



**TRABAJO DE GRADO**  
**Opción Seminario-Diplomado.**

Analizando Datos

Corporación Universitaria Remington.

Ingeniería de sistemas

Visualización de datos

Estudiante:

Jonathan Fernando Calvache

Profesor:

John Fredy Mira Mejía

Noviembre de 2023

## Tabla de Contenido

Resumen.....	3
Objetivo General.....	5
Marco contextual .....	7
Que son los datos .....	8
Tipos de datos .....	8
Como se guardan los datos en la actualidad .....	9
DAS.....	9
NAS.....	9
Analizar Datos .....	10
Que se puede encontrar en los datos .....	11
Herramientas en el análisis de datos .....	11
Métodos de análisis de datos.....	12
Análisis de datos e inteligencia artificial .....	13
Perfil del analista de datos .....	13
Desarrollo e implementación del aprendizaje.....	14
Ejercicio 1 .....	15
Ejercicio 2 .....	21
Imágenes .....	36
Conclusiones.....	38
Bibliografía .....	39

## Resumen

En el siguiente documento revisaremos que es el análisis de datos, iniciando con la breve explicación, para que sirven porque es importante tener datos y analizarlos, explicaremos como se guardan, que tipos de datos se manejan cual es la diferencia entre ellos y como se guardan en los equipos electrónicos; aquellos que se utilizan mucho hoy en día, explicaremos que hay muchas aplicaciones o herramientas para analizar los datos, que es importante guardarlos en un software especializado; que nos ordene la información y nos de la facilidad de extraerlos sin realizar ningún cambio, al realizar dicha operación.

Las bases de datos encargada de guardar la información de una forma estructurada nos agilizan el tema de análisis, la información debe ser clara concisa, hace que el análisis sea más claro.

Como en la mayoría de las ciencias, la capacidad de analizar datos tiene sus procesos y métodos; que nos ayudan a encontrar más rápido la información que deseamos, explicaremos en que consiste el método ya que este nos ayuda a trazar una ruta específica, en el documento; se observa dos ejercicios de análisis donde se explica el funcionamiento de dos de los programas que se utilizan para analizar datos, Excel y power bi, nos muestra cómo podemos encontrar de forma rápida información relevante y nos permite crear diagramas de fácil explicación.

Tanto Excel como Power bi, herramientas de Microsoft son aplicaciones que pueden trabajar separadas y juntas volviéndose una de tantas herramientas poderosas.

El tiempo que se gasta en analizar datos es ínfimo en comparación a la información relevante que encuentras en los datos.

**Palabras clave**

**Pronostico:** *“Es la predicción de lo que sucederá con un elemento determinado dentro del marco de un conjunto dado de condiciones. Se diferencia del presupuesto porque este último es el resultado de decisiones encaminadas a generar las condiciones que propiciarán un nivel deseado de dicho elemento”*  
(Gestiopolis, s.f.)

**Cuantitativo:** *“El concepto cuantitativo tiene relación directa con cantidad, por lo tanto, sus variables son siempre medibles. El concepto cualitativo tiene relación directa con calidad, por lo tanto, sus variables son siempre interpretativas”*.  
(Significados, s.f.)

**Cualitativo** *“Una cualidad es una propiedad que existe en cualquier objeto, individuo, entidad o estado, la cual se puede analizar comparándolo con otro semejante. Así, lo cualitativo depende de la percepción social, cultural o subjetiva del objeto”* (Significados, s.f.)

**Método** *“Se entiende entonces como una serie de pasos que se deben seguir para cumplir un objetivo.  
En términos amplios, el método puede definirse como la forma en la que una persona está habituada a realizar una tarea. Es decir, puede tratarse del orden de acciones que sigue Roberto para prepararse todas las mañanas y estar listo para salir al trabajo”*. (economipedia.com, s.f.)

**Estadística** *“Es una rama de las matemáticas que te permite recopilar, organizar y analizar datos según la necesidad que tengas, por ejemplo: obtener un resultado, comparar información, tomar mejores decisiones, entre muchas cosas más”*  
(gcfglobal, s.f.)

