



UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

# **HERRAMIENTA INTERACTIVA PARA APLICAR TÉCNICAS DE VISUALIZACIÓN EN ANÁLISIS DE DATOS**

**Gonzalo Mauricio Zapata Ocampo**

Universidad Nacional de Colombia  
Facultad de Administración, Departamento de informática y Computación  
Manizales, Colombia  
2021

# **HERRAMIENTA INTERACTIVA PARA APLICAR TÉCNICAS DE VISUALIZACIÓN EN ANÁLISIS DE DATOS**

**Gonzalo Mauricio Zapata Ocampo**

Tesis presentada como requisito parcial para optar el título de:

**Magister en Administración de Sistemas Informáticos**

Director:

PhD. Néstor Darío Duque Méndez

Línea de Profundización:

Sistemas de Información

Grupo de Investigación:

Grupo de Ambientes Inteligentes Adaptativos – GAIA

Universidad Nacional de Colombia

Facultad de Administración, Departamento de informática y Computación

Manizales, Colombia

2021



**Dedicatoria**

*No importa cuán despacio vayas mientras no te detengas.*

*Confucio*



## **Agradecimientos**

A Dios que es el pilar de todo lo que hago y del cual no se puede realizar nada sin que él esté presente.

Al director y docente Néstor Darío Duque Méndez por su paciencia y dedicación ya que gracias a su colaboración. A mi familia por todas las enseñanzas que me han dado, a mi mamá Nelly Ocampo Osorio por su constancia y dedicación hacia mí y nunca desfallecer, mi tío Néstor de Jesús Ocampo Osorio por enseñarme el amor por las matemáticas, a mi abuela Marina Ocampo Osorio por instruirme el respeto y el valor de las personas, a mi Madrina Blanca Miriam Ocampo Osorio por instruirme el respeto hacia Dios.

## **HERRAMIENTA INTERACTIVA PARA APLICAR TÉCNICAS DE VISUALIZACIÓN EN ANÁLISIS DE DATOS**

### **Resumen**

Con la revolución de los datos que evidenciamos día a día, nos enfrentamos al reto de desarrollar herramientas que almacenen y procesen información, con el fin de poder tomar decisiones bien fundamentadas. En este sentido, la visualización de la información compilada de grandes archivos o bases de datos requiere de programas especializados y habilidades en lenguajes de programación, o de incurrir en un alto costo de adquisición de aplicaciones o de licencias de uso temporales, sumado a la necesidad de equipos de cómputo acondicionados para hacer este tipo de operaciones gráficas de manera depurada y que se traduzca en datos relevantes, que sean de fácil entendimiento para los responsables de análisis.

Estas bases masivas de información se deben poder analizar por todos los sectores productivos y de servicios, seleccionando la mejor opción según sus necesidades y parámetros establecidos, ya sean estos, rangos de presupuestos, límites de tiempos o disponibilidades de materiales, entre otros. Para ello toda la información, datos y archivos, deben reposar en un sistema que posea la habilidad de procesar la información según las necesidades de acuerdo con las restricciones propias de su entorno. El trabajo presentado en este documento pretende diseñar, desarrollar e implementar un Software Interactivo de Visualización de Datos, que posea las características de: *i)* gratuidad; *ii)* facilidad de ejecución, sin requerir del usuario experiencia previa en lenguajes de programación; y *iii)* apto para ser utilizado por quien lo necesite en línea, con una interfaz amigable e intuitiva para el usuario, que haga de su operatividad una tarea fácil de realizar o aprender, aportando nuevas herramientas a esta área del conocimiento, conocida como Data Analysis o Análisis de Datos.

**Palabras clave:** Análisis de Datos, Procesamiento de Información, Software Interactivo, Toma de Decisiones, Visualización gráfica.

# **INTERACTIVE TOOL TO APPLY VISUALIZATION TECHNIQUES IN DATA ANALYSIS**

## **Abstract**

Every day there is a data revolution, so we face the challenge of developing tools to store and process information so that we can make informed decisions. Therefore, viewing the information collected from large files or databases requires specialized programs and skills in programming languages, or incurring the high cost of acquiring applications or licenses for temporary use, in addition to the need for improved computer equipment to graphical operations in a refined way and translated into relevant data, easily understood by those responsible for the analysis.

Massive information databases must be able to be analyzed by all productive and service sectors, selecting the best option according to their needs and established parameters, whether they are budget ranges, deadlines or material availability, among others. For this, all the information, data and files must be in a system that has the capacity to process the information according to the needs according to the restrictions of its environment. The objective of the work presented in this document is to design, develop and implement Interactive Data Visualization Software, which has the following characteristics: i) free; ii) ease of execution, without requiring the user to have previous experience in programming languages; and iii) suitable for use by those who need it online, with a friendly and intuitive interface for the user, which makes its operation an easy task to perform or learn, providing new tools to this area of knowledge, known as Analysis of Data or data analysis.

**Keywords:** Data Analysis, Information Processing, Interactive Software, Decision Making, Graphic Visualization.

# Contenido

<b>INTRODUCCION.....</b>	15
<b>1. DESCRIPCION DEL PROBLEMA.....</b>	17
<b>1.1. Planteamiento Del Problema.....</b>	17
<b>1.2. Preguntas de Investigación Emergentes.....</b>	22
<b>1.3. Objetivos.....</b>	23
<b>1.3.1. Objetivos General.....</b>	23
<b>1.3.2. Objetivos Específicos.....</b>	23
<b>1.4. Alcance.....</b>	24
<b>1.5. Metodología.....</b>	25
<b>1.6. Cumplimiento de Objetivos.....</b>	27
<b>1.7. Principales Contribuciones Logradas.....</b>	29
<b>2. MARCO TEÓRICO Y CONCEPTUAL.....</b>	30
<b>2.1. Marco teórico.....</b>	30
<b>2.1.1. Análisis de Datos.....</b>	31
<b>2.1.2. Minería de Datos.....</b>	32
<b>2.1.3. Técnica de Visualización en Análisis de Datos.....</b>	33
<b>2.2 Revisión de las características del software de visualización de datos y selección de componentes de interés.....</b>	37
<b>2.2.1 MINITAB.....</b>	37
<b>2.2.2 Software SAS.....</b>	39
<b>2.2.3 Software TABLEAU.....</b>	43
<b>2.2.4 Software RAPID MINER.....</b>	53

<b>2.2.5 Software Orange.....</b>	54
<b>2.3 Trabajos Relacionados.....</b>	58
<b>2.4 ¿Qué Es Un Framework?.....</b>	61
<b>2.5 ¿Qué es Laravel?.....</b>	61
<b>3 DESARROLLO E IMPLEMENTACIÓN DE UN SOFTWARE INTERACTIVO DE VISUALIZACIÓN DE DATOS.....</b>	63
<b>3.1 Descripción del Aplicativo.....</b>	63
<b>3.2 Funcionalidad.....</b>	64
<b>3.3 Tipo de Información por cada gráfica.....</b>	71
<b>3.4 Mercado Objetivo.....</b>	74
<b>3.5 Metodología de análisis UML.....</b>	75
<b>3.4.1 Análisis y diseño de la herramienta.....</b>	75
<b>3.4.1.1 Análisis de requerimientos.....</b>	75
<b>3.4.2 Modelado del dominio.....</b>	77
<b>3.4.3 Casos de uso.....</b>	78
<b>3.4.4 Diagrama de actividades.....</b>	80
<b>3.4.5 Diagrama de clases.....</b>	81
<b>3.4.6 Diagrama de secuencia.....</b>	82
<b>3.4.7 Diagrama de estados.....</b>	83
<b>3.4.8 Diagrama de componentes.....</b>	84
<b>3.4.9 Diagrama de despliegue.....</b>	85
<b>3.6 Características específicas de diseño.....</b>	86
<b>3.7 Manual de Usuario.....</b>	87
<b>4 VALIDACIÓN FUNCIONAL DE LA HERRAMIENTA.....</b>	98
<b>4.1 Caracterización de la Muestra.....</b>	98

<b>4.2 Resultados de la Muestra.....</b>	<b>100</b>
<b>5 CONCLUSIONES.....</b>	<b>105</b>
<b>6 BIBLIOGRAFÍA.....</b>	<b>107</b>
<b>ANEXOS.....</b>	<b>110</b>

## Lista de Gráficos o Figuras

<b>GRÁFICO 1 HERRAMIENTAS DE MINITAB.....</b>	38
<b>GRÁFICO 2 PANTALLA INICIAL DE SAS.....</b>	40
<b>GRÁFICO 3 PRINCIPALES SUB-VENTANAS SAS.....</b>	41
<b>GRÁFICO 4 GRÁFICO DE BARRAS SIMPLES EN TABLEAU.....</b>	44
<b>GRÁFICO 5 EJEMPLO GRAFICA DE BARRA CON GAMA DE COLORES.....</b>	45
<b>GRÁFICO 6 EJEMPLO DE BARRAS APILADAS EN TABLEAU.....</b>	46
<b>GRÁFICO 7 EJEMPLO GRÁFICO DE LÍNEAS SIMPLES EN TABLEAU.....</b>	47
<b>GRÁFICO 8 EJEMPLO DE GRÁFICO DE LÍNEAS MÚLTIPLES EN TABLEAU.....</b>	48
<b>GRÁFICO 9 EJEMPLO DE GRÁFICO CON LÍNEAS CON ETIQUETA EN TABLEAU.....</b>	49
<b>GRÁFICO 10 EJEMPLO DE GRÁFICO CIRCULAR SIMPLE EN TABLEAU.....</b>	50
<b>GRÁFICO 11 EJEMPLO DE GRÁFICO CIRCULAR DETALLADO EN TABLEAU.....</b>	51
<b>GRÁFICO 12 EJEMPLO DE GRÁFICO CIRCULAR DETALLADO CON EL NOMBRE DEL FABRICANTE EN TABLEAU.....</b>	52
<b>GRÁFICO 13 PRINCIPALES FUNCIONALIDADES DE ORANGE.....</b>	55
<b>GRÁFICO 14 MENÚ DE NAVEGACIÓN.....</b>	63
<b>GRÁFICO 15 TIPOS DE GRÁFICOS DISEÑADOS.....</b>	64
<b>GRÁFICO 16 BARRAS HORIZONTALES.....</b>	65
<b>GRÁFICO 17 BARRAS HORIZONTALES CON PORCENTAJES.....</b>	65
<b>GRÁFICO 18 BARRAS APILADAS.....</b>	66
<b>GRÁFICO 19 BARRAS AGRUPADAS CON ETIQUETAS.....</b>	66
<b>GRÁFICO 20 HISTOGRAMA BÁSICO.....</b>	67
<b>GRÁFICO 21 HISTOGRAMA ACUMULATIVO.....</b>	67
<b>GRÁFICO 22 HISTOGRAMA CON MÚLTIPLE CONJUNTO DE DATOS.....</b>	68
<b>GRÁFICO 23 MÚLTIPLE HISTOGRAMAS.....</b>	68
<b>GRÁFICO 24 DIAGRAMA DE RADAR.....</b>	69
<b>GRÁFICO 25 DIAGRAMA POLAR POR ÁREA.....</b>	69
<b>GRÁFICO 26 DIAGRAMA CIRCULAR BÁSICO.....</b>	70
<b>GRÁFICO 27 DIAGRAMA LÍNEAS APILADAS.....</b>	70
<b>GRÁFICO 28 MODELADO DE DOMINIO.....</b>	77

<b>GRÁFICO 29 DIAGRAMA DE CASO DE USO.....</b>	78
<b>GRÁFICO 30 DIAGRAMA DE ACTIVIDADES.....</b>	80
<b>GRÁFICO 31 DIAGRAMA DE CLASES.....</b>	81
<b>GRÁFICO 32 DIAGRAMA DE SECUENCIA.....</b>	82
<b>GRÁFICO 33 DIAGRAMA DE ESTADOS.....</b>	83
<b>GRÁFICO 34 DIAGRAMA DE COMPONENTES.....</b>	84
<b>GRÁFICO 35 DIAGRAMA DE DESPLIEGUE.....</b>	85
<b>GRÁFICO 36 INICIO DE SESIÓN.....</b>	87
<b>GRÁFICO 37 INGRESO ADMINISTRADOR.....</b>	87
<b>GRÁFICO 38 ADMINISTRADOR DE USUARIOS.....</b>	88
<b>GRÁFICO 39 REGISTRO DE USUARIO.....</b>	89
<b>GRÁFICO 40 LISTA DE PERMISOS.....</b>	89
<b>GRÁFICO 41 INICIO DE SESIÓN COMO USUARIO.....</b>	90
<b>GRÁFICO 42 SESIÓN USUARIO PRIMERA VEZ.....</b>	90
<b>GRÁFICO 43 MENÚ DE NAVEGACIÓN.....</b>	91
<b>GRÁFICO 44 WEB PARA SUBIR EL ARCHIVO.....</b>	91
<b>GRÁFICO 45 ARCHIVO SUBIDO POR EL USUARIO.....</b>	92
<b>GRÁFICO 46 LISTA DE IMÁGENES.....</b>	93
<b>GRÁFICO 47 EJEMPLO HISTOGRAMA CON MÚLTIPLE CONJUNTO DE DATOS.....</b>	94
<b>GRÁFICO 48 ARCHIVO PARA GRAFICAR O ELIMINAR.....</b>	94
<b>GRÁFICO 49 OPCIONES DE LAS GRÁFICAS QUE PUEDE REALIZAR EL SISTEMA.....</b>	95
<b>GRÁFICO 50 RESULTADO DE LA GRÁFICA MÚLTIPLE HISTOGRAMA.....</b>	96
<b>GRÁFICO 51 OPCIÓN DE DESCARGA DE IMAGEN.....</b>	96
<b>GRÁFICO 52 DESCARGA DE LA IMAGEN EN FORMATO .PNG.....</b>	97
<b>GRÁFICO 53 VENTA EMERGENTE EN DONDE SE MUESTRA LA UBICACIÓN EN DONDE SE ALMACENAR LA IMAGEN.....</b>	97
<b>GRÁFICO 54 ANÁLISIS DE LA ENCUESTA POR SEXO.....</b>	98
<b>GRÁFICO 55 ANÁLISIS DE LA ENCUESTA POR EDAD.....</b>	99
<b>GRÁFICO 56 ANÁLISIS POR PROFESIÓN DE LA MUESTRA.....</b>	99
<b>GRÁFICO 57 RESULTADOS DEL PILAR FACILIDAD DEL INGRESO.....</b>	100
<b>GRÁFICO 58 RESULTADOS DEL PILAR DISEÑO DE LA APLICACIÓN.....</b>	101

XIV Herramienta Interactiva Para Aplicar Técnicas De Visualización  
En Análisis De Datos

<b>GRÁFICO 59 RESULTADOS DEL PILAR USO DE LAS VISUALIZACIONES.....</b>	101
<b>GRÁFICO 60 PILAR RENDIMIENTO.....</b>	102
<b>GRÁFICO 61 PILAR SATISFACCIÓN DEL USUARIO.....</b>	103

## **Lista de Tablas**

<b>TABLA 1 REVISIÓN DE HERRAMIENTAS PARA VISUALIZACIÓN DE DATOS.....</b>	<b>21</b>
<b>TABLA 2 METODOLOGÍA CONSOLIDADA.....</b>	<b>26</b>
<b>TABLA 3 ESQUEMAS PARA EL ANÁLISIS DE DATOS SAS.....</b>	<b>43</b>
<b>TABLA 4 REQUERIMIENTOS FUNCIONALES.....</b>	<b>76</b>
<b>TABLA 6 DESCRIPCIÓN CASO DE USO.....</b>	<b>80</b>

## INTRODUCCION

La actual explosión de los datos y de la información explicada por el desarrollo de tecnologías de almacenamiento y el extensivo uso del internet ha potenciado la necesidad del desarrollo de técnicas para el análisis inteligente de los datos (Viktor & Paquet, 2006) (Han, Kamber, & Pei, 2012).

Entre estas técnicas desarrolladas se pueden resaltar: *i)* la minería de datos y *ii)* la visualización de información. Estas ofrecen varias técnicas que se complementan entre sí para apoyar el descubrimiento de patrones y relaciones en los datos. En este sentido, es importante resaltar que las técnicas tradicionales (algorítmicas) analizan los datos automáticamente mientras que las técnicas de visualización pueden aprovechar el proceso de extracción de datos al proporcionar una plataforma para comprender los datos y generar hipótesis sobre los datos basados en capacidades humanas, como el conocimiento, la percepción y el dominio y la creatividad (Tan, Steinbach, & Kumar, 2006)

Actualmente existen varias herramientas que permiten realizar análisis de datos y generar visualizaciones, en gráficas predeterminadas. Estas herramientas incluyen técnicas de clasificación como: *i)* árboles de decisión; *ii)* programas de inducción de reglas y redes neuronales; *iii)* enfoques de reglas de asociación o agrupación, entre otros (Viktor & Paquet, 2006). Sin embargo, estas herramientas cuentan restricciones de uso, principalmente: *i)* un alto costo y *ii)* requisitos de un alto nivel de conocimiento en programación (Viktor & Paquet, 2006).

Con lo anterior, el problema que se observa es la existencia de herramientas gratuitas que requieren un alto nivel de conocimiento para poderlas trabajar o que se puedan utilizar por un periodo de tiempo, lo que limita el proceso de aprendizaje del usuario. Entonces la pregunta de análisis de este trabajo de Maestría es: ¿Cómo diseñar, desarrollar e implementar un Software Interactivo de Visualización de Datos? que posea las características de: *i*) gratuidad; *ii*) facilidad de ejecución, sin requerir del usuario experiencia previa en lenguajes de programación; y *iii*) apto para ser utilizado por quien lo necesite en línea.

Teniendo en cuenta el problema que se plantea, en este documento se propone desarrollar una herramienta interactiva que permita de manera fácil y sencilla aplicar las técnicas de visualización de datos orientada a usuarios no programadores. Para lo que se utilizará como Framework LAVAREL y como librería especializada en PHP para este tipo de herramientas APEX CHART.

Finalmente, con esta investigación se evidencia la importancia del estudio de este tipo de herramientas de visualización en el país, pues no solo permiten el procesamiento masivo de información, sino que contribuyen a un análisis profundo de relaciones entre variables y patrones de comportamiento, lo que es de gran utilidad en el sector empresarial y microempresarial en Colombia. En especial, porque la concentración de software diseñado para suplir esta necesidad es de pago o requieren conocimientos avanzados de programación. De igual forma, la literatura relacionada permite observar que es un área del conocimiento en la que aún se puede generar valor agregado. Pues aún es experimental y requiere más desarrollo.

# 1 DESCRIPCION DEL PROBLEMA

En esta sección del documento se presenta el planteamiento del problema y la pregunta de investigación, así como las preguntas emergentes de la investigación, los objetivos y la metodología planteada. Finalmente, se presentan los resultados y logros obtenidos.

## 1.1 Planteamiento Del Problema

La visualización de datos es una forma de transformar información abstracta en gráficos o imágenes. Lo anterior, implica no solamente la presentación tradicional de tablas y texto, requiere optimizar los procesos, interpretar la información y proveer herramientas para la toma de decisiones y la evaluación de procesos (Yan & Wang, 2017).

En este sentido, la tecnología diseñada para la visualización de datos se concentra en la generación de herramientas que permitan desarrollar gráficas e imágenes interactivas, intuitivas y fáciles de analizar. Principalmente, porque este tipo de herramientas permiten que el conocimiento detrás de los datos sea interpretado y le provee al usuario la habilidad de determinar señales características para tomar decisiones, así como una vista más asertiva de las relaciones y variaciones entre los datos optimizando el tiempo y los recursos computacionales que tiene a su disposición. Es así como se puede afirmar que la visualización de los datos es un plus al análisis de datos (Viktor & Paquet, 2006) (Qiang , Wei , & Hanfei, 2010).

Bajo estos argumentos, la visualización de datos es una consecuencia natural del análisis de datos en bruto, minería de datos y otras técnicas y aplicativos desarrolladas para el análisis de datos.

Sin embargo, en la actualidad se observan algunas restricciones naturales en estas herramientas. Principalmente, su uso está limitado a las reglas de decisión que contienen por defecto, la codificación de su algoritmo, en particular, para extraer relaciones, patrones, desviaciones atípicas, y la estandarización de parámetros. Por ello es indispensable contar con conocimientos en programación para cambiar las reglas de análisis estadísticos en sus árboles de decisión, asimismo, para hacer los análisis de la información y obtener los resultados que le brinden certeza a sus búsquedas (Nouanesengsy, Woodring, Patchett, Myers, & Ahrens, 2014).

Ahora bien, estos datos que en un espacio bidimensional se analizan como estructura de dos proyecciones, x y, pueden ser tratados con reglas simples de correlación y graficados en formatos de diagramas, de dispersión, que muestran las tendencias centrales y los puntos de variación y desviación, siendo graficas muy simples, pero que al poder contemplar los datos visualmente, pueden clarificar y ser comprendidos de una forma más adaptativa por parte del usuario, que puede reconocer en sus líneas, otros tipos nuevos de relaciones y depurarlas aún más con el uso de estos parámetros incluidos en las librerías (Han, Kamber, & Pei, 2012).

Entonces, la cuestión es enlazar de manera apropiada el análisis de datos y su visualización en un contexto donde el usuario, no deba usar programación para crear sus propias reglas y parámetros, sino que estos estén disponibles. Para lograr esto, se han definido objetos de comandos establecidos, que poseen las jerarquías de clases, métodos y atributos, que conforman las librerías, que a su vez le dan libertad al usuario para relacionar sus búsquedas con diversos parámetros que están codificados y su uso deductivo (Chen & Zhou, 2016).

Actualmente existen varias herramientas que permiten realizar un análisis de datos y generar visualizaciones, en gráficas predeterminadas, sin embargo, como se ha mencionado en el documento, estas herramientas cuentan con las dificultades de su programación, pues se deben tener conocimientos en lenguajes de programación para su uso y modificación, llevar esto al terreno de los usuarios fuera de los contextos científicos, tiene un antecedente en el uso de hojas de cálculo que son inductivamente más fáciles de trabajar y de programar sus gráficas, con la ayuda de bloques de fórmulas en diferentes campos que el usuario puede utilizar y cambiar a su medida de su interés (Viktor & Paquet, 2006).

Entre estas herramientas se encuentra los lenguajes de programación como R y PYTHON, que presentan las siguientes características: *i)* son gratuitos y *ii)* requiere un conocimiento avanzado de programación. El primero es uno de los más utilizados para el análisis de datos científicos, permite graficarlos y mostrarlos de una manera más entendible. El segundo es muy utilizado para la visualización de datos, en especial porque viene por defecto con librerías ya desarrolladas para la visualización de datos, lo cual facilita el análisis de los datos. Como se mencionó ambos requieren tener un alto conocimiento de programación (Rossant, 2015) (Adams, 2014).

