



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

1 de 2

Neiva, 7 de febrero de 2023

Señores

CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN

UNIVERSIDAD SURCOLOMBIANA

Ciudad Neiva

El (Los) suscrito(s):

Lisseth Tatiana López Méndez, con C.C. No. 1003803377,

Jennifer Torres Sánchez, con C.C. No. 1007677701,

Mercy Lili Peña Morales con C.C. No. 55165585.

Autor(es) de la tesis y/o investigación del semillero Investigando y Reflexionando en Investigación Matemática. Semillero (INREMAT).

Titulado Analisis de datos de calidad agua usando la herramienta Excel

presentado y aprobado en el año 2023 como requisito para optar al título de

Licenciados en Matemáticos;

Autorizo (amos) al CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN de la Universidad Surcolombiana para que, con fines académicos, muestre al país y el exterior la producción intelectual de la Universidad Surcolombiana, a través de la visibilidad de su contenido de la siguiente manera:

- Los usuarios puedan consultar el contenido de este trabajo de grado en los sitios web que administra la Universidad, en bases de datos, repositorio digital, catálogos y en otros sitios web, redes y sistemas de información nacionales e internacionales “open access” y en las redes de información con las cuales tenga convenio la Institución.
- Permita la consulta, la reproducción y préstamo a los usuarios interesados en el contenido de este trabajo, para todos los usos que tengan finalidad académica, ya sea en formato Cd-Rom o digital desde internet, intranet, etc., y en general para cualquier formato conocido o por conocer, dentro de los términos establecidos en la Ley 23 de 1982, Ley 44 de 1993, Decisión Andina 351 de 1993, Decreto 460 de 1995 y demás normas generales sobre la materia.
- Continúo conservando los correspondientes derechos sin modificación o restricción alguna; puesto que, de acuerdo con la legislación colombiana aplicable, el presente es un acuerdo jurídico que en ningún caso conlleva la enajenación del derecho de autor y sus conexos.

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional [www.usco.edu.co](http://www.usco.edu.co), link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

2 de 2

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

EL AUTOR/ESTUDIANTE:

Firma:

LISETH TATIANA LOPEZ MENDEZ  
ESTUDIANTE DE LA LIC. EN MATEMATICAS

EL AUTOR/ESTUDIANTE:

Firma:

JENNIFER TORRES SANCHEZ  
ESTUDIANTE DE LA LIC. EN MATEMATICAS

Firma:

MERCY LILI PEÑA MORALES  
DIRECTORA DEL SEMILLERO INREMAT



**TÍTULO COMPLETO DEL TRABAJO:** “Análisis de datos de calidad de agua usando la herramienta Excel”

**AUTOR O AUTORES:**

Primero y Segundo Apellido	Primero y Segundo Nombre
López Méndez	Lisseth Tatiana
Torres Sánchez	Jennifer

**DIRECTOR Y CODIRECTOR SEMILLERO:**

Primero y Segundo Apellido	Primero y Segundo Nombre
Peña Morales	Mercy Lili

**ASESOR (ES):**

Primero y Segundo Apellido	Primero y Segundo Nombre
NO APLICA	

**PARA OPTAR AL TÍTULO DE:** Licenciados en Matemáticas

**FACULTAD:** Educación

**PROGRAMA O POSGRADO:** Licenciatura en Matemáticas

**CIUDAD:** Neiva

**AÑO DE PRESENTACIÓN:** 2023

**NÚMERO DE PÁGINAS:** 59

**TIPO DE ILUSTRACIONES** (Marcar con una X):

Diagramas\_\_\_ Fotografías\_x\_ Grabaciones en discos\_\_\_ Ilustraciones en general\_x\_ Grabados\_\_\_  
Láminas\_\_\_ Litografías\_\_\_ Mapas\_x\_ Música impresa\_\_\_ Planos\_\_\_ Retratos\_\_\_ Sin ilustraciones\_\_\_  
Tablas o Cuadros\_x\_

**SOFTWARE** requerido y/o especializado para la lectura del documento: Ninguno

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional [www.usco.edu.co](http://www.usco.edu.co), link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



<b>CÓDIGO</b>	<b>AP-BIB-FO-07</b>	<b>VERSIÓN</b>	<b>1</b>	<b>VIGENCIA</b>	<b>2014</b>	<b>PÁGINA</b>	<b>2 de 4</b>
---------------	---------------------	----------------	----------	-----------------	-------------	---------------	---------------

**MATERIAL ANEXO:** NO

**PREMIO O DISTINCIÓN** (En caso de ser LAUREADAS o Meritoria): N/A

**PALABRAS CLAVES EN ESPAÑOL E INGLÉS:**

<u>Español</u>	<u>Inglés</u>
1. Analisis de datos	Analysis of data
2. Datos de la vida real	Real life data
3. Manual	Manual
4. Excel	Excel
5. Educación	Education
6. Matemáticas	Math

**RESUMEN DEL CONTENIDO:** (Máximo 250 palabras)

*La presente investigación, analizó una base datos de parámetros de calidad de agua tomados en las Islas Vírgenes de los Estados Unidos; conformada por las Isla San Croix (STC), San John (STJ) y San Thomas (STT), en un periodo de tiempo desde el año 1975 hasta el 2021; con el propósito de establecer un manual de guía para el análisis de datos, utilizando el software Excel.*

*La metodología de este estudio fue de carácter cuantitativo con enfoque descriptivo. Los datos de calidad de agua fueron facilitados por un investigador de las Islas Vírgenes de los Estados Unidos. La muestra de estudio correspondió a tres periodos de tiempo, desde el año 1975 al 2000, 2006 al 2010 y 2011 al 2021.*

*El análisis de la información se realizó en VIII etapas, que incluyeron: la organización de datos; separación y conversión de datos; depuración de variables; comparación de datos con variables de referencia; registro de datos en tablas; graficas; análisis y presentación de resultados. En cada etapa se estableció el paso a paso para analizar datos.*



*Como resultado de la investigación se logró establecer un manual de guía para el análisis de datos usando como modelo el análisis de calidad de agua de las Islas Vírgenes, San Croix, San John y San Thomas, a través de la herramienta Excel.*

**ABSTRACT:** (Máximo 250 palabras)

*The present investigation analyzed a database of water quality parameters taken in the Virgin Islands of the United States; made up of San Croix Island (STC), San John (STJ) and San Thomas (STT), in a period of time from 1975 to 2021; with the purpose of establishing a guide manual for data analysis, using Excel software.*

*The methodology of this study was quantitative with a descriptive approach. Water quality data was provided by a researcher from the US Virgin Islands. The study sample corresponded to three periods of time, from the year 1975 to 2000, 2006 to 2010 and 2011 to 2021.*

*The information analysis was carried out in VIII stages, which included: data organization; data separation and conversion; variable debugging; comparison of data with reference variables; data recording in tables; graphics; analysis and presentation of results. In each stage, the step by step was established to analyze data.*

*As a result of the investigation, it was possible to establish a guide manual for data analysis using the analysis of water quality of the Virgin Islands, Saint Croix, Saint John and Saint Thomas as a model, through the Excel tool.*



### APROBACION DEL SEMILLERO

Comité de currículo del programa:

MARTHA CECILIA MOSQUERA URRUTIA  
JEFE DE PROGRAMA LIC. EN MATEMÁTICAS

MERCY LILI PEÑA MORALES  
DIRECTORA DEL SEMILLERO INREMAT

HERBERT ENRIQUE QUINTERO FONSECA  
INTEGRANTE DEL GRUPO RECOMACI PROGRAMA LIC. EN MATEMATICAS

**ANÁLISIS DE DATOS DE CALIDAD DE AGUA USANDO LA HERRAMIENTA  
EXCEL.**

**JENNIFER TORRES SÁNCHEZ**

**LISSETH TATIANA LOPEZ MENDEZ**

**UNIVERSIDAD SURCOLOMBIANA  
FACULTAD DE EDUCACIÓN  
PROGRAMA DE LICENCIATURA EN MATEMÁTICAS**

**NEIVA- HUILA**

**2023**

**ANALISIS DE DATOS DE CALIDAD DE AGUA USANDO LA HERRAMIENTA EXCEL**

**Trabajo presentado como requisito de grado para optar al título de Licenciados en  
Matemáticas por:**

**JENNIFER TORRES SANCHEZ**

**20182174026**

**LISSETH TATIANA LOPEZ MENDEZ**

**20182174066**

**Asesora:**

**DRA. MERCY LILI PEÑA MORALES**

**UNIVERSIDAD SURCOLOMBIANA**

**FACULTAD DE EDUCACIÓN**

**PROGRAMA DE LICENCIATURA EN MATEMÁTICAS**

**NEIVA- HUILA**

**2023**

## AGRADECIMIENTOS

Llevar a cabo este proceso de investigación y de formación nos permitió lograr un sueño y alcance de metas profesionales, que nos hacen sentir orgullosas y agradecidas con Dios por iluminarnos.

Gracias a nuestra familia que siempre nos ha brindado su apoyo y motivación incondicional para poder cumplir todos nuestros objetivos personales y académicos.

Le gratificamos muy profundamente a nuestros asesores Mercy Lili Peña Morales y Herbert Enrique Quintero Fonseca por su dedicación y paciencia, sin sus palabras y correcciones precisas no habiéramos podido lograr llegar a esta instancia tan anhelada.

Son muchos los docentes que han sido parte de nuestra formación académica, y a todos ellos les queremos agradecer por transmitir sus conocimientos, también agradecemos a la Universidad Surcolombiana que nos ha exigido tanto, pero al mismo tiempo nos permitió alcanzar grandes metas.

Por último, agradecemos al Departamento de Planeación y Recursos Naturales de las Islas Vírgenes de los Estados Unidos (DPNR), por facilitarnos la base de datos de calidad de agua.

## Resumen

La presente investigación, analizó una base datos de parámetros de calidad de agua tomados en las Islas Vírgenes de los Estados Unidos; conformada por las Isla San Croix (STC), San John (STJ) y San Thomas (STT), en un periodo de tiempo desde el año 1975 hasta el 2021; con el propósito de establecer un manual de guía para el análisis de datos, utilizando el software Excel.

La metodología de este estudio fue de carácter cuantitativo con enfoque descriptivo. Los datos de calidad de agua fueron facilitados por un investigador de las Islas Vírgenes de los Estados Unidos. La muestra de estudio correspondió a tres periodos de tiempo, desde el año 1975 al 2000, 2006 al 2010 y 2011 al 2021.

El análisis de la información se realizó en VIII etapas, que incluyeron: la organización de datos; separación y conversión de datos; depuración de variables; comparación de datos con variables de referencia; registro de datos en tablas; graficas; análisis y presentación de resultados. En cada etapa se estableció el paso a paso para analizar datos.

Como resultado de la investigación se logró establecer un manual de guía para el análisis de datos usando como modelo el análisis de calidad de agua de las Islas Vírgenes, San cruz, San John y Santo Tomas, a través de la herramienta Excel.

***Palabras claves:*** Análisis de datos, Excel, manual de guía.

### **Abstract**

This research analyzed a database of water quality parameters taken in the Virgin Islands of the United States; which included San Croix Island (STC), San John (STJ) and San Thomas (STT), from 1975 to 2021; with the purpose of establishing a guidelines manual for data analysis, using the Microsoft Excel software.

The methodology of this study was quantitative with a descriptive approach. Water quality data was provided by a researcher from the US Virgin Islands. The study sample corresponded to three periods of time, from the year 1975 to 2000, 2006 to 2010 and 2011 to 2021.

The data analysis was carried out in VIII stages, which included: data organization; data separation and conversion; variable debugging; comparison of data with reference variables; data recording in tables; graphics; analysis and presentation of results. Each stage was described step by step to analyze the data set.

As a result of the investigation, it was possible to establish a guidelines manual for data analysis using the analysis of water quality of the Virgin Islands, Saint Croix, Saint John and Saint Thomas as a model, using the Microsoft Excel as a software tool.

**Keywords:** Data analysis, Excel, manual.

## TABLA DE CONTENIDO

Resumen.....	4
1. Introducción .....	8
2. Objetivos.....	9
<b>2.1 Objetivo General</b> .....	9
<b>2.2 Objetivos Específicos</b> .....	9
3. Planteamiento del Problema .....	10
<b>3.1 Descripción del Problema</b> .....	10
<b>3.2 Pregunta de Investigación</b> .....	10
4. Justificación .....	11
5. Marco teórico .....	12
<b>5.1 Análisis de Datos</b> .....	12
<b>5.2 Lineamientos para el Análisis de Datos</b> .....	12
<b>5.3 La Importancia de Usar Excel en el Análisis de Datos</b> .....	13
<i>5.3.1 Ventajas del programa Excel en el análisis de datos</i> .....	14
<b>5.4 Análisis de Datos en la Educación Matemática</b> .....	14
<b>5.5 ¿Qué es un Manual?</b> .....	15
<b>5.6 Parámetro de Calidad de Agua pH.</b> .....	15

<b>5.7 ¿Qué es el Oxígeno Disuelto (OD)?</b> .....	15
6. Marco Metodológico.....	17
7. Resultados .....	37
8. Análisis de resultados .....	50
9. Conclusiones.....	55
10. Sugerencias .....	56
11. Referencias.....	57

## 1. Introducción

Hoy en día saber analizar datos es de gran importancia en diferentes ámbitos como lo es en la educación; pero en los contenidos no se evidencia la enseñanza explícita de como analizar datos, en consecuencia, se refleja la falta de conocimientos en trabajar con datos situados de la vida real. Por lo tanto, es importante determinar un manual de guía de análisis de datos en los que se evidencie un paso a paso de los procesos que se deban realizar.