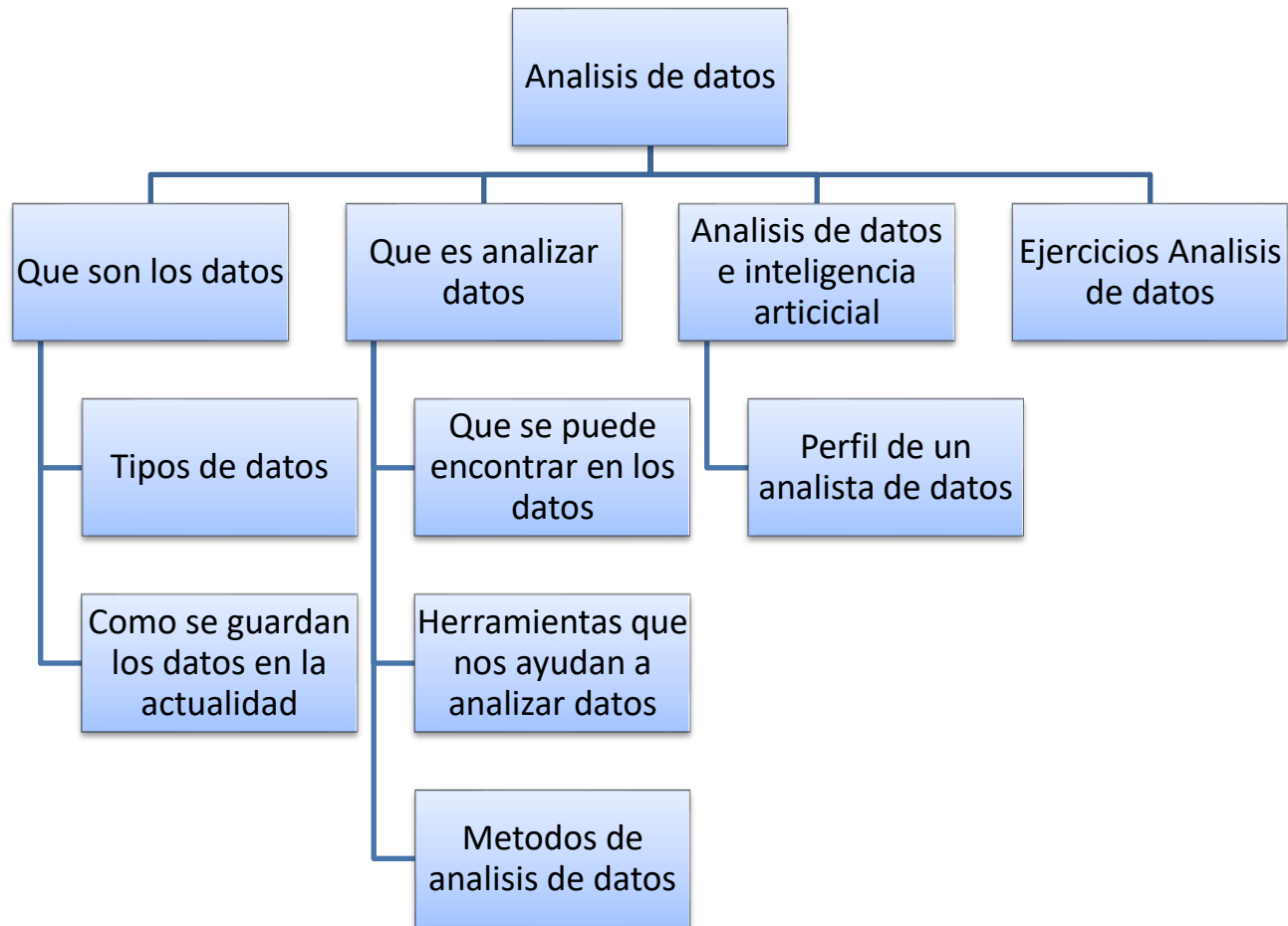
## **Objetivo General**

Comprender que son los datos, entender porque guardar los datos solo es un paso en el mundo actual, Analizar porque es importante darle valor a los datos que tenemos a nuestro alrededor.