Por su parte, SAS, ORANGE, MINITAB, entre otros, su principal desventaja, es que son de pago por licencia de uso y que debe renovarse constantemente (mensual, anual). Estas herramientas ofrecen una versión de prueba de corto tiempo y utilidades limitadas que impiden que los analistas puedan utilizarla, eventualmente.

Como hemos destacado anteriormente, en la mayoría de las herramientas, el analista debe tener un conocimiento en programación o realizar un estudio previo a través de los tutoriales lo cual conlleva mucho tiempo. La Tabla 1

expone las herramientas de visualización de datos que actualmente son más conocidas, relacionando su costo, tiempo de prueba y la referencia pertinente.

*Tabla 1 Revisión de Herramientas para Visualización de Datos*

<b>Herramienta</b>	<b>Costo/año</b>	<b>Tiempo</b>	<b>Referencia</b>
<b>MINITAB</b>	\$1.595 dólares	30 días de prueba gratis	<a href="http://www.minitab.com/es-mx/">www.minitab.com/es-mx/</a>
<b>SAS</b>	X	No tiene	<a href="http://www.sas.com/es_es/home.html">www.sas.com/es_es/home.html</a>
<b>TABLEAU</b>	\$ 70 dólares	No tiene	<a href="https://www.tableau.com/es-es">https://www.tableau.com/es-es</a>
<b>RAPID MINER</b>	De 2.500 a 10.000 Dólares	30 días de prueba gratis	<a href="https://rapidminer.com/">https://rapidminer.com/</a>
<b>ITALASSI</b>	Empresa dedicada a la venta de software	No tiene	<a href="https://provalisresearch.com/es_products/italassi/">https://provalisresearch.com/es_products/italassi/</a>
<b>ORANGE</b>	Free		<a href="https://orange.biolab.si/">https://orange.biolab.si/</a>

Fuente: Elaboración propia con referencia a las publicaciones en Internet. Ver Referencias.

Con lo anterior y en un contexto general estas herramientas de software de análisis y visualización de datos tienen un elemento en común que son los períodos de prueba, lo que permite la utilización de este aplicativo por un determinado tiempo, esto genera ventajas y desventajas para el usuario, ya que se puede interactuar y conocer los elementos que se utilizaran, pero cuando se acaba el tiempo límite se observan restricciones, ya que para continuar usándola, tiene que pagar un costo de licencia, lo cual es un gran impedimento, en particular, para las pequeñas empresas.

Finalmente, se destaca la dificultad que tienen los usuarios no programadores de pequeñas empresas para aplicar técnicas de visualización, dado que las herramientas con interfaz amigable y amplias opciones generalmente son de costo, y las opciones libres requieren de conocimiento de programación para usar las librerías.

Con lo anterior, el problema que se observa es la existencia de herramientas gratuitas que requieren un alto nivel de conocimiento para poderlas trabajar o que se puedan utilizar por un periodo de tiempo, lo que limita el proceso de aprendizaje del usuario. Entonces la pregunta de análisis de este trabajo de Maestría es: ¿Cómo diseñar, desarrollar e implementar un Software Interactivo de Visualización de Datos? que posea las características de: *i)* gratuidad; *ii)* facilidad de ejecución, sin requerir del usuario experiencia previa en lenguajes de programación; y *iii)* apto para ser utilizado por quien lo necesite en línea.

Teniendo en cuenta el problema que se plantea, la solución propuesta es realizar una herramienta interactiva que permita de manera fácil y sencilla aplicar las técnicas de visualización de datos orientada a usuarios no programadores.

## **1.2 Preguntas de Investigación Emergentes**

- ¿Qué componentes debe contar una herramienta interactiva para la técnica de visualización de datos de software libre, orientada a usuarios no programadores?
- ¿Cuáles librerías podrían ser incluidas para aplicar la técnica de visualización de datos en esa herramienta?
- ¿Cómo se desarrollará e implementarán los componentes y librerías de este Software Interactivo de Visualización de Datos?
- ¿Cómo se validará que la herramienta es útil para usuarios no programadores?

## **1.3 Objetivos**

### **1.3.1 Objetivos General**

Diseñar, desarrollar e implementar una herramienta interactiva para la técnica de visualización de datos con software libre, orientada a usuarios no programadores.

### **1.3.2 Objetivos Específicos**

- Determinar los componentes libres que se deben incluir en la herramienta interactiva para la aplicación de la técnica de visualización de datos.
- Seleccionar las librerías para aplicar la técnica de visualización de datos que podrían ser incluidas en la herramienta a desarrollar.
- Implementar los componentes y librerías en una aplicación web.
- Validar con usuarios no programadores la funcionalidad de la herramienta desarrollada.

## **1.4 Alcance**

Diseñar, desarrollar e implementar una herramienta interactiva para la aplicación de técnicas de visualización en análisis de datos, orientada a usuarios no expertos en el uso de técnicas de análisis de datos.

En este documento no se evaluará el desempeño del análisis de los datos.

## 1.5 Metodología

**Tabla 2 Metodología Consolidada**

OBJETIVOS	ETAPAS	ACTIVIDADES
<b>Objetivo 1</b> Determinar los componentes libres que se debe incluir en la herramienta interactiva para la aplicación de la técnica de visualización de datos.	<b>Etapa 1</b> Revisión de las características de los softwares de visualización de datos y selección de componentes de interés.	1 Determinar los softwares a revisar. 2 Revisar la interfaz de usuario, navegación y herramientas gráficas del software de visualización de datos. 3 Analizar las funciones y opciones que poseen cada uno de los aplicativos. 4 Definir los componentes relevantes y las características que debería tener un software de visualización de datos, que sea intuitivo y de interfaz amigable para los usuarios no programadores.
<b>Objetivo 2</b> Seleccionar las librerías para aplicar la técnica de visualización de datos que podrían ser incluidas en la herramienta a desarrollar.	<b>Etapa 2</b> Selección de librerías libres y abiertas orientadas a la técnica de visualización de datos.	1 Exploración y evaluación de librerías libres para visualización de datos. 2 Selección de librerías orientadas a la técnica de visualización de datos.
<b>Objetivo 3</b> Implementar los componentes y librerías en una aplicación web.	<b>Etapa 3</b> Diseño de los componentes del software para la técnica de visualización de datos.  <b>Etapa 4</b> Desarrollo de los componentes del software para la técnica de visualización de datos en la aplicación web.	1 Diseño de componentes de la herramienta interactiva para su posterior desarrollo 2 Diseño de la base de datos. 3 Maquetación de los componentes funcionales y operatividad del software. 4 Construcción de los módulos con las funcionalidades establecidas para la herramienta. 5 Construcción de la base de datos según los requerimientos. 6 Integración y prueba de los componentes en la aplicación web.

OBJETIVOS	ETAPAS	ACTIVIDADES
<b>Objetivo 4</b> Validar con usuarios no programadores la funcionalidad de la herramienta desarrollada.	<b>Etapa 5</b> Pruebas de análisis de funcionalidad, del software de visualización de datos, llevadas a cabo con usuarios no programadores.	<ol style="list-style-type: none"><li>1 Diseño de pruebas de testeo para evaluar el comportamiento y funcionalidad del software con usuarios no programadores.</li><li>2 Definición de los usuarios para pruebas.</li><li>3 Aplicar las pruebas de testeo en ambientes controlados.</li><li>4 Analizar los resultados de las pruebas y hacer los ajustes necesarios.</li></ol>

## 1.6 Cumplimiento de Objetivos

**Objetivo 1:** Determinar los componentes libres que se debe incluir en la herramienta interactiva para la aplicación de la técnica de visualización de datos.

Para la realización de este objetivo se estudiaron en profundidad diferentes herramientas disponibles y los componentes utilizados, así como una gran variedad de literatura relacionada con la generación de aplicativos o frameworks de visualización de datos.

**Objetivo 2:** Seleccionar las librerías para aplicar la técnica de visualización de datos que podrían ser incluidas en la herramienta a desarrollar.

Se realizó una revisión de las principales librerías usadas por herramientas similares a las desarrolladas y se determinó que se utilizaría APEX CHAR. A razón de que es una librería especializada para el análisis de datos y visualización de gráficas.

**Objetivo 3:** Implementar los componentes y librerías en una aplicación web.

Se desarrollo el aplicativo web utilizando Laravel, el framework de PHP más popular en el mundo informático. Laravel es una plataforma de código abierto que destaca por ofrecer una sintaxis elegante, y, crear códigos con una amplia variedad de funciones de manera relativamente sencilla (LARAVEL, 2021). El desarrollo de este objetivo se detallará en el capítulo 3 de este documento.

**Objetivo 4:** Validar con usuarios no programadores la funcionalidad de la herramienta desarrollada.

Se utilizaron usuarios sin conocimiento en programación para hacer pruebas de la aplicación, sin previa capacitación, con el fin de evaluar la facilidad e intuitividad de la herramienta. Para lo que se diseñaron encuestas de prueba (Ver Anexo).

## **1.7 Principales Contribuciones Logradas**

Con este proyecto se logró diseñar, desarrollar e implementar un Software Interactivo de Visualización de Datos con las características de: *i)* gratuitad; *ii)* facilidad de ejecución, sin requerir del usuario experiencia previa en lenguajes de programación; y *iii)* apto para ser utilizado por quien lo necesite en línea.

La funcionalidad de este aplicativo web es generar graficas científicas de datos numéricos de una manera sencilla y gratuita, solamente subiendo un archivo en el enlace indicado y se obtendrá una gran variedad de visualizaciones gráficas, lo que genera una evaluación más flexible por parte del usuario.

---

## **2 MARCO TEÓRICO Y CONCEPTUAL**

### **2.1 Marco teórico**

En un contexto más general y para realizar una evaluación de las herramientas que se usan actualmente, hay que analizar las ventajas y desventajas que tiene cada herramienta. Esto implicaría un estudio a profundidad de una de ellas lo cual no es pertinente para este documento, ya que en función del tipo de empresa y del tipo de información que se maneja, se pueden tener varias opciones según la preferencia del usuario y la que más se adapten al esquema empresarial. Por esto, en esta sección se pretende entender de manera global, las fortalezas y debilidades sin dar preferencia a un software en específico.

Por otro lado, hay una gran variedad de software desarrollado para el análisis y visualización de datos, por lo que explicarlos todos no es pertinente. Sin embargo, para tener una visión amplia del referente actual, se hará una aproximación a cada herramienta y conceptos necesarios para esta investigación.

### **2.1.1 Análisis de Datos**

Análisis de Datos (*Data Analysis*, o *DA*) es la ciencia que examina datos en bruto con el propósito de sacar conclusiones sobre la información. El análisis de datos es usado en varias industrias para permitir que las compañías y las organizaciones tomen decisiones empresariales y también es usado en las ciencias para verificar o reprobar modelos o teorías existentes (Adams, 2014).

El análisis de datos se distingue de la extracción de datos por su alcance, su propósito y su enfoque sobre el análisis. Los extractores de datos clasifican inmensos conjuntos de datos usando software sofisticado para identificar patrones no descubiertos y establecer relaciones escondidas. El análisis de datos se centra en la inferencia, el proceso de derivar una conclusión basándose solamente en lo que conoce el investigador (Adams, 2014).

---

## 2.1.2 Minería de Datos

Según (Tan, Steinbach, & Kumar, 2006) la minería de datos es una tecnología que combina los métodos de análisis de datos tradicionales con sofisticados algoritmos para procesar grandes volúmenes de datos. La minería de datos también es conocida como Generación de Conocimiento a partir de Datos o *Knowledge Discovery from Data* (KDD), que es el proceso general de conversión de datos sin procesar en información útil. Este proceso consiste en una serie de pasos de transformación, desde el preprocesamiento de datos hasta el pos-procesamiento de los resultados de la extracción de datos (Han, Kamber, & Pei, 2012).

Con ello se ha ampliado el rango de oportunidades para explorar y analizar nuevos tipos de datos, así como, para analizar tipos antiguos de datos de nuevas maneras. En este sentido, se puede afirmar que la minería de datos es el proceso de descubrir automáticamente información útil en grandes repositorios de datos. Las técnicas de minería de datos se implementan para explorar grandes bases de datos con el fin de encontrar patrones novedosos y útiles que de otro modo podrían seguir siendo desconocidos (Han, Kamber, & Pei, 2012).

Es importante destacar que las técnicas tradicionales de análisis de datos a menudo han encontrado dificultades prácticas para enfrentar los desafíos planteados por los nuevos conjuntos de datos. Los siguientes son algunos de los desafíos específicos que motivaron el desarrollo de la minería de datos: *i)* Escalabilidad; *ii)* Alta dimensionalidad; *iii)* Datos heterogéneos y complejos; *iv)* Propiedad de Datos y Distribución; *v)* Análisis no tradicional (Tan, Steinbach, & Kumar, 2006).

### **2.1.3 Técnica de Visualización en Análisis de Datos**

La explosión actual de datos e información, causada principalmente por las tecnologías de almacenamiento de datos, así como el uso extensivo de Internet y las tecnologías relacionadas, ha aumentado la necesidad urgente del desarrollo de técnicas para el análisis inteligente de datos. La minería de datos, que se refiere al descubrimiento y la extracción de fragmentos de conocimiento de grandes repositorios de datos, tiene como objetivo atender esta necesidad (Viktor & Paquet, 2006).

La minería de datos y la visualización de información ofrecen varias técnicas que se complementan entre sí para apoyar el descubrimiento de patrones en los datos. Las técnicas tradicionales (algorítmicas) analizan los datos automáticamente mientras que las técnicas de visualización pueden aprovechar el proceso de extracción de datos al proporcionar una plataforma para comprender los datos y generar hipótesis sobre los datos basados en capacidades humanas, como el conocimiento, la percepción y el dominio y la creatividad (Tan, Steinbach, & Kumar, 2006).

En la actualidad se desarrolla la era de la información, donde la gran cantidad de datos que se dispone es abrumadora y cada vez es más relevante el tratamiento de la información dentro de las organizaciones, que buscan tomar las decisiones acertadas optimizando el tiempo y los recursos disponibles (Han, Kamber, & Pei, 2012).

Con el desarrollo de los Procesadores de Datos, se abrió un área del conocimiento con un variado número de posibilidades, en particular, para el análisis multi variable y multi factorial de la información. Es por eso que se

---

empezaron a proveer soluciones como: Excel, la hoja de cálculo de Microsoft, lo que inicio una revolución en el procesamiento de los datos, con un adelanto en la facilitación de los análisis de la información en una amplia gama de disciplinas, desde los análisis matemáticos formulados, hasta las funciones más complejas de estadística, todo con la posibilidad de ver gráficamente los datos agrupados en cuantas jerarquías como desee el usuario, según la capacidad de programación y formulación (Han, Kamber, & Pei, 2012) (Tan, Steinbach, & Kumar, 2006).

Con lo anterior se ha desarrollado nuevos formatos para procesar grandes cantidades de información y se está abriendo camino en cada una de las disciplinas de análisis, desde las investigaciones científicas hasta en el sector empresarial, incluyendo la relación de variables en los procesos que antes no se lograban correlacionar (Han, Kamber, & Pei, 2012). Esto permite realizar proyecciones y predicciones más acertadas.

Posteriormente, se incluyeron procesos de depuración de la información, la posibilidad de visualizar los datos, que encuentra aquellas relaciones y resalta la información importante dentro de una nube inmensa de datos, mostrando las tendencias, valores anómalos y eliminando capas de datos, facilitando el entendimiento de los fenómenos analizados y brindando el sentido que se requiere para cada situación, así como las posibilidades de proyectarlos y pronosticarlos, para generar las decisiones adecuadas con los menores costes de inversión posible (Han, Kamber, & Pei, 2012).

Con lo anterior, una pregunta importante a resolver es ¿Por qué es importante visualizar la información? Al respecto, generar un gráfico, a partir de un conglomerado de datos, con el fin de un mejor entendimiento

de las relaciones que se producen entre la información, es una herramienta de incalculable valor, que va más allá de adornar la parte "informativa" de un reporte o para darle mayor comprensión visual a una infografía.

Como lo confirman una gran variedad de autores (Adams, 2014) (Chen & Zhou, 2016) (Khlebnikov, Kainz, Steinberger, & Schmalstieng, 2013) (Murray, 2017) (Viktor & Paquet, 2006) (Yan & Wang, 2017) producir visualizaciones eficaces de los datos que puedan entregar un entendimiento más profundo de los fenómenos analizados, es una compleja combinación, entre la forma, la función y los requerimientos de análisis. Elegir el gráfico que entrega la mejor visualización de los datos depurados y como representar las relaciones y dependencias, es una habilidad que se desarrolla a través de la práctica y está fundamentado en las capacidades que ofrece el software, para amalgamar la información en un panorama representativo, que muestre todos los aspectos relevantes del análisis.

Por lo que en el mercado actual hay gran cantidad de herramientas para la visualización y el análisis de datos, que poseen cualidades y características distintas, desde lo simple a lo complejo, desde lo intuitivo a lo obtuso, con alta complejidad a las interfaces netamente amigables con el usuario, pero no todos los aplicativos se pueden estandarizar y utilizar en cualquier organización. Es importante tener en cuenta las dependencias o necesidades de las organizaciones.

Con todo, el viejo aforismo, de que una “imagen vale más que mil palabras”, es aplicable a esta rama del conocimiento pues la visualización de los datos y la exposición grafica de estos, ayudan a mejorar la comprensión y por ende la explicación de los fenómenos estudiados. Lo que impone un reto a

---

los futuros profesionales en el desarrollo de habilidades para el tratamiento de los datos, eligiendo las características y variables adecuadas para cada necesidad de evaluación de los datos y aquella información que se requiere encontrar (Han, Kamber, & Pei, 2012).

## **2.2 Revisión de las características del software de visualización de datos y selección de componentes de interés**

### **2.2.1MINITAB**

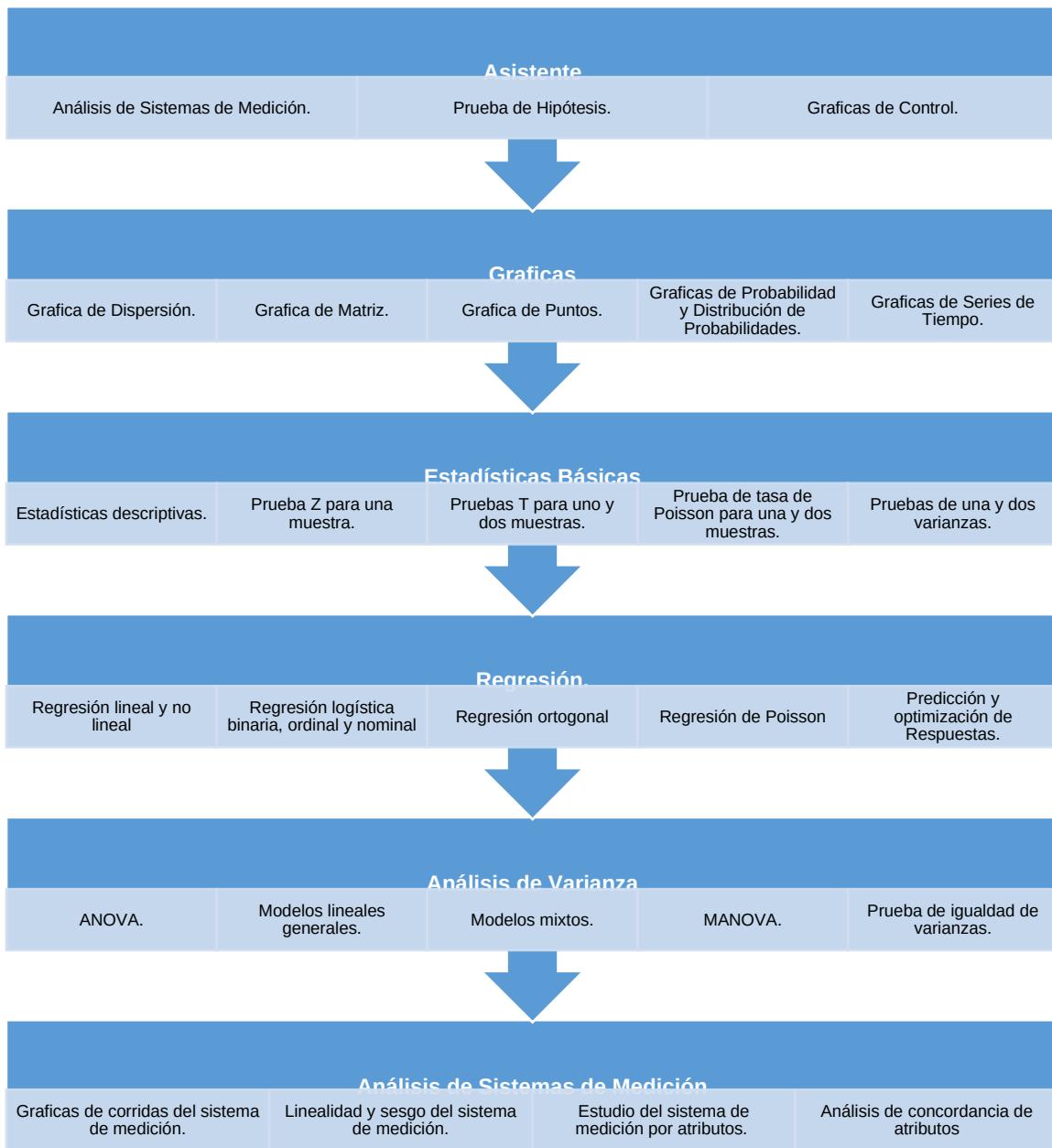
Minitab Statistical Software (su nombre comercial), es un aplicativo de análisis estadístico de datos, que además ofrece análisis predictivos y de negocio, con visualizaciones graficas en varios formatos, simples y complejos, que son de gran valor para la toma asertiva de decisiones en todas las áreas de las organizaciones generando mayores niveles de productividad y colaboración (MINITAB, 2021).

Esta es una herramienta con demasiada complejidad, dada la amplia variedad de funcionalidades que dificultan su comprensión y usabilidad. Asimismo, su esquema de diseño ofrece la posibilidad de trabajar colaborativamente en varios tableros complementarios simultáneamente.

#### **2.2.1.1 Que funcionalidad tiene**

Tiene una amplia gama de esquemas estadísticos y gráficas para desarrollar soluciones y dependiendo del tipo de necesidad de análisis a realizar, posee funciones especiales para análisis de graficas de control y posibilidad de modelado de proyectos. Para poder analizar estas graficas se tendría que trabajarlas y entender el funcionamiento de la herramienta de Minitab ya que esta posee una gran variedad como se observa en el gráfico 1 (MINITAB, 2021).

