

Biblioteca Visualización de Datos Avanzada

Importancia del impacto visual, la creatividad y la libertad en la creación de gráficos personalizados



Contenido

1.	Ir	ntroducción	3
2.	J	ustificación	5
3.	С	Objetivos	6
4.	Α	lcance	8
5.	M	larco teórico	9
Ę	5.1.	Análisis de herramientas disponibles	9
Ę	5.2.	Importancia del impacto visual, el arte y la personalización de gráficos	15
6.	D	iseño e implementación de la biblioteca	17
6	3.1	Herramientas y tecnologías utilizadas para la elaboración de la biblioteca	17
6	5.2	Estructura del proyecto	20
6	5.3	Guía de usuario casos de uso	22
7.	С	Conclusiones	23
8.	Ρ	Perfeccionamiento y oportunidades de Expansión	24
9.	В	sibliografía	25



1. Introducción

El presente documento expone el caso práctico desarrollado como parte de mi proceso de titulación en la Maestría en Ciencia de Datos del Instituto Tecnológico Autónomo de México (ITAM). Este trabajo amplía y detalla los resultados obtenidos durante mi Estancia de Investigación, llevada a cabo bajo la supervisión del profesor José Ezequiel Soto Sánchez.

El proyecto consiste en la creación de una biblioteca de visualización de datos enfocada en apoyar el entendimiento y la presentación de datos en contextos educativos y profesionales con el fin de permitir a los usuarios crear gráficos personalizados, trascendiendo las limitaciones de herramientas predefinidas. El origen de esta biblioteca radica en la necesidad de un recurso didáctico para la materia de Visualización de Datos Avanzada, impartida por primera vez en el ITAM durante el semestre enero-mayo 2024. Este curso planteaba el reto de enseñar técnicas avanzadas de visualización a estudiantes con antecedentes diversos, incluyendo aquellos sin experiencia previa en lenguajes de programación como JavaScript. Así, uno de los objetivos principales del producto, es servir como un puente entre la teoría y la práctica, facilitando tanto la comprensión de conceptos complejos como la exploración creativa en la presentación de datos.

A diferencia de otras herramientas de visualización existentes, VDA permite a los usuarios diseñar gráficos hechos a la medida de sus necesidades, eliminando la necesidad de ajustar los datos a formatos predefinidos. Este enfoque fomenta la libertad creativa y permite explorar el arte como un componente esencial en la comunicación visual, haciendo que cada gráfico no solo sea funcional, sino también estéticamente impactante y adecuado al contexto.



El desarrollo de la biblioteca requirió un dominio de tecnologías web como HTML5, CSS y JavaScript, así como conocimientos en diseño gráfico, teoría del color y experiencia de usuario. En cada etapa del proyecto, busqué integrar estos elementos para garantizar que VDA sea intuitiva, accesible y escalable, proporcionando un recurso educativo que inspire a futuros científicos de datos a innovar en la forma en que presentan y exploran la información.

Más allá de ser una solución técnica o un recurso académico, esta biblioteca es una invitación a repensar las herramientas de visualización de datos. Al analizar sus capacidades frente a otras opciones del mercado, también se destacan sus limitaciones, permitiendo identificar áreas de mejora y oportunidades de expansión. Considero que el impacto de este proyecto no se limita a su aplicación en el aula, sino que tiene el potencial de convertirse en una plataforma que fomente la creatividad y la personalización en el ámbito de la ciencia de datos.

Mediante este proceso, no solo he consolidado parte de los conocimientos adquiridos durante mi maestría, sino que también he respondido a la necesidad de brindar un recurso efectivo para la materia de Visualización de Datos Avanzada y me alegra poder apoyar en la educación de futuros estudiantes.



2. Justificación

La visualización de datos es una herramienta esencial para comunicar información de manera efectiva, al transformar datos complejos en representaciones claras y visualmente atractivas. Si bien existen diversas herramientas en el mercado que facilitan este proceso, muchas de ellas presentan limitaciones importantes para ciertos usuarios y escenarios. A menudo, estas herramientas están diseñadas con gráficos predefinidos que, aunque funcionales, restringen la capacidad de personalización. Esto obliga a los usuarios a ajustar sus datos al formato establecido, en lugar de permitirles crear representaciones totalmente adaptadas a sus necesidades específicas.