El manejo de datos de situaciones cotidianas es importante porque se relaciona con actividades, significados y contextos implicados en la producción de la información estadística que ha de servir para el análisis de datos.

En el análisis de datos se adquieren importantes beneficios como la organización de datos; permitiendo determinar las variables que se van a estudiar; otro beneficio es la identificación de datos atípicos, donde se realiza la depuración de dichos datos. Por último, la relación de las variables en tablas.

En este sentido, la pregunta de investigación buscó establecer un manual de guía para analizar datos a través de la herramienta de Excel que permitieron identificar factores y variables aplicadas a la estadística y se utilizó datos reales provenientes de la toma de parámetros de calidad de agua en las Islas Vírgenes de los Estados Unidos; Santa Cruz, San Juan y Santo Tomas.

La investigación se desarrolló en ocho etapas que permitió la organización de datos; separación, conversión de datos y depuración de variables.

## **2. Objetivos**

### **2.1 Objetivo General**

Establecer un manual de guía de análisis de datos, usando datos de calidad de agua de las Islas Vírgenes a través de la herramienta Excel.

### **2.2 Objetivos Específicos**

- Caracterizar datos de calidad de agua recolectados de las Islas Vírgenes utilizando el programa Excel.
- Identificar, clasificar y estandarizar las variables de calidad de agua para elegir el parámetro a estudiar.
- Describir los parámetros del pH y el oxígeno disuelto (OD) evidenciando las etapas relacionadas con el análisis de datos.

### **3. Planteamiento del Problema**

#### **3.1 Descripción del Problema**

Actualmente analizar datos es de gran importancia, tanto en lo laboral como en lo educativo, sin embargo, las competencias que tenemos para analizar datos son muy pocas.

También siempre que analizamos datos usamos datos descontextualizados de la vida real, en donde se observa la falta del manejo de datos de situaciones diarias. Además, requerimos de las competencias tecnológicas para analizar datos, ya que en el trascurso de nuestra formación; se evidencia la falta de experiencias en el manejo de las herramientas para usar programas que nos facilite el análisis de datos.

Por otro lado, en los cursos específicamente los de la Licenciatura en Matemáticas, no se evidencia la importancia de saber analizar datos, es por esto que en los contenidos curriculares hace falta incorporar algunos cursos que trabajen específicamente el análisis de datos; así adquiriendo habilidades y competencias necesarias para trabajar con datos de la vida real.

Debido a estas carencias que tenemos, nos surge la siguiente pregunta de investigación.

#### **3.2 Pregunta de Investigación**

¿Cómo unos datos suministrados de calidad de agua nos permiten establecer un manual de guía para analizar datos?

#### 4. Justificación

Aprender a analizar datos es importante porque influye en diferentes contextos. De acuerdo con Santillán (2013), “el análisis de datos está prácticamente en todos los aspectos de nuestra vida, porque a partir de todas nuestras actividades es posible recopilar datos que, después de ser analizados, nos permiten tomar decisiones” (p.1).

Por lo anterior, es importante que las personas sean capaces de interpretar la información que se encuentra en su entorno y así realizar un correcto análisis de los datos para dar solución a las diferentes situaciones que se presentan.

En la educación es importante que exista un manual de guía de análisis de datos, para que sea usado en talleres, conferencias y capacitaciones de investigaciones. Sin embargo, hasta el momento ha sido complejo que se cumpla esta meta ya que en los establecimientos educativos no se evidencia la importancia del análisis de datos.

Por otro lado, Sabino señala que:

El análisis de los datos no es tarea que se improvisa. Por el contrario, todo investigador debe dominar su tema y trabajar con rigurosidad para tener una idea precisa de cuáles serán los lineamientos principales del análisis que se deben de efectuar antes de comenzar a analizar datos. (2003, p.103).

Teniendo como referente lo anterior, en esta investigación se determinó un manual de guía de análisis de datos, con la ayuda del software Excel, el cual es un programa apropiado para realizar cálculos, filtrar información, realizar diferentes gráficas entre otros. Este software está orientado básicamente al campo contable, estadístico y financiero; permitiendo explorar y comprender conceptos matemáticos básicos (Gamboa, 2007 y citado en Torres, s. f. ).

## **5. Marco teórico**

### **5.1 Análisis de Datos**

Como afirma Encinas (1993, citado en Alva Santos, s. f.) “los datos en sí mismos tienen limitada importancia, es necesario "hacerlos hablar", en ello consiste, en esencia, el análisis e interpretación de los datos” (p.1). Es por esto que la interpretación es fundamental en el análisis de datos, ya que ayuda a identificar el contexto de los datos y a la vez a determinar el objeto de estudio. Además, Peña (2017) expone que el análisis de datos “integra distintas operaciones en la que el investigador o analista somete ciertos datos, bien sea de orden cuantitativo o cualitativo, a una serie de análisis, lecturas e interpretaciones, según sea el enfoque de su investigación o requerimiento informativo” (p.30).

El análisis de datos es un trabajo extenso, que puede ser optimizado por diferentes herramientas tecnológicas, como: Microsoft power BI, Programación en R, SAS, Python, Excel, Tableau Public, Rapid Miner, Apache Spark, Qlik View (KeepCoding, 2022).

En este estudio se trabajó con el software Excel, que es ampliamente utilizado para trabajar con datos tabulares, ya que podremos ordenarlos, filtrarlos, hacer una búsqueda sobre ellos, o utilizar herramientas avanzadas para analizar dicha información.

### **5.2 Lineamientos para el Análisis de Datos**

Durante el proceso de esta investigación se tuvo en cuenta los siguientes lineamientos determinados por Santos (s. f, p.4) :

- Los datos recogidos para una investigación tienen poco significado si no se les ordena o clasifica siguiendo algún sistema. Además, es la única forma de poder

realizar el análisis e interpretación de los datos en forma adecuada y con economía de tiempo y esfuerzo.

- La tabulación, como técnica, consiste en ordenar y situar los datos en tablas.
- En cambio, si se ordenan los puntajes en una distribución de frecuencias se puede examinar la configuración general y determinar las formas de distribución de los puntajes y su significado.
- En la tabulación de los datos un concepto básico es el de frecuencia, es decir el número de casos que poseen una característica determinada.
- La distribución de frecuencias se realiza estableciendo un determinado número de intervalos de clase en los cuales se marcan las puntuaciones. (En estadística existen principios para determinar el número de intervalos y la amplitud de éstos.)
- Después que se han anotado todas las puntuaciones se cuentan las marcas para hallar la frecuencia o número de casos que corresponde a cada intervalo. La suma de todas las frecuencias es igual al número de casos estudiados.

Los anteriores lineamientos se tuvieron en cuenta de forma general; sin embargo, el objetivo de este estudio es determinar un manual de guía de análisis de datos que, describa el paso a paso de cada etapa realizada.

### **5.3 La Importancia de Usar Excel en el Análisis de Datos**

El análisis de datos es un trabajo extenso y riguroso, que puede ser optimizado por diferentes herramientas tecnológicas, las cuales tienen varias funciones como lo es el software Excel. Ramirez (2017) expone que “Excel es ampliamente utilizado para trabajar con datos

tabulares, ya que podremos ordenarlos, filtrarlos, hacer una búsqueda sobre ellos, o utilizar herramientas avanzadas para analizar dicha información” (p.20). Actualmente, Excel es un programa muy importante para el análisis de datos, porque ofrece análisis rápidos y eficaces, y a la vez tiene bastantes herramientas para reducir el tiempo de trabajo.

### ***5.3.1 Ventajas del programa Excel en el análisis de datos***

Excel es un programa de fácil acceso y muy eficaz; con él se pueden realizar buenos trabajos, como lo es, el análisis de datos en investigaciones; cuenta con diversas ventajas que facilitan los procesos a realizar, dentro de este orden de ideas. Ramírez (2017) nos menciona algunas de las ventajas de Excel las cuales son: “Organizar datos, importar datos, crear gráficas, explorar conjuntos de datos masivos, almacenar datos, realizar cálculos complejos, ordenar información alfabéticamente y alfanumérica” (p.21).

### **5.4 Análisis de Datos en la Educación Matemática**

En el análisis de datos Batanero afirma que:

En esta clase de estudios se investigan las actividades, significados y contextos implicados en la producción de la información estadística que ha de servir de base para el análisis. Estos estudios se interesan por la observación empírica directa de la conducta y de las prácticas de los productores de datos estadísticos y de los entornos de investigación naturales en los que se producen y usan los mismos (1992, p.6).

### **5.5 ¿Qué es un Manual?**

El manual es una guía, el cual se crea para obtener una información detallada, ordenada, sistemática e integral que contiene todas las instrucciones, responsabilidades e información sobre funciones, sistemas y procedimientos de las distintas operaciones o actividades que se realizan (Palma, 20005, citado por Vivanco, 2017).

### **5.6 Parámetro de Calidad de Agua pH.**

Según Cifuentes, Torres y Frías:

El conocimiento del pH del agua del mar tiene importancia en biológica, ya que muchos fenómenos biológicos pueden estar regulados por el mismo; parece ser que incluso puede haber una influencia del pH en las migraciones de diversas especies de animales marinos; por tal razón es de interés su determinación y valoración. El agua de mar es ligeramente alcalina y el valor de su pH está entre 7,5 y 8,4 unidades de pH y varía en función de la temperatura; si ésta aumenta, el pH disminuye y tiende a la acidez; también, puede variar en función de la salinidad, la presión o profundidad y la actividad vital de los organismos marinos (1997, p.76).

### **5.7 ¿Qué es el Oxígeno Disuelto (OD)?**

El oxígeno disuelto es importante para los organismos acuáticos necesitan oxígeno para sobrevivir y crecer; algunas especies requieren de niveles elevados de oxígeno disuelto. La unidad (mg/l) representa los miligramos por litro. Se tuvo en cuenta la muestra de variación entre 1 mg/l a 20 mg/l en esta investigación.

Al respecto ,Peña (2007), nos dice que:

El oxígeno disuelto (OD) es la cantidad de oxígeno que está disuelta en el agua y que es esencial para los riachuelos y lagos saludables. El nivel de oxígeno disuelto puede ser un indicador para saber la contaminación del agua. Generalmente, un nivel más alto de oxígeno disuelto indica agua de mejor calidad. Si los niveles de oxígeno disuelto son demasiado bajos, algunos peces y otros organismos no pueden sobrevivir (2002, p.2).

## 6. Marco Metodológico

Este trabajo de investigación consiste en el análisis de datos de calidad de agua, con el propósito de configurar un manual de guía para analizar datos, usando una metodología de enfoque cuantitativo descriptivo. En este sentido, el método cuantitativo de Hernandez, Fernández y Bapista manifiestan que:

Usan la recolección de datos para probar hipótesis, con base en la medición numérica y el análisis estadístico, para establecer patrones de comportamiento y probar teorías. Además, señalan que este enfoque es secuencial y probatorio, cada etapa precede a la siguiente y no podemos “brincar o eludir” pasos, el orden es riguroso, aunque desde luego, podemos redefinir alguna fase y parte de una idea, que va acotándose y, una delimitada, se derivan objetivos y preguntas de investigación, se revisa la literatura y se construye un marco o una perspectiva teórica. (2010, p.1).

Por lo tanto, la recolección de los datos implica ser orientada para lograr proveer un buen entendimiento de las diferentes variables; el investigador empieza; aprender, observar y manipular los distintos datos a medida que se va avanzando la investigación.

## 6.1 Población

Los datos de esta investigación son datos recolectados de las islas vírgenes de los Estados Unidos, el territorio está conformado por tres Islas, St. Croix (STC), St. John (STJ) y St. Thomas (STT). Tal y como se muestra en la figura 1. La población de este estudio corresponde a los datos de calidad de agua que fueron tomados desde el año 1975 hasta el año 2021.

### Figura 1

*Mapa de las Islas Vírgenes de los Estados Unidos*



*Nota: Islas San Croix, San Jhon, San Thomas. (peterhermesfurian, ( s. f. ))*

## 6.2 Muestra

La muestra de este estudio corresponde a los datos tomados en las islas St. Croix, St. John y St. Thomas, fueron tomados utilizando el instrumento EXO<sup>2</sup> para la toma de parámetros como

se observa en la figura 2. Los datos fueron facilitados en tres documentos de Excel del año 1975 al 2000, 2006 al 2010 y 2011 al 2021.

## **Figura 2**

*Toma de muestreo de parámetros de calidad de agua*



*Cortesía de Amber Packar (Directora de laboratorio de calidad de agua, Universidad de las Islas Vírgenes).*

La cantidad de datos pertenecientes al periodo del 1975 al 2000 fue inicialmente de 45236 datos, para el periodo del 2006 al 2010 eran 119262 datos y del 2010 al 2021 se contó con 150273 datos, esto equivale al 71.2 % de la población. Los datos representaban diferentes parámetros de calidad de agua, fueron tomados en distintos sitios alrededor de las islas.

### **6.3 Método de Recolección de Datos**

Los datos trabajados en esta investigación son facilitados por un investigador de las Islas Vírgenes de los Estados Unidos; fueron suministrados mediante un documento de Excel.