### **Objetivos específicos**

1. Como se guardan los datos, los tipos de datos y como se guardan en la actualidad
2. Porque es importante analizar los datos y que podemos encontrar en ellos
3. Conocer las herramientas para analizar datos y los métodos que se utilizan para analizarlos.
4. Conocer la función de un analista de datos y como la inteligencia artificial puede mejorar el análisis en los datos

Imagen 1. Mapa conceptual



### **Marco contextual**

Al escuchar la palabra “Análisis de datos” nos parece una actividad muy dura de realizar o que son muy pocas las personas que hacen esta actividad y que se debe tener un amplio conocimiento. Aunque la persona que se dedica a este proceso debe tener buenas bases en manejo de información, analizar datos no es nada del otro mundo como se escucha, Este procedimiento como toda ciencia; tiene su investigación sus métodos, esta ciencia que nace de la necesidad de conocer las actividades explorar los datos que se crean a diario en equipos informáticos, que pueden ser de una empresa o actividad y que junto con la evolución de la tecnología afianza temas como estadística pronósticos varianza promedios entre otros que hacen que analizar los datos se vuelva una experiencia muy comfortable.

Para analizar los datos en el mundo actual tenemos muchas herramientas, aplicaciones que nos permiten detallar mejor la información; analizarla entenderla, encontrar información relevante que todas las personas interesadas la comprendan; es deber del analista mostrar la información que forma fácil, y no tan saturada, generalmente los analistas presentan informes a sus lideres compañeros en forma de imagen, donde las imágenes tienen el trabajo de mostrar una información veraz y sencilla fácil de entender.

## Que son los datos

Según Wikipedia: *“Un dato es una representación simbólica (numérica, alfabética, algorítmica, espacial, etc.) de un atributo o variable cuantitativa o cualitativa. Los datos describen hechos empíricos, sucesos y entidades. Es un valor o referente que recibe el computador por diferentes medios, los datos representan la información que el programador manipula en la construcción de una solución o en el desarrollo de un algoritmo”* (Wikipedia, 18)

En breve descripción los datos, no son más que una información comprensible para las personas que guardan algo; en un código conocido por ellos, en varios casos puede ser la escritura, a través de los libros, las historias entre otras, que ayudan a mantener la información que ocurrió en una época anterior y se quiere comprender o enseñar.

Actualmente con herramientas electrónicas podemos decir que esos datos están guardados en códigos entendibles para una maquina y traducidos para las personas en un lenguaje comprensible.

## Tipos de datos

Aunque las máquinas electrónicas son codificadas para entender los pulsos electrónicos y guardarlas como información, los seres humanos definimos y separamos como guardar los datos, estos son:

1. Texto
2. Numérico
3. Lógico

Cada uno de ellos con dependencias inferiores son el conjunto de datos que se guarda en la actualidad y se separan para un mejor entendimiento.

## **Como se guardan los datos en la actualidad**

Aunque existen muchas formas de guardar los datos, en este documento nos centraremos específicamente en conocer la forma electrónica y software de guardar datos.

Podemos resumir que actualmente toda la información es guardada en herramientas electrónicas con pulsos eléctricos muchas de estas herramientas conocidos como discos duros, mecánicos ópticos. USB tarjetas entre otras, están configurados con códigos basados en pulsos electrónicos un ejemplo será código binario, código hexadecimal, configuraciones ya realizadas en la máquina para que pueda entender que tipo de información guarda y pueda ser devuelta de la misma forma.

Actualmente existen 2 tipos de almacenamiento de datos electromecánicamente

### **DAS**

Sigla de "Direct Attached Storage", en español "Almacenamiento adjunto directo". Esta categoría incluye todo el almacenamiento en el que existe una conexión física entre el hardware donde se almacenan los datos y la computadora que los lee. Es el caso de los discos duros (HD), memorias USB, SSD, entre otros.

### **NAS**

Sigla de "Network Attached Storage" o "Almacenamiento adjunto en la Red". En este caso, el hardware en el que se almacena la información es diferente al que se accede, lo que permite que varias computadoras accedan a los mismos datos de forma remota.

(rockcontent.com, 2023)

Guardar información de manera desordenada no lleva a nada, solo a tener un montón de información que no puedes utilizar, un ejemplo claro de no organizar y de no tener software especializado para guardar los datos sería; guardar facturas de una empresa en formato pdf en varias carpetas, estas guardando datos importantes para la empresa, pero encontrar una factura en específico sería cada vez más complejo a medida que vaya aumentando las ventas.