Adicionalmente, muchas de estas herramientas requieren conocimientos avanzados en programación o licencias costosas, lo que puede excluir a quienes no poseen estas competencias o recursos. Por otro lado, su enfoque práctico, aunque eficiente, a menudo descuida el aspecto artístico y de personalización en la visualización de datos, ignorando que el diseño visual impacta directamente en la interpretación y comprensión de la información.



3. Objetivos

Accesibilidad:

Una de las principales limitaciones de muchas herramientas de visualización de datos en el mercado es su falta de flexibilidad y accesibilidad para una amplia gama de usuarios. VDA intenta aminorar esta brecha proporcionando una biblioteca de visualización de datos diseñada para ser intuitiva, escalable y accesible incluso para aquellos con poca experiencia en programación.

Escalabilidad

A medida que los proyectos crecen en complejidad, VDA permite integrar funcionalidades avanzadas, esto es posible gracias a su enfoque modular, que facilita la adición de nuevos elementos y características sin que el rendimiento o la usabilidad se vean comprometidos.

Flexibilidad y personalización:

La herramienta busca no solo garantizar que los gráficos sean funcionales, sino que también ofrece flexibilidad para personalizar el aspecto visual y funcional de las gráficas, permitiendo a los usuarios diseñar gráficos adaptados a sus datos y necesidades. Esto se logra proporcionando opciones de personalización que permiten crear gráficos que comuniquen eficazmente la información.



Arte y creatividad en la visualización de datos

A diferencia de otras herramientas que vienen con gráficos predefinidos que limitan la creatividad del usuario, VDA permite una mayor personalización de los gráficos según las necesidades específicas de cada proyecto. De forma análoga, VDA promueve la integración del componente artístico como un elemento fundamental en la representación de datos, con el objetivo de captar la atención del lector, estimular la creatividad y subrayar la relevancia del diseño en la comunicación visual.

Así, se busca elevar tanto el estándar estético como el comunicativo de las gráficas, facilitando la creación de visualizaciones únicas que destaquen por su claridad e impacto.

Apoyo a estudiantes y profesionales

Recordemos que la biblioteca nace a raíz de la materia de Visualización de Datos Avanzada, con el propósito de ser un recurso práctico y flexible que acompañe a los estudiantes en su proceso de aprendizaje, facilitando el desarrollo de habilidades en visualización de datos. Sin embargo, ha evolucionado para adaptarse también a las necesidades del entorno profesional. De esta manera, VDA se presenta como un puente entre la teoría y la práctica, favoreciendo el desarrollo de habilidades tanto para futuros como para actuales científicos de datos, permitiendo que la herramienta sea útil tanto en el ámbito académico como profesional.



4. Alcance

El proyecto tiene como objetivo desarrollar una biblioteca de visualización de datos que abarque tanto gráficos esenciales como ejemplos avanzados. Se estima incluir 13 gráficos clásicos o fundamentales, tales como gráficos de barras, líneas, burbujas y mapas, que serán diseñados para cubrir las necesidades más comunes en la representación de datos. Además, se desarrollarán 5 ejemplos avanzados que combinarán elementos artísticos, matemáticos y técnicas de visualización, mostrando cómo la biblioteca puede ser utilizada para crear representaciones visuales únicas, personalizadas y de gran alcance.

La biblioteca será diseñada para ser 100 % funcional y accesible, con un enfoque en la facilidad de uso, la adaptabilidad y escalabilidad. Contará con una documentación completa y detallada que permitirá a los usuarios comprender cómo integrar e implementarla en diferentes proyectos. Esta documentación incluirá una guía de usuario que describirá paso a paso el uso de la biblioteca, desde su instalación hasta la creación de visualizaciones personalizadas mediante casos de uso documentados, además de detallar las necesidades técnicas y de implementación requeridas.

Para garantizar la accesibilidad del proyecto, este se almacenará en un repositorio público de GitHub, donde estarán disponibles tanto el código fuente como la guía de usuario y los casos de uso. Estos ejemplos servirán como modelos para ilustrar las capacidades de la biblioteca en contextos reales, permitiendo a los usuarios adaptar la biblioteca a sus necesidades específicas.



5. Marco teórico

5.1. Análisis de herramientas disponibles

En el campo de la visualización de datos, existe una gran variedad de herramientas diseñadas para generar representaciones gráficas a partir de datos. Estas herramientas se pueden clasificar en dos categorías: herramientas de "propósito general", como Tableau, y bibliotecas de "programaciones especializadas", como Plotly. Cada clasificación tiene sus propios objetivos, así como sus propias ventajas y limitaciones, lo que influye en su uso con ependencia de las necesidades de los usuarios.