Estos datos estaban distribuidos en 63 columnas de la columna A la columna BK, cada una de ellas tenía un nombre y representaba una variable asociada a la calidad de agua, como se observa en la tabla 1; algunas columnas se encontraban vacías. Para facilitar el análisis se

procedió a seleccionar y organizar las columnas que se consideraron relevantes para el propósito de este estudio.

**Tabla 1**

*Columna con su respectivo parámetro*

Columna	Nombre.
A	M
B	OrganizationFormalName
C	ActivityIdentifier
D	ActivityTypeCode
E	ActivityMediaName
F	ActivityMediaSubdivisionName
G	ActivityStartDate
H	ActivityStartTime/Time
I	ActivityStartTime/TimeZoneCode
J	ActivityEndDate
K	ActivityEndTime/Time
L	ActivityEndTime/TimeZoneCode
M	ActivityDepthHeightMeasure/MeasureValue
N	ActivityDepthHeightMeasure/MeasureUnitCode
O	ActivityDepthAltitudeReferencePointText
P	ActivityTopDepthHeightMeasure/MeasureValue
Q	ActivityTopDepthHeightMeasure/MeasureUnitCode
R	ActivityBottomDepthHeightMeasure/MeasureValue
S	ActivityBottomDepthHeightMeasure/MeasureUnitCode
T	ProjectIdentifier
U	ActivityConductingOrganizationText

Columna	Nombre.
V	IdentificardordeUbicacióndeMonitoreo
W	ActivityCommentText
X	SampleAquifer
Y	HydrologicCondition
Z	HydrologicEvent
AA	SampleCollectionMethod/MethodIdentifier
AB	SampleCollectionMethod/MethodIdentifierContext
AC	SampleCollectionMethod/MethodName
AD	SampleCollectionEquipmentName
AE	ResultDetectionConditionText
AF	SampleCollectionEquipmentName
AG	ResultSampleFractionText
AH	ResultMeasureValue
AI	ResultMeasure/MeasureUnitCode
AJ	MeasureQualifierCode
AK	ResultStatusIdentifier
AL	StatisticalBaseCode
AM	ResultValueTypeName
AN	ResultWeightBasisText
AO	ResultTimeBasisText
AP	ResultTemperatureBasisText
AQ	ResultParticleSizeBasisText
AR	PrecisionValue
AS	ResultCommentText
AT	USGSPCode
AU	ResultDepthHeightMeasure/MeasureValue

Columna	Nombre.
AV	ResultDepthHeightMeasure/MeasureUnitCode
AW	ResultDepthAltitudeReferencePointText
AX	SubjectTaxonomicName
AY	SampleTissueAnatomyName
AZ	ResultAnalyticalMethod/MethodIdentifier
BA	ResultAnalyticalMethod/MethodIdentifierContext
BB	ResultAnalyticalMethod/MethodName
BC	MethodDescriptionText
BD	MethodDescriptionText
BE	AnalysisStartDate
BF	ResultLaboratoryCommentText
BG	DetectionQuantitationLimitTypeName
BH	DetectionQuantitationLimitMeasure/MeasureValue
BI	DetectionQuantitationLimitMeasure/MeasureUnitCode
BJ	PreparationStartDate
BK	ProviderName

Para el análisis de los datos se determinaron las siguientes VI etapas

### ***6.3.1 Etapa I: Organización de los datos***

Se inicio realizando una copia de los datos originales en una nueva hoja de Excel, luego se hizo una filtraron por la columna D; así se ordenaron por el tipo de actividad y se usó un color diferente para diferenciar los tipos de actividad (Figura 3).

**Figura 3**

*Filtración de la columna (actividad)*

*Nota.* Diferencia de tipo de actividad en colores. Elaboración propia.

Posteriormente se copia esta hoja y se repite el procedimiento de ordenar los datos por la característica usando la filtración de la columna AF; se determinan nuevas características, se proporcionan un color para cada una de ellas (Figura 4).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	M	Organiz	Activity	Actividad	Activity	Activity	Activity	Activity	Activity	Activity	Activity	Activity	Activity	Activity	Activity
52633	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						0.5	m	
52635	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						0.5	m	
52636	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						0.5	m	
52637	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						0.5	m	
52638	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						0.5	m	
56725	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						3	m	
56726	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						3	m	
56727	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						3	m	
56728	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						3	m	
56729	USVIST_W	Governmente	USVIST_W	Field Msr/Obs	Water		2008-08-101:07:00						3	m	
100343	USVIST_W	Governmente	USVIST_W	Sample-Other	Water		2009-11-C 00:05:00						1	ft	
100368	USVIST_W	Governmente	USVIST_W	Sample-Other	Water		2009-11-C 00:15:00						1	ft	
100416	USVIST_W	Governmente	USVIST_W	Sample-Other	Water		2009-11-C 00:02:00						1	ft	
100419	USVIST_W	Governmente	USVIST_W	Sample-Other	Water		2009-11-C 00:15:00						1	ft	
100449	USVIST_W	Governmente	USVIST_W	Sample-Other	Water		2009-11-C 00:05:00						1	ft	
100493	USVIST_W	Governmente	USVIST_W	Sample-Other	Water		2009-11-C 00:52:00						1	ft	
100769	USVIST_W	Governmente	USVIST_W	Sample-Routine	Water		2006-06-101:12:00						1	ft	
101002	USVIST_W	Governmente	USVIST_W	Sample-Routine	Water		2006-06-C 01:27:00						1	ft	
101079	USVIST_W	Governmente	USVIST_W	Samole-Routine	Water		2006-05-301:07:00						1	ft	

**Figura 3**

*Filtración de la columna AF (característica)*



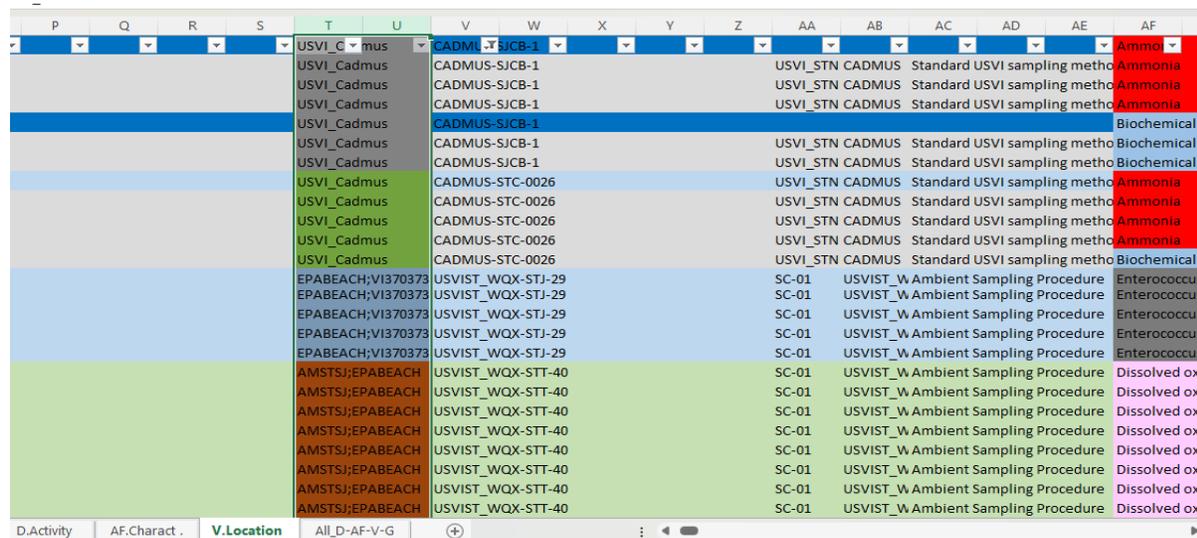
AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	A
1	Sample	Sample	Result	Charact	Results	Result	Result	Measur	Results	Statisti	Result	Result	Result	Result	Precisid	Result	USGSP	Result	Result	Result	Subje
3				Ammonia	0.021	mg/l															
13	Standard USVI sampling method			Ammonia	0.009	mg/l															
27	Standard USVI sampling method			Ammonia	0.012	mg/l															
56	Standard USVI sampling method			Ammonia	0.013	mg/l															
64	Standard USVI sampling method			Ammonia	0.01	mg/l															
234	Standard USVI sampling method			Biochemical oxygen c	2.94	mg/l															
262				Carbon	0.169	mg/l															
268	Standard USVI sampling method			Carbon	0.134	mg/l															
272	Standard USVI sampling method			Carbon	0.192	mg/l															
279	Standard USVI sampling method			Carbon	0.128	mg/l															
281	Standard USVI sampling method			Carbon	0.19	mg/l															
287	Standard USVI sampling method			Carbon	0.19	mg/l															
317	Standard USVI sampling method			Carbon	0.179	mg/l															
322	Standard USVI sampling method			Carbon	0.312	mg/l															
325	Standard USVI sampling method			Carbon	0.162	mg/l															
328	Standard USVI sampling method			Carbon	0.0655	mg/l															
353	Standard USVI sampling method			Chlorophylla, correct	0.61	ug/l															
369	Standard USVI sampling method			Chlorophylla, correct	0.61	ug/l															
547	Standard USVI sampling method			Chlorophylla, uncorrect	0.69	ug/l															
560	Standard USVI sampling method			Chlorophylla, uncorrect	0.69	ug/l															
1170	Stony Coral Reef Assessment: Q			Coral Bleached Cover	0	%															
2366	Stony Coral Reef Assessment: Q			Coral Bleached Cover	0	%															
2367	Stony Coral Reef Assessment: Q			Coral Bleached Cover	0	%															
3532	Stony Coral Reef Assessment: Q			Coral Disease Cover	0	%															
4728	Stony Coral Reef Assessment: Q			Coral Disease Cover	0	%															

*Nota.* Se filtra la columna AF para diferenciar con un color cada característica. Elaboración propia.

Por último, en una nueva hoja, se filtran los datos de la columna V, se ordenan por el identificador de ubicación de monitoreo, la cual representa la isla y su ubicación. Debido a la cantidad de sitios en la columna, se decide caracterizar cuatro sitios; los cuales se codifican como STC, STJ, STT y datos no definidos. Se establecieron colores para cada uno de estos sitios, coloreando las columnas T y la columna U (Figura 5).

**Figura 5**

*Sitios de ubicación*



P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF
				USVI_Cadmus	CADMUS-SJCB-1	USVI_STN	CADMUS	Standard USVI sampling metho	Ammonia							
				USVI_Cadmus	CADMUS-SJCB-1	USVI_STN	CADMUS	Standard USVI sampling metho	Ammonia							
				USVI_Cadmus	CADMUS-SJCB-1	USVI_STN	CADMUS	Standard USVI sampling metho	Ammonia							
				USVI_Cadmus	CADMUS-SJCB-1	USVI_STN	CADMUS	Standard USVI sampling metho	Ammonia							
				USVI_Cadmus	CADMUS-STC-0026	USVI_STN	CADMUS	Standard USVI sampling metho	Ammonia							
				USVI_Cadmus	CADMUS-STC-0026	USVI_STN	CADMUS	Standard USVI sampling metho	Ammonia							
				USVI_Cadmus	CADMUS-STC-0026	USVI_STN	CADMUS	Standard USVI sampling metho	Ammonia							
				USVI_Cadmus	CADMUS-STC-0026	USVI_STN	CADMUS	Standard USVI sampling metho	Ammonia							
				USVI_Cadmus	CADMUS-STC-0026	USVI_STN	CADMUS	Standard USVI sampling metho	Biochemical							
				EPABEACH;VI370373	USVIST_WQX-STJ-29	SC-01	USVIST_W	Ambient Sampling Procedure	Enterococcus							
				EPABEACH;VI370373	USVIST_WQX-STJ-29	SC-01	USVIST_W	Ambient Sampling Procedure	Enterococcus							
				EPABEACH;VI370373	USVIST_WQX-STJ-29	SC-01	USVIST_W	Ambient Sampling Procedure	Enterococcus							
				EPABEACH;VI370373	USVIST_WQX-STJ-29	SC-01	USVIST_W	Ambient Sampling Procedure	Enterococcus							
				EPABEACH;VI370373	USVIST_WQX-STJ-29	SC-01	USVIST_W	Ambient Sampling Procedure	Enterococcus							
				AMSTS;EPABEACH	USVIST_WQX-STT-40	SC-01	USVIST_W	Ambient Sampling Procedure	Dissolved ox							
				AMSTS;EPABEACH	USVIST_WQX-STT-40	SC-01	USVIST_W	Ambient Sampling Procedure	Dissolved ox							
				AMSTS;EPABEACH	USVIST_WQX-STT-40	SC-01	USVIST_W	Ambient Sampling Procedure	Dissolved ox							
				AMSTS;EPABEACH	USVIST_WQX-STT-40	SC-01	USVIST_W	Ambient Sampling Procedure	Dissolved ox							
				AMSTS;EPABEACH	USVIST_WQX-STT-40	SC-01	USVIST_W	Ambient Sampling Procedure	Dissolved ox							
				AMSTS;EPABEACH	USVIST_WQX-STT-40	SC-01	USVIST_W	Ambient Sampling Procedure	Dissolved ox							
				AMSTS;EPABEACH	USVIST_WQX-STT-40	SC-01	USVIST_W	Ambient Sampling Procedure	Dissolved ox							

*Nota.* Se diferencia los cuatros sitios de ubicación; sitios no definidos, STC, STJ y STT. Elaboración propia.