Por tal razón, aunque es importante guardar los datos, también es importante tener el software especializado para guardar esos datos.

Una gran cantidad de datos es llamada colección de datos o bases de datos, y se puede guardar en programas que se especializan en ello, estas pueden ser:

- Base de datos relacional vs no relacional
- Bases de datos orientadas a documentos
- Bases de datos de clave-valor
- Bases de datos orientadas a grafos
- Bases de datos columnares

Todas ellas tienen algo en común, tienen la capacidad de dejar extraer en forma rápida la información que se necesita de forma ordenada.

### **Analizar Datos**

Como bien lo vimos en la parte superior de este documento, los datos de una empresa, personas o sistemas, son guardadas en aplicaciones especializadas, tener esos datos guardados sirven para revisar historiales en el ejemplo de una empresa la cantidad de ventas realizadas en un día en específico, los datos guardados se convierten en una información vital si se analiza el comportamiento de cada uno de ellos, dependiendo de las necesidades de una empresa, esta puede apalancarse en la información que tenga guardada

en sus bases de datos para realizar tomas de decisiones frente a sus procesos, Analizar datos es encontrar información útil que le pueda servir a alguien.

*El análisis de datos es un estudio que usa información en su estado natural, la analiza y extrae ideas útiles para la toma apropiada de decisiones dentro de una empresa. (ebac.mx)*

### **Que se puede encontrar en los datos**

Al analizar los datos; podemos encontrar diferente información vital para la empresa entre ellos tenemos tendencias lo que ocurrió porque sucedió, puede volver a pasar, diagnóstico de eventos para repetir o no repetir, cuando tenemos datos cuantitativos podemos encontrar objetivos medibles, realizar estudios estadísticos gráficos, Análisis Predictivos, Prescriptivos.

### **Herramientas en el análisis de datos**

Como el análisis de datos es encontrar información relevante; de datos agrupados en una base de datos en la mayoría de los casos, La persona que realiza esta función debe ayudarse con software especializado para este fin, Aunque un análisis se puede realizar sin una herramienta especializada para tal fin, tener una que sea dedicada para este trabajo ayudara a realizar más rápido el trabajo.

Alguna de las herramientas que ayudan a analizar datos

1. *El cuadrante mágico de Gartner como punto de partida*
2. *Microsoft Power BI*
3. *Tableau*
4. *Qlik*

5. *ThoughtSpot*
6. *MicroStrategy*
7. *Looker*
8. *Tibco Spotfire*
9. *Sisense*
10. *SAS*
11. *Salesforce*
12. *Microsoft Excel*
13. *Google hojas de calculo*

(IMF blog de Tegnologia, 2023)

### **Métodos de análisis de datos**

Tener un método para analizar los datos es una guía fundamental para encontrar buenos resultados.

1. Determinar dirección
2. Establecer parámetros
3. Datos agregados
4. Seguimiento del progreso
5. Eliminar datos irrelevantes
6. Ejecutar análisis estadístico
7. Establecer prácticas de administración de dato
8. Involucrar la tecnología
9. Emplear visualización de datos
10. construcción de narrativas

(zipreporting, 2023)

## **Análisis de datos e inteligencia artificial**

El ingreso de la Inteligencia artificial a los negocios es imparable, y el ingreso de esta herramienta al análisis de datos no es la excepción, esta herramienta nos ayudara a automatizar más los procesos que tenemos y mejorarlos.

Según el sitio web (geekflare.com) la inteligencia artificial tiene las siguientes ventajas frente a los analistas experimentados.

*Sin enfoque de código para el análisis de datos.*

*Cualquier persona con conocimientos básicos de matemáticas y estadísticas, junto con ideas de gestión empresarial, puede generar conocimientos prácticos a partir de herramientas de análisis de datos de IA.*

*Estas herramientas de IA reducen la necesidad de analistas de datos y científicos dedicados a unos pocos para que pueda ahorrar en costos de personal.*

*Las herramientas de análisis de datos de IA pueden explorar patrones ocultos e información procesable de los datos de su negocio en segundos. Esto llevaría meses a los científicos de datos humanos.*

La inteligencia artificial no sacara del juego en el momento a las personas que trabajan con análisis de datos, esta es una herramienta adicional que se puede utilizar para mejorar los procesos y dar mejores resultados.