Las herramientas de "propósito general" están diseñadas para un público amplio, con interfaces gráficas que permiten crear visualizaciones de forma rápida y sencilla sin necesidad de programación, sin embargo, suelen estar limitadas en términos de personalización y comúnmente requieren licencias con costo monetario.

Por otro lado, las bibliotecas de "programaciones especializadas" como Plotly están orientadas a desarrolladores y científicos de datos. Este tipo de herramientas ofrecen un mayor grado de personalización y flexibilidad, no obstante, su uso puede representar un desafío para usuarios sin experiencia en programación, y su curva de aprendizaje suele ser más pronunciada.

Ambos enfoques, aunque efectivos, tienen limitaciones. Minetras que algunas herramientas tienden a ser rígidas en términos de personalización y estilos artísticos, las otras, pueden ser menos accesibles para usuarios sin habilidades técnicas avanzadas. Adicionalmente, en ambos casos, se presenta un problema de escalabilidad ya que no cuentan con un esquema o guía para poder ser modificadas o añadir funcionalidades especificas por el usuario de manera sencilla.



Ante este panorama, VDA pretende combinar la accesibilidad de las herramientas de "propósito general" con la flexibilidad y personalización de las bibliotecas, además de documentar y guiar al usuario para poder escalar la biblioteca personalizándola con nuevas funcionalidades y enfatizando el impacto artístico y visual de los gráficos.

A continuación, se presenta una breve descripción de algunas de las herramientas encontradas en el mercado.

Tableau

Tableau es unaherramienta de "propósito general" que permite a los usuarios conectar, analizar y visualizar grandes volúmenes de datos de manera intuitiva. Es especialmente útil para generar dashboards interactivos.

Power BI

Desarrollado por Microsoft, es una herramienta que combina análisis de datos y visualización, también ofrece integración con otras herramientas del ecosistema Microsoft, como Excel y Azure, facilitando la manipulación de datos y la creación de informes interactivos y visualizaciones gráficas.

Plotly

Es una biblioteca de código abierto para la creación de gráficos interactivos en Python, R y JavaScript. Es bastante flexible y capaz de generar visualizaciones avanzadas, incluyendo gráficos en 3D, mapas y gráficos científicos.



Seaborn

Se trata de una biblioteca basada en Matplotlib que simplifica la creación de gráficos estadísticos en Python. Ofrece una interfaz amigable y un diseño visualmente atractivo, haciendo énfasis en gráficos estadísticos como distribuciones, relaciones y tendencias de datos.

Estas herramientas son potentes y ampliamente utilizadas, pero suelen estar diseñadas para casos específicos y presentan limitaciones. Mientras Tableau y Power BI son ideales para usuarios no programadores, su capacidad de personalización es limitada, especialmente en diseños artísticos. Por otro lado, Plotly y Seaborn ofrecen más flexibilidad a programadores, pero requieren conocimientos técnicos avanzados y no siempre facilitan la creación de gráficos estéticamente impactantes.

A continuación, se presenta una tabla comparativa de las herramientas anteriormente descritas, describiendo sus beneficios y limitaciones.



	Facilidad de uso	Usuario objetivo	Personalización	Escalabilidad	Enfoque	Costo
Tableau	De fácil uso e interfaz intuitiva.	No técnicos.	Ofrece un conjunto limitado de plantillas y opciones de personalización	No permite integrar o modificar su código.	Análisis de información y presentación "básica" de gráficos. Aunque generan gráficos atractivos, su enfoque principal es la funcionalidad.	Tiene un modelo de licencias por usuario.
Power BI	De fácil uso e interfaz intuitiva.	No técnicos.	Ofrece un conjunto limitado de plantillas y opciones de personalización	No permite integrar o modificar su código.	Análisis de información y presentación "básica" de gráficos. Aunque generan gráficos atractivos, su enfoque principal es la funcionalidad.	Incluye versiones gratuitas limitadas y opciones de pago.
Plotly	Es una solución más técnica que requiere de conocimientos de programación.	Orientada a programadores.	Brinda mayor control de estilos y configuraciones técnicas, pero su falta de documentación de código limita su adaptabilidad en casos específicos.	Modificar características suele ser complejo y requiere conocimientos avanzados de su arquitectura interna.	Provee más control técnico que las dos primeras, pero carece de un enfoque de impacto visual.	Tiene una versión de código abierto gratuita, pero su versión empresarial con características avanzadas es de pago.