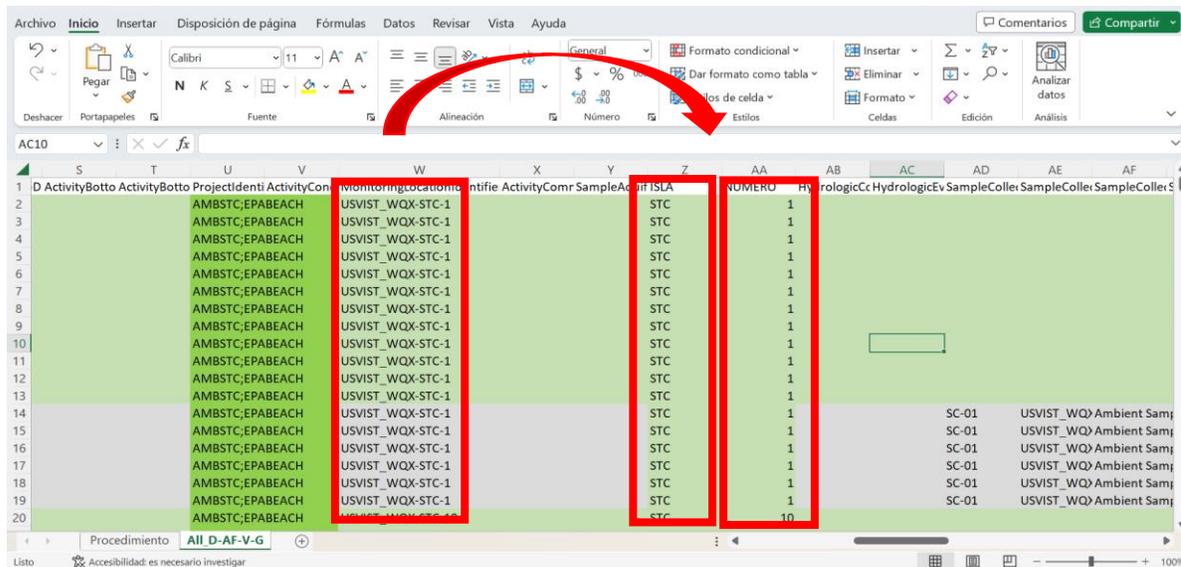
Para una organización final, se realiza la copia de los datos en una nueva hoja, se filtran por la columna D (actividad), AF (característica), V (ubicación) y G (isla y fecha); se procede a guardar el documento.

**6.3.2 Etapa II: separación y conversión de datos**

Se inicio a trabajar en nuevo documento de Excel con los datos de la última hoja realizada en el documento anterior; Posteriormente, se separa la información de la columna V, la cual contiene la información de la isla y su ubicación, se insertan dos columnas, una se nombran isla y la otra como ubicación; para un mejor manejo de los datos, se inició el proceso filtrando los datos de una de las islas, en ese caso la isla San Croix (STC). Se filtró la información de la columna V; se copiaron y pegaron los datos en las nuevas columnas, este proceso se realizó con las otras islas STJ y STT (Figura 7).

**Figura 6**

*Separación del parámetro Isla*

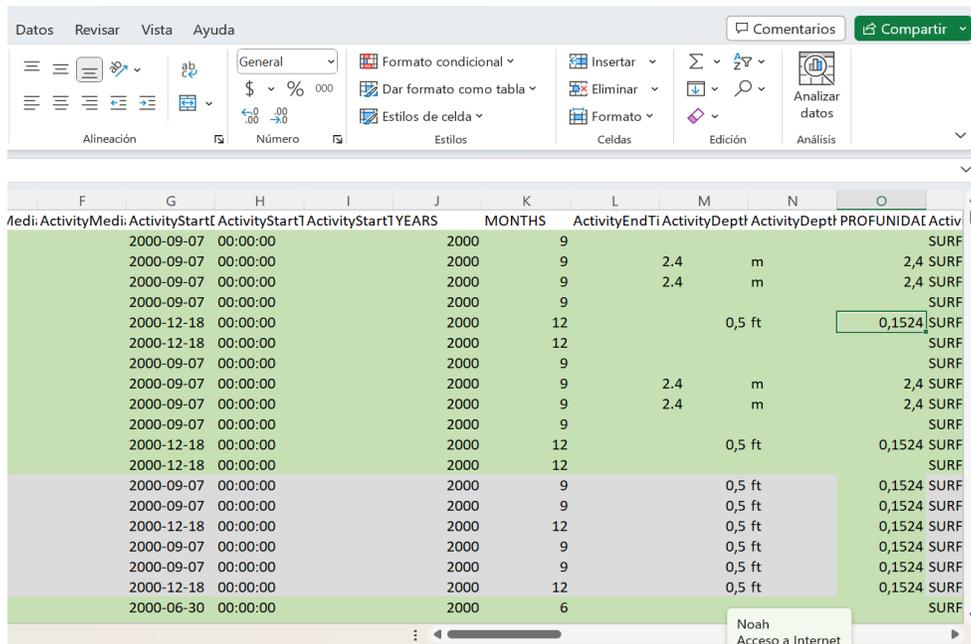


*Nota:* la separación de datos mejora la compresión de la investigación. Elaboración Propia

Seguidamente se realizó el procedimiento de separar la información de la columna G, la cual contiene el año, mes y día en que fueron tomados los datos, solo se tuvo en cuenta el año y el mes; por consiguiente, se insertan dos columnas, una se nombran año y la otra como mes. Se realizó nuevamente el mismo procedimiento anterior filtrando la información de la columna nueva (isla), iniciando con la isla STC; se identifican los años y los meses en que se tomaron los datos en esa Isla, se copian y pegan en las nuevas columnas. El proceso se hizo de igual forma con las islas STJ y STT (Figura 7).

**Figura 7**

*Separación del parámetro año*



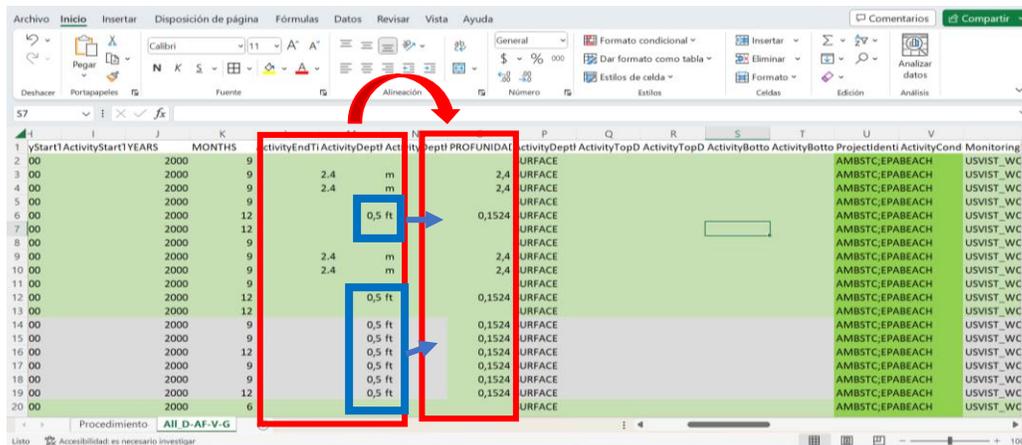
Medi: ActivityMedi:	ActivityStart	ActivityStart	ActivityStart	YEARS	MONTHS	ActivityEndTi	ActivityDept	ActivityDept	PROFUNDIDAD	Activi
	2000-09-07	00:00:00		2000	9					SURF
	2000-09-07	00:00:00		2000	9	2.4	m		2,4	SURF
	2000-09-07	00:00:00		2000	9	2.4	m		2,4	SURF
	2000-09-07	00:00:00		2000	9					SURF
	2000-12-18	00:00:00		2000	12		0,5 ft		0,1524	SURF
	2000-12-18	00:00:00		2000	12					SURF
	2000-09-07	00:00:00		2000	9					SURF
	2000-09-07	00:00:00		2000	9	2.4	m		2,4	SURF
	2000-09-07	00:00:00		2000	9	2.4	m		2,4	SURF
	2000-09-07	00:00:00		2000	9					SURF
	2000-12-18	00:00:00		2000	12		0,5 ft		0,1524	SURF
	2000-12-18	00:00:00		2000	12					SURF
	2000-09-07	00:00:00		2000	9		0,5 ft		0,1524	SURF
	2000-09-07	00:00:00		2000	9		0,5 ft		0,1524	SURF
	2000-12-18	00:00:00		2000	12		0,5 ft		0,1524	SURF
	2000-09-07	00:00:00		2000	9		0,5 ft		0,1524	SURF
	2000-09-07	00:00:00		2000	9		0,5 ft		0,1524	SURF
	2000-12-18	00:00:00		2000	12		0,5 ft		0,1524	SURF
	2000-06-30	00:00:00		2000	6					SURF

*Nota.* La separación del parámetro año mejora la compresión de la investigación. Elaboración Propia

En los datos se tenía la información de la profundidad en que fueron tomados y estaban registrados en la columna M. En la columna N, la unidad de medida de la profundidad, estaba en Metros (m) y en Pies (ft), para tener un manejo de los datos en una sola unidad de medida, se decide realizar la conversión de los datos que están en Pies a Metros. Se realiza la conversión en una columna nueva por lo tanto se nombró profundidad (m) en la que se indicó por medio de una formula los valores que se convirtieron, teniendo en cuenta el valor de Pie en Metros, luego con los beneficios que nos brinda Excel se realiza la conversión automáticamente con cada cantidad de datos que está en Pies, estos datos quedan registrados directamente en la columna profundidad(m), (Figura 8).

**Figura 8**

*Conversión*



ActivityStart	ActivityStart	YEARS	MONTHS	ActivityEnd	ActivityDepth	ActivityDepth	Profundidad	ActivityDepth	ActivityTopD	ActivityTopD	ActivityBottD	ActivityBottD	ProjectIdent	ActivityCond	Monitoring
2 00	2000	9		2,4	m		2,4	URFACE					AMBSTC;EPABEACH	USVIST_WC	
3 00	2000	9		2,4	m		2,4	URFACE					AMBSTC;EPABEACH	USVIST_WC	
4 00	2000	9		2,4	m		2,4	URFACE					AMBSTC;EPABEACH	USVIST_WC	
5 00	2000	9		2,4	m		2,4	URFACE					AMBSTC;EPABEACH	USVIST_WC	
6 00	2000	12		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
7 00	2000	12		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
8 00	2000	9		2,4	m		2,4	URFACE					AMBSTC;EPABEACH	USVIST_WC	
9 00	2000	9		2,4	m		2,4	URFACE					AMBSTC;EPABEACH	USVIST_WC	
10 00	2000	9		2,4	m		2,4	URFACE					AMBSTC;EPABEACH	USVIST_WC	
11 00	2000	9		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
12 00	2000	12		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
13 00	2000	12		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
14 00	2000	9		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
15 00	2000	9		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
16 00	2000	12		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
17 00	2000	9		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
18 00	2000	9		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
19 00	2000	12		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	
20 00	2000	6		0,5 ft			0,1524	URFACE					AMBSTC;EPABEACH	USVIST_WC	

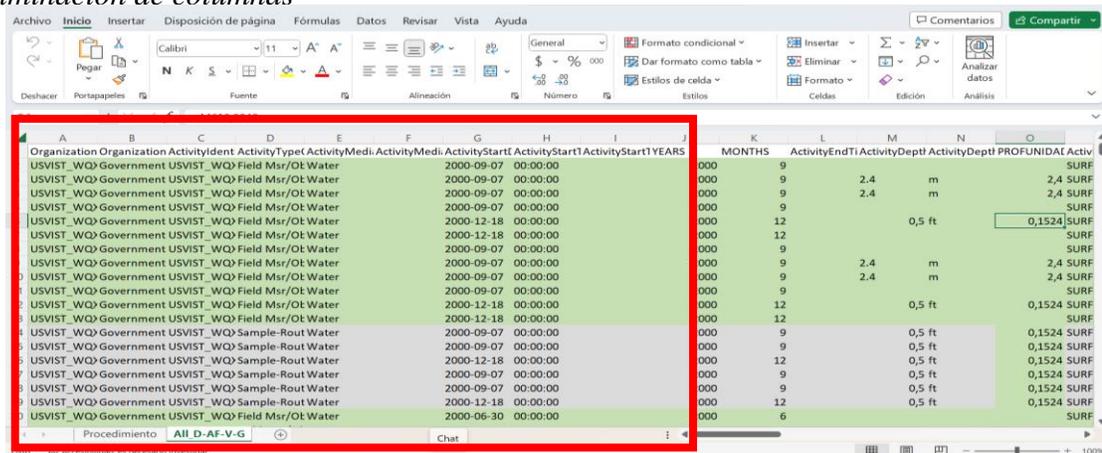
*Nota.* Conversión de datos de Pies a Metros. Elaboración propia.

**6.3.3 Etapa III: depuración de variables**

En una nueva hoja de Excel, se realizó la copia de los datos que se guardaron en el último documento; se eliminaron las columnas que no tienen información útil (Figura 9); se dejaron las columnas en la que se especifica el año, mes, profundidad(m), isla, ubicación, la característica, el valor de medida y la unidad de medida; los datos se filtran y ordenan respecto a la característica, isla, año y mes. Este documento quedó listo para trabajar cualquier parámetro que se quiera analizar. En este estudio se elige el parámetro de calidad de agua pH y el oxígeno disuelto.