### **Perfil del analista de datos**

Un analista de datos es aquella persona que recoge gran cantidad de información de una empresa o sistema y bajo el análisis de esa información permite crear estrategias y tomas de decisiones a futuro, impactando en la eficiencia y rentabilidad de la empresa.

Funciones generales:

- ❖ Recolección de datos
- ❖ Preparación y limpieza de datos
- ❖ Modelación de datos
- ❖ Interpretación de patrones o tendencias
- ❖ Presentación y visualización de los datos
- ❖ Presentación de tendencias y formulación de hipótesis

(unir.net)

En la actualidad no existe un estudio específico para ser analista de datos, pero existen recomendaciones para que el analista se desempeñe bien en su campo.

Según la página web (unir.net) un analista de datos debería tener conocimientos en:

1. Matemática Computacional
2. Ingeniería Informática
3. Física
4. Ciencia de Datos
5. Ingeniería de Telecomunicaciones
6. Economía
7. Administración y Dirección de Empresas
8. Estadística

Los conocimientos básicos en cada una de las antes mencionadas nos hará más fácil crear estrategias y dar a entender a los ejecutivos de la empresa los resultados que se obtuvieron en un análisis para la toma de decisiones.

### **Desarrollo e implementación del aprendizaje**

En este ejercicio seguiremos el método antes mencionado; para obtener los mejores resultados, para esta solución tomaremos una base de datos del curso realizado en la plataforma recomendada por la Institución Uniremington.

Y de los cuales daremos recomendaciones como hacer el proceso de analisis de acuerdo al curso realizado.

Por lo general una aplicación especializada en guardar informacion, tiene la opcion de descargar la informacion guardada en varios formatos como Json, xml, formato excel y CSV, de estos formatos se recomienda exportar los datos en este ultimo mencionado, pues este formato estan configurado varias aplicaciones de Analisis y no presenta ningun problema, todos los archivos de este formato esta precedido por el nombre y termina en .csv

### **Ejercicio 1**

#### **Determinar direccion:**

El aplicativo que vamos a utilizar para analizar datos sera la hoja de calculo excel en este caso seran las ventas de videojuegos realizadas en un determinado tiempo .

Cuando se van analizar datos, se recomienda tener la mayor cantidad de datos posibles referentes a la informacion que se necesita, pues bajo esa informacion se revisa tendencias se aplica estadistica que dan una informacion mas veraz y entendible, como lo muestra en la imagen inferior en total tenemos en nuestra base de estudio 59517 datos por analizar.