	Facilidad de uso	Usuario objetivo	Personalización	Escalabilidad	Enfoque	Costo
Seaborn	Es una solución más técnica que requiere de conocimientos de programación.	Orientada a programadores.	Brinda mayor control de estilos y configuraciones técnicas, pero su falta de documentación de código limita su adaptabilidad en casos específicos.	Modificar características suele ser complejo y requiere conocimientos avanzados de su arquitectura interna.	Provee más control técnico que las dos primeras, pero carece de un enfoque de impacto visual.	Es una solución más técnica que requiere de conocimientos de programación.
VDA	Combina facilidad de uso con la flexibilidad necesaria para personalizar gráficos.	Accesible tanto para estudiantes en proceso de aprendizaje como para profesionales con experiencia en programación.	Permite una personalización complete, desde colores, tipografías, hasta formas y animaciones.	Su documentación y ejemplos documentados así como su diseño modular la hacen escalable y adaptable.	Diseñada para destacar el aspecto artístico y el impacto visual en las representaciones gráficas, sin comprometer la claridad en la comunicación de los datos. Su principal objetivo es ofrecer un alto nivel de personalización en los gráficos, permitiendo adaptarlos a las necesidades específicas de cada usuario.	Es de código abierto y gratuita.



En conclusión, la biblioteca VDA, a diferencia de las herramientas previamente señaladas, se centra en ofrecer personalización y escalabilidad a nivel de código y funcionalidades. Además, reconociendo que el proyecto surge de una necesidad identificada por los alumnos de la Maestría en Ciencia de Datos, la biblioteca busca ser una solución accesible para quienes no poseen un amplio conocimiento en programación. Su propósito es facilitar la integración gradual de estos usuarios en el campo de la visualización de datos, al mismo tiempo que proporciona personalización, escalabilidad y un enfoque visual impactante, todo ello en una herramienta gratuita, intuitiva y versáti.



5.2. Importancia del impacto visual, el arte y la personalización de gráficos

En el ámbito de la visualización de datos, el impacto visual es fundamental para captar la atención de los usuarios y facilitar la comprensión de información compleja. Un gráfico bien diseñado no solo presenta datos de manera clara, sino que también crea una conexión visual que atrae y retiene la atención del observador. La biblioteca VDA reconoce esta necesidad al priorizar la creación de gráficos con un alto impacto visual, asegurando que cada visualización no solo comunique información, sino que también genere un efecto duradero en la audiencia.

El arte en la visualización de datos va más allá de la estética; es una herramienta esencial para potenciar la comunicación efectiva. La incorporación de elementos artísticos en los gráficos permite resaltar patrones, jerarquías y mensajes clave de forma más accesible e intuitiva. La biblioteca VDA se inspira en esta idea al ofrecer opciones para integrar componentes artísticos en las visualizaciones, creando gráficos que no solo informan, sino que también emocionan y comunican con mayor profundidad.

La personalización es otro aspecto clave en la visualización de datos, ya que las necesidades y contextos de los usuarios varían ampliamente. Las herramientas que ofrecen opciones limitadas a menudo obligan a los usuarios a adaptarse a plantillas predefinidas, lo que puede dificultar la representación precisa de los datos. VDA, por el contrario, pretende que los usuarios tengan el control total sobre sus gráficos, permitiendo modificar colores, formas, estilos y otros elementos clave para reflejar con exactitud sus ideas.



En conjunto, el impacto visual, el arte y la personalización se combinan para transformar gráficos comunes en herramientas poderosas de comunicación. VDA no solo busca cumplir con estos aspectos, sino trasladar esto a un nivel accesible, permitiendo que los usuarios, incluso aquellos sin experiencia en programación, puedan adentrarse paulatinamente a la biblioteca y llegar a crear representaciones visuales que sean claras, impactantes y personalizadas.



6. Diseño e implementación de la biblioteca

6.1 Herramientas y tecnologías utilizadas para la elaboración de la biblioteca

En el desarrollo de la biblioteca VDA se hizo uso de diversas herramientas de programación y presentación de contenido web, a continuación, se menciona cada una de ellas.