**Figura 9**

*Eliminación de columnas*



*Nota:* Se depuraron las columnas que no tenían información. Elaboración propia.

Luego se depuran los sitios no definidos eliminando cierta cantidad de filas que no tenían información relevante para la investigación. Para la realización de la depuración, se filtra el color con el que se identificó el sitio; se tomó la cantidad de los datos que se eliminaban de cada característica y se escribían en la hoja de procedimientos para determinar cuántos datos se eliminaban en total, (Figura 10). Por último, se realizó una tabla de comparación de números de filas en total y el número de filas eliminadas.

**Figura 10**

*Tabla de cantidad de datos eliminados en cada característica.*

21										TOTAL
22	AMMONIA					22	5			27
23	Ammonia-nitrogen							45		45
24	Biochemical oxygen demand, standard conditions					23	5	46		74
25	CARBON					23	5			28
26	Chlorophyll a, corrected for pheophytin					23	5	25		53
27	Chlorophyll a, free of pheophytin							84		84
28	Chlorophyll a, uncorrected for pheophytin					23	5			28
29	Coral Bleached Cover				2362					2362
30	Coral Diseased Cover				2362					2362
31	Coral Percent Cover				2362					2362
32	Coral Percent Live				1374					1374
33	Count				46027					46027
34	Depth, Bottom					5	1			6
35	Depth, Secchi disk depth					21	5	37		63
36	Diameter				1374					1374
37	Dissolved oxygen (DO)					43	9	340		392
38	Dissolved oxygen saturation									0
39	Enterococcus					23	5	27		55
40	Fecal Coliform					23	5	27		55
41	Height				1374					1374
42	Inorganic carbon					23	5			28

*Nota.* Se registran la depuración de datos no definidos. Elaboración propia.

### 6.3.4 Etapa IV: comparación de datos con variables de referencia

En el documento anterior se filtra los datos correspondientes a la característica del pH y el oxígeno disuelto. Se realizó la copia de los datos en un nuevo documento, en el que se inicia la eliminación de los datos que no corresponde a la medida estándar de cada una de ellas. (Figura 11).

**Figura 11**

*Eliminación de filas*

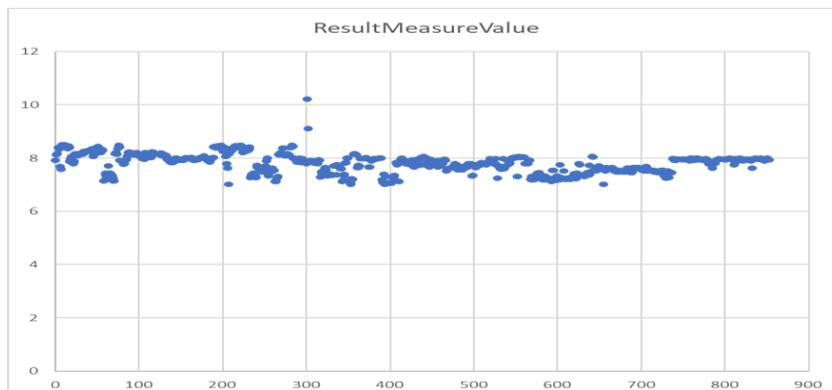
A	B	C	D	E	F	G	H	I	J
Year	Month	PROFUNDID/ ISLA	NUMERO	Characteristi	ResultSampl	ResultMeasu	ResultMeasure/	Measure/	
2008	3	2,9 STC		1	pH		5,43	None	
2008	3	0,12 STC		1	pH		4,33	None	
2008	3	0,3 STC		10	pH		4,51	None	
2008	3	1,7 STC		10	pH		4,45	None	
2008	3	0,5 STC	11B		pH		4,26	None	
2008	3	1,5 STC	11B		pH		4,47	None	
2008	3	0,4 STC		12	pH		5,26	None	
2008	3	2,6 STC		12	pH		5,2	None	
2008	3	1,79 STC		13A	pH		4,7	None	
2008	3	0,3 STC		13A	pH		4,73	None	
2008	3	0,3 STC		13A	pH		4,5	None	
2008	3	1,5 STC		13A	pH		4,5	None	
2008	3	1,8 STC		14A	pH		3,51	None	
2008	3	0,3 STC		14A	pH		3,45	None	
2008	3	0,8 STC		14B	pH		3,76	None	
2008	3	0,14 STC		14B	pH		3,51	None	
2008	3	2,5 STC		15	pH		3,48	None	
2008	3	0,1 STC		15	pH		3,12	None	
2008	3	0,2 STC		15A	pH		4,26	None	
2008	3	3,2 STC		15A	pH		4,27	None	
2008	3	7,7 STC		16	pH		4,58	None	
2008	3	0,1 STC		16	pH		4,83	None	

*Nota:* Eliminación de filas que no contaban con datos importantes en la investigación.  
Elaboración propia

El pH se determina de acuerdo con su medida estándar que debe ser entre 7 y 10,21. La eliminación se realizó filtrando los datos que se son menores de 7 y los que son mayores a 10,21; a su vez se eliminan las filas que se encontraban vacías, también se realiza el mismo procedimiento con el parámetro de oxígeno disuelto teniendo en cuenta que su medida estándar esta entre 1 mg/l a 20 mg/l. Por último, se realizó una tabla la cual contiene la cantidad de datos que se eliminaron y la cantidad de datos que quedaron. Para verificar que los datos con los que se iba a continuar trabajando estuvieran en la unidad de medida del pH, se realizó una gráfica en la que se evidenciara la dispersión de los datos (Figura 12).

**Figura 12**

*Grafica de Dispersión*



*Nota.* Se visualiza la variación estándar del pH, se encuentra entre 7 y 10,21. Elaboración propia

### 6.3.5 Etapa V: registro de datos en tablas

En una tabla se registra los datos de los años en los que se encontraban, los años, los meses, la cantidad de datos, el promedio, desviación estándar y el coeficiente de variación, de la característica del pH; se inicia filtrando la isla STC, como los datos fueron tomados en años diferentes se realiza la filtración del año y del mes, se hizo con el fin de calcular el promedio, la desviación estándar y el coeficiente de variación, esto se realizó por medio de una de las utilidades que trae Excel en la que nos brindaba las fórmulas y solo se debía seleccionar los datos para la realización de los cálculos, cada resultado se iba registrando en la tabla, este proceso también se realizó con las islas STJ y STT (Figura 13, 14 y 15).

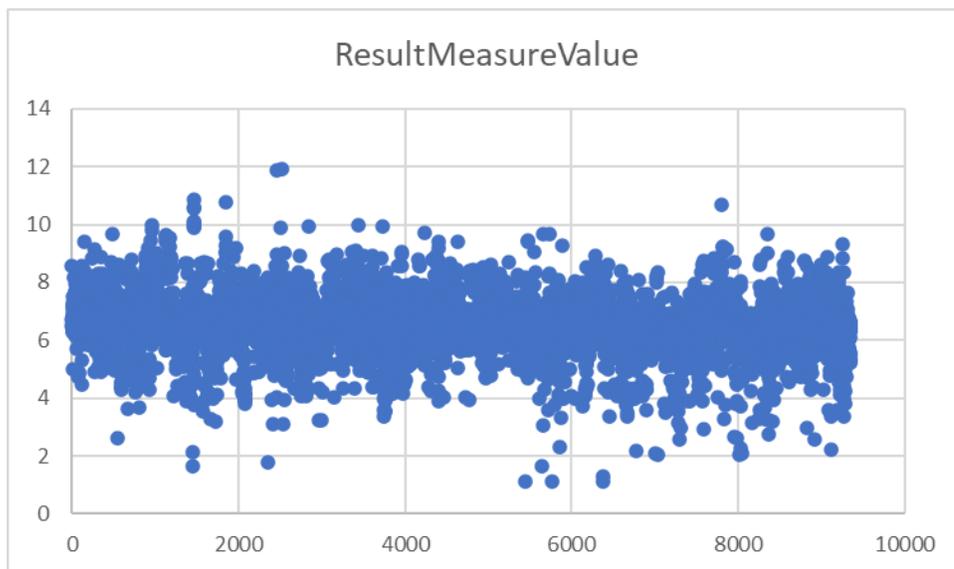
**Figura 4**

*Registro de datos del pH.*

AÑO	ISLA	MES	CANTIDAD	PROMEDIO	DESVEST.M	COEFICIENTE
2008	STC	3	1	7,90		
	STC	12	56	8,21	0,20	2,42
	STJ	6	2	9,66	0,78	89,13
	STT	3	1	7,18		
	STT	6	1	7,08		
	STT	8	1	7,38		
	STT	12	18	7,21	0,06	0,86
TOTAL			80			
2009	STC	9	6	7,27	0,12	1,59
	STC	10	8	7,33	0,19	2,54
	STC	11	6	8,28	0,14	1,70
	STC	12	6	7,85	0,06	0,76
	STJ	3	14	7,86	0,04	0,53
	STJ	6	26	7,49	0,15	2,03
	STJ	9	4	7,25	0,14	1,88
	STJ	10	10	7,43	0,42	5,60
	STJ	11	4	8,14	0,02	0,21
	STJ	12	4	7,67	0,04	0,55
	STT	3	112	7,77	0,14	1,75
	STT	6	120	7,51	0,30	4,05
	STT	9	82	7,54	0,09	1,19
	STT	10	12	7,42	0,10	1,32

**Figura 5**

*Grafica de dispersión Oxígeno Disuelto.*



*Nota.* se visualiza la variación estándar del oxígeno, se encuentra entre 1gm/l y 20gm/l.

Elaboración propia

**Figura 11**

*Tablas donde se relaciona varios parámetros*

AÑO	ISLA	MES	PROMEDIO	CANTIDAD	DESVEST M	COEFICIENTE
2010	STC	7	6,99	106	0,66	9,45
		10	6,72	44	0,48	7,20
	TOTAL			150		
2011	STC	4	6,82	84	0,61	8,92
				88		
	TOTAL					
2012	STC	3	6,69	52	0,74	11,01
		4	6,90	64	0,63	9,09
		5	6,89	116	0,60	8,70
		6	6,56	72	0,70	10,69
		7	6,75	26	0,35	5,15
	TOTAL	9	6,58	140	0,79	12,04
2013	STC	9	5,95	120	0,74	12,42
		12	6,78	116	0,76	11,22
	STJ	9	6,71	26	0,68	10,07
		12	6,81	26	0,33	4,84
	STT	9	6,74	116	0,69	10,19
2013		12	7,15	120	0,86	11,98
				524		
	TOTAL					
	STC	6	6,12	114	0,40	6,59
	9	6,18	118	0,82	13,26	
	12	5,63	114	0,65	11,55	
	5	6,88	28	0,64	9,63	

MES	STC	STJ	STI	TOTAL	STC
1					
2					
3					
4					8,58
5					
6		7,94	26	7,92	96
7				20	18,27
8	8,02	106			8,02
9					
10	8,30	44			8,30
11					44
12					

*Nota:* En la tabla izquierda se obtiene desviación estándar y coeficiente en cada año, isla y mes; y en la tabla derecha se relaciona el año, el mes, cantidad y promedio. Elaboración propia.

### 6.3.6 Etapa VI: gráficas.

Se realiza una tabla de forma horizontal en la cual contiene los años en los que se tomaron los datos, el promedio y cantidad de cada uno de los meses en cada isla y el total del promedio de cada mes en las islas durante los años, luego otra tabla en la que se registra el promedio total de cada mes en los años del 1998 al 2021 y la cantidad de datos de cada mes. Estas tablas se realizaron en la misma hoja trabajada en la etapa anterior.

La tabla de la característica del pH y del oxígeno disuelto, se realizó desde el año 1998 hasta el año 2021, (Figura 16, Figura 17), otro investigador registro información de los años 2001 hasta el 2005, la cual fueron facilitados para su debida grafica.