## **Gráfico 1 Herramientas de MINITAB**



Fuente: Elaboración propia con información de (MINITAB, 2021)

### **2.2.1.2 Que falencias tiene**

Las desventajas prácticas observadas del software MINITAB, principalmente se relacionan con: *i)* los procesos de descarga; *ii)* sus posibilidades de pago y pruebas gratis, que brinda acceso a todo su potencial, o la muestra gratis, que se ve limitada por el periodo de prueba de 30 días, siendo poco tiempo para lograr explorar todas sus capacidades; y *iii)* el moqueteado de la gráfica es muy complejo ya que tiene sub cuadros de edición que pueden confundir al usuario y no limitar las especificaciones pertinentes que la gráfica requiere (MINITAB, 2021).

## **2.2.2Software SAS**

### **2.2.2.1 Qué características tiene.**

El aplicativo estadístico y de visualización de datos de SAS (SAS, 2021), se desarrolló como un programa para procesar, analizar, orientar y visualizar gran cantidad de información, en un amplio espectro de posibilidades estadísticas, que prometen depurar los datos, con el objetivo de mejorar la gestión organizacional, en la toma de decisiones, respaldadas por información que agregue valor, utilizando modelos predictivos, descriptivos, de simulación y optimización, para facilitar la gestión de todas las áreas de la compañía, entregando herramientas que les ayudan a generar soluciones asertivas a las demandas de sus departamentos.

Las características más distintivas del aplicativo SAS tienen que ver con su interfaz gráfica demasiado básica, es un marco de ventanas con múltiples opciones de tratamiento de los datos. Estas ventanas funcionan en conjunto, lo

---

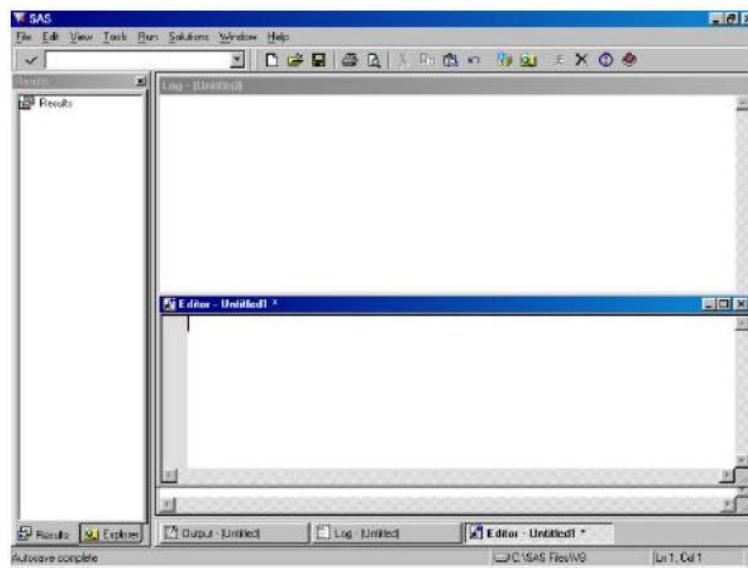
cual hace que el usuario tenga un mayor campo de trabajo para su utilización en el análisis de datos.

### 2.2.2.2 Que funcionalidad tiene.

Con la información suministrada en (SAS, 2021) se pueden destacar las siguientes funcionalidades:

**Simultaneidad:** Tiene la ventaja de trabajarse con varias ventanas de procesos diferentes, simultáneamente en un solo proyecto, sea colaborativo con más puntos en la red de la organización, o con un punto externo con acceso directo desde la nube a un usuario autorizado, haciendo que la colaboración entre áreas diferentes o cargos, pueda optimizar los tiempos de desarrollo, de los análisis, brindándoles libertad de elegir la visualización de los datos elegidos en diversos formatos de gráfico, que se ajustan a las necesidades específicas de cada cliente organizacional, aun trabajando en conjunto.

**Gráfico 2 Pantalla Inicial de SAS**



Fuente: *(SAS, 2021)*

**Interfaz:** Su configuración demasiado básica e intuitiva, la interfaz de usuario del programa SAS, es una vista que se asemeja a Windows, lo que hace que su entorno de trabajo, en sus diversas opciones de funciones, le facilite bastante el trabajo al usuario, al estar familiarizado con este tipo de disposiciones y de arreglos intuitivos en su manejo, tal como se muestra en la Gráfica 2.

A continuación, se realiza una breve explicación de las sub-ventanas que tiene a su disposición el programa, para realizar los procesos de análisis y de visualización de los datos, según sea el criterio del usuario o a quienes vaya dirigido el análisis (SAS, 2021).

---

**Gráfico 3 Principales Sub-Ventanas SAS**



Fuente: Elaboración propia con información de (SAS, 2021)

Otras características que tienen a su disposición el software mencionado es la capacidad de manejar, diferentes esquemas para el análisis de datos, los cuales se mencionan en la Tabla 3.

**Tabla 3 Esquemas para el Análisis de Datos SAS**

Modelo Predictivo	SAS Minería Visual de Datos y Aprendizaje Automático	Acelere la exploración de datos y el desarrollo de modelos con los últimos algoritmos estadísticos, de análisis de texto y de aprendizaje automático en un entorno en memoria.
Analista Textual	Análisis Visual de Texto	Obtenga el máximo valor de sus datos no estructurados utilizando la potencia combinada del procesamiento del lenguaje natural, el aprendizaje automático y las reglas lingüísticas.
Programador de Código Abierto	Programación SAS Viya	Los desarrolladores y científicos de datos tienen una oportunidad única de intentar codificar con SAS utilizando el lenguaje de su elección: SAS, Python o R. Todo lo que necesita es un navegador.
Analista de Negocios	Análisis Visual de SAS	Proporcione a los usuarios de todos los niveles de conocimientos la capacidad de explorar visualmente los datos mientras utilizan tecnologías en memoria para realizar cálculos analíticos y descubrimientos más rápidos.
Analista de Datos Estadísticos	SAS Estadística Visual	Capacite a estadísticos, científicos de datos y analistas de negocios para que descubran perspectivas clave mediante el análisis de datos y la creación y el perfeccionamiento interactivo de modelos descriptivos y predictivos.

Fuente: (Valls and Bandiella, 2018) y (SAS, 2021)

### 2.2.2.3 Que falencias tiene

Con lo anterior se evidencia que una de las principales desventajas es la necesidad de una alta capacitación en programación, por parte de los usuarios, debido a que se deben programar las tareas de análisis de datos a través de un gran número de comandos, módulos de programación, funciones numéricas, funciones aleatorias y funciones alfanuméricas, que una vez establecidos, se pueden reutilizar en otros procesamientos de datos (SAS, 2021).

---

## **2.2.3Software TABLEAU**

### **2.2.3.1 Qué características tiene**

Según la documentación provista por (TABLEAU, 2021) este software cuenta con una amplia gama de herramientas para ser utilizadas, en los análisis de datos, permitiendo una gran selección de componentes, para la gestión y depuración de los datos, en una interfaz gráfica, bastante amigable con los usuarios. Entrega al usuario, posibilidades para enriquecer sus análisis de datos y ofrece la posibilidad de trabajo colaborativo en tiempo real. Asimismo, permite la posibilidad de mezclar varias fuentes de datos simultáneamente, en una sola vista. Con lo que se puede concluir que es un aplicativo de fácil utilización, que no requiere capacitación en programación, debido a que sus comandos de procesamiento, ya se encuentran decodificados en sus funciones (Henschen, 2018) (Salmre, 2021).

### **2.2.3.2 Que funcionalidad tiene**

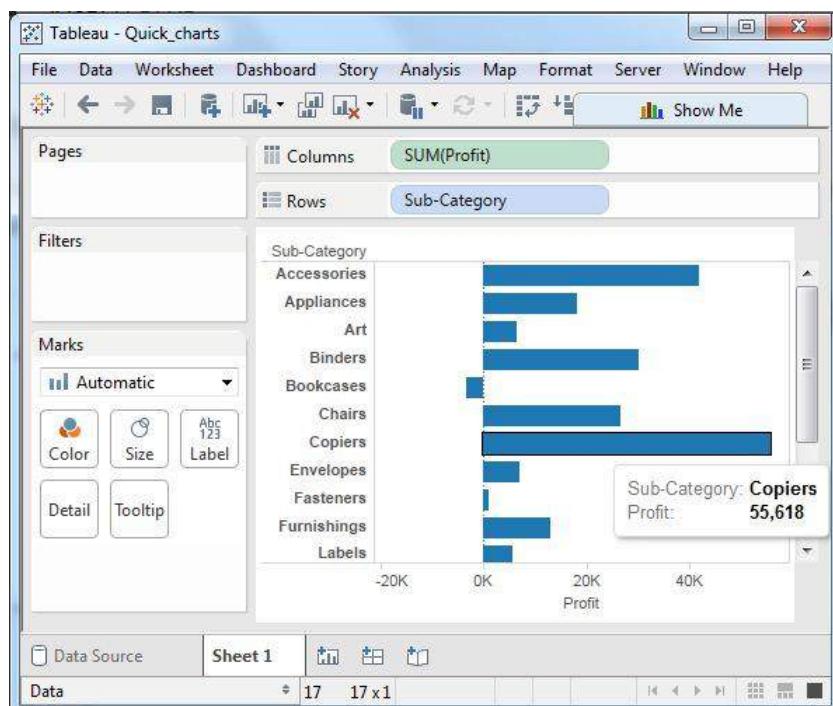
Con lo anterior se puede afirmar que es un software práctico, debido a su entorno gráfico, que le entrega al usuario una vista de iconos, con las funciones del programa, con la capacidad de combinar cuadros informativos de los análisis y graficas en un solo espacio, permitiendo cargar varias fuentes de datos en una sola vista, y procesar bases de datos y hojas de cálculo en cualquier tamaño, con la posibilidad de trabajo cooperativo y edición simultánea en cada punto de acceso (TABLEAU, 2021) (Salmre, 2021).

A continuación, se dará un repaso breve por las posibilidades de graficación de datos en el Tableau (TABLEAU, 2021) y las visualizaciones más comunes, usadas para los informes.

- **TABLEAU GRÁFICO DE BARRAS:**

Un gráfico de barras representa datos en barras rectangulares, con la longitud de la barra proporcional al valor de la variable. Tableau genera automáticamente un gráfico de barras cuando arrastra una dimensión al estante Fila y mide al estante Columna. También se puede utilizar la opción de gráfico de barras presente en el botón Mostrarme, si los datos no son apropiados para el gráfico de barras, esta opción se atenuará automáticamente. En Tableau, se pueden crear varios tipos de gráficos de barras mediante el uso de una dimensión y una medida (TABLEAU, 2021).

**Gráfico 4 Gráfico de Barras Simples en Tableau**



Fuente: (TABLEAU, 2021)

---

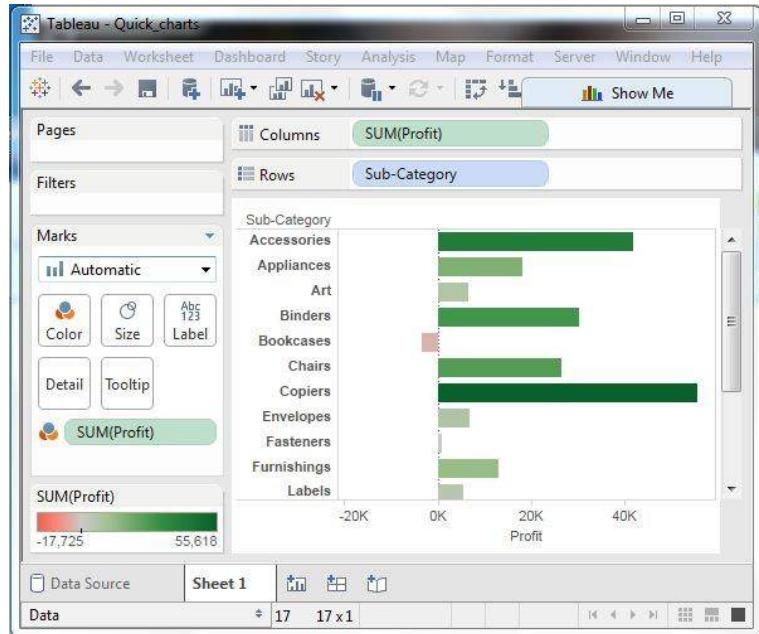
- **GRÁFICO DE BARRAS SIMPLES**

Al procesar los datos y elegir las opciones de visualización, se obtiene un gráfico de barras simples, que automáticamente degrada los colores, según sea la magnitud de los valores de las columnas, como se muestra en el gráfico 4.

- **GRÁFICO DE BARRAS CON GAMA DE COLORES**

Las barras más largas obtienen sombras más oscuras y las barras más pequeñas obtienen las sombras más claras, accediendo en el campo de beneficio, paleta de colores, bajo del Panel de Marcas, puede hacerse cuantos cambios de color se deseé, en cada barra individual, se produce un color diferente para las barras negativas. Como se muestra en la Gráfico 5 (TABLEAU, 2021).

***Gráfico 5 Ejemplo Grafica de Barra con Gama de Colores***

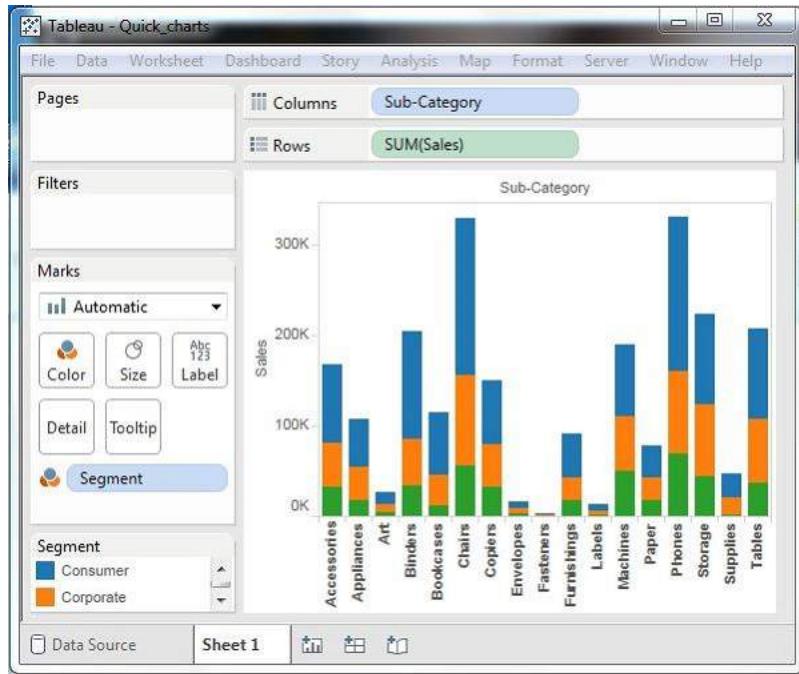


Fuente: (TABLEAU, 2021)

- **GRÁFICO DE BARRAS APIADAS**

Con el fin agregar otra dimensión al gráfico de barras, para mejorar la comprensibilidad en una comparación, el programa entrega la posibilidad de graficar barras sobre barras, con contrastes de colores, como aparece en la figura cuatro, el cuadro que muestra la distribución de cada segmento en cada barra. Como se muestra en el gráfico 6.

**Gráfico 6 Ejemplo de Barras apiladas en Tableau**



Fuente: (TABLEAU, 2021)

- **TABLEAU – GRÁFICO DE LÍNEAS**

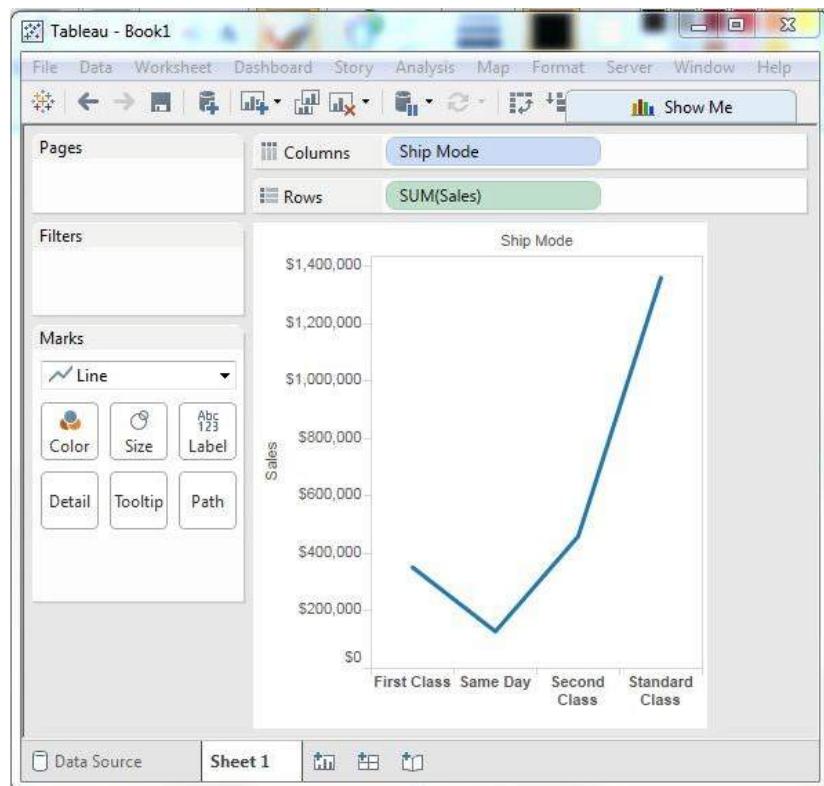
Permite la posibilidad de visualizaciones simples, para análisis de tendencia entre puntos, tomando una medida y una dimensión a lo largo de los dos ejes, el par de valores para cada observación se convierte en un punto y la unión de todos estos puntos crea una línea que muestra la variación o relación entre las dimensiones y las medidas elegidas (TABLEAU, 2021).

- **GRÁFICO DE LÍNEAS SIMPLES**

Se elige una dimensión y una medida para crear un gráfico de líneas simple, para graficar este ejemplo se usará un archivo de ventas y modos de envío, que entregara directamente la variación, arrastrando la dimensión Modo de envío al Estante de columnas y Ventas al estante Filas, eligiendo el gráfico de líneas de la tarjeta de marcas, se visualiza el gráfico de líneas, que muestra la

variación de las ventas para los diferentes modos de envío. Como se muestra en el gráfico 7 (TABLEAU, 2021).

**Gráfico 7 Ejemplo Gráfico de Líneas Simples en Tableau**

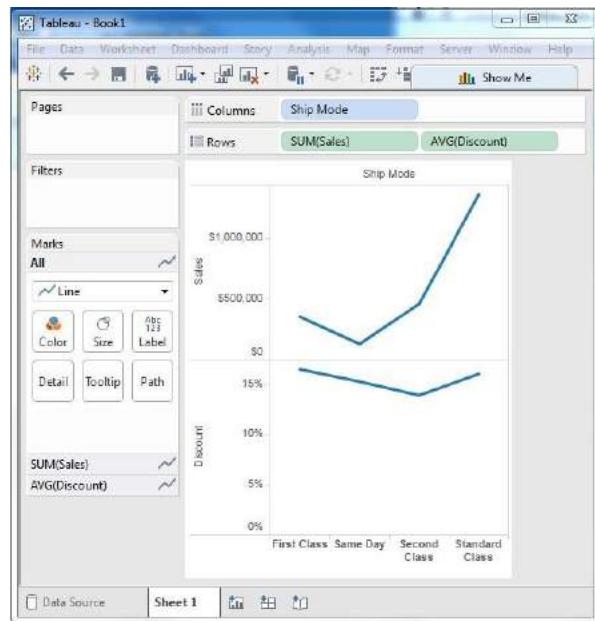


Fuente: (TABLEAU, 2021)

- **GRÁFICO DE LÍNEAS DE MEDIDAS SIMPLES**

Puede usar una dimensión con dos o más medidas en un gráfico de líneas. Esto producirá múltiples gráficos de líneas, cada uno en un panel. Cada panel representa la variación de la dimensión con una de las medidas. Como se muestra en el gráfico 8 (TABLEAU, 2021).

**Gráfico 8 Ejemplo de Gráfico de Líneas Múltiples en Tableau**



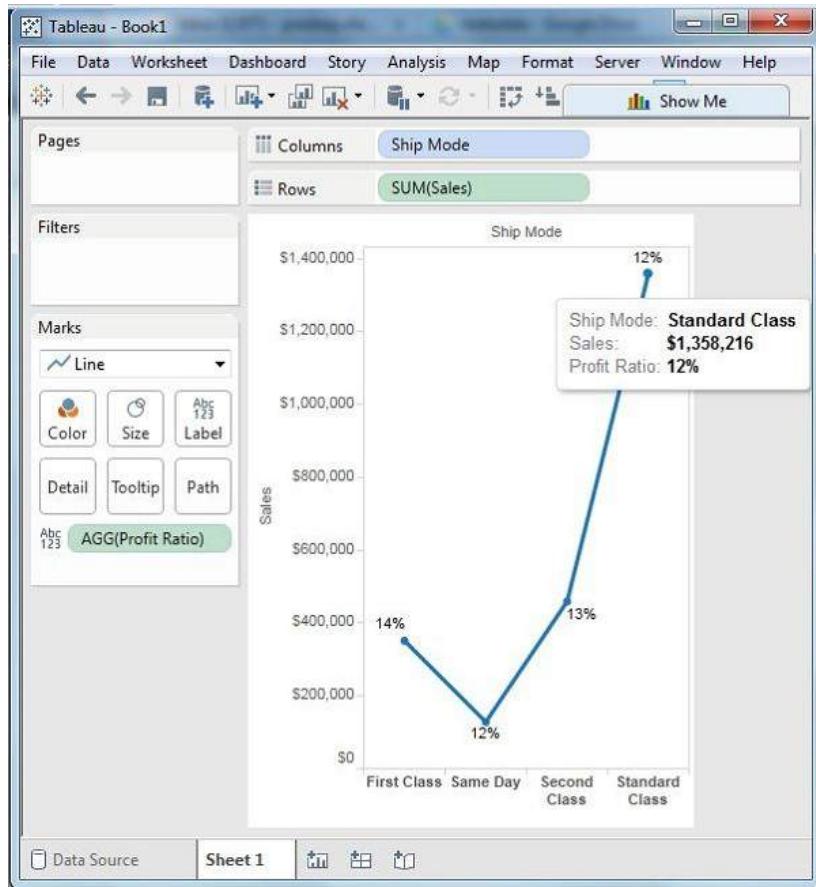
**Figura 1. Gráfico de Líneas Múltiples**

Fuente: (TABLEAU, 2021)

- **GRÁFICO DE LÍNEAS CON ETIQUETA**

Cada uno de los puntos que forman el gráfico de líneas, se puede etiquetar para hacer visibles los valores de la medida. Usando la medida de Relación de beneficio, en el panel de etiquetas en la tarjeta de Marcas. En este caso particular, el promedio, obteniendo el siguiente cuadro con etiquetas (TABLEAU, 2021).