Imagen 2. Primera tabla por analizar

Item	Nombre del Videojuego	Plataforma	Año	Categoría	Desarrollador	Ventas EEUU	Ventas Europ.	Ventas Japón	Otras Regiones	Ventas Globales
5391	Mushi Bugyou	3DS	2013	Acción	Namco Bandai Games	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5392	Patrician IV: Gold Edition	PC	2011	Familiar	Kalypso Media	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01
5393	Transformers: Fall of Cybertron	PC	2012	Acción	Activision	\$ 0,01	\$ -	\$ -	\$ -	\$ 0,01
5394	Mario & Luigi: Paper Jam & Mario Kart 7 Double	3DS	2015	Familiar	Nintendo	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5395	Driving Simulator 2011	PC	2011	Carreras	N/A	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01
5396	The Treasures of Mystery Island 3 Pack - Save M	PC	2011	Pasatiempos	Unknown	\$ 0,01	\$ -	\$ -	\$ -	\$ 0,01
5397	Hyakki Yagyou Kaidan Romance	PSP	2012	Acción	Quinrose	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5398	Hakuoki: Reimeiroku - Omouhase Kara	PSV	2015	Acción	Idea Factory	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5399	God Eater Off Shot: Tachibana Sakuya-hen Twin	PS4	2016	Acción	Namco Bandai Games	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5400	Bound By Flame	X360	2014	Rol	N/A	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01
5401	Codename: Panzers Complete Collection	PC	2016	Estrategia	Nordic Games	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01
5402	Dramatical Murder Re:code	PSV	2014	Aventura	Nitroplus	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5403	Hanasaku Manimani	PSV	2014	Aventura	5pb	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5404	Mighty No. 9	XOne	2016	Plataforma	Deep Silver	\$ 0,01	\$ -	\$ -	\$ -	\$ 0,01
5405	Teslagrad	PSV	2015	Plataforma	Rain Games	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01
5406	End of Nations	PC	2012	Estrategia	Trion Worlds	\$ 0,01	\$ -	\$ -	\$ -	\$ 0,01
5407	Fujiko F. Fujio Characters: Great Assembly! Slight	3DS	2014	Acción	Namco Bandai Games	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5408	Resident Evil 4 HD	XOne	2016	Disparos	Capcom	\$ 0,01	\$ -	\$ -	\$ -	\$ 0,01
5409	Farming 2017 - The Simulation	PS4	2016	Simulación	UIG Entertainment	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01
5410	Grisaia no Kajitsu: La Fruit de la Grisaia	PSP	2013	Aventura	Prototype	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5411	Rugby Challenge 3	XOne	2016	Deportes	Alternative Software	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01
5412	Outdoors Unleashed: Africa 3D	3DS	2011	Deportes	Mastiff	\$ 0,01	\$ -	\$ -	\$ -	\$ 0,01
5413	Fit & Fun	Wii	2011	Deportes	Unknown	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01
5414	Breach	PC	2011	Disparos	Destineer	\$ 0,01	\$ -	\$ -	\$ -	\$ 0,01
5415	Chou Ezaru wa Akai Hana: Koi wa Tsuki ni Shirub	PSV	2016	Acción	dramatic create	\$ -	\$ -	\$ 0,01	\$ -	\$ 0,01
5416	Know How 2	DS	2010	Pasatiempos	7G//AMES	\$ -	\$ 0,01	\$ -	\$ -	\$ 0,01