### Figura 6

La tabla de la característica del pH y del oxígeno disuelto

	1998			1999				2000				2001				2002			
	STJ	STT	TOTAL	STC	STJ	STT	TOTAL	STC	STJ	STT	TOTAL	TOTAL	STC	STJ	STT	TOTAL	STC	STJ	
1																			
2																			
3																			
4																			
5																			
6																			
7																			
8																			
9																			
10																			
11																			
12																			
13																			
14																			
15																			
16																			
17																			
18																			
19																			
20																			
21																			
22																			
23																			
24																			
25																			
26																			
27																			
28																			
29																			
30																			
31																			
32																			

**Figura 7**

La tabla de la característica del pH y del oxígeno disuelto

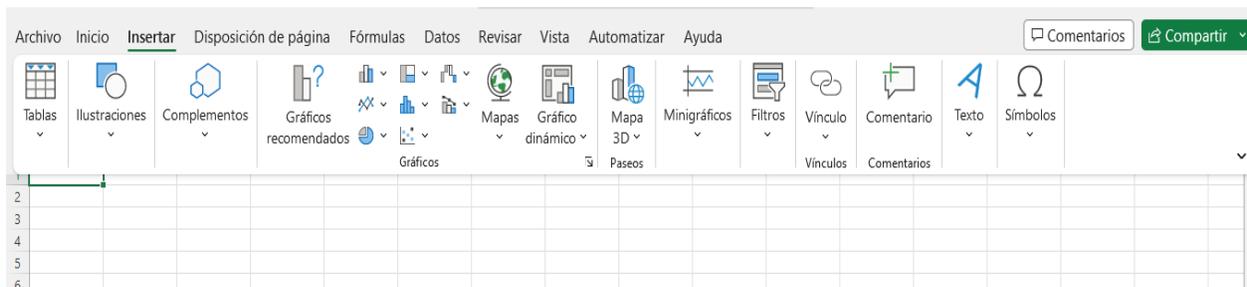
MES	1999						2000						2001				2002																																																																																																																																																																																																																																																																																		
	STJ	STI	TOTAL	STC	STJ	STI	TOTAL	TOTAL	STC	STJ	STI	TOTAL	STC	STJ	STI	TOTAL	STC	STJ	STI	TOTAL																																																																																																																																																																																																																																																																															
1																																																																																																																																																																																																																																																																																																			
2									6,34	300																																																																																																																																																																																																																																																																																									
3	6,97	38	5,74	38	6,35	76		6,35			6,12	12	6,34	300			6,06	318	5,81	72	5,83	279	5,94	669																																																																																																																																																																																																																																																																											
4																																																																																																																																																																																																																																																																																																			
5																																																																																																																																																																																																																																																																																																			
6	5,79	24	5,45	113	5,51	137	5,91	248	3,97	6	5,86	254	5,74	6,09	228	6,06	114	5,77	353	5,92	695																																																																																																																																																																																																																																																																														
7							5,71	312			5,71	312	5,71	5,79	138			5,55	234			5,64	372																																																																																																																																																																																																																																																																												
8												0		5,92	180							5,92	180																																																																																																																																																																																																																																																																												
9												0				5,80	108					5,80	108	5,99	306	4,03	66	5,64	372																																																																																																																																																																																																																																																																						
10								5,78	76	5,50	240	5,57	316	5,57																																																																																																																																																																																																																																																																																					
11								5,76	96	5,88	6	5,77	102	5,77																																																																																																																																																																																																																																																																																					
12							6,51	150			5,63	264	5,95	414	5,85		5,29	66	5,20	186	5,22	252	3,50	84	7,59	102	7,66	375	7,02	561																																																																																																																																																																																																																																																																					
21	<table border="1"> <thead> <tr> <th rowspan="2">MES</th> <th colspan="6">1999</th> <th colspan="6">2000</th> <th colspan="4">2001</th> <th colspan="4">2002</th> </tr> <tr> <th>STC</th> <th>STJ</th> <th>STI</th> <th>TOTAL</th> <th>STC</th> <th>STJ</th> <th>STI</th> <th>TOTAL</th> <th>TOTAL</th> <th>STC</th> <th>STJ</th> <th>STI</th> <th>TOTAL</th> <th>STC</th> <th>STJ</th> <th>STI</th> <th>TOTAL</th> <th>STC</th> <th>STJ</th> <th>STI</th> <th>TOTAL</th> </tr> </thead> <tbody> <tr> <td>ene</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> <tr> <td>feb</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> <tr> <td>mar</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> <tr> <td>abr</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> <tr> <td>may</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> <tr> <td>jun</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> <tr> <td>jul</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> <tr> <td>ago</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> <tr> <td>sep</td> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> </tbody> </table>																								MES	1999						2000						2001				2002				STC	STJ	STI	TOTAL	STC	STJ	STI	TOTAL	TOTAL	STC	STJ	STI	TOTAL	STC	STJ	STI	TOTAL	STC	STJ	STI	TOTAL	ene																									feb																									mar																									abr																									may																									jun																									jul																									ago																									sep																								
MES	1999						2000						2001				2002																																																																																																																																																																																																																																																																																		
	STC	STJ	STI	TOTAL	STC	STJ	STI	TOTAL	TOTAL	STC	STJ	STI	TOTAL	STC	STJ	STI	TOTAL	STC	STJ	STI	TOTAL																																																																																																																																																																																																																																																																														
ene																																																																																																																																																																																																																																																																																																			
feb																																																																																																																																																																																																																																																																																																			
mar																																																																																																																																																																																																																																																																																																			
abr																																																																																																																																																																																																																																																																																																			
may																																																																																																																																																																																																																																																																																																			
jun																																																																																																																																																																																																																																																																																																			
jul																																																																																																																																																																																																																																																																																																			
ago																																																																																																																																																																																																																																																																																																			
sep																																																																																																																																																																																																																																																																																																			

## 7. Resultados

Se presentan los resultados obtenidos de las etapas realizadas en la investigación para analizar los datos de calidad de agua, recolectados en las Islas Vírgenes de los Estados Unidos, El programa que se utilizó para analizar los datos fue el software Excel, herramienta que tiene un gran número de funciones que se pueden aprovechar según la necesidad del análisis de los datos. Las funciones de Excel que se utilizaron fueron: la filtración de datos, la eliminación de grandes cantidades de datos, conversiones, la realización de operaciones mediante fórmulas, elaboración de tablas y la realización de gráficas (Figura 18).

### Figura 18

#### *Barra de herramientas*



*Nota.* Un componente de la interfaz gráfica de un programa. Elaboración propia

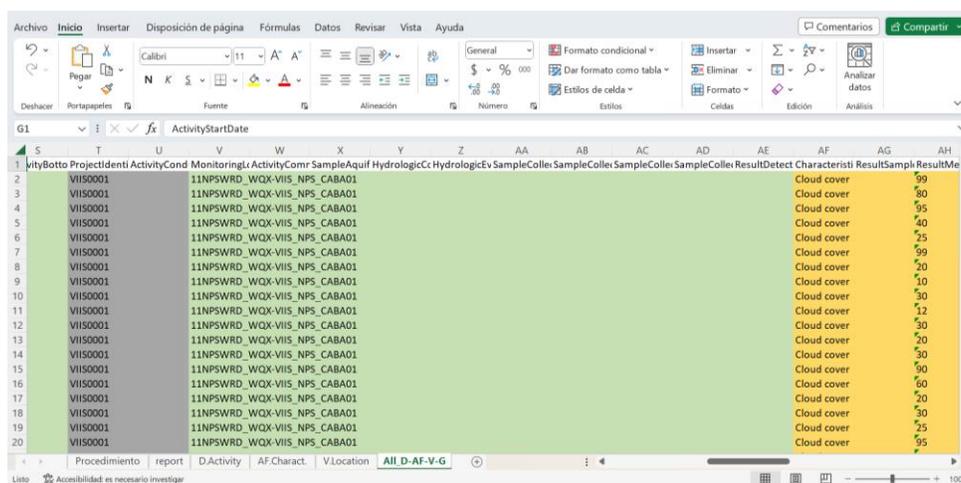
Durante cada etapa se realizaban diferentes procedimientos con el objetivo de determinar las variables que se consideraron relevantes para el propósito de este estudio; a continuación, se presenta el análisis de los datos establecidos en las VI etapas que se presentaron en la metodología.

## Etapa I. Organización de datos.

La organización de los datos se realizó en 4 hojas de Excel y se le dio un nombre a dicho documento. En la cuarta hoja, los datos quedan organizados en las siguientes columnas: columna D (actividad), AF (característica), V (ubicación) y G (isla y fecha) como se muestra en la figura 19.

**Figura 19**

*Organización de columnas*



The screenshot shows an Excel spreadsheet with the following data structure:

	D	AF	V	G
1	lityBotto ProjectIdenti	MonitoringL	ActivityComr	SampleAquif
2	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		HydrologicC; HydrologicEv
3	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		SampleCollec
4	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		SampleCollec
5	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		SampleCollec
6	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		SampleCollec
7	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		ResultDetect
8	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		Characteristi
9	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		ResultSampl
10	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		ResultMe
11	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
12	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
13	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
14	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
15	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
16	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
17	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
18	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
19	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		
20	VIIS0001	11NPSWRD_WQX-VIIS_NPS_CABA01		

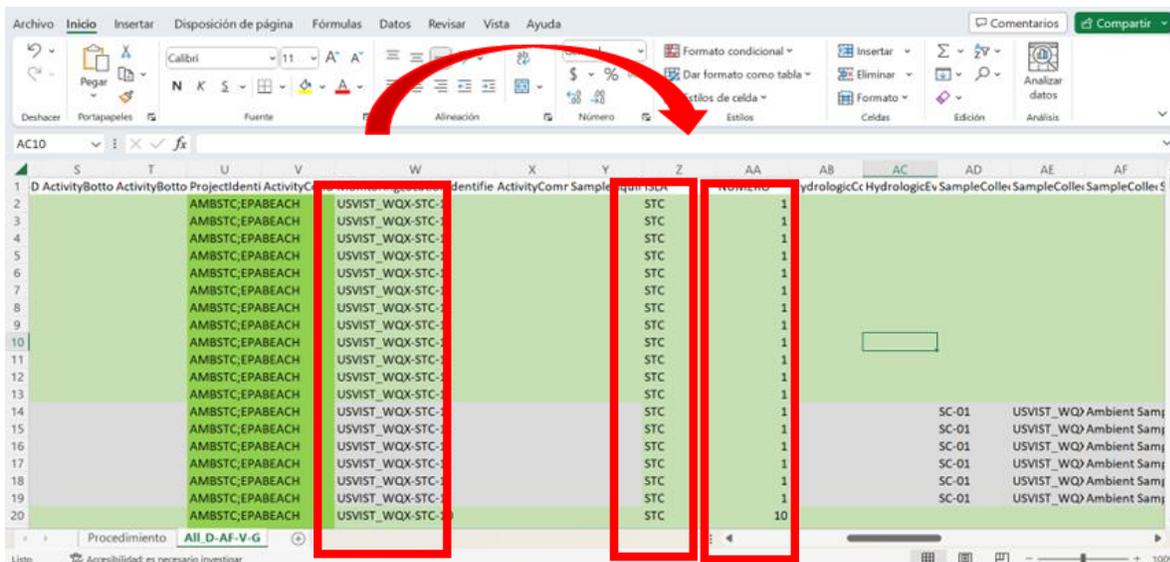
*Nota.* Todas las columnas D, AF, V y G se encontraban organizadas en una sola hoja del Excel.  
Elaboración propia.

## Etapa II: Separación y conversión de datos.

En esta etapa se obtuvo la separación de las variables por año e isla como se observa en la figura 20, se hizo una conversión en metros (m) a todos los datos que se encontraba en otra medida distinta, con el fin de tener los datos estandarizados.

**Figura 20**

*Separación de variables*



*Nota.* Se realiza la separación de la información de la columna W. Elaboración propia.

### **Eta III: Depuración de variables.**

Al determinar la cantidad de datos con los que se continuó trabajando y su debida estandarización en cada periodo, se eliminaron las columnas que no contaban con información necesaria para la investigación o que se encontraban vacías. El documento contaba con 63 columnas y se eliminaron 55 columnas; por lo anterior quedaron 8 columnas como se muestra en la tabla 2.

**Tabla 2**

*Columna Finales*

Columna	Nombre
A	Año
B	Mes
C	Profundidad (m)
D	Isla
E	Ubicación
F	Característica
G	Valor de medida
H	Unidad de medida

Este documento quedó listo para trabajar cualquier parámetro que se quiera analizar.

Para esta etapa se continuó trabajando con las columnas organizadas; se logró minimizar los datos depurando los sitios no definidos de las Islas Vírgenes. Teniendo en cuenta los datos eliminados por característica, se realizó una tabla de comparación de números de datos iniciales, número de datos eliminados y número de datos finales para cada periodo, Como se observan en la tabla 3.

**Tabla 3**

Años	Datos iniciales	Datos eliminados	Datos finales
1975-2000	45236	38629	6607
2006-2010	119261	60267	58994
2011-2021	139963	19038	120925
Total	304460	117934	186526

*Nota.* El total de datos analizados correspondió al 43.6 % de la población.

#### **Etapa IV: Comparación de datos con datos de referencia.**

En este estudio se eligió trabajar con las características del pH y oxígeno disuelto, teniendo en cuenta el valor de medida de cada una de ellas; para el pH el rango estándar es entre 7 a 10,21 y el del oxígeno disuelto entre 1 mg/l a 20 mg/l. Teniendo en cuenta los rangos estándar de pH y oxígeno disuelto, se eliminaron las filas que no contenían datos en esos intervalos, se realizaron unas tablas de comparación de números de datos iniciales , número de datos eliminados y número de datos finales para cada periodo con cada característica como se muestra en las tablas 4,5,6,7,8 y 9.

#### **Característica del pH.**

**Tabla 4**

*Periodo 1975 al 2000.*

Datos	Cantidad
Iniciales	876
Eliminados	2
Finales	874

**Tabla 5**

*Periodo 2006 al 2010.*

Datos	Cantidad
Iniciales	1975
Eliminados	902
Finales	853

**Tabla 6**

*Periodo 2011 al 2021.*

Datos	Cantidad
Iniciales	1475
Eliminados	283
Finales	1192

**Característica Oxígeno Disuelto**

**Tabla 7**

*Periodo 1975 al 2000.*

Datos	Cantidad
Iniciales	1611
Eliminados	0
Finales	1611

**Tabla 8**

*Periodo 2006 al 2010.*

Datos	Cantidad
Iniciales	645
Eliminados	129
Finales	516

**Tabla 9**

*Periodo 2011 al 2021.*

Datos	Cantidad
Iniciales	9375
Eliminados	21
Finales	9354

### **Etapa V: Registro de datos en tablas**

Al realizar la eliminación de filas del pH y el oxígeno disuelto, se realizó una tabla en la que se organizó la cantidad de datos que tenía cada Isla que incluyó el año, el mes, el promedio, la desviación estándar y el coeficiente de variación para las dos características pH y oxígeno disuelto. Una vez se obtuvo los parámetros del pH y oxígeno disuelto, se identificó que en el pH en los años 2006 y 2007 no se encontraban datos (Figura 21). En el parámetro de oxígeno en los años 1975 hasta al 1998 tampoco se encontraron datos (Figura 22). Estas tablas quedaron listas para la realización de la gráfica de los años 1999 al 2021.