**Gráfico 9 Ejemplo de Gráfico Con Líneas con Etiqueta en Tableau**



Fuente: (TABLEAU, 2021)

- **TABLEAU – GRÁFICO CIRCULAR**

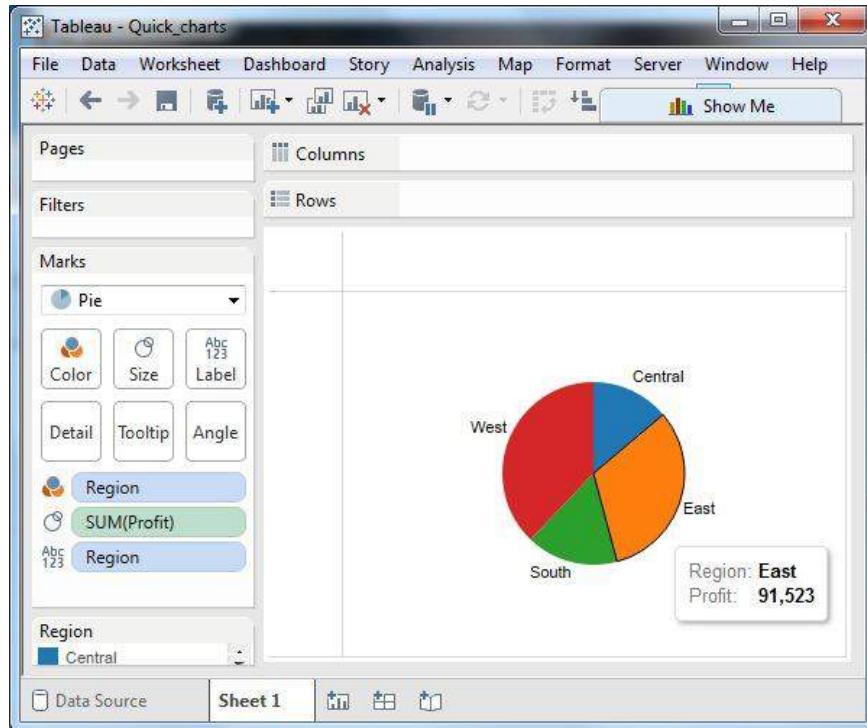
Representa los datos como segmentos de un círculo con diferentes tamaños y colores. Las secciones se etiquetan y sus valores se representan en el gráfico (TABLEAU, 2021).

- **GRÁFICO CIRCULAR SIMPLE**

Al seleccionar una dimensión y una medida, por ejemplo, la dimensión denominada región con la medida denominada beneficio. Soltarla en la dimensión Región en los colores y las marcas de etiqueta, se ajusta la medida de beneficio en la marca de tamaño y escoger el tipo de gráfico como Pie.

Aparece el siguiente cuadro que muestra las 4 regiones en diferentes colores. Como se muestra en el gráfico 10 (TABLEAU, 2021).

**Gráfico 10 Ejemplo de Gráfico Circular Simple en Tableau**

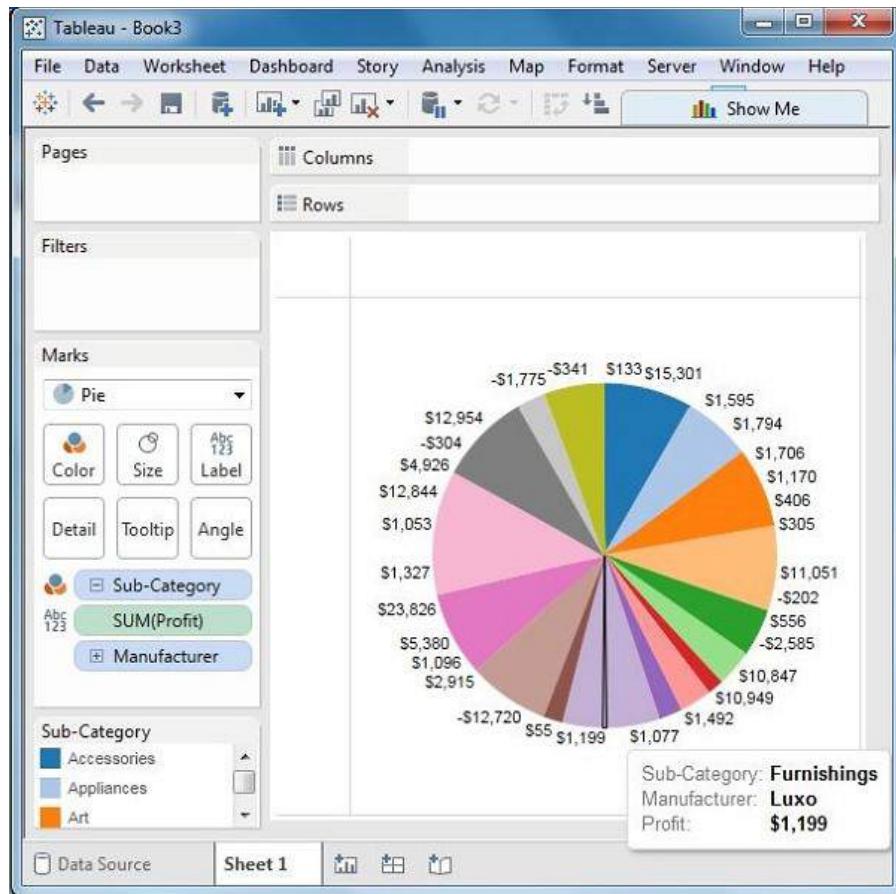


Fuente: (TABLEAU, 2021)

- **GRÁFICO CIRCULAR DETALLADO**

Etiquetado de datos de un gráfico circular, en este ejemplo, se usa la dimensión Subcategoría con dos niveles más: Fabricante y Nombre del producto. Se obtiene la medida beneficio y usándola en la marca Etiquetas. Da como resultado el gráfico circular que muestra los valores para cada segmento. Como se muestra en el gráfico 11 (TABLEAU, 2021).

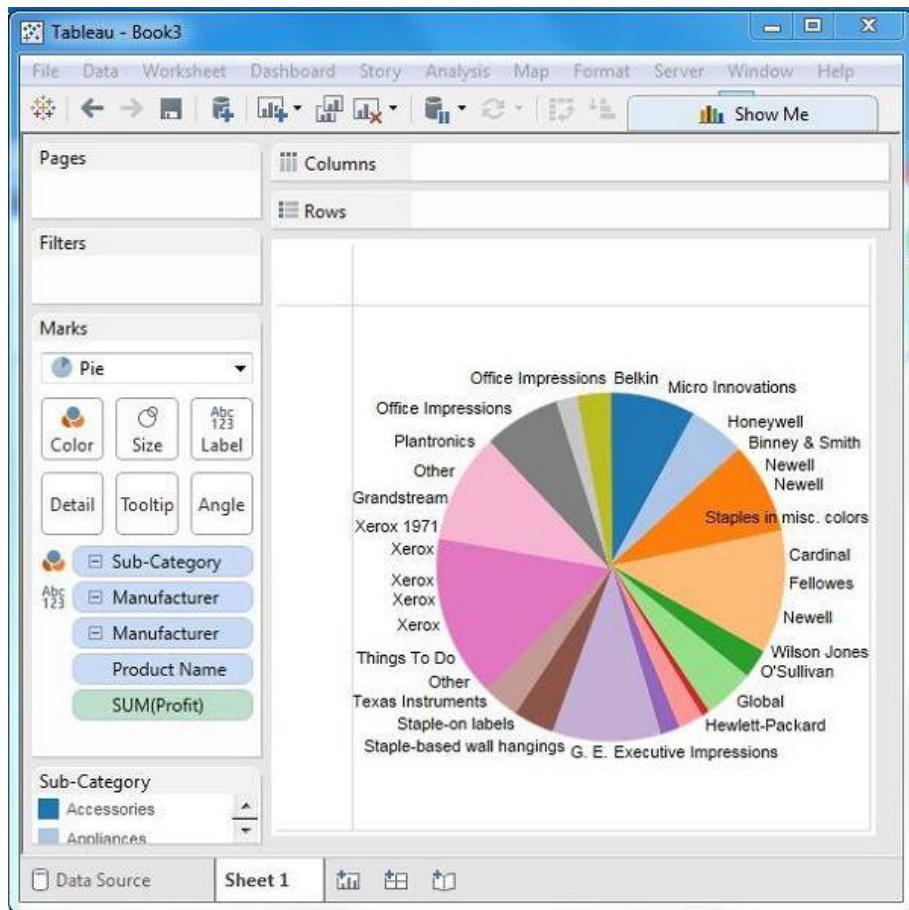
**Gráfico 11 Ejemplo de Gráfico Circular Detallado en Tableau**



Fuente: (TABLEAU, 2021)

Al pasar un nivel más en la jerarquía, obtenemos al fabricante a medida que la etiqueta y el gráfico circular anterior cambian al siguiente. Como se muestra en el gráfico 12.

**Gráfico 12 Ejemplo de Gráfico Circular Detallado con el Nombre del Fabricante en Tableau**



Fuente: (TABLEAU, 2021)

### 2.2.3.3 Que Falencias Tiene

Con el anterior análisis se puede concluir que este software presenta las siguientes desventajas: *i)* alto costo, para acceder a toda la gama de sus posibilidades; y *ii)* la disposición de los iconos, en el área de trabajo, para los usuarios, tiene tal cantidad de opciones para los análisis, que le pueden desorientar, por lo que sin el entrenamiento o la experiencia adecuada, el usuario puede desorientarse del objetivo del análisis de la información (TABLEAU, 2021).

### 2.2.4 Software RAPID MINER

### **2.2.4.1 Qué Características tiene**

Es un programa de software para el análisis, procesamiento, visualización y minería de datos, que ejecuta los procesos de análisis de datos por medio del encadenamiento de operadores preinstalados a través de un entorno gráfico. Entre sus beneficios se incluyen, la integración de datos, la transformación, el aprendizaje automático e integración de aplicaciones con enfoque unificado. RapidMiner agiliza el aprendizaje, mejora la estandarización y simplifica el mantenimiento y la extensibilidad, lo que da como resultado una gran productividad y eficiencia (RAPIDMINER, 2021).

### **2.2.4.2 Qué Funcionalidad Tiene**

Presenta un conjunto integrado de herramientas y características robustas muy potentes, con cada componente se observa una interfaz fácil de usar que ayuda a los usuarios a obtener mayor productividad desde el primer momento. Su herramienta de diseño de flujo de trabajo visual ofrece a los usuarios un entorno visual, fácil de usar que les permite diseñar, crear e implementar procesos analíticos, presentaciones visuales y modelos. Con RapidMiner, los datos despejados, desorganizados y aparentemente inútiles se vuelven muy valiosos. El sistema simplifica el acceso y el administrador de datos, permitiéndole acceder, cargar y evaluar todo tipo de datos, incluidos textos, imágenes y pistas de audio. RapidMiner le permite estructurarlos de una manera que sea fácil de comprender para usted y su equipo (RAPIDMINER, 2021).

### **2.2.4.3 Qué Falencias tiene**

---

Para realizar los análisis de datos se tiene por obligación que trabajar con los componentes por defecto, lo cual facilita el trabajo, sin embargo, cada una de estas cajas ya están predeterminadas y se quiere modificar algún parámetro requiere un conocimiento más detallado del proceso. Además, para personas que no tienen el concepto de la herramienta es difícil de entender.

## **2.2.5Software Orange**

### **2.2.5.1 Qué Características Tiene**

Orange se trata de visualizaciones de datos que ayudan a descubrir patrones de datos ocultos, proporcionando una mayor intuición detrás de los procedimientos de análisis de datos, asimismo, permite fomentar la comunicación entre científicos de datos y expertos en dominios. Los widgets de visualización incluyen diagrama de dispersión, diagrama de caja e histograma, y visualizaciones específicas del modelo como dendrograma, diagrama de silueta y visualizaciones de árbol, entre otras (Orange, 2021).

Esta aplicación ofrece la exploración interactiva de datos, para unos análisis cualitativos rápidos, con visualizaciones simples, fáciles de comprender y explicar, con una interfaz gráfica, que permite concentrarse en los análisis exploratorios y no en la programación, los valores predeterminados con inteligencia artificial hacen que la creación rápida de prototipos de un flujo de trabajo de análisis de datos sea extremadamente fácil (Orange, 2021).

### **2.2.5.2 Que Funcionalidad Tiene**

Dado que ORANGE, es una de las herramientas más prácticas analizadas en este documento, a continuación, se detallará las principales funcionalidades:

**Gráfico 13 Principales Funcionalidades de Orange**



## MINERIA DE TEXTO

Procesado de Texto  
Cuerpo de la Red  
Bolsa de Palabras  
Análisis de Sentimientos  
Modelado de Temas

## BIOINFORMATICA

Expresión Diferencial  
Ir al Navegador  
Análisis de Conglomerados  
Grafica de Volcán  
Actualización de Bases De Datos

## UNICELULAR

Cargar Datos  
Conjunto de Datos de Celda Única  
Filtrado  
Proceso de Celda Única  
Alinear Conjunto de Datos

## ANALISIS DE IMÁGENES

Importar imágenes  
Visor de Imágenes  
Incrustación de Imágenes  
Cuadricula de Imagen  
Guardar Imágenes

## REDES

Archivo de Red  
Explorador de Redes  
Generador de Redes  
Análisis de Red  
Red de Grupos

## EDUCACIONAL

K-Medidas Interactivas  
Descenso de Gradiante  
Regresión Polinomial  
Clasificación Polinomial  
Datos Aleatorios

## SERIES DE TIEMPO

Interpolar  
Transformación en Movimiento  
Evaluación del Modelo  
Ajuste Estacional  
Porción de Tiempo

Fuente: Elaboración propia con fuente de información (Orange, 2021)

### **2.2.5.3 Que Falencias Tiene**

La nueva versión 3.27.1 de ORANGE comparada con el primer software instalado para esta investigación (versión 3.23), se evidencio su evolución y se puede concluir que esta herramienta es un software que no cuenta desventajas significativas o notables tanto en diseño, herramientas de análisis y funcionalidad.

---

## 2.3 Trabajos Relacionados

Se necesitaron varias décadas y un progreso sustancial en la ingeniería de software y en los paradigmas de interfaz de usuario para crear suites modernas de minería de datos, que ofrezcan simplicidad en la implementación, integración de excelentes herramientas de visualización para la minería de datos exploratoria y la flexibilidad de diseñar nuevas formas de análisis para los usuarios. En particular, con una gran flexibilidad para lograr adaptarse a las necesidades particulares. En este orden de ideas, existe un sinnúmero de estudios relacionados, en esta sección se realizará una aproximación a estas investigaciones y se detallará las técnicas usadas y el alcance obtenido.

En su publicación (Ma & Jones, 2005) utilizan herramientas de visualización, técnicas de modelación y un modelo de IBM Data Explorer, el cual es un paquete científico de visualización de datos para lograr analizar la distribución espacial de los sedimentos generados por la contaminación ambiental en New York y New Jersey. Con lo que concluye que este tipo de herramientas de modelación y visualización son realmente útiles y permiten obtener información sobre el alcance y la distribución de la polución, como variable analizada.

Por su parte, (Smeulders & Heijs, 2005) desarrolla una visualización en 3D con los datos del mercado financiero, utilizando de 30 a 100 variables de 25.000 a 1.000.000 de clientes. De esta forma, usando técnicas de visualización (MiniSet y Iris Explorer) correlacionan las variables y determinan de una manera más sencilla patrones en los datos. La visualización interactiva desarrollada permite a los investigadores de mercado realizar una exploración rápida y efectiva de las condiciones del mercado y permite identificar aquellos

patrones, que con otras herramientas de análisis de datos requeriría más tiempo.

De igual forma, (Rubel, et al., 2010) desarrolla un marco integrado de visualización y análisis de datos en 3D, que permite la agrupación de datos guiados por el usuario, esto con el objetivo de potenciar la exploración de nuevos conjuntos de datos, en particular en el área de la biología genética. Utiliza un método de clústeres que conduce a una visualización mejorada y permite un análisis detallado de la información. Para esto utilizan una librería open source de diseño de clústeres, llamada Clúster 3.0. Lo anterior permitió crear un GUI dedicado.

(Qiang , Wei , & Hanfei, 2010) utiliza minería de datos y técnicas de visualización con el objetivo de determinar la mejor forma de graficar datos espaciales, los autores llegan a la conclusión de que este es un campo del conocimiento joven y que requiere mucha más dedicación e investigación. (Grotten , Krone , Muller, Reina, & Ertl, 2015) crea un prototipo de un marco genérico de visualización que optimiza la capacidad del GPU y genera gráficos con un alto rendimiento. Este software, denominado MegaMol permite la visualización de datos basados en datos particionados, lo que permitió crear una herramienta flexible a las estructuras de datos y con un diseño funcional.

(Khlebnikov, Kainz, Steinberger, & Schmalstieng, 2013) desarrolla un método para visualizar simultáneamente datos multivariados en tiempo real y en 3D. Para ello, usan el método NVIDIA CUDA, lo que permite una menor tasa de error en la lectura de la información. (Chen & Zhou, 2016) confirma las ventajas de utilizar técnicas de visualización 3D en el análisis interactivo de la visualización de datos. (Beran, Van Ingen, Zaslavsky, & Valentine, 2008) desarrolla una aproximación al análisis de colecciones de datos con la

---

herramienta OLAP (Online Analytical Processing), que usa los servicios de SQL Server 2008, Excel y Virtual Earth para visualizar los resultados.

(Liu & Chen, Fast Data Analysis with integrated statistical metadata in Scientific Datasets, 2012) usan librerías específicas como HDF5 y PnetCDF para proveer una herramienta de análisis de datos. Posteriormente, estos autores (Liu & Chen, Improving Data Analysis Performance for High-Performance Computing with Integrating Statistical Metadata in Scientific Datasets, 2012) seleccionan otras librerías para su estudio HDF5, ADIOS Y PnetCDF, con el objetivo de determinar cuál optimiza la visualización de los resultados de la mejor forma. Concluyendo que el conjunto de librerías disponibles no es suficiente para mejorar el rendimiento.

(Nouanesengsy, Woodring, Patchett, Myers, & Ahrens, 2014) presenta una visualización ADR (Analysis-Driven Refinement) inspirada en AMR (Adaptive Mesh Refinement) para el desarrollo de un marco de visualización masiva de datos. Con ello logran particionar la información por espacio, tiempo y variable usando métodos de priorización. Finalmente, (Westrip, 2010) desarrolla un software libre para edición y validación de archivos CIFs, disponible para Linux, MacOS y Windows. Este software provee una gran variedad de herramientas de visualización y validación de datos.

Con lo anterior se puede evidenciar una fuerte tendencia en el desarrollo de herramientas de visualización y análisis en 3D. Sin embargo, como muchos autores antes mencionados afirmaron este campo esta aún incipiente, pues exige de un mayor estudio e investigación que permita desarrollar técnicas avanzadas de análisis masivos de datos, en particular, open source y con una interfaz de usuario fácil de utilizar.

## **2.4 ¿Qué Es Un Framework?**

Es una herramienta que facilita el desarrollo de software, el cual viene por defecto con una estructura ya desarrollada que permite la optimización del código que se está desarrollando (Riasetiawan & Mahmood, 2010). El cual proporciona bibliotecas de código para funciones de uso común. Para el desarrollo de este software se utilizó un framework PHP, a razón de:

- Desarrollo más rápido.
- Menos código que escribir.
- Bibliotecas para tareas comunes.
- Hace que el código sea más limpio.
- Mejor trabajo en equipo.
- Más fácil de mantener.

## **2.5 ¿Qué es Laravel?**

El framework de PHP más popular en el mundo informático, Laravel es una plataforma de código abierto que destaca por ofrecer una sintaxis elegante, y, permite crear códigos con una amplia variedad de funciones de manera relativamente sencilla.

Este “framework de PHP para artesanos web” (tal como se describe en el sitio web oficial de Laravel), dispone de un diseño altamente intuitivo y múltiples herramientas para asumir todo tipo de desafíos (LARAVEL, 2021).

---

Entre sus características podemos resaltar:

- Extensa librería o biblioteca con archivos de terceros
- Motor de plantillas Blade
- Sistema de ruteo (RESTful)
- Eloquent ORM (Object Relational Mapping)
- Peticiones de Fluent
- Basado en Composer
- Soporte para MVC y caché
- Utiliza componentes de Symfony

### **3 DESARROLLO E IMPLEMENTACIÓN DE UN SOFTWARE INTERACTIVO DE VISUALIZACIÓN DE DATOS**

#### **3.1 Descripción del Aplicativo**

El software, Herramienta Interactiva Para La Visualización en Análisis de Datos, es un sistema que consta de un inicio de sesión por el cual se ingresa a la plataforma para su utilización, ya dentro de la misma se puede empezar a trabajar. Al ingresar por primera vez se observa una página web, en la cual se ve en un espacio de trabajo en blanco con un título que dice Archivos Subidos por el Usuario en el caso de que ya tenga almacenada información. El menú de navegación que maneja el sitio web es básico para su utilización y funcionalidad, consta de tres opciones, como se aprecia en el gráfico 14.

**Gráfico 14 Menú de Navegación**



Fuente: Elaboración Propia

---

### 3.2 Funcionalidad

La funcionalidad de este proyecto web es el de generar graficas científicas de datos numéricos una manera sencilla y gratuita, ya que las herramientas que hay en el mercado cumplen con la funcionalidad de generar las gráficas, pero se requiere un conocimiento muy elevado de programación.

Este conocimiento en el ámbito investigativo es rápidamente adoptado, pero para usuarios sin esta preparación es muy complejo, pues les tomaría mucho tiempo entender y apropiar los conocimientos necesarios de programación para lograr el objetivo final que es crear una gráfica. En cambio, con el desarrollo web de esta aplicación, lo que se logra es que solamente al subir y guardar un archivo en el enlace indicado, se pueda obtener diferentes tipos de Gráficos, como se muestra en el gráfico 15, esta herramienta cuenta con una variedad de estilos para visualizar los datos.