JavaScript

Es un lenguaje de programación ampliamente utilizado para el desarrollo web, reconocido por su versatilidad y capacidad para ejecutar código del lado del cliente en navegadores. Es la base para la creación de interfaces interactivas y aplicaciones dinámicas.

De acuerdo con Amazon Web Services (AWS), "JavaScript es un lenguaje de programación que los desarrolladores utilizan para hacer páginas web interactivas. Desde actualizar fuentes de redes sociales a mostrar animaciones y mapas interactivos, las funciones de JavaScript pueden mejorar la experiencia del usuario de un sitio web. Como lenguaje de scripting del lado del servidor, se trata de una de las principales tecnologías de la World Wide Web. Por ejemplo, al navegar por Internet, en cualquier momento en el que vea un carrusel de imágenes, un menú desplegable "click-to-show" (clic para mostrar), o cambien de manera dinámica los elementos de color en una página web, estará viendo los efectos de JavaScript." \(^1\) (AWS, s.f.).

La elección de JavaScript se debe a su papel central en el desarrollo web y su capacidad para interactuar con Canvas, HTML y CSS. Además, permite integrar funcionalidades avanzadas, como la manipulación de datos y animaciones, esenciales para la biblioteca.



D3.js

D3.js (Data-Driven Documents) es una biblioteca de código abierto que permite crear gráficos interactivos y visualizaciones complejas utilizando estándares web como SVG, HTML y CSS.

D3 no es una biblioteca de gráficos en el sentido tradicional, ya que no tiene el concepto de 'gráficos' como tal; cuando se visualizan datos con D3, más bien se hace uso de una variedad de primitivas para generar un gráfico más complejo.

En este caso, D3 fue seleccionada por su potencia y flexibilidad en la creación de gráficos personalizados. Su enfoque en la manipulación directa del DOM y la conexión con datos hace que sea ideal para las necesidades de la biblioteca VDA.

HTML

HTML (HyperText Markup Language) es el lenguaje estándar para estructurar y presentar contenido en la web, funge como el esqueleto de las páginas web, definiendo la disposición de elementos como texto, imágenes y gráficos.

Entre los elementos de HTML, particularmente se hace uso de HTML5 Canvas, el cual permite renderizar gráficos, imágenes y animaciones de manera dinámica en una página web. Este componente proporciona un espacio en blanco sobre el cual es posible dibujar mediante programación, utilizando el contexto de dibujo de JavaScript, además es muy versátil, permitiendo crear desde gráficos simples como líneas y círculos hasta gráficos complejos y animaciones interactivas, lo que lo convierte en una herramienta muy poderosa para visualización de datos y gráficos interactivos.



Con base en los argumentos presentados, HTML es imprescindible para la construcción de la biblioteca, pues representa la base y estructura sobre la cual ésta se construye, haciendo posible la integración de visualizaciones en una página web.

CSS

CSS (Cascading Style Sheets) es el lenguaje de diseño para la personalización, diseño y colocación de estilos en páginas web mediante el control de aspectos como colores, tipografías, espaciado y disposición visual.

CSS complementa las capacidades de VDA al permitir estilos personalizados para las visualizaciones, lo que garantiza que los gráficos no solo sean funcionales, sino también estéticamente atractivos. Esta capacidad de personalización es clave para cumplir con el objetivo de fomentar el arte, el diseño y el impacto visual en la representación de datos.



6.2 Estructura del proyecto

El proyecto de la biblioteca VDA adopta una estructura modular diseñada para facilitar su mantenimiento, escalabilidad y comprensión, con lo cual, se pretende ofrecer una descripción de cada método proporcionado en la biblioteca, incluyendo su propósito, entradas, salidas y casos de uso.

A continuación, se describe la estructura de carpetas y archivos principales del proyecto, junto con su funcionalidad:

En la carpeta raíz se encuentra el archivo principal de la biblioteca, vda.js, el cual encapsula todos los métodos y el código necesario para su uso. Al mismo nivel, se ubica la carpeta data, que contiene diversas bases de datos en formato CSV utilizadas para los casos de uso. Además, en el mismo nivel, se organizan varias carpetas, cada una correspondiente a un caso de uso que ejemplifica la aplicación de los distintos métodos de la biblioteca, tal como se observa en la siguiente imagen.

