespaciales.
5. Visualiza datos jerárquicos y/o conectados

Tabla 8.8. Escala para medir el criterio fuente de datos

Escala	Descripción
0	No tiene ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene las cinco características anteriores.

Características del criterio “Archivos de exportación”

La herramienta:

1. Ofrece la opción de exportar las visualizaciones en PDF.
2. Permite la exportación de los gráficos en formatos estáticos como PNG o JPEG.
3. Posibilita la exportación a formatos vectoriales como SVG o EPS.
4. Admite la exportación en formatos interactivos como HTML.
5. Exporta los datos a formatos de hoja de cálculo como Excel o CSV.

Tabla 8.9. Escala para medir el criterio archivos de exportación

Escala	Descripción
0	No tiene ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene las cinco características anteriores.

Características del criterio “Costo”

La herramienta:

1. No necesita de una licencia de pago inicial para su uso.
2. No tiene costos de suscripción mensuales o anuales.
3. No cuenta con cargos extras por añadir nuevos usuarios o colaboradores al sistema.
4. No cobra cargos para instalar actualizaciones o acceder a características más avanzadas.
5. No tiene costos asociados al volumen de datos que puede manejar.

Tabla 8.10. Escala para medir el criterio “Costo”

Escala	Descripción
0	No tiene ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene las cinco características anteriores.

Características del criterio “Renderizado rápido de resultados”

La herramienta:

1. Genera gráficos con pequeños conjuntos de datos de manera eficiente.
2. Renderiza gráficos con múltiples capas, detalles o niveles de personalización sin afectar el tiempo de respuesta.
3. Grafica grandes volúmenes de datos sin que el proceso de visualización se vuelva lento.
4. Actualiza los gráficos rápidamente cuando se realizan cambios en los datos o se aplican filtros.
5. Garantiza tiempos de renderizado rápidos en diferentes entornos como navegadores web, dispositivos móviles o software de escritorio.

Tabla 8.11. Escala para medir el renderizado de resultados

Escala	Descripción
0	No tiene ninguna de las características anteriores.
1	Tiene una de las características anteriores.
2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene las cinco características anteriores.

Características del criterio “Calidad de la visualización”

La herramienta:

1. Construye gráficos nítidos y fáciles de interpretar.
2. Permite ajustar diversos aspectos de los gráficos, como color, fuentes, grosor de línea y escalas.
3. Genera gráficos atractivos, bien organizados y que cumplen con los principios de diseño visual.
4. Ofrece una variedad de tipos de gráficos avanzados, como gráficos 3D, mapas geoespaciales, gráficos interactivos y otros.
5. Mantiene la calidad visual sin importar el dispositivo o el tamaño de pantalla.

Tabla 8.12. Escala para medir la calidad de la visualización

Escala	Descripción
0	La herramienta no ofrece gráficos claros ni bien diseñados.
1	Tiene una de las características anteriores.

ANEXOS

2	Tiene dos de las características anteriores.
3	Tiene tres de las características anteriores.
4	Tiene cuatro de las características anteriores.
5	Tiene las cinco características anteriores.