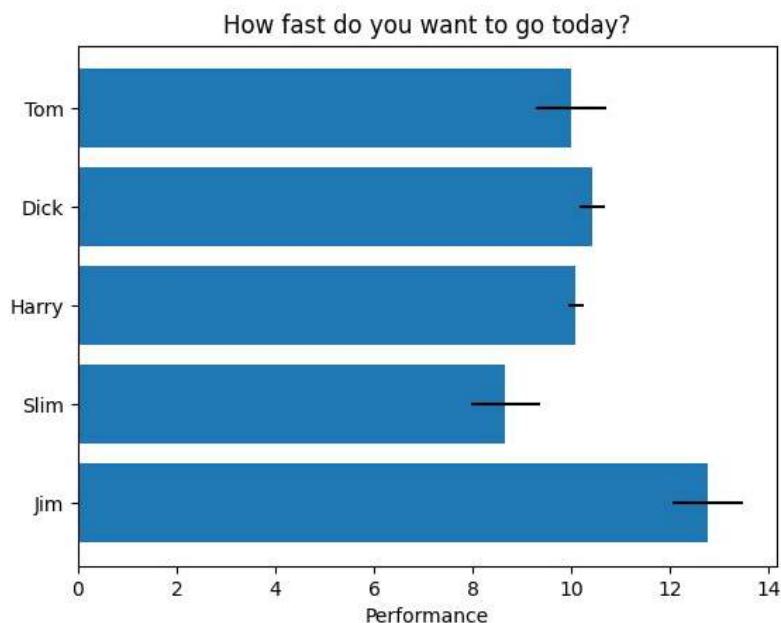
**Gráfico 15 Tipos de Gráficos diseñados**



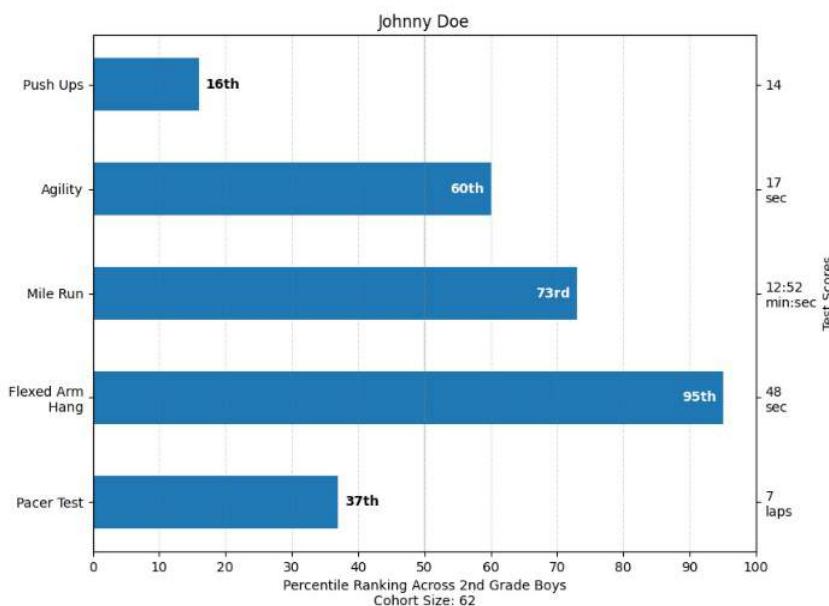
Fuente: Elaboración propia

A continuación, se visualizarán los tipos de gráficos que la herramienta puede suministrar al usuario:

**Gráfico 16 Barras Horizontales**

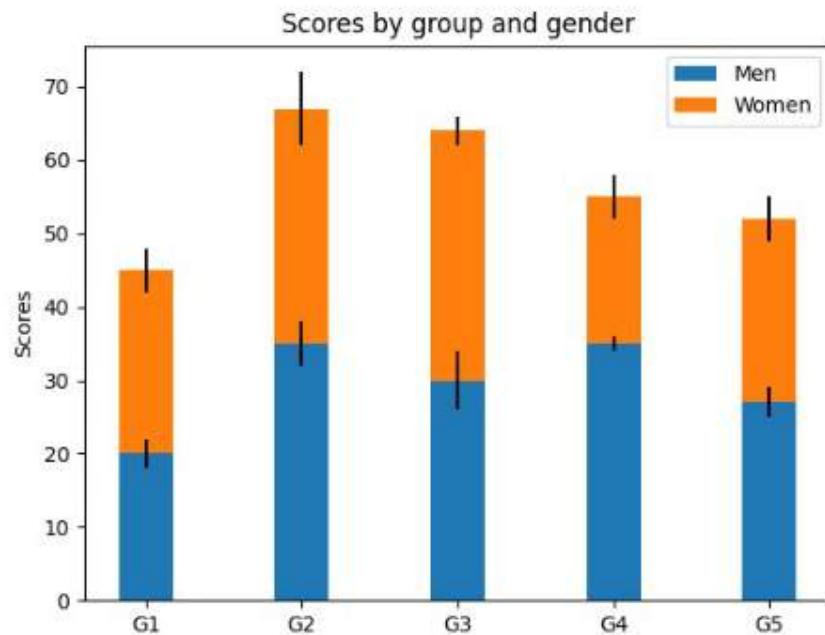


**Gráfico 17 Barras Horizontales con Porcentajes**

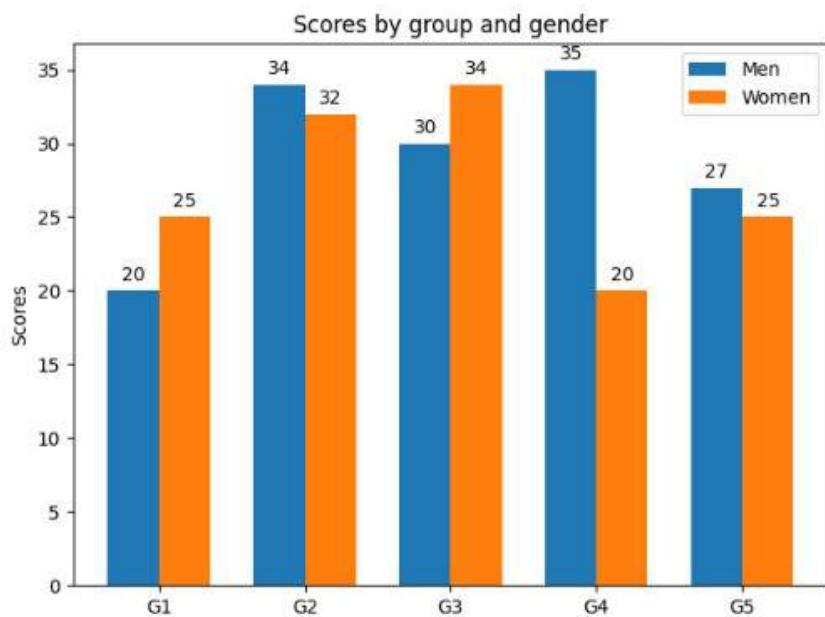


---

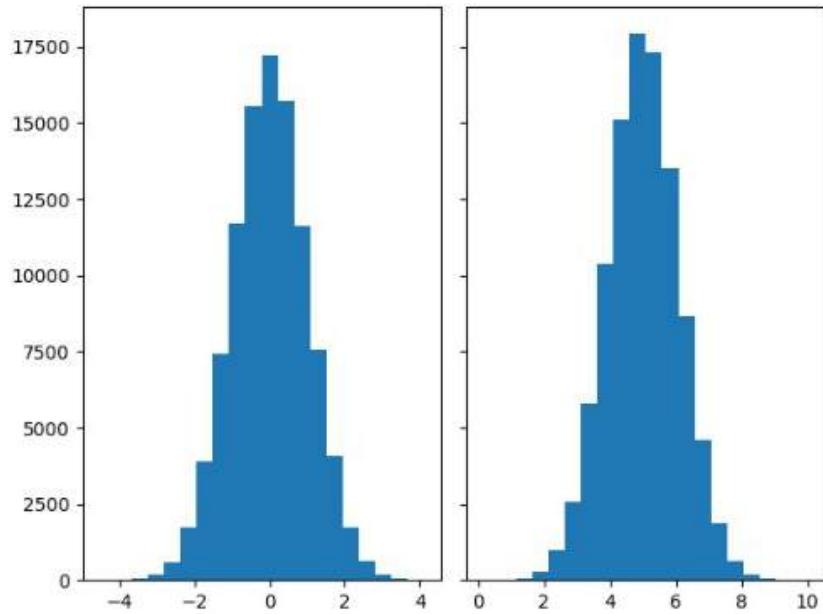
**Gráfico 18 Barras Apiladas**



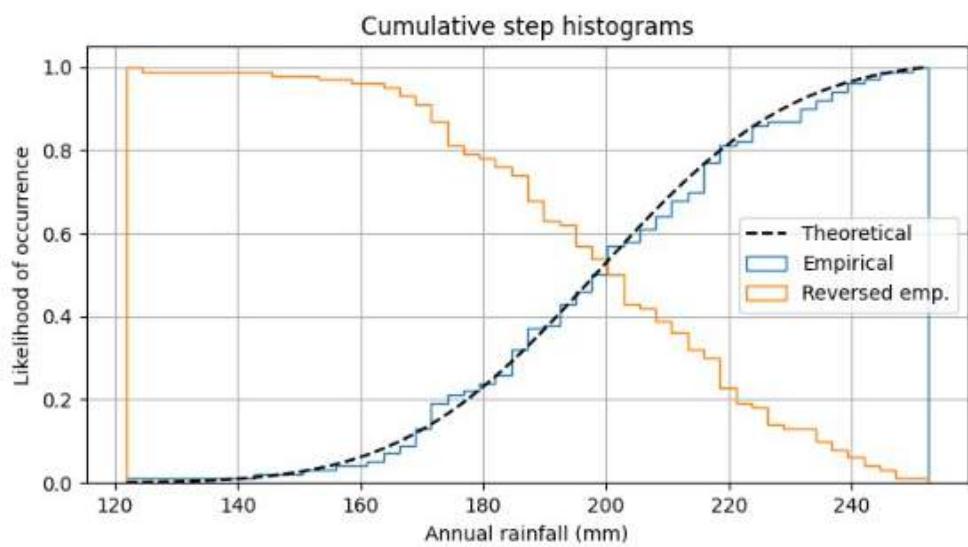
**Gráfico 19 Barras Agrupadas con Etiquetas**



**Gráfico 20 Histograma Básico**



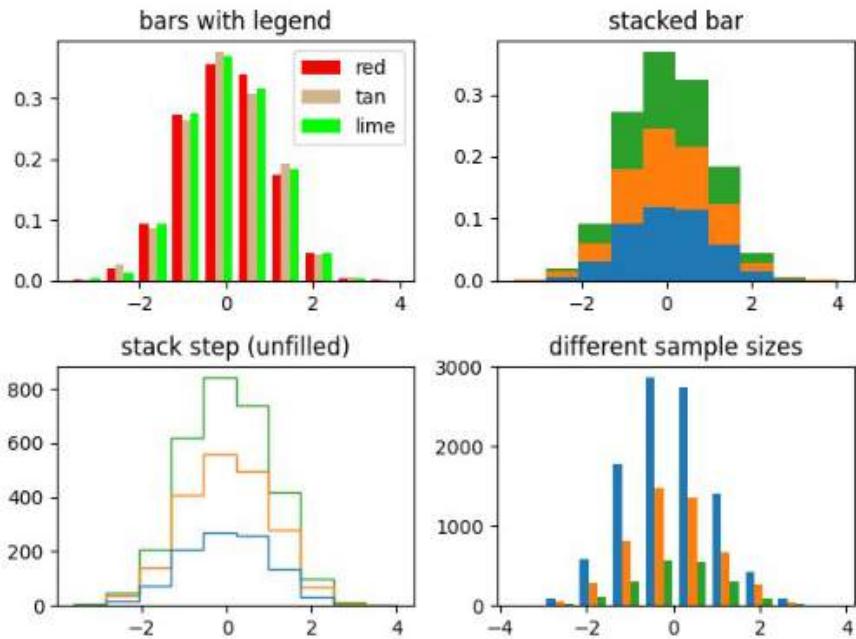
**Gráfico 21 Histograma Acumulativo**



---

**Gráfico 22 Histograma con Múltiple Conjunto de Datos**

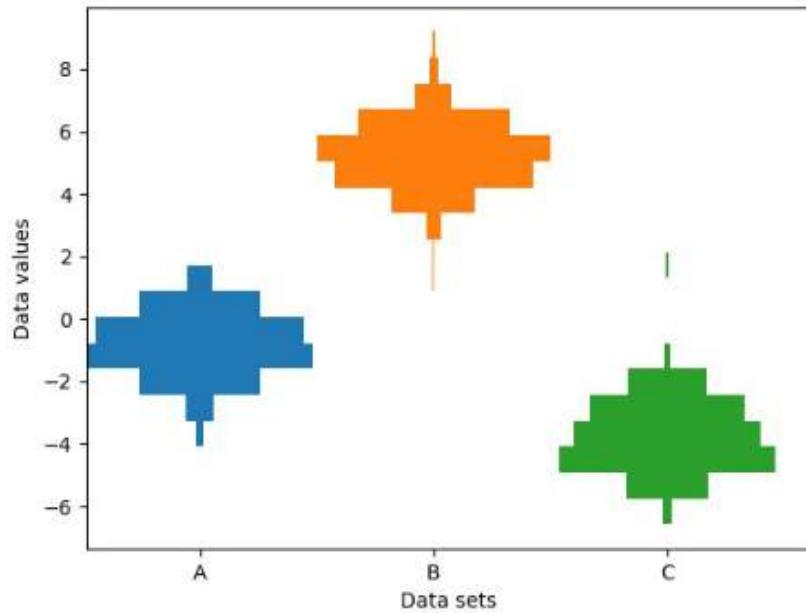
---



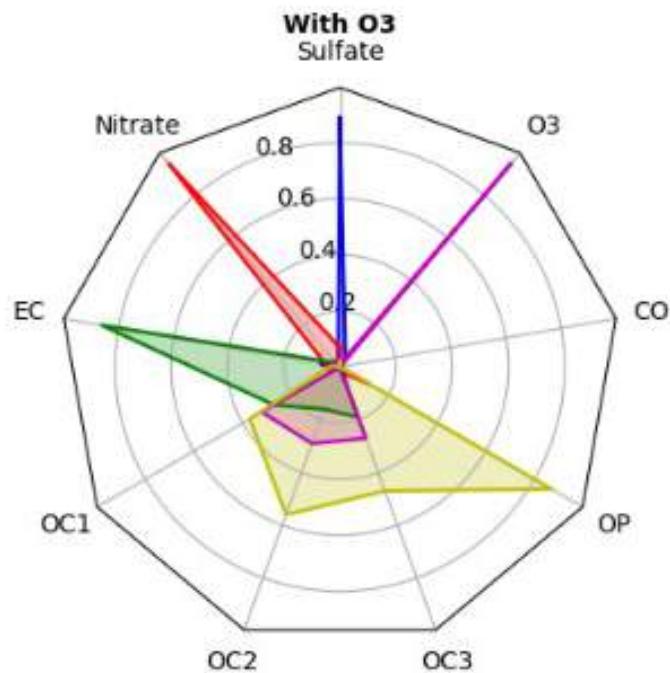
---

**Gráfico 23 Múltiple Histogramas**

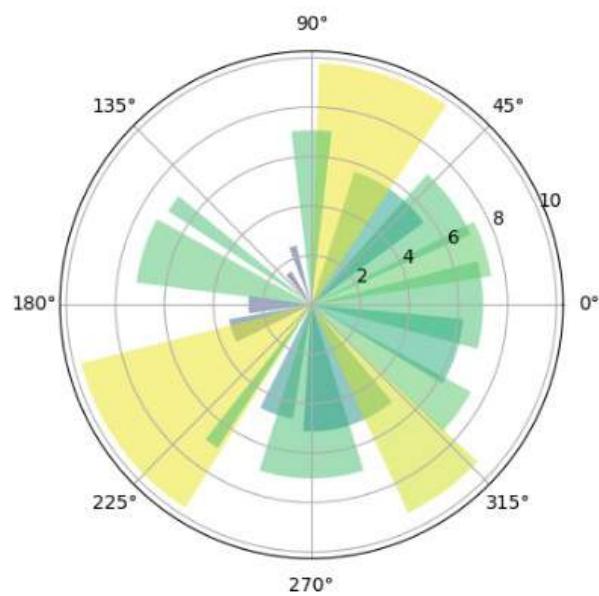
---



**Gráfico 24 Diagrama de Radar**

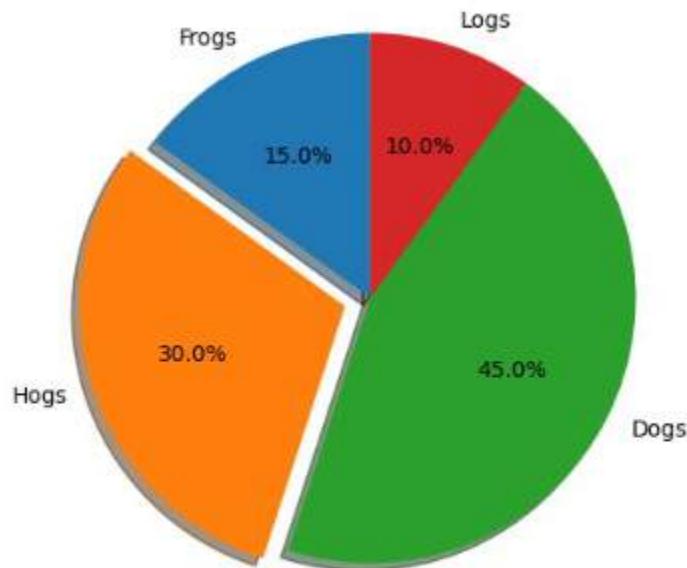


**Gráfico 25 Diagrama Polar por Área**

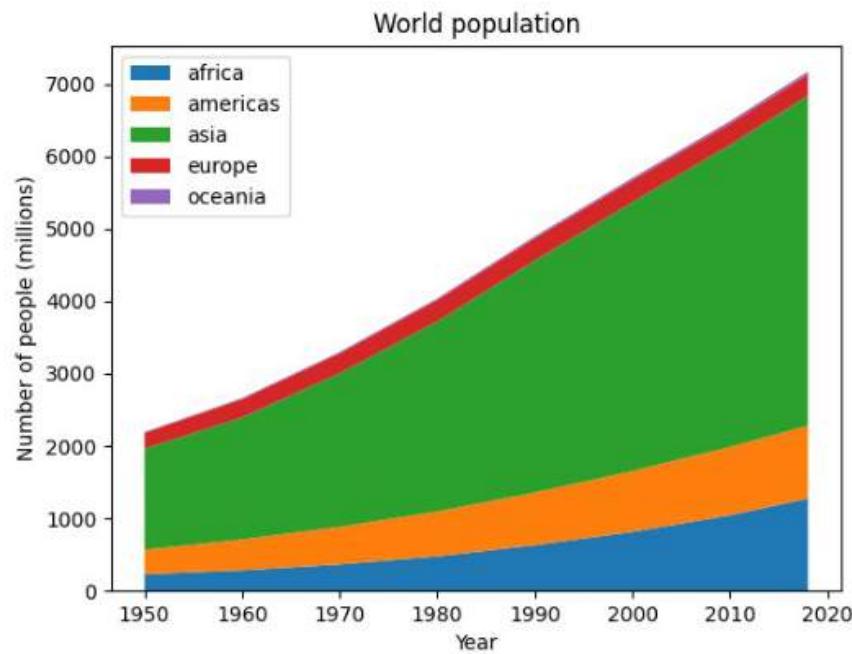


---

**Gráfico 26 Diagrama Circular Básico**



**Gráfico 27 Diagrama Líneas Apiladas**



### **3.3 Tipo de Información por cada gráfica**

#### **Barras Horizontales y Barras Horizontales con porcentaje**

Este tipo de gráficas permite comparar valores visualizar sobre una línea de tiempo. Las variables de medición deben tener características similares en la información a comparar. Entre más información tenga la fuente de datos, es más difícil encontrar valores divergentes. Con este tipo de gráficos, es posible encontrar con mayor agilidad estos valores.

Las barras con porcentaje, al igual que las barras de valores, se usan para comparar variables. La diferencia radica en que los porcentajes permiten establecer un criterio de comparación general preciso situando al analista en una perspectiva global.

#### **Barras Apiladas Con o Sin Etiquetas**

Esta gráfica facilita la comparación de dos o más variables sobre un contador temporal. La diferencia entre este tipo de barras y las barras horizontales es que en este tipo se gráfica una encima de otra y de esa manera permite visualizar la relación de cada variable con respecto a la cantidad total. No es sugerido para graficar muchos segmentos, ya que entre más se tengan, es más difícil de analizar.

#### **Barras Agrupadas**

Este gráfico es similar a las barras horizontales y tienen su mismo uso. La diferencia radica en que este tipo de gráfica muestra las barras de manera vertical. Como todos los anteriores tipos de gráficos en barra, permite mostrar datos en columnas separadas. Su uso principal es la comparación de cantidades entre sí con el objetivo de encontrar tendencias en esta información.

---

## **Histograma Básico**

Este tipo le muestra al usuario dos gráficas similares. El objetivo es que pueda realizar acercamientos, selecciones y comparaciones diferentes en cada gráfico aportando mayor valor al análisis. Cada uno de los dos gráficos mostrados, es una gráfica de líneas donde al manipular uno de los ejes en cualquiera de ellas, ambas gráficas reaccionan de manera similar para facilitar la comparación de los datos presentados.

## **Histograma Acumulativo**

Este tipo de gráficos presenta una gráfica de líneas punteadas diferentes junto con su respectiva pendiente, permitiendo presentar dos variables con diferentes que no necesariamente tienen la misma escala de valores.

## **Histograma Con Múltiple Conjunto de Datos**

Similar al histograma básico, este tipo presenta cuatro gráficas diferentes: barras horizontales, barras apiladas, líneas y barras verticales. Cuando el analista selecciona un rango o un segmento en cualquiera de las cuatro gráficas, inmediatamente las otras reaccionan ante esta selección de información mostrando los mismos datos de diferentes maneras.

## **Múltiple Histograma**

Este tipo de gráficas sirve para mostrar tres variables en un plano de dos dimensiones. Dos variables que forman los ejes y la otra variable es la que da la amplitud en el gráfico y permite visualmente entender cuál es la diferencia de los datos en sus tres variables al mismo tiempo. Cada elemento de graficado

en este histograma aparece representado en círculos o áreas cuyos tamaños varían de acuerdo con la importancia relativa dada por la tercera variable.