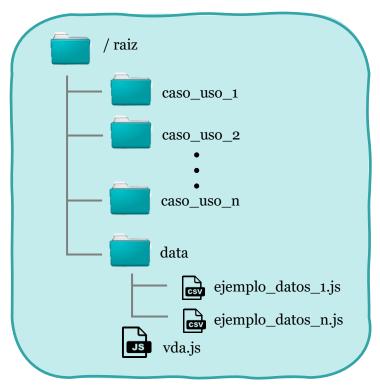


Ilustración 1. Arquitectura de carpetas de la biblioteca VDA



Cada carpeta de caso de uso, se encuentra formada por los siguientes archivos:

- Archivo HTML: Define la capa de presentación para el caso de uso específico.
- Archivo CSS: Contiene los estilos aplicados al archivo HTML para mejorar su diseño visual.
- Archivo JS: Gestiona la lógica del caso de uso, incluyendo la lectura de datos, su procesamiento y la generación de visualizaciones.

En la imagen siguiente podemos apreciar la arquitectura previamente descrita.

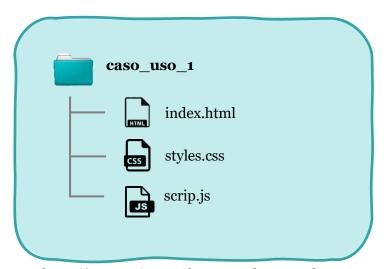


Ilustración 2. Arquitectura de carpetas de un caso de uso



6.3 Guía de usuario casos de uso

Se desarrolló una guía de usuario diseñada para facilitar el aprendizaje y uso efectivo de la biblioteca VDA. Esta guía aborda las configuraciones iniciales necesarias para integrar y ejecutar la biblioteca, además incluye una descripción detallada de cada uno de los métodos que la componen.

En la guía se documentan ampliamente los casos de uso diseñados para ilustrar las capacidades de la biblioteca. Cada caso de uso incluye:

- **Descripción de la implementación:** Se explica el caso de uso, y los métodos utilizados.
- Datos de entrada y salida: Se especifican los parámetros de entrada y salida para los métodos y los tipo de datos de dichas variables, asegurando que el usuario tenga un entendimiento completo de cada método.
- HTML y JavaScript: Se desglosa la estructura y contenido del archivo HTML utilizado para la capa de presentación, así como el código JavaScript que implementa la lógica necesaria para la visualización.
- Referencias visuales: Para cada caso de uso, se han incluido capturas de pantalla que muestran el resultado final en el navegador. Estas imágenes permiten al usuario identificar si la implementación ha sido exitosa y si los resultados cumplen con las expectativas planteadas.
- Código documentado: Adicionalmente, cada sección del código fuente está debidamente comentada para facilitar su comprensión, garantizando que los usuarios puedan modificar y ampliar la funcionalidad según sus necesidades.

La guía está diseñada pensando en la accesibilidad y claridad, haciendo uso de un lenguaje técnico pero comprensible, para que tanto usuarios con experiencia en programación como principiantes puedan beneficiarse de ella. Al cubrir cada método y funcionalidad de la biblioteca, segura que los usuarios tengan las herramientas necesarias para recrear los ejemplos propuestos y realizar sus propias visualizaciones personalizadas.



7. Conclusiones





8. Perfeccionamiento y	oportunida (des de Ex	pansión
------------------------	--------------	-----------	---------



9. Bibliografía

- 1. D3.js. (s.f.), What is D3?, recuperado el 23 de noviembre del 2024 de https://d3js.org/what-is-d3
- Mozilla Developer Network (s.f.), Conceptos básicos de HTML, recuperado el 23 de noviembre del 2024 de https://developer.mozilla.org/es/docs/Learn/Getting_started_with_the_web/HTML_basics
- 3. Wikipedia contributors. (s.f.), Canvas element, recuperado el 23 de noviembre de 2024 de https://en.wikipedia.org/wiki/Canvas_element
- 4. GeeksforGeeks (s.f.), CSS Introduction, recuperado el 23 de noviembre de 2024 de https://www.geeksforgeeks.org/css-introduction/
- 5. ClickUp (s.f.). The 15 best data visualization tools in 2023, recuperado el 30 de noviembre de 2024, de https://clickup.com/blog/data-visualization-tools/
- QuestionPro (s. f.). Herramientas de visualización de datos, recuperado el 30 de noviembre de 2024 de https://www.questionpro.com/blog/es/herramientas-de-visualizacion-de-datos/