**Figura 21**

*Tabla de la característica del pH 2006 al 2010*

AÑO	ISLA	MES	CANTIDAD	PROMEDIO	DESVEST.M	COEFICIENTE
2008	STC	3	1	7,90		
	STC	12	56	8,21	0,20	2,42
	STJ	6	2	9,66	0,78	89,13
	STT	3	1	7,18		
	STT	6	1	7,08		
	STT	8	1	7,38		
	STT	12	18	7,21	0,06	0,86
TOTAL			80			
2009	STC	9	6	7,27	0,12	1,59
	STC	10	8	7,33	0,19	2,54
	STC	11	6	8,28	0,14	1,70
	STC	12	6	7,85	0,06	0,76
	STJ	3	14	7,86	0,04	0,53
	STJ	6	26	7,49	0,15	2,03
	STJ	9	4	7,25	0,14	1,88
	STJ	10	10	7,43	0,42	5,60
	STJ	11	4	8,14	0,02	0,21
	STJ	12	4	7,67	0,04	0,55
	STT	3	112	7,77	0,14	1,75
	STT	6	120	7,51	0,30	4,05
	STT	9	82	7,54	0,09	1,19
	STT	10	12	7,42	0,10	1,32

*Nota.* No se encuentra datos en los años 2006 al 2007. Elaboracion propia.

**Figura 22**

*Tabla del parámetro Oxígeno Disuelto 1975 al 2000*

AÑO	ISLA	MES	PROMEDIO	CANTIDAD	DESVEST.M	COEFICIENTE
1999	STJ	3	7,91	38	0,07	0,88
	STJ	6	8,06	24	0,07	0,81
	STJ	11	7,91	58	0,17	2,12
	STT	3	8,12	38	0,08	0,93
	STT	4	8,06	74	0,12	1,55
	STT	6	7,97	110	0,06	0,77
	STT	10	7,90	3	0,00	0,00
	STT	11	7,99	60	0,13	1,59
	STT	12	7,92	112	0,14	1,78
TOTAL				531		
2000	STJ	10	8,00	34	0,00	0,00
	STJ	11	8,26	48	0,07	0,85
	STT	6	8,20	3	0,00	0,00
	STT	8	8,15	3	0,00	0,00
	STT	10	8,02	120	0,09	1,18
	STT	11	8,10	3	0,00	0,00
	STT	12	8,16	132	0,09	1,09
TOTAL				343		

*Nota.* No se encuentra datos en los años 1975 al 1998 . Elaboracion propia.

## Etapa VI: Graficas

Para la realización de las gráficas del parámetro de calidad de agua del pH y el oxígeno Disuelto, se realizaron dos tablas; una donde se relacionaba la Isla, el año, el mes, y el promedio, de cada parámetro y en la otra se relacionaba el promedio total de cada mes durante los años 1998 al 2021 con su respectiva cantidad de datos. (Figura 23, Figura 24)

**Figura 23**

*Promedio y cantidad total de los meses.*

CARACTERISTICA pH		
MES	PROMDEIO	CANTIDAD
ENERO	8,05	640
FEBRERO	8,05	1248
MARZO	8,03	2774
ABRIL	8,02	539
MAYO	8,13	1802
JUNIO	8,06	2702
JULIO	8,01	1542
AGOSTO	8,09	1286
SEPTIEMBRE	8,14	2343
OCTUBRE	8,01	769
NOVIEMBRE	8,05	1350
DICIEMBRE	8,21	2515

**Figura 24**

*Promedio y cantidad total de los meses.*

CARACTERISTICA OXIGENO DISUELTO		
MES	PROMEDIO	CANTIDAD
ENERO	7,52	745
FEBRERO	7,62	1793
MARZO	7,29	3451
ABRIL	9,96	539
MAYO	9,42	1934
JUNIO	6,93	4037
JULIO	7,25	2337
AGOSTO	8,32	1555
SEPTIEMBRE	7,21	3001
OCTUBRE	7,28	910
NOVIEMBRE	7,53	1341
DICIEMBRE	6,95	3385

### Interpretacion de graficas.

#### Parámetro pH.

- Los valores de la característica del pH en los años del 1998 al 2021 alcanzaron un valor mínimo de 7,31 en Julio del 2015 y un valor máximo de 9,00 en diciembre del 2014. Figura 25
- Los datos oscilaron mucho más entre noviembre del año 2005 hasta septiembre del 2015. Figura 25
- La gráfica de la figura 26, nos muestra el promedio total de todos los meses, de los años del 1998 al 2021, alcanzando un valor mínimo de 8,00 en marzo y un valor máximo de 8,21 en diciembre; se observa que los datos se encuentran en promedio de 8,00.

Figura 25

Grafica histórica característica pH.

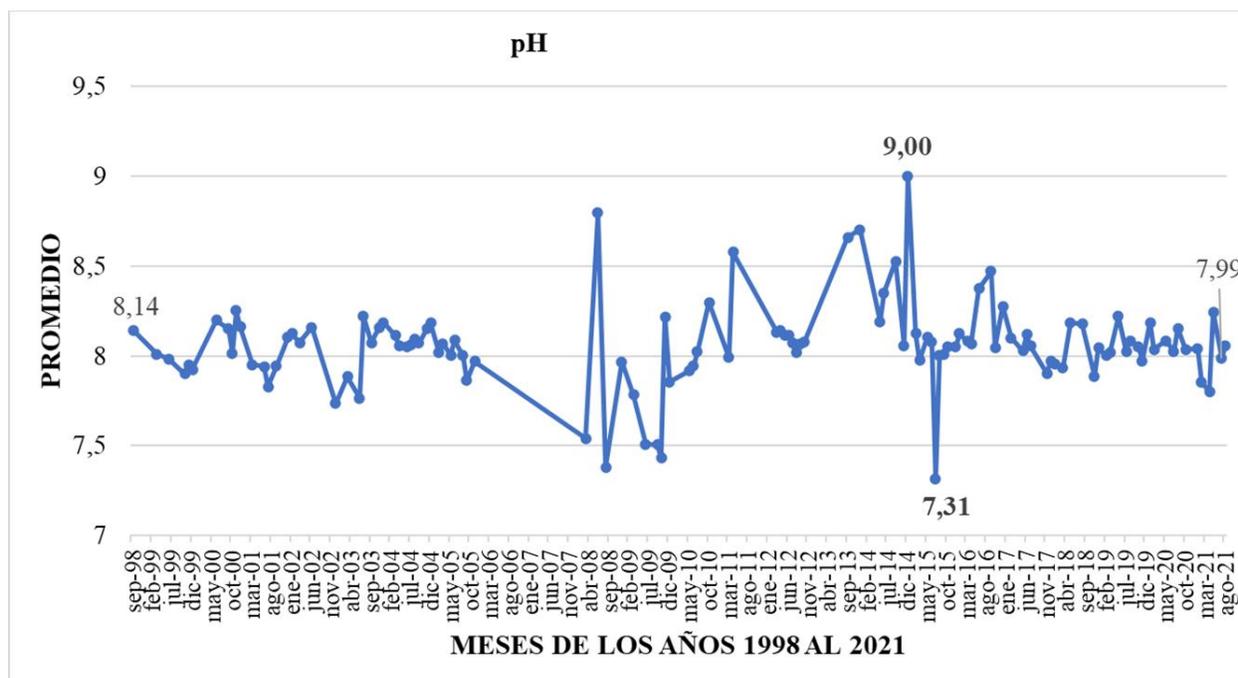
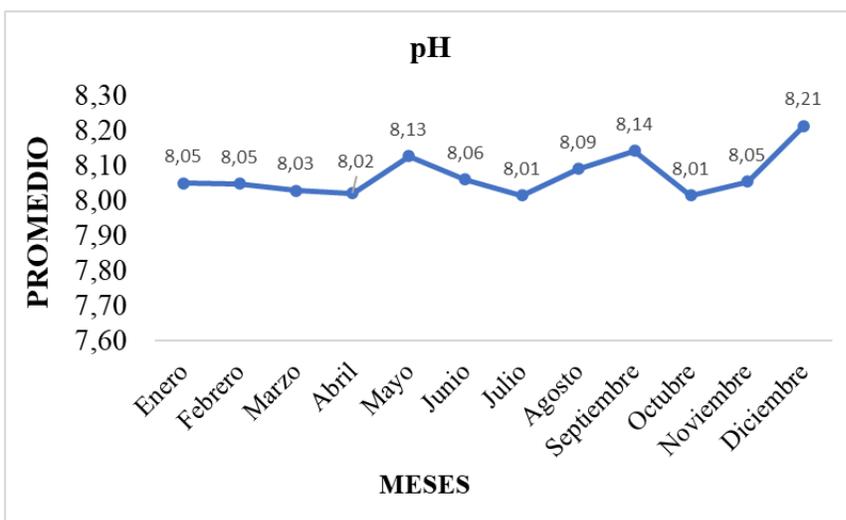


Figura 26

Grafica de pH promedio mensual para el periodo 1998-2021



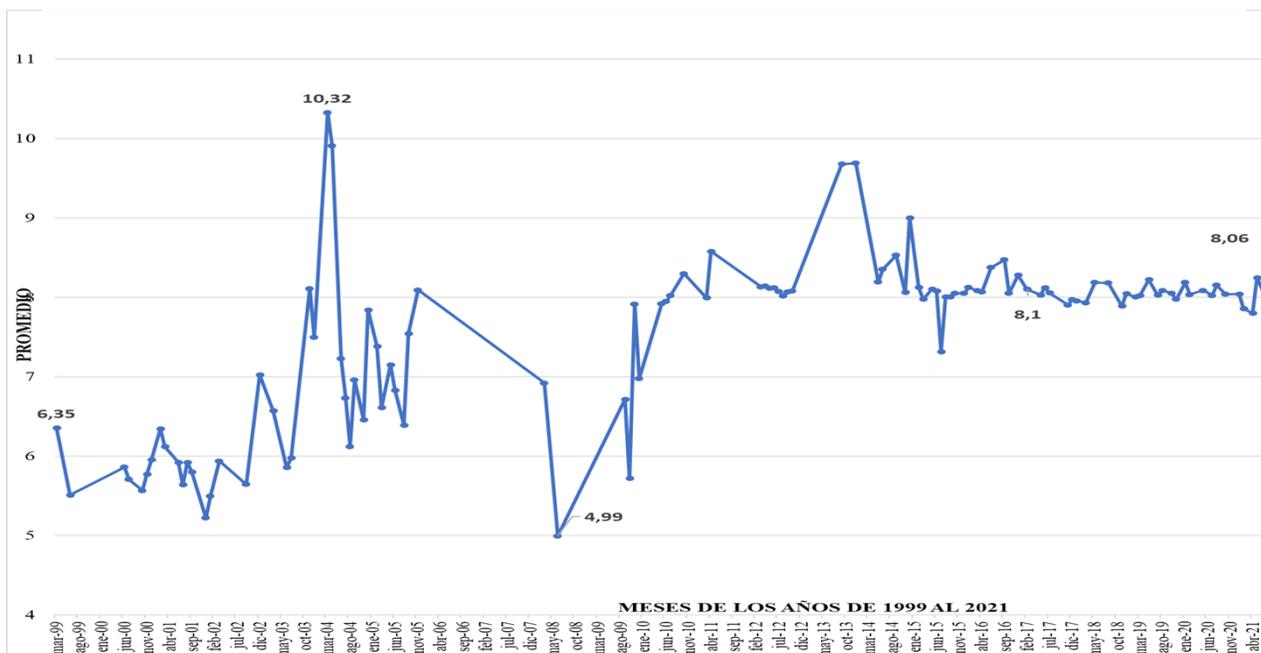
El promedio del pH encontrado en este estudio fue entre 8,01 a 8,21 unidades; similar al reportado por (Delvalls, 1999, pág. 359). En zonas distantes de la costa el cual encontró valores de pH de 8,3 unidades.

### **Parámetro Oxígeno Disuelto**

- Los valores de la característica del Oxígeno Disuelto en los años del 1999 al 2021 alcanzaron un valor mínimo de 4,99 mg/l en junio del 2008 y un valor máximo de 10,32 mg/l en marzo del 2004 como se muestra en la figura 27.
- Los datos oscilaron mucho más desde marzo del año 1999 hasta diciembre del 2016.
- Desde febrero del año 2017 hasta agosto del año 2021, los datos tienen menos variación por lo tanto se encuentran en promedio sobre 8 mg/l. figura 27.
- La gráfica de la figura 28, nos muestra el promedio total de todos los meses, de los años del 1998 al 2021, alcanzando un valor mínimo de 6,93 en junio y un valor máximo de 9,96 en abril; también se observa que de marzo a agosto los datos oscilaron más y que desde enero a marzo y de septiembre a diciembre tienen menor variación. En general, los datos se encuentran en promedio de 7,21 mg/l a 7,62 mg/l.