## **Diagrama de Radar**

Esta herramienta de análisis sirve para encontrar las brechas entre diferentes variables pertenecientes a cada elemento a analizar. Su forma es parecida a una telaraña con puntas en su extremo correspondiente a cada variable. Cada una de las redes que componen este gráfico es la escala de las variables y sobre estas redes es que se realiza una marcación de color para cada elemento.

## **Diagrama de Polar Área**

Este tipo de gráficos consiste en varios círculos por cada valor en la escala del tipo de datos y se usa para mostrar tendencias de estos valores. Se puede usar para mostrar variables en función de una dirección. Al igual que en el diagrama de radar, cada conjunto de datos se dibuja sobre el entramado circular.

## **Diagrama Circular Básico**

El gráfico circular es un círculo dividido en dos o más secciones y resaltan la posición de cada dato con respecto al todo. Además de mostrar la proporción, permite entender la relación de cada elemento comparado.

## **Diagrama de Líneas Apiladas**

Este tipo de gráfico resalta las diferencias entre las variables permitiendo observar su crecimiento y divergencia con respecto a las otras. Este tipo de gráfico es efectivo cuando la información que se representa es creciente o decreciente con respecto a las otras.

---

### **3.4 Mercado Objetivo**

Este desarrollo de software está orientado para aquellas personas que no poseen el tiempo ni el interés de aprender a programar, ya que su objetivo es la obtención de una gráfica para el análisis de datos. El software de análisis de datos trabaja para la simplificación de los resultados, con el fin de visualizar los valores obtenidos en una gráfica para su estudio y en función del resultado obtenido poder determinar, si la gráfica resultante es la adecuada para su comprensión.

## 3.5 Metodología de análisis UML

### 3.4.1 Análisis y diseño de la herramienta

En esta sección, se encuentra el análisis de la aplicación basado en la metodología mencionada previamente.

#### 3.4.1.1 Análisis de requerimientos

Este trabajo de investigación es la fuente de la información principal.

**Requerimientos funcionales:** En la tabla 4 se encuentran los requerimientos funcionales.

**Tabla 4 Requerimientos Funcionales**

<b>Id</b>	<b>Requerimiento</b>	<b>Categoría</b>
R1	Administrar usuarios y sus credenciales	Administración Usuarios
R2	Permitir el acceso a los usuarios	Administración Usuarios
R3	Administrar los permisos de la aplicación	Administración Usuarios
R4	Permitir subir archivos CSV y realizar su administración	Administración Archivos
R5	Generar gráficas de los archivos	Gráficas
R6	Descargar gráficas generadas	Gráficas

*Fuente: Elaboración propia*

---

## **Requerimientos no funcionales:**

### ***i.1 Interfaz***

Teniendo en cuenta la interfaz más intuitiva para los usuarios no expertos, la herramienta realizada debe tener una interfaz web.

### ***i.2 Seguridad***

La herramienta solicitada requerirá acceso por usuario y contraseña al iniciar. Se deberá realizar validación a los campos ingresados por los usuarios con el fin de evitar ataques de seguridad de la información. No se llevará control de los gráficos realizados.

### ***i.3 Comunicaciones***

La interfaz gráfica debe poder accederse a través del protocolo HTTP usando código HTML para ser interpretado por el navegador de los usuarios.

### ***i.4 Portabilidad***

La herramienta desarrollada deberá funcionar en los navegadores Firefox, Chrome y Edge en sus últimas versiones.

### ***i.5 Lenguaje de programación***

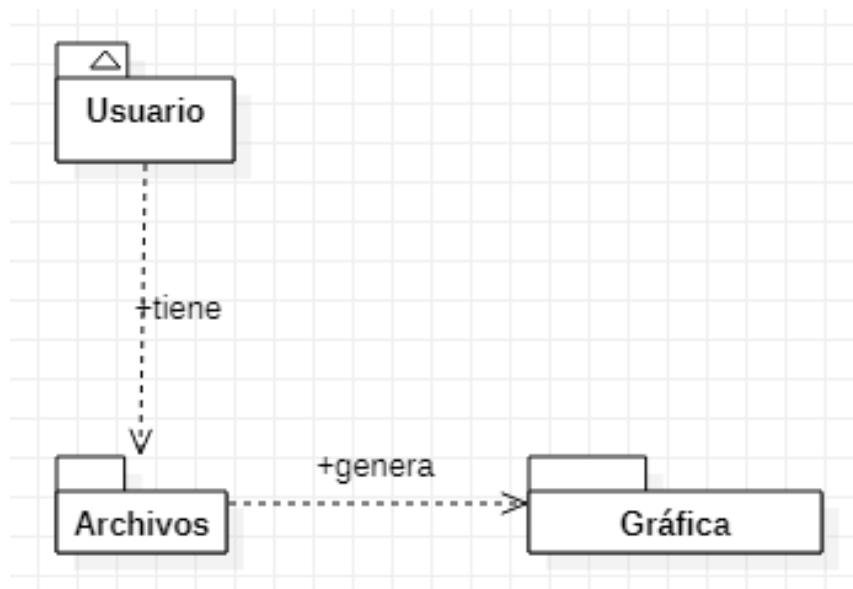
Teniendo en cuenta que no se cuenta con presupuesto para compra de licencias, se debe usar lenguajes de programación que tenga herramientas

libres para desarrollar. Además, se debe poder realizar un sitio web liviano y de programación ágil. Por esta razón, se elige PHP como lenguaje de programación para realizar la nueva herramienta.

### 3.4.2 Modelado del dominio

El modelo del gráfico 28 muestra el dominio de la aplicación.

Gráfico 28 Modelado de Dominio

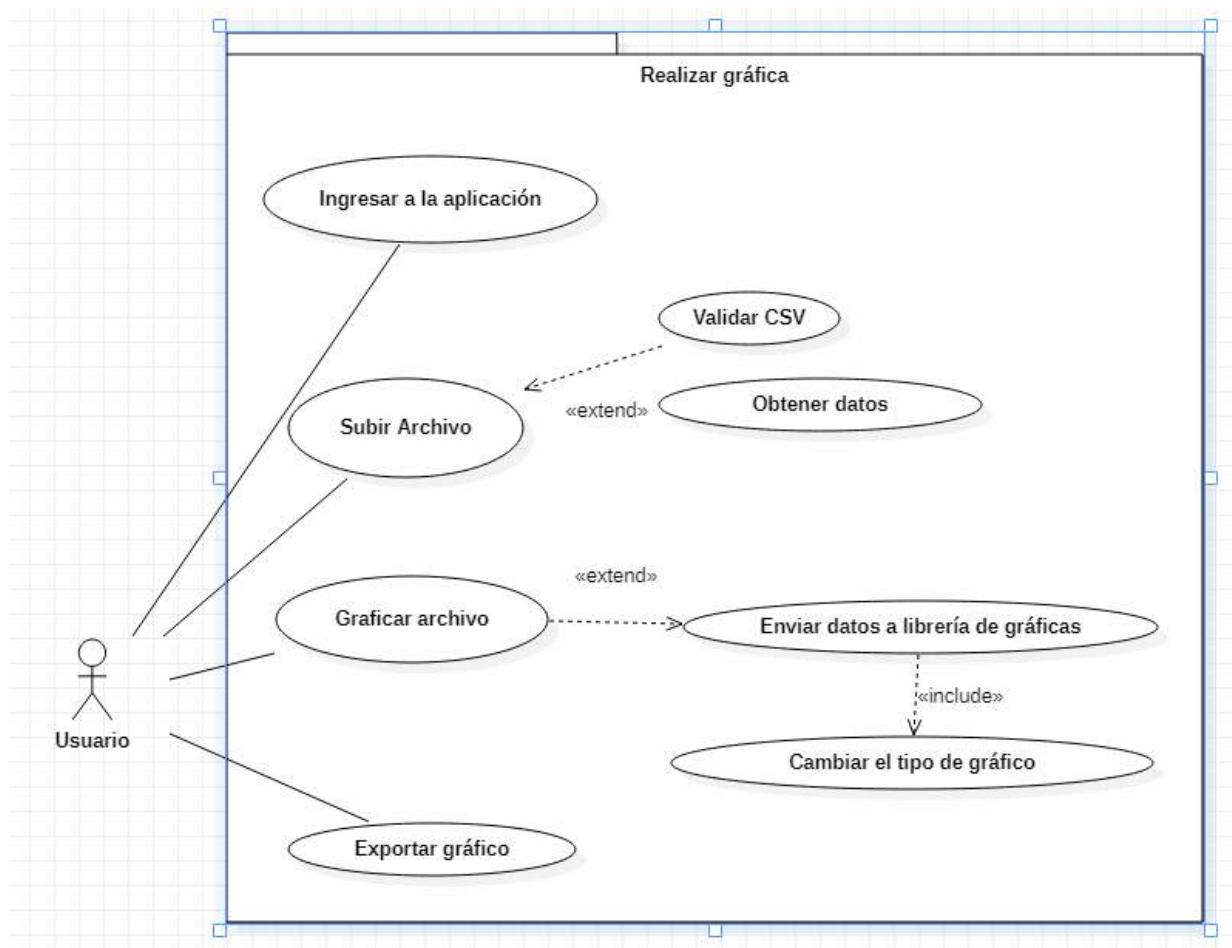


Fuente: Elaboración Propia

### 3.4.3 Casos de uso

El uso del sistema de información será la generación de gráficas, se realiza el diagrama de este caso de uso en el gráfico 29.

**Gráfico 29 Diagrama de caso de uso**



Fuente: Elaboración Propia

## Descripción

En el gráfico 28, el usuario interactúa con la aplicación para la generación de la gráfica. En la tabla 5 está la descripción.

**Tabla 5 Descripción caso de uso**

<b>Caso de uso</b>	<b>1</b>
<b>Nombre del caso de uso</b>	Generación de gráfica
<b>Actores</b>	Usuario del sistema
<b>Descripción</b>	Un usuario ingresa a la aplicación, sube el archivo y genera la gráfica.
<b>Tipo</b>	Primario y esencial
<b>Referencias cruzadas</b>	R1, R2, R3, R4, R5, R6
<b>Pre-Condición</b>	El usuario tiene el archivo CSV.
<b>Post-Condición</b>	La gráfica generada
<b>Escenario principal</b>	El usuario inicia la aplicación con su usuario y contraseña y sube el archivo que desea graficar. El sistema valida el archivo y lo almacena en el servidor. A partir de los datos del archivo, el sistema los organiza de acuerdo con los tipos de gráficos solicitados por el usuario. El sistema envía estos datos ordenados a la librería que realiza las gráficas y muestra el resultado permitiendo descargar al usuario las gráficas generadas.
<b>Escenario alternativo</b>	El usuario decide guardar el archivo, pero

---

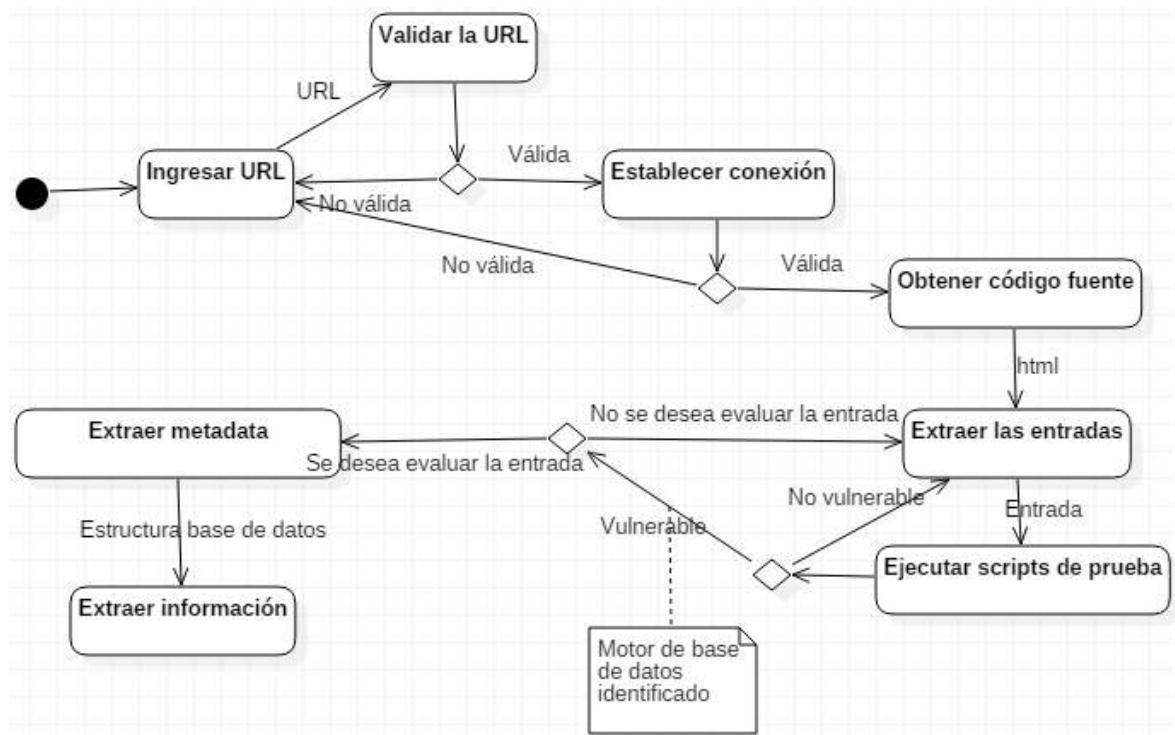
no graficarlo en ese momento.

*Fuente: Elaboración Propia*

### **3.4.4 Diagrama de actividades**

Para establecer las actividades del caso de uso se realiza el diagrama de actividades mostrado en el gráfico 30.

**Gráfico 30 Diagrama de actividades**



*Fuente: Elaboración Propia*

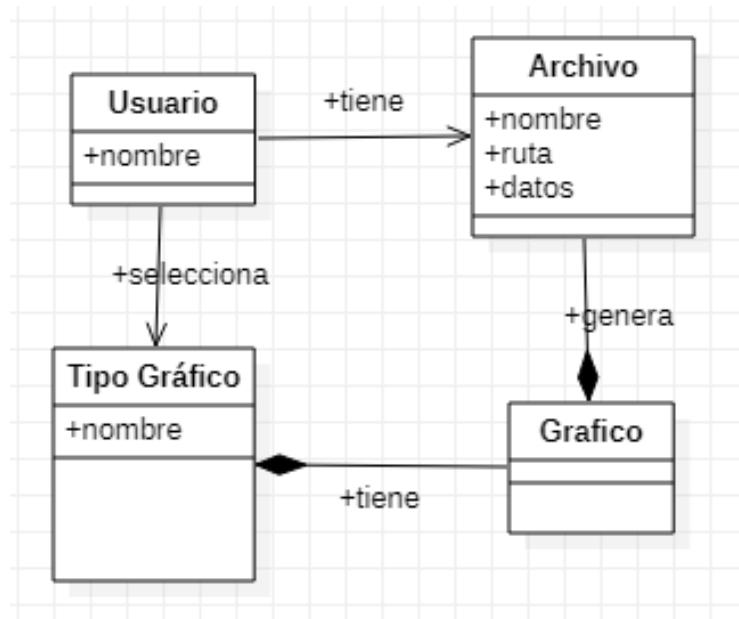
Con base a este diagrama de actividades, se procede a realizar el diagrama de clases.

### 3.4.5 Diagrama de clases

---

El diagrama de clases contiene una descripción de la arquitectura de la herramienta. Este diagrama se encuentra en el gráfico 31.

**Gráfico 31 Diagrama de clases**

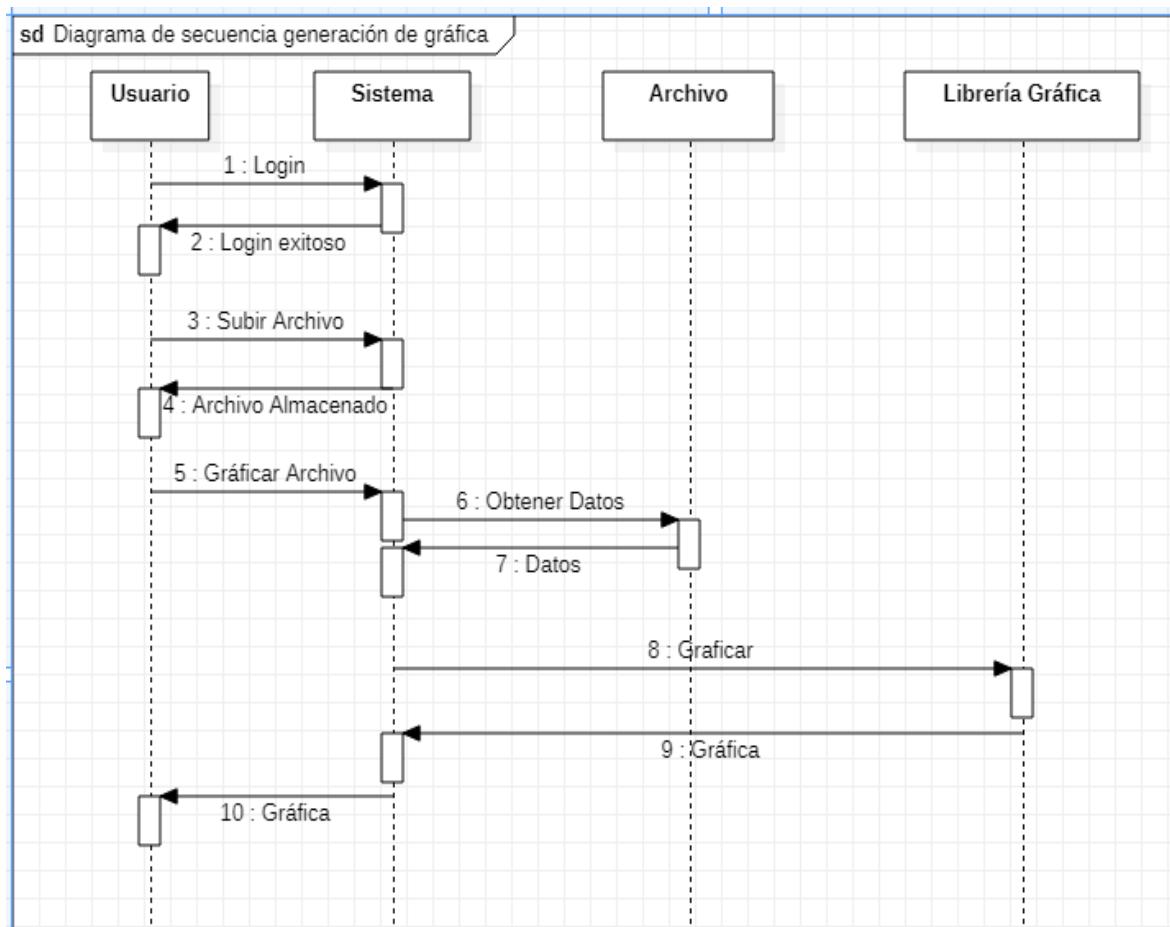


*Fuente: Elaboración propia*

### 3.4.6 Diagrama de secuencia

El diagrama de secuencia muestra la interacción interna entre los componentes. Este diagrama se muestra en el gráfico 32

**Gráfico 32 Diagrama de secuencia**



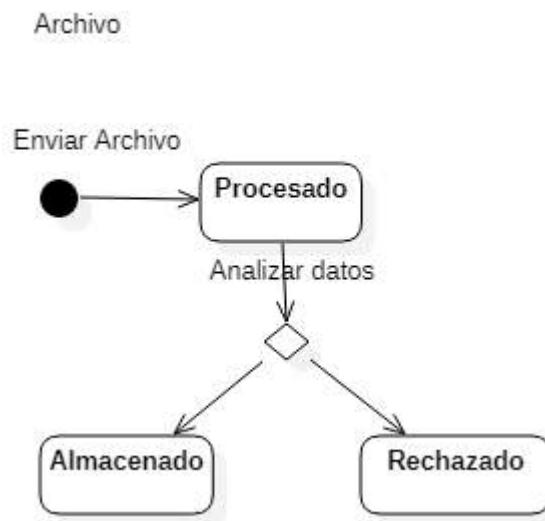
Fuente: Elaboración Propia

### 3.4.7 Diagrama de estados

---

Los archivos pueden ser almacenados o rechazados por el análisis de los datos almacenados. Este componente se encuentra en el gráfico 33.

**Gráfico 33 Diagrama de estados**

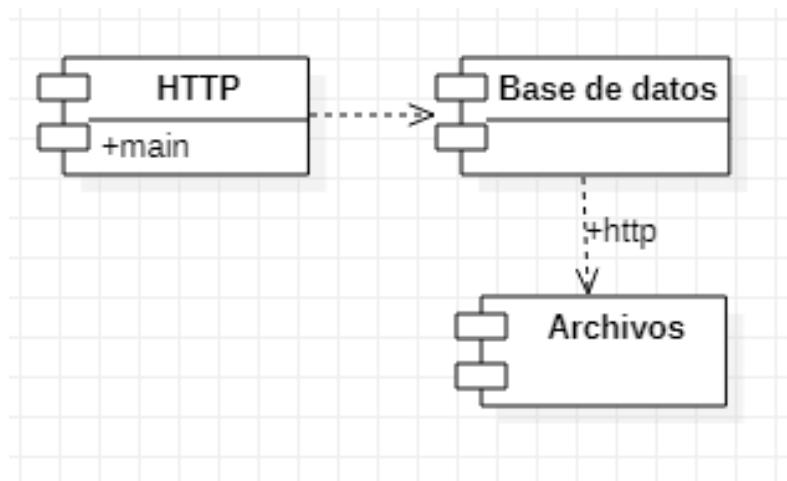


*Fuente: Elaboración propia*

### 3.4.8 Diagrama de componentes

El diagrama de componentes permite observar la conectividad del sistema. Se encuentra en el gráfico 34.

Gráfico 34 Diagrama de componentes



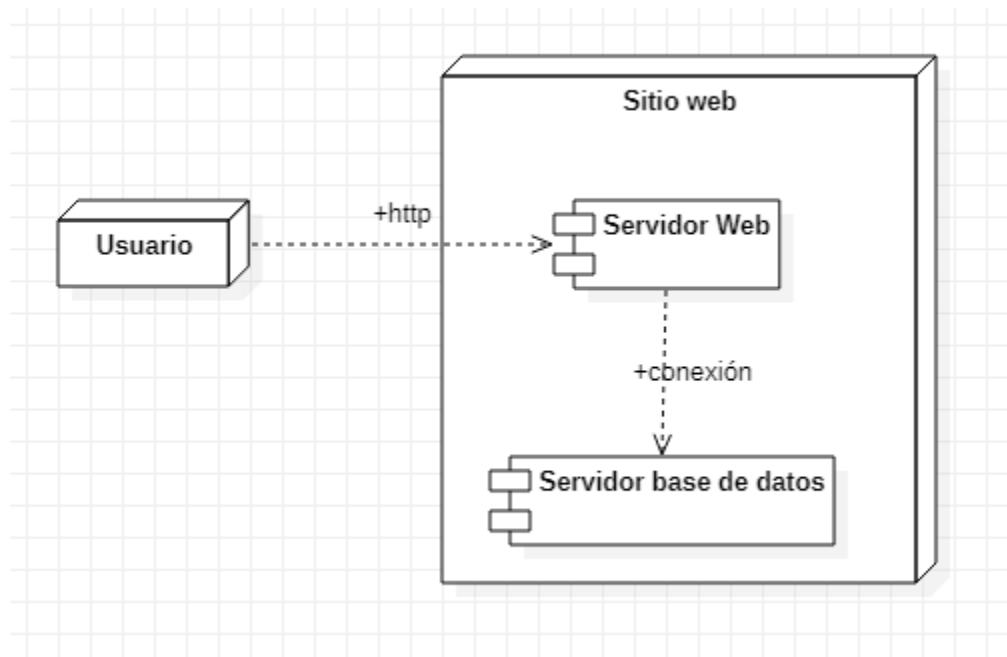
Fuente: Elaboración propia

---

### 3.4.9 Diagrama de despliegue

Este diagrama demuestra el despliegue del sistema de información. (Gráfico 35)

**Gráfico 35 Diagrama de despliegue**



*Fuente: Elaboración propia*

### **3.6 Características específicas de diseño**

El archivo está planteado para que sea en formato Excel con una terminación punto csv (.csv), el cual debe estar separado por comas, lo que permite que la lectura de los datos en el sistema sea más fácil de analizar y así poder generar las gráficas más fácilmente.

Para determinar las características de desarrollo de software, se requiere un entendimiento de lo que conforma cada elemento de desarrollo y cómo se utilizaron estas herramientas para llevar a cabo el objetivo final de la generación del proyecto. A continuación, se detallan algunos conceptos básicos utilizados en el desarrollo.

---

## 3.7 Manual de Usuario

### 3.7.1 Inicio de Sesión

En el gráfico 36, se muestra el pantallazo inicial, en el cual pueden ingresar los perfiles de tipo administrador y usuario.

**Gráfico 36 Inicio de Sesión**



La imagen es una captura de pantalla de un sitio web. En la parte superior, hay un logo de la Universidad Nacional de Colombia Sede Manizales y el texto "HERRAMIENTA INTERACTIVA PARA LA VISUALIZACION EN ANALISIS DE DATOS". Abajo de esto, se encuentra un cuadro de diálogo titulado "Acceso" que contiene campos para "E-Mail" y "Password", un checkbox para "Recordarme" y un botón "Login".

Fuente: Elaboración propia

### 3.7.2 Sesión Administrador

El administrador es el usuario que tiene el acceso al sistema principal del aplicativo, y es el que puede generar los permisos y los accesos a las personas que se quieren registrar en el sistema (Ver gráfico 37).

**Gráfico 37 Ingreso Administrador**



La imagen es una captura de pantalla similar a la del Gráfico 36, pero con datos prellenados. El campo "E-Mail" contiene "admin@gmail.com" y el campo "Password" contiene ".....". Los demás elementos del formulario son los mismos que en el Gráfico 36.

Fuente: Elaboración propia

Después de haber iniciado sesión, el administrador entra en su perfil de usuario como administrador y en el pantallazo inicial, aparece el listado de los usuarios que se han registrado en el sistema, la cual está conformada por.

- Nombre del Usuario Registrado
- Email de Usuario
- Fecha en que se registro
- El rol que posee el usuario
- Tipo de operación que quiere realizar, actualizar o eliminar.
- Creación de un nuevo usuario

**Gráfico 38 Administrador de Usuarios**

The screenshot shows a user management interface titled 'Administrador de Usuarios'. At the top right are 'Permisos' and 'Roles' buttons. Below is a table with columns: Nombre, Email, Fecha Registro, Usuario Roles, and Operaciones (Edit and Delete buttons). Three users are listed: Gonzalo Mauricio Zapata O., Nelly Ocampo Osorio, and Gonzalo Zapata. A 'Nuevo Usuario' button is at the bottom left, and 'Activar Windows' with a link to configuration is at the bottom right.

Nombre	Email	Fecha Registro	Usuario Roles	Operaciones
Gonzalo Mauricio Zapata O.	admin@gmail.com	September 07, 2020	Administrador	<a href="#">Editar</a> <a href="#">Eliminar</a>
Nelly Ocampo Osorio	cliente@gmail.com	September 08, 2020		<a href="#">Editar</a> <a href="#">Eliminar</a>
Gonzalo Zapata	gonzalemauricio@man.com	March 23, 2021		<a href="#">Editar</a> <a href="#">Eliminar</a>

Fuente: Elaboración propia

Después de mostrar la lista de usuarios, el perfil de administrador detalla sus capacidades, ya que es el que genera el registro de usuarios, el cual tiene los campos de nombre, email, perfil, contraseña, entre otros (Ver gráfico 38).

---

### **Gráfico 39 Registro de Usuario**

---

 Nuevo Usuario

Nombre

Email

Administrador

Password

Confirmar Password

Fuente: Elaboración propia

Después de la creación del nuevo usuario, el administrador tiene la opción de crear los permisos y roles, cuando crea los permisos es dirigido al enlace de permisos, en el cual hay una lista de los permisos creados en el sistema. La creación de un nuevo permiso tiene el mismo mecanismo que la creación de usuario, en el que se puede crear un permiso, editarlo y eliminarlo. Como se muestra en el gráfico 40. El administrador tiene la facultad de crear los roles que va a tener cada usuario.

---

### **Gráfico 40 Lista de Permisos**

---

 Available Permissions

Permissions	Operation
Administrador	<input type="button" value="Edit"/> <input type="button" value="Delete"/>
Administrador de Roles y Permisos	<input type="button" value="Edit"/> <input type="button" value="Delete"/>

Activar Windows

Fuente: Elaboración propia

### 3.6.2 Sesión Usuario

Como se dijo anteriormente en el administrador del sistema, para poder ingresar como usuario en el sistema se tiene que iniciar sesión. Como se muestra en el gráfico 41.

**Gráfico 41 Inicio de sesión como usuario**



Fuente: Elaboración propia

Después de iniciar sesión el usuario registrado entra en su perfil, el cual muestra el menú de navegación y el título Archivos subidos por el usuario. Ya que es la primera vez que entra en el sistema, no debe tener ningún archivo subido. Como se muestra en el gráfico 42.

**Gráfico 42 Sesión usuario primera vez**



ARCHIVOS SUBIDOS POR EL USUARIO

Fuente: Elaboración propia

---

Ya dentro del sistema el usuario puede hacer uso de este y recorrer los menús de navegación. Los cuales le permiten interactuar con el sistema. Como se muestra en el gráfico 43.

**Gráfico 43 Menú de Navegación**



Fuente: Elaboración propia

El menú de navegación Subir Archivo, dirige a la sección para subir el archivo. Como se muestra en el gráfico 44.

**Gráfico 44 Web para subir el archivo**

A screenshot of a web form titled "SUBIR ARCHIVO CSV". It has two input fields: "Nombre Del Archivo:" (File Name:) with a placeholder "Nombre Archivo" and "Seleccione Un Archivo:" (Select File) with a placeholder "Seleccionar archivo Ningún archivo seleccionado". At the bottom are two buttons: "Guardar Archivo" (Save File) and "Activar Wir" (Activate Wir).

Fuente: Elaboración propia.

Como se ve en el gráfico 44, se observa un formulario para subir el archivo CSV, que se quiere analizar, la primera opción del formulario es la de darle un nombre al archivo que se quiere subir, la segunda opción es la de seleccionar el archivo que se quiere subir y, por último, está la opción de guardar archivo que es la que permite realizar dicho proceso.

Después de realizar la alternativa de Guardar Archivo, devuelve al inicio de sesión en donde aparece el archivo subido, como se muestra en el gráfico 45.

**Gráfico 45 Archivo subido por el usuario**

The screenshot shows a web interface for data visualization. At the top, the logo of the Universidad Nacional de Colombia Sede Manizales is displayed. Below it, the text "HERRAMIENTA INTERACTIVA PARA LA VISUALIZACION EN ANALISIS DE DATOS" is centered. A user profile "Nelly Ocampo Osorio" is visible on the right. Below this, a navigation bar includes "Inicio", "Subir Archivo", and "Opciones De Gráficas". The main content area is titled "ARCHIVOS SUBIDOS POR EL USUARIO". A table lists the uploaded file "Ingresos" (with a CSV icon) by user "Nelly Ocampo Osorio". The table has columns: Archivo, Nombre Archivo, Nombre Usuario, Eliminar Archivo, and Graficar. The "Eliminar" button is red, and the "Graficar" button is blue.

Archivo	Nombre Archivo	Nombre Usuario	Eliminar Archivo	Graficar
Ingresos IngresosIZIRCeBvBgQZo8yTTNJ8.csv	Ingresos	Nelly Ocampo Osorio	<a href="#">Eliminar</a>	<a href="#">Graficar</a>

Fuente: Elaboración propia.

Ya teniendo el archivo subido, el usuario tiene la opción de graficar el archivo que tiene subido o ver las opciones que tiene para hacer la gráfica.

Si tomo la decisión de ver las opciones de gráfica, se muestra una lista con las gráficas que puede ver para así tomar la mejor decisión de que grafica quiere hacer. Como se muestra en el gráfico 46.

## Gráfico 46 Lista de imágenes

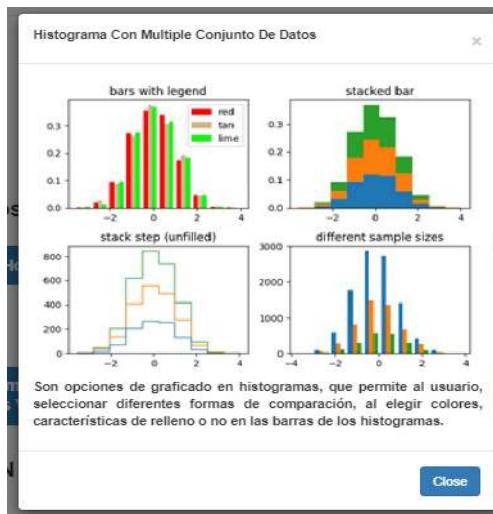


Fuente: Elaboración propia

Al ver la lista de imágenes que tiene el sistema, el usuario puede hacer clic en cualquiera de las opciones, lo que le mostrará una imagen de la gráfica que selecciono y una explicación de esta.

Para observar el funcionamiento, se toma la opción de Histograma Con Múltiple Conjunto de Datos, lo cual muestra la imagen y la explicación de la imagen. Como se muestra en el gráfico 47.

### **Gráfico 47 Ejemplo Histograma Con Múltiple Conjunto de Datos**



Fuente: Elaboración propia

Así el usuario tiene una perspectiva más amplia de lo que puede o no hacer con el archivo que subió. Es en este instante que ya puede tomar la determinación de hacer la gráfica con respecto al archivo que subió y los datos que tiene el mismo.

Para llevar a cabo el análisis de los datos que subió el usuario, nos devolvemos al inicio, que es donde se encuentra el archivo ya subido. En este punto el usuario puede tomar la decisión si quiere realizar el análisis del archivo o eliminarlo. Como se muestra en el gráfico 48.

### **Gráfico 48 Archivo para graficar o eliminar**

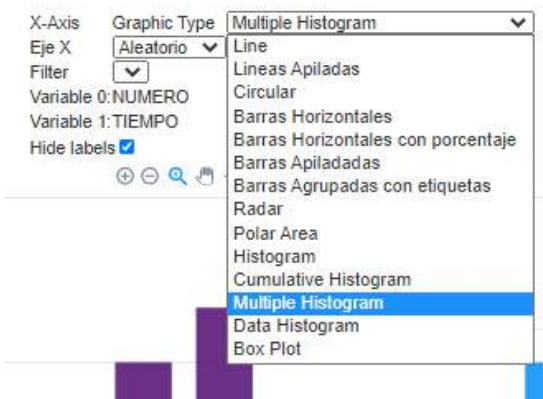
	EjemploNestorHoja2	Nelly Ocampo Osorio	<b>Eliminar</b>	<b>Graficar</b>
EjemploNestorHoja2gXOC0kS3km279u6V0.csv				

Fuente: Elaboración propia

---

Si se da clic en el botón de graficar, se toman los datos y se muestran las gráficas que se pueden realizar con los datos que tiene el archivo seleccionado. Como se muestra en el gráfico 49.

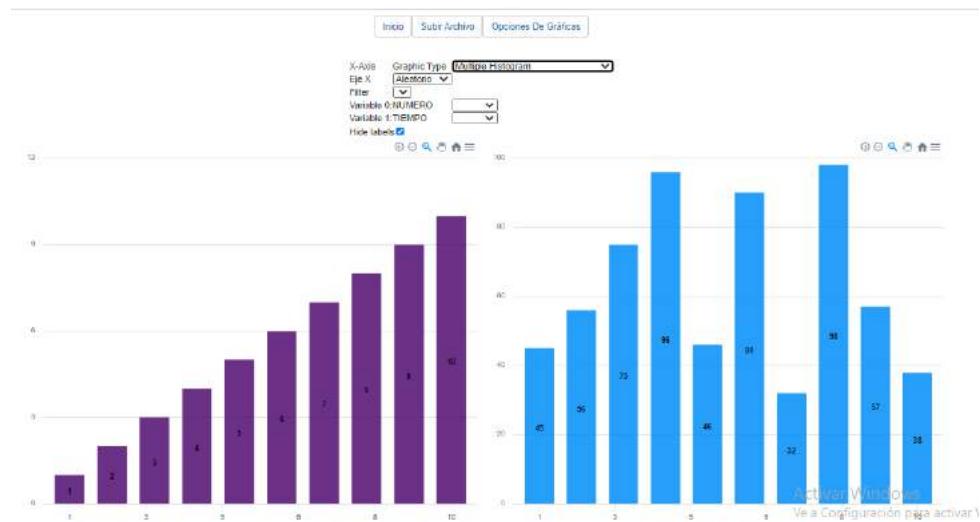
### Gráfico 49 Opciones de las gráficas que puede realizar el sistema



Fuente: Elaboración propia

Al estar ya en la sección en que se puede realizar la gráfica, se tiene un menú desplegable, en el cual, se muestran todas las grafica que puede realizar el usuario con el archivo seleccionado. Después de seleccionar la opción más adecuada para la realización de la gráfica se da clic en la opción que quiere graficar el usuario y se genera la gráfica. Como se muestra en el gráfico 50. En la cual se escoge Múltiple Histograma.

## Gráfico 50 Resultado de la gráfica Múltiple Histograma



Fuente: Elaboración propia

Al ver el resultado de la selección de la gráfica, puede descargarla y guardarla en el equipo. Tal como se muestra en el gráfico 51.

## Gráfico 51 Opción de descarga de imagen



Fuente: Elaboración propia

Después de tomar la decisión en que formato de descarga, se da clic en la opción y se descarga. Como se muestra en el gráfico 52. La opción de descarga es png.

---

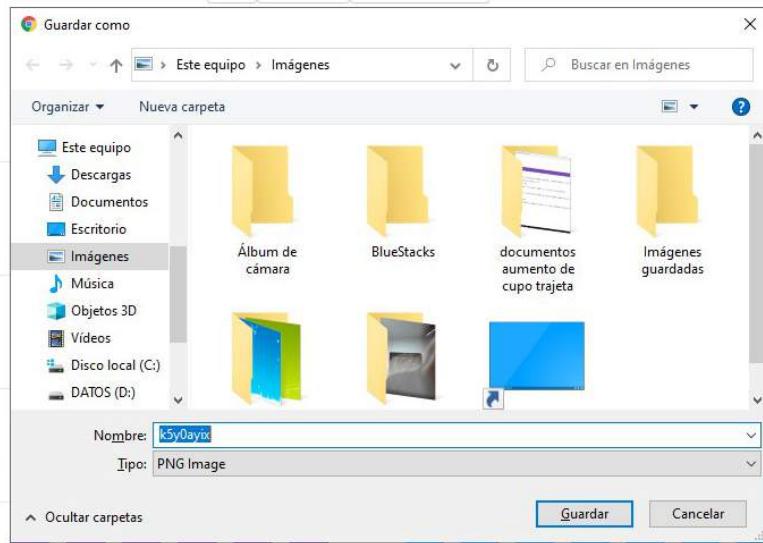
**Gráfico 52 Descarga de la imagen en formato .png**



Fuente: Elaboración propia

Al dar clic en la opción de descarga de formato .png se abre una venta emergente, que indica el nombre que tiene la imagen por defecto y el destino donde quiere que se guarde la imagen. Como se muestra en el gráfico 53.

**Gráfico 53 Venta emergente en donde se muestra la ubicación en donde se almacenar la imagen**



Fuente: Elaboración propia

## 4 VALIDACIÓN FUNCIONAL DE LA HERRAMIENTA

Para realizar la validación funcional de la herramienta se seleccionó una muestra de 37 (el 60% de los usuarios eran hombres y el 40% eran mujeres) usuarios y entre 20 y 70 años, sin conocimientos de lenguaje de programación y se les solicitó diligenciar la encuesta de evaluación de la aplicación (Ver Anexo). A continuación, se realizará un breve análisis de los resultados obtenidos.

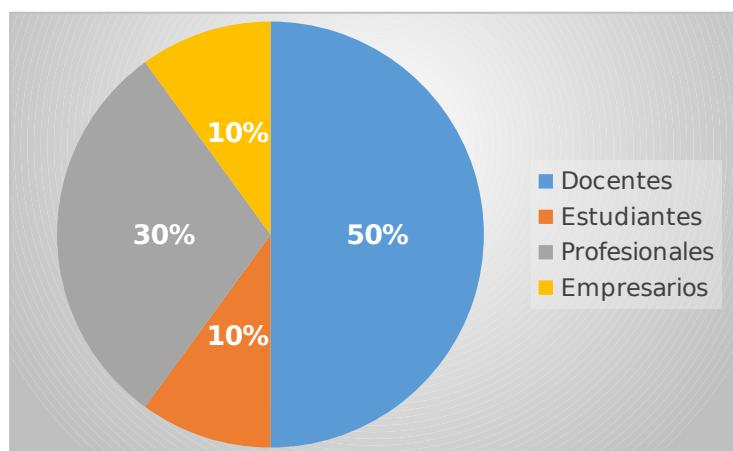
### 4.1 Caracterización de la Muestra

La encuesta se realizó a un grupo de usuarios con las siguientes características:

**Definición de la muestra:**

a.i El 40% de la población muestral era mujeres, donde el 50% se encontraban dentro del rango de edad de 40-60 años, y el 60% hombres con participaciones iguales según rango de edad (Ver gráfico 54);

**Gráfico 54 Análisis de la Encuesta por Sexo**

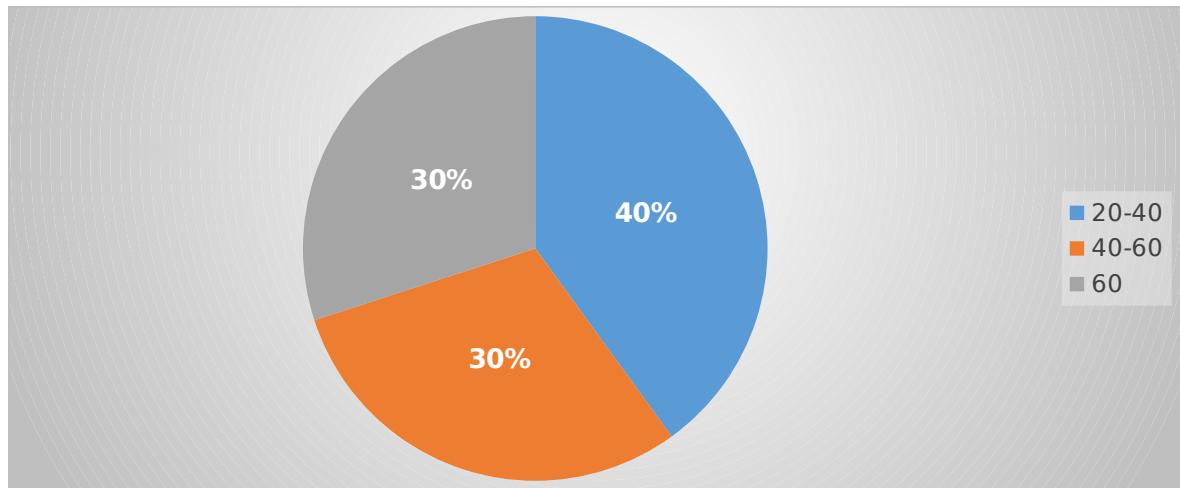


Fuente: Elaboración propia

---

a.ii Los rangos de edad de los encuestados giraban en torno a los 20 y 70 años, en este orden de ideas, el 40% estaba entre los 20-40 años, el 30% entre los 40-60 y finalmente otro 30% para los mayores de 60 años (Ver gráfico 55);

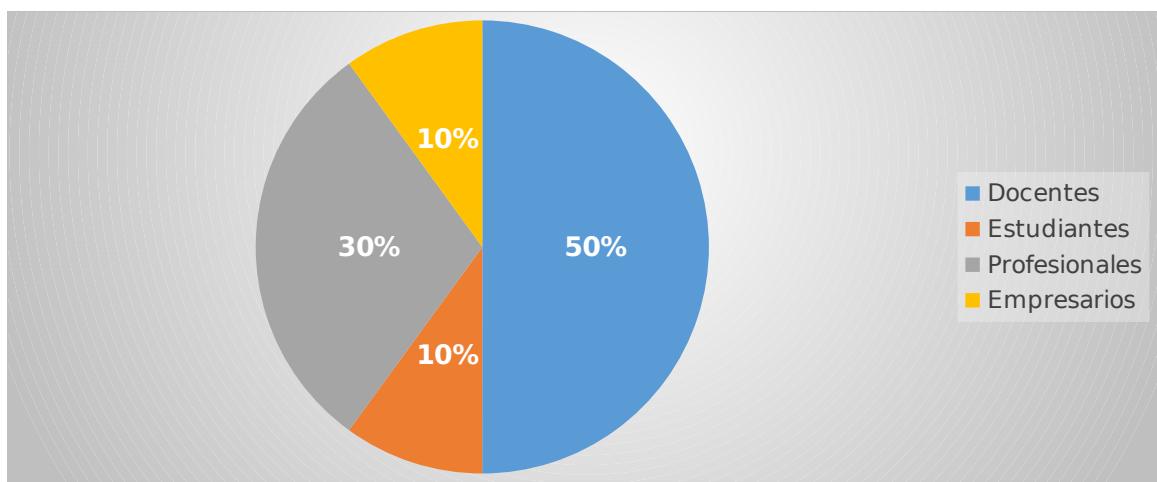
**Gráfico 55 Análisis de la Encuesta por Edad**



Fuente: Elaboración propia

a.iii El 50% eran docentes de diferentes áreas del saber, el 30% eran profesionales, 10% eran Estudiantes y Empresarios (Ver gráfico 56).