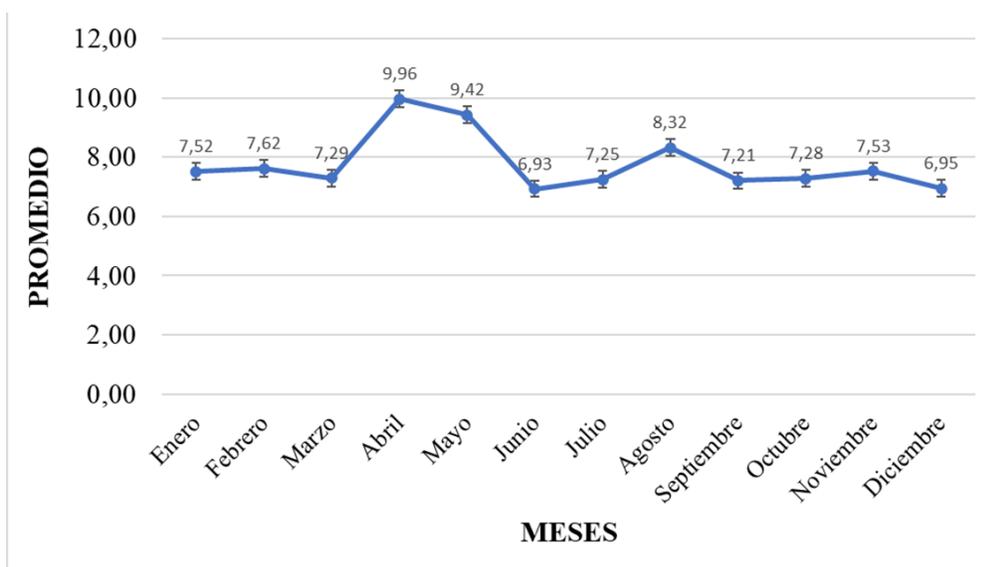
**Figura 27**

*Grafica histórica característica oxígeno disuelto (OD), 1999-2021*



**Figura 28**

*Grafica de oxígeno disuelto (OD) promedio mensual para el periodo 1999-2021*



## 8. Análisis de resultados

El presente capítulo de la investigación muestra los análisis de los resultados del estudio referente a los datos suministrados de calidad de agua de las islas vírgenes de Estados Unidos, haciendo uso del programa Excel donde se mostró gráficas y tablas, de tal forma que permite alcanzar una valoración total de dicho estudio.

El uso del software Excel nos permitió avanzar rápido mediante procesos que están basados en fórmulas como sumas, promedios, coeficiente de variación, entre otras. Se caracteriza por ser un medio matemático que permite el análisis de variabilidad de información numérica y categórica.

Es conveniente expresar que la estadística juega un papel muy importante en el análisis de datos, debido a esto, no se puede estudiar una base de datos sin un buen conocimiento de la estadística.

En esta investigación se obtuvo que el lineamiento para analizar los datos está comprendido en VIII etapas, las cuales fueron:

- I. Organización de datos.
- II. Separación y conversión de datos.
- III. Depuración de variables.
- IV. Comparación de datos con variables de referencia.
- V. Registro de datos en tablas.
- VI. Graficas.
- VII. Analisis.
- VIII. Presentación de resultados.

A continuación, se describen cada una de las etapas que se deben de tener en cuenta para el análisis de datos:

## I ORGANIZACIÓN DE DATOS

En esta primera etapa la función es organizar las variables que se van a utilizar durante todo el análisis, con el fin de observar los cambios que se irán realizando, además es importante usar estrategias como colores, donde se identifiquen las diferentes variables; se recomienda guardar una copia de seguridad en cada etapa.

## II SEPARACIÓN Y CONVERSIÓN DE DATOS

La importancia de depurar los datos fue identificar qué valores ausentes no se tienen en cuentan en el análisis; también las celdas que se encuentran vacías, teniendo en cuenta que la base de datos se reduce.

En la separación de los datos fue importante reestructurar columnas de datos que contenían dos variables; como por ejemplo en el parámetro fecha de actividad, donde está registrado el año y mes de las diferentes islas, se separó el año en una nueva columna y el mes en otra columna. La separación se realizó con el fin de determinar las debidas funciones en ellas.

Es fundamental realizar una estandarización de datos para trabajar los datos con una sola unidad de medida y evitar errores profundos en el análisis trabajando datos con diferentes medidas; por ejemplo, pies y metros.

### III

### DEPURACIÓN DE VARIABLES

Es importante reconocer qué variables no se van a utilizar en la investigación para eliminarlas y así minimizar la base de datos.

### IV

### COMPARACIÓN DE DATOS CON VARIABLES DE REFERENCIA

En esta etapa se escogen la característica que se quiera analizar de calidad de agua. En este estudio se realizaron algunos procedimientos como:

- Filtraciones para mirar las celdas que se encontraban vacías.
- Eliminación de fallos en la transcripción o medidas estándares de las características. Se debe tener en cuenta el contexto donde se encuentra la variable a analizar, de acuerdo con el estándar definido para cada variable. por ejemplo, la medida estándar del Ph es de 7,00 a 10,21 unidades y algunos datos registraban pH de 0,1 unidades.

### V

### REGISTRO DE DATOS EN TABLAS

Es muy importante tener en cuenta la relación de las variables para registrar los datos en tablas, así teniendo información clara y concisa de lo que se quiere graficar.

## VI GRAFICAS

Es una herramienta que permiten representar de manera accesible información facilitando la comprensión de las variables. Se debe identificar las variables dependiente e independiente para graficar. Ejemplo en este estudio se relacionó la variable independiente(meses) y la variable dependiente(promedio).

Para el análisis de datos se consideraron importante también las siguientes dos tapas:

## VII ANÁLISIS

Es un proceso fundamental que realiza el investigador donde deberá describir en detalle los comportamientos de los datos y de los parámetros que se tienen en cuenta para analizar.

## VIII PRESENTACIÓN DE RESULTADOS

Ya obtenido los análisis y resultados de la investigación, es importante socializarlos mediante ponencias, talleres, póster y conferencias para dar a conocer los logros obtenidos en el estudio. Los lineamientos que se obtuvieron en la investigación se sustentan en el estudio

(Exploratorio, 2021, pág. 6); Los cuales son:

1. Análisis descriptivo
2. Ajustes de tipo de variable
3. Detección y tratamiento de datos ausentes
4. Identificación de datos atípicos
5. Correlación de variables

Teniendo en cuenta los lineamientos anteriores, se evidencia que tiene 5 etapas a seguir en el análisis de datos, y en nuestro estudio se determinaron 8 etapas.

En la primera etapa se diferencia que el autor inicia con una descripción de datos, a diferencia de nuestra etapa, la cual se inicia ordenando los datos y a su vez haciendo su debida descripción. Las otras etapas realizadas por el autor se tuvieron en cuenta en el estudio de esta investigación; sin embargo, no se determinaron en el mismo orden.

## 9. Conclusiones

- Se logró establecer un manual de guía para el análisis de datos usando como modelo el análisis de calidad de agua de las Islas Vírgenes, San Croix, San John y San Thomas, a través de la herramienta Excel.
- Se caracterizaron los datos de calidad de agua en ocho variables las cuales fueron: Año, mes, isla, ubicación, profundidad, característica, valor de medida y unidad de medida. En la caracterización emergen los lineamientos del análisis de datos que incluyeron la organización; separación y conversión de datos y depuración de variables.
- Se obtuvo la identificación, clasificación y estandarización de las variables que se estudiaron en los parámetros del pH y el oxígeno disuelto y se realizó la comparación de datos con valores de referencia.
- Se hizo la representación gráfica de las características de calidad de agua del pH y el oxígeno disuelto, evidenciando las etapas realizadas en el análisis de datos.

## 10. Sugerencias

- Tener los datos originales, guardar copia de cada procedimiento realizado.
- Identificar las variables que se están manipulando mediante estrategias pedagógicas (por ejemplo, colores para identificar cambios en las variables).
- Realizar un registro de los pasos que se van efectuando en cada etapa.

## 11. Referencias

- Centro de Escritura Javeriano. (2020). *Normas APA, séptima edición*. Pontificia Universidad Javeriana, seccional Cali. <https://www2.javerianacali.edu.co/centro-escritura/recursos/manual-de-normas-apa-septima-edicion#gsc.tab=0%C2%A0>
- Batanero, J. D. (1992). El análisis de datos como útil y objeto de la didáctica de la matemática. 3. Obtenido de <https://www.ugr.es/~batanero/pages/ARTICULOS/ANALISIS.pdf>
- Cifuentes. (Torres & Frías ,1997). *EL OCÉANO Y SUS RECURSOS. II. LAS CIENCIAS DEL MAR: OCEANOGRAFÍA GEOLÓGICA Y OCEANOGRAFÍA QUÍMICA*. Obtenido de <http://www.bio-nica.info/biblioteca/Cifuentes1997b.pdf>
- Delvalls. (1999). MEDIDAS EN CONTINUO DE pH EN CONDICIONES DE AFLORAMIENTO: LA CORRIENTE DE CALIFORNIA. 359. Obtenido de <https://www.redalyc.org/pdf/480/48025303.pdf>
- Exploratorio, A. (2021). Guía práctica de introducción al Análisis Exploratorio de Datos. 6. Obtenido de [https://datos.gob.es/sites/default/files/doc/file/analisis\\_exploratorio\\_de\\_datos\\_2021\\_v6.pdf](https://datos.gob.es/sites/default/files/doc/file/analisis_exploratorio_de_datos_2021_v6.pdf)
- Gamboa. (2007, citado en Torres). PROPUESTA METODOLÓGICA PARA MEJORAR EL APRENDIZAJE DE ESTADÍSTICA UTILIZANDO EL SOFTWARE EXCEL. 17. Obtenido de <https://dspace.ups.edu.ec/bitstream/123456789/16029/1/UPS-CT007771.pdf>
- Hernandez. (Fernández, Bapista, 2010). *Biblioteca virtual*. Obtenido de <https://www.eumed.net/tesis->

doctorales/2012/eal/metodologia\_cuantitativa.html#:~:text=La%20metodolog%C3%ADa%20cuantitativa%20de%20acuerdo,o%20fen%C3%B3meno%20objeto%20de%20estudio

KeepCoding, R. (19 de Enero de 2022). *Keepcoding*. Obtenido de <https://keepcoding.io/big-data/herramientas-de-data-analytics-mas-usadas/>

Palma. (20005, citado por Vivanco, 2017). Obtenido de <http://scielo.sld.cu/pdf/rus/v9n3/rus38317.pdf>

Palma. (2005 citado en alva santos, 2017). Análisis de los datos e interpretación de los resultados. 1. Obtenido de [https://cmapspublic2.ihmc.us/rid=1177276899217\\_1477413697\\_5143/analisdatosinterpretac-1.pdf](https://cmapspublic2.ihmc.us/rid=1177276899217_1477413697_5143/analisdatosinterpretac-1.pdf)

Peña. (26 de Junio de 2007). CALIDAD DE AGUA OXIGENO DISUELTO (OD). 2. Obtenido de <https://www.dspace.espol.edu.ec/bitstream/123456789/6162/5/Investigacion.pdf>

Peña, S. (2017). *Análisis de datos*. Fundación Universitaria del Área Andina. Obtenido de <https://core.ac.uk/download/pdf/326425169.pdf>

peterhermesfurian. (( s. f. )). *123RF*. Obtenido de 123RF: [https://es.123rf.com/photo\\_38616580\\_islas-v%C3%ADrgenes-de-los-estados-unidos-mapa-pol%C3%ADtico-un-grupo-de-islas-en-el-caribe-que-es-un-%C3%A1rea-insular-d.html](https://es.123rf.com/photo_38616580_islas-v%C3%ADrgenes-de-los-estados-unidos-mapa-pol%C3%ADtico-un-grupo-de-islas-en-el-caribe-que-es-un-%C3%A1rea-insular-d.html)

Ramirez, A. M. (2017). *USOS Y APLICACIONES DEL EXCEL*. Obtenido de

[https://repository.icesi.edu.co/biblioteca\\_digital/bitstream/10906/83888/1/TG02061.pdf](https://repository.icesi.edu.co/biblioteca_digital/bitstream/10906/83888/1/TG02061.pdf)

Sabino, C. (2003). *El Proceso de Investigación*. Editorial Panapo de Venezuela. 103. Obtenido

de [http://paginas.ufm.edu/sabino/ingles/book/proceso\\_investigacion.pdf](http://paginas.ufm.edu/sabino/ingles/book/proceso_investigacion.pdf)

Santillán, M. L. (23 de Junio de 2013). *Publimetro* . Obtenido de

<https://www.publimetro.com.mx/mx/columnaunam/2013/06/24/estadistica-nuestra-vida-diaria.html>

Santos. ((s.f.)). *ANÁLISIS DE LOS DATOS E INTERPRETACIÓN DE LOS RESULTADOS*.

Obtenido de

[https://cmapspublic2.ihmc.us/rid=1177276899217\\_1477413697\\_5143/analisdatosinterpretac-1.pdf](https://cmapspublic2.ihmc.us/rid=1177276899217_1477413697_5143/analisdatosinterpretac-1.pdf)