**Gráfico 56 Análisis por profesión de la muestra**



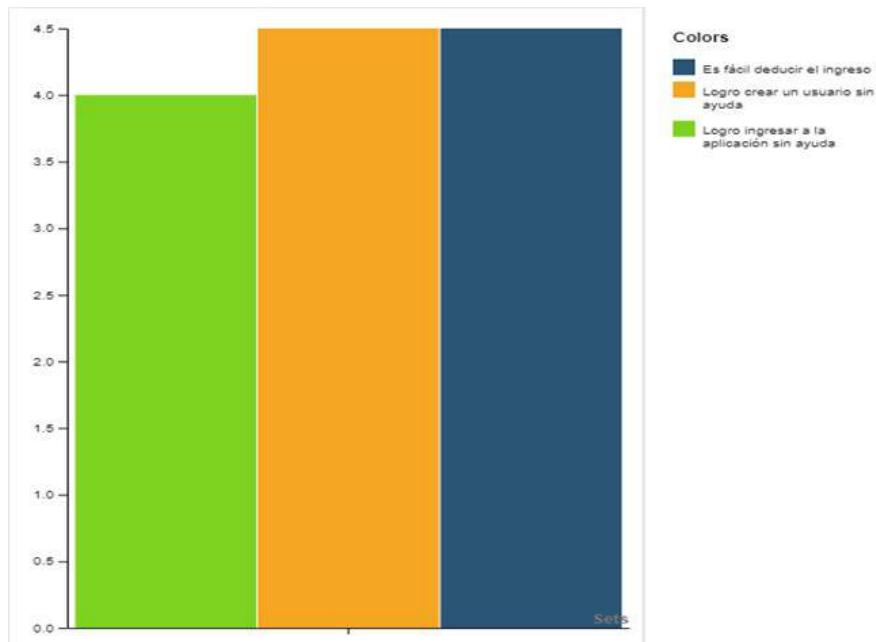
Fuente: Elaboración propia

**Recolección de datos y Análisis de Resultados:** A partir de la muestra obtenida, se realizó una revisión de cada una de las encuestas, revisando los datos generales y los diferentes puntajes seleccionados según la temática. Posteriormente se agruparon de acuerdo con cada área. La información obtenida se tabuló y analizó estadísticamente, mediante la utilización de cuadros con porcentajes.

## 4.2 Resultados de la Muestra

En la sección que evalúa el pilar de la facilidad para ingresar en la aplicación (siendo 5 excelente y 1 muy deficiente) se observó que la media de los usuarios calificó que la aplicación refleja una facilidad de ingreso, sin embargo, para algunos usuarios presentaron inconvenientes con la sección del Administrador. En este sentido, la mayoría de los usuarios lograron crear un usuario e ingresar sin ayuda (Ver gráfica 57).

**Gráfico 57 Resultados del Pilar Facilidad del Ingreso**

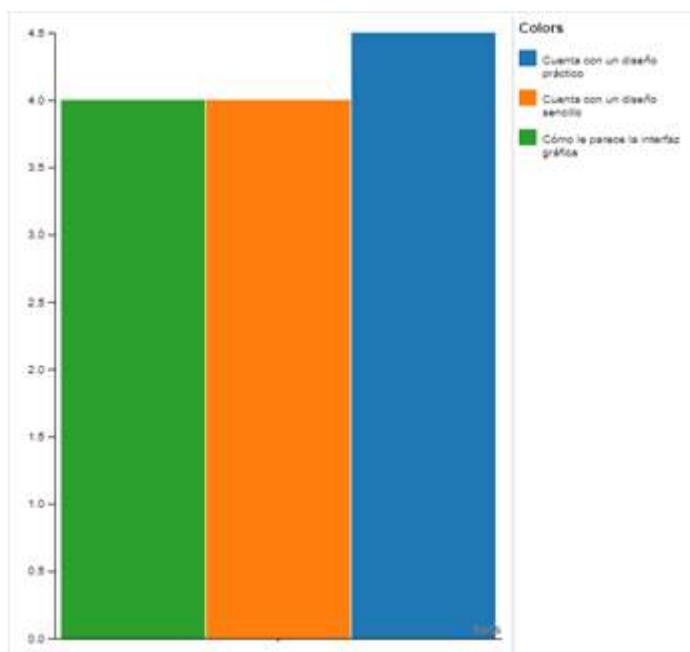


---

Fuente: Elaboración propia

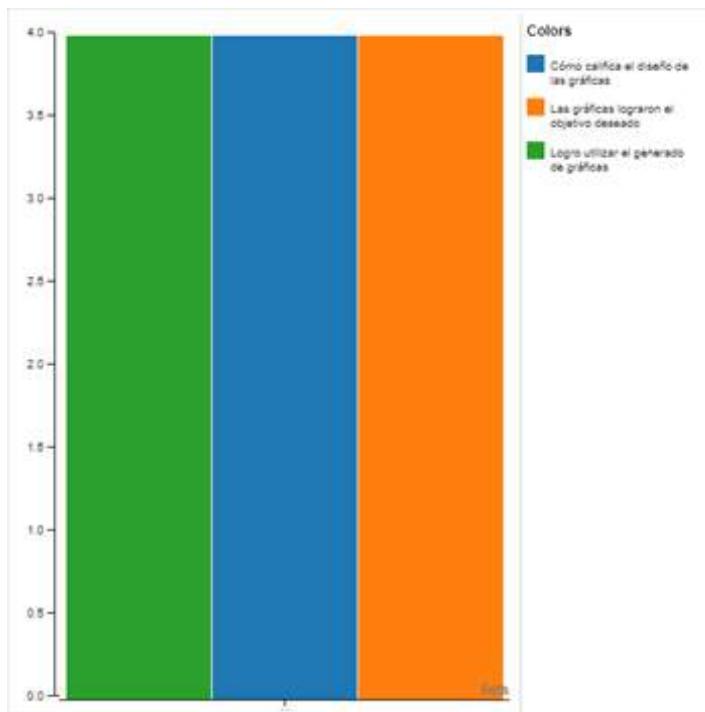
En la sección que evalúa el diseño de la aplicación (siendo 5 excelente y 1 muy deficiente) se observó que de los usuarios lograron crear una gráfica y les gustó la variedad de visualizaciones disponibles. Encontraron que el diseño era amigable y práctico, sin aglutinar información para el usuario (Ver gráfico 58). Sin embargo, se encontró usuarios que no lograron diseñar una visualización.

**Gráfico 58 Resultados del Pilar Diseño de la Aplicación**



Fuente: Elaboración propia

**Gráfico 59 Resultados del Pilar Uso de las Visualizaciones**

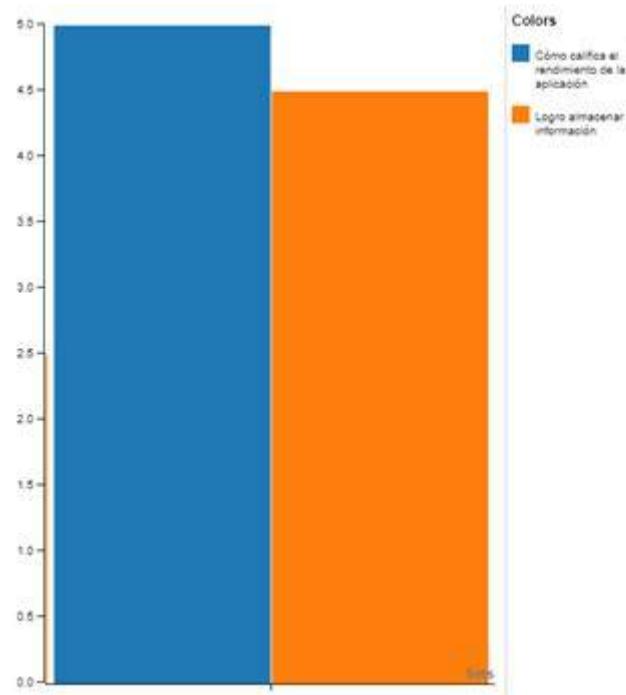


Fuente: Elaboración propia

Para evaluar el rendimiento se realizaron dos preguntas relacionadas con el rendimiento y el almacenamiento (siendo 5 excelente y 1 muy deficiente) se observó que de los usuarios identificaron que la aplicación cumple con sus expectativas de rendimiento y velocidad, así como identificaron la capacidad de almacenamiento como un valor agregado de la aplicación, en particular, para futuros análisis (Ver gráfico 60).

---

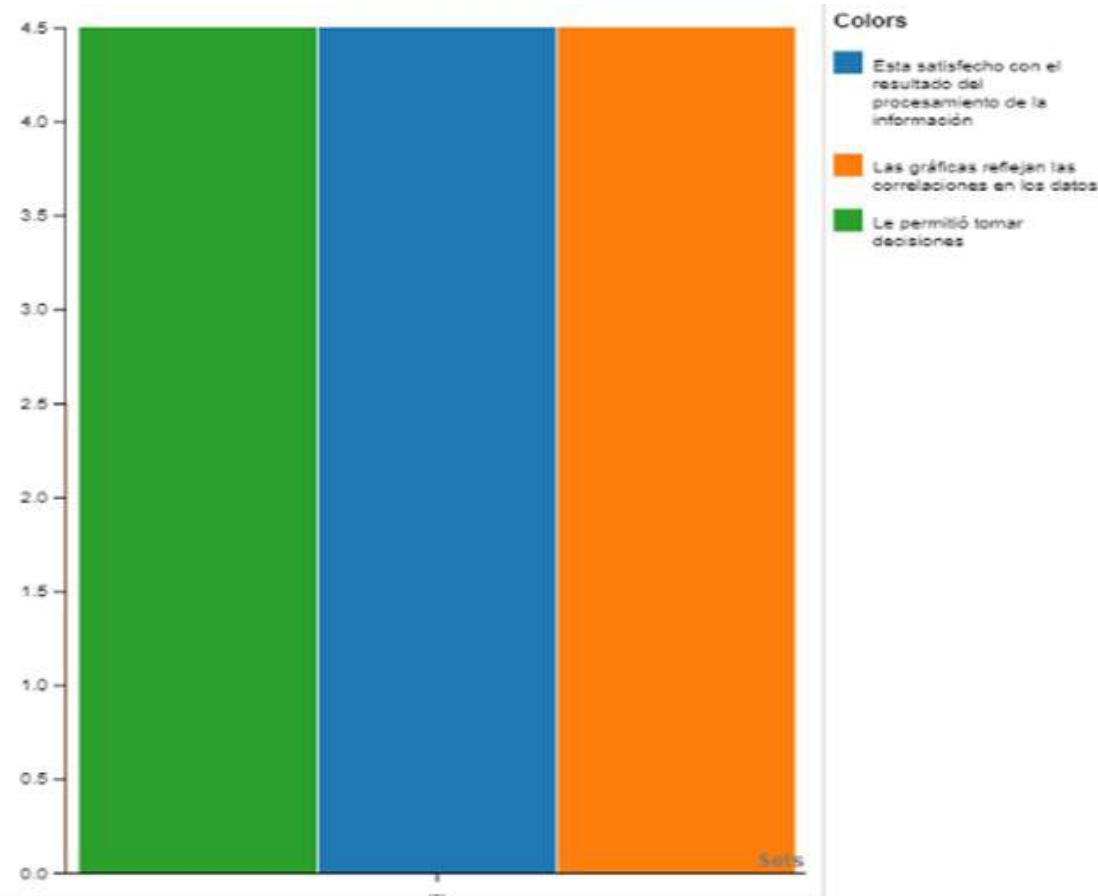
### Gráfico 60 Pilar Rendimiento



Fuente: Elaboración propia

Para evaluar la satisfacción del usuario se incluyeron tres preguntas relacionadas, principalmente, con la utilidad para tomar decisiones en base de las correlaciones reflejadas en las visualizaciones y del resultado del procesamiento de los datos por parte de la aplicación. De nuevo, los resultados reflejan la gran utilidad cumpliendo uno de los objetivos propuestos en este trabajo (Ver gráfico 60).

**Gráfico 61 Pilar Satisfacción del Usuario**



Fuente: Elaboración propia

Para evaluar la satisfacción del usuario se incluyeron tres preguntas relacionadas, principalmente, con la utilidad para tomar decisiones en base de las correlaciones reflejadas en las visualizaciones y del resultado del procesamiento de los datos por parte de la aplicación. De nuevo, los resultados reflejan la gran utilidad cumpliendo uno de los objetivos propuestos en este trabajo (Ver gráfico 61).

---

## 5 CONCLUSIONES

Con esta investigación se evidencia la importancia del estudio de este tipo de herramientas de visualización en el país, pues no solo permiten el procesamiento masivo de información, sino que contribuyen a un análisis profundo de relaciones entre variables y patrones de comportamiento, lo que es de gran utilidad en el sector empresarial y microempresarial en Colombia. En especial, porque la concentración de software diseñado para suplir esta necesidad es de pago y requieren conocimientos avanzados de programación.

Como se puede ver en la revisión de herramientas realizada, las aplicaciones actuales para la visualización y análisis de datos son de pago o son enfocadas en personas con altos conocimientos técnicos. El valor agregado de esta aplicación es que no tiene costo y es sencilla para que la puedan usar los usuarios sin conocimiento técnicos avanzados.

En las encuestas realizadas, se destaca por la mayoría de los usuarios que el uso de la herramienta es simple y les genera el valor agregado que estaban buscando. El hecho de realizar la herramienta en el Framework Laravel permite tener escalabilidad y añade una capa de seguridad que se traduce en estabilidad para la aplicación. Además, el uso de la librería escogida para graficar, ApexChart, permitió realizar el cambio de gráficas de una manera rápida y sin realizar un refresco de la página, facilitando el uso y la comparación de las gráficas por parte de los usuarios. Los gráficos aportados por esta librería son suficientes para realizar una visualización avanzada de la información. Los gráficos radar y polar permitieron visualizar información que

otros tipos de gráficos no mostraban, facilitando la identificación de los puntos donde se producía una desviación significativa.

Igualmente, los usuarios destacaron la utilidad de la aplicación que permite guardar los archivos para posterior uso. Esto permite realizar trabajos a largo plazo y añade persistencia a la aplicación.

Una de las dificultades que se encontró es la mala calidad de los datos que suben los usuarios, debido a que mucha de esta información proviene de encuestas realizadas y quienes las diligenciaron seleccionaban información con otra codificación o sin el formato apropiado. Por lo tanto, es necesario realizar como trabajo futuro un validador de datos que depure la información de manera automática sin alterar el resto que está correctamente especificada.

En este orden de ideas, con el aplicativo desarrollado se puede identificar posibles mejoras que permitan la optimización y el análisis de datos en 3D, pues es una tendencia mundial en el análisis y visualización de datos. Sin embargo, este no era el objetivo inicial de esta investigación, por lo que se propone para estudios futuros dedicados específicamente en esta área del desarrollo y minería de datos.

---

## 6 BIBLIOGRAFÍA

- Adams, C. (2014). *Learning Python Data Visualization*. UK: Packt Publishing Ltd.
- Beran, B., Van Ingen, C., Zaslavsky, I., & Valentine, D. (2008). OLAP Cube Virtualization of Environmental Data Catalogs. *San Diego Supercomputer Center*.
- Chen , L., & Zhou, H. (2016). Research and Application of Dynamic and Interactive data Visualization based on D3. *IEEE*.
- Grotten , S., Krone , M., Muller, C., Reina, G., & Ertl, T. (2015). MegaPol-A prototyping Framework for Particle-Based Visualization. *IEEE Vol. 21 No. 2*.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques*. USA: Morgan Kaufmann - Elsevier .
- Henschen, D. (2018). A look at Tableau´s intelligent data prep, discovery, recommendation and natural language query capabilities. *Constellation Research*.
- Khlebnikov, R., Kainz, B., Steinberger, M., & Schmalstieng, D. (2013). Noise-Based Volume Rendering for the Visualization of Multivariate Volumetric Data. *IEEE Vol. 19 No. 12*.
- LARAVEL. (15 de 04 de 2021). *LARAVEL*. Obtenido de <https://laravel.com/>
- Liu, J., & Chen, Y. (2012). Fast Data Analysis with integrated statistical metadata in Scientific Datasets.
- Liu, J., & Chen, Y. (2012). Improving Data Analysis Performance for High-Performance Computing with Integrating Statistical Metadata in Scientific Datasets. *High Performance Computing, Networking Storage and Analysis*.

- Ma, H., & Jones, K. W. (2005). Scientific Visualization and Data Modeling of Scattered Sediment Contaminant in New York/New Jersey Estuaries.
- MINITAB. (17 de 04 de 2021). *MINITAB*. Obtenido de Minitab® Statistical Software: <https://www.minitab.com/es-mx/products/minitab/>
- Murray, S. (2017). *Interactive Data Visualization for the Web*. United States: O'Reilly Media Inc.
- Nouanesengsy, B., Woodring, J., Patchett, J., Myers, K., & Ahrens, J. (2014). ADR Visualization: A generalized Framework for Ranking Large-Scale Scientific Data using Analysis-Driven Refinement. *IEEE*.
- Orange. (17 de 04 de 2021). Obtenido de <https://orangedatamining.com/>
- Otoo, E. J., & Rotem, D. (2006). Efficient storage allocation of Large-Scale Extendible Multi-dimentional Scientific Datasets. *University of California*.
- Qiang , X., Wei , Y., & Hanfei, Z. (2010). Application of Visualization Tecnology in Spatial Data Mining. *Computing, Control and Industrial Engineering*.
- RAPIDMINER. (17 de 4 de 2021). Obtenido de <https://rapidminer.com/>
- Riasetiawan, M., & Mahmood, A. K. (2010). Science-Forge: A collaborative Scientific Framework. *IEEE*.
- Rossant, C. (2015). *Learning IPython for Interactive Computing and Data Visualization* . UK: Packt Publishing Ltd.
- Rubel, O., Weber, G. H., Huang, M.-Y., Bethel, E., Biggin, M. D., Fowlkes, C. C., . . . Hamann, B. (2010). Integrating Data Clustering and Visualization for the Analysis of 3D Gene Expression Data. *IEEE Vol. 7 No. 1*.
- Salmre, I. (16 de 04 de 2021). *Tableau: Overview and Proof Points*. Obtenido de [https://www.tableau.com/sites/default/files/whitepapers/tableau\\_online\\_scalability\\_wp\\_0.pdf](https://www.tableau.com/sites/default/files/whitepapers/tableau_online_scalability_wp_0.pdf)
- SAS. (15 de 04 de 2021). *Documentación* . Obtenido de <https://support.sas.com/en/documentation.lang.html>

- 
- SAS. (15 de 4 de 2021). *SAS: Analytics, Artificial Intelligence and Data Management*. Obtenido de [https://www.sas.com/en\\_us/home.html](https://www.sas.com/en_us/home.html)
- Schulz, H.-J., Angelini, M., Santucci, G., & Schumann, H. (2016). An Enhanced Visualization Process Model for Incremental Visualization. *IEEE Vol. 22 No. 7.*
- Simmhan, Y., Barga, R., Van Ingen, C., Lazowska, E., & Szalay, A. (2008). On building scientific workflow system for Data Management in Cloud. *IEEE.*
- Smeulders, R., & Heijs, A. (2005). Interactive Visualization of High Dimentional Marketing Data in the Financial Industry. *IEEE .*
- TABLEAU. (16 de 04 de 2021). *TABLEAU SOFTWARE, LLC, A SALESFORCE COMPAN*. Obtenido de <https://www.tableau.com/>
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining*. United States: Pearson Education Inc.
- Viktor, H., & Paquet, E. (2006). Visualization Techniques for Data Mining. *NRC Publications*, 1190-1195.
- Westrip, S. P. (2010). pubICIF: Software for editing, validating and formating crystallographic information files. *Applied Crystallography*.
- Yan, H., & Wang, J. (2017). Research and Application of the Test Data Visualization. *Data Science in Cyberspace*.
- Yang, E., Ross, D., Nagella, S., & Turner, M. (2015). Data optimised computing for heterogeneous Big Data Computing Applications. *IEEE.*

## ANEXOS

Diseño de la Encuesta de Evaluación de la Aplicación a usuarios sin conocimientos en programación:

# CUESTIONARIO

<b>Nombre y apellidos:</b> <hr/>		
<b>Edad:</b> <hr/>	<b>Sexo:</b> <hr/>	<b>Profesión:</b> <hr/>
<b>Fecha de Uso:</b> ___ - ___ - ___	<b>Fecha de encuesta:</b> ___ - ___ - ___	<b>Calificación:</b> <hr/>

Marque con una X la puntuación que considere más acorde (1 muy deficiente,  
5 excelente)

INGRESO					
	1	2	3	4	5
<b>Logro Ingresar a la aplicación sin ayuda</b>					
<b>Logro crear un usuario sin ayuda</b>					
<b>Es fácil de deducir los pasos de ingreso</b>					
<b>Comentarios</b>					

DISEÑO		1	2	3	4	5
¿Cómo le parece la Interfaz gráfica?						
¿Cuenta con un diseño sencillo?						
¿Cuenta con un diseño práctico?						
<b>Comentario</b>						
S						

USO DE LAS GRÁFICAS		1	2	3	4	5
¿Logro utilizar el generador de gráficas?						
¿Cómo califica el diseño de las gráficas?						
En su opinión, ¿las gráficas lograron el objetivo deseado?						
<b>Comentario</b>						
S						

RENDIMIENTO		1	2	3	4	5
¿Como califica el rendimiento de la aplicación?						
¿Logro almacenar información?						
<b>Comentario</b>						
S						

---

## ANÁLISIS DE DATOS

	1	2	3	4	5
¿Está satisfecho con el resultado del procesamiento de la información?					
¿Las gráficas reflejan las correlaciones en los datos?					
¿Le permitió tomar decisiones?					
Comentarios					

¿Había utilizado otra herramienta de visualización de datos?

¿En cuál?

¿Cómo calificaría nuestra aplicación de forma global?