vgsales

+

Promedio: 665,7738345

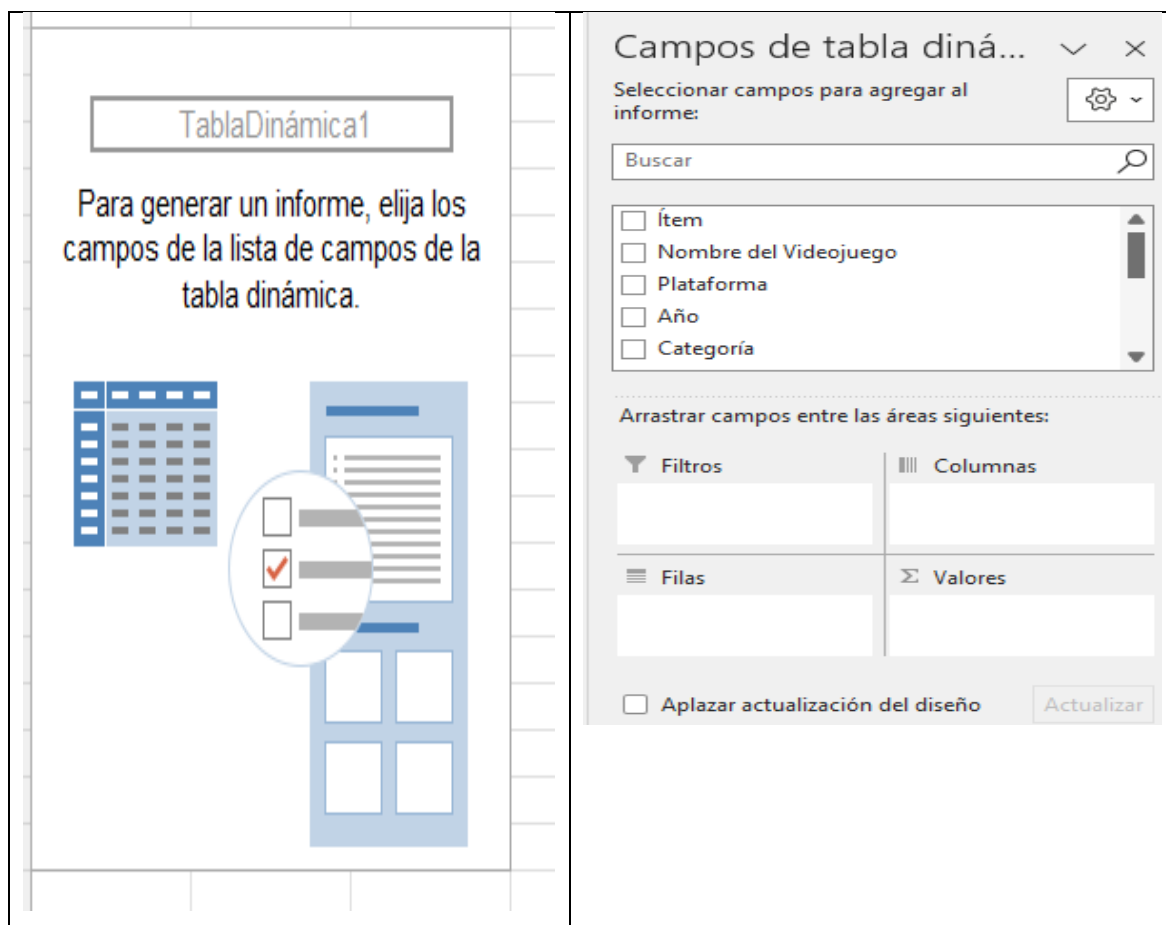
Recuento: 59587

Suma: 25071711,06

## Establecer parametros

Al tener un cantidad de datos de varias filas y columnas nos podemos dar cuenta que la informacion puede ser ordenada y filtrada en informacion que solo necesitemos, la hoja de excel tiene una herramienta muy poderosa que se llama tablas dinamicas, esta caracteritica nos permite filtrar y gerenar una tabla como mejor nos paresca.

*Imagen 3. Opciones de tablas dinámicas*



## Datos agregados

Adicional a los datos que nos genera la base que tenemos, en nuestra hoja podemos hacer calculos adicionales que nos ayudaran en el analisis de los datos.

Con una tabla dinamica estos datos se realizan automaticamente, se pueden encontrar promedios sumas recuento, entre varias utiles para tener respuestas estadisticas

*Imagen 4 Tabla dinamica sencilla*

Plataforma	(Todas)
<b>Etiquetas de fila</b>	<b>Suma de Ventas Globales</b>
+ Acción	702
+ Aventura	65
+ Carreras	129
+ Combate	87
+ Deportes	346
+ Disparos	474
+ Estrategia	38
+ Familiar	250
+ Pasatiempos	24
+ Plataforma	123
+ Rol	309
+ Simulación	73
<b>Total general</b>	<b>2.621</b>

## Seguimiento del proceso eliminar datos irrelevantes

Al enfocarnos en las repuestas que nosotros necesitemos podemos ubicar los datos que mas nos interesan, en este caso que plataforma vendio mas a nivel mundial, cual fue el tipo de juego que mas se vendio, podemos colocar los paises con mas ventas y menores ventas, esta informacion puede servir a una empresa desarrolladora pueda elegir que tipo de juego crear y si esta se encuentra dentro de los paises analizados

Imagen 5. Tabla completa suma de ventas en videoguegos

Plataforma	(Todas)			
Año	(Todas)			
Etiquetas de fila	Suma de Ventas Japón	Suma de Ventas Europa	Suma de Ventas EEUU	Suma de Ventas Globales
⊕ Acción	72	245	304	702
⊕ Aventura	15	19	24	65
⊕ Carreras	7	58	48	129
⊕ Combate	14	22	42	87
⊕ Deportes	16	121	168	346
⊕ Disparos	14	169	235	474
⊕ Estrategia	7	13	14	38
⊕ Familiar	25	70	133	250
⊕ Pasatiempos	4	7	11	24
⊕ Plataforma	16	40	57	123
⊕ Rol	102	74	108	309
⊕ Simulación	13	26	28	73
<b>Total general</b>	<b>306</b>	<b>864</b>	<b>1.172</b>	<b>2.621</b>

La hoja de calculo nos permite separar datos estadisticos que nos ayudan a entender mejor el comportamiento de las ventas en diferentes paises de forma

Imagen 6. Tabla de ventas Maximos en Japon Estados Unidod y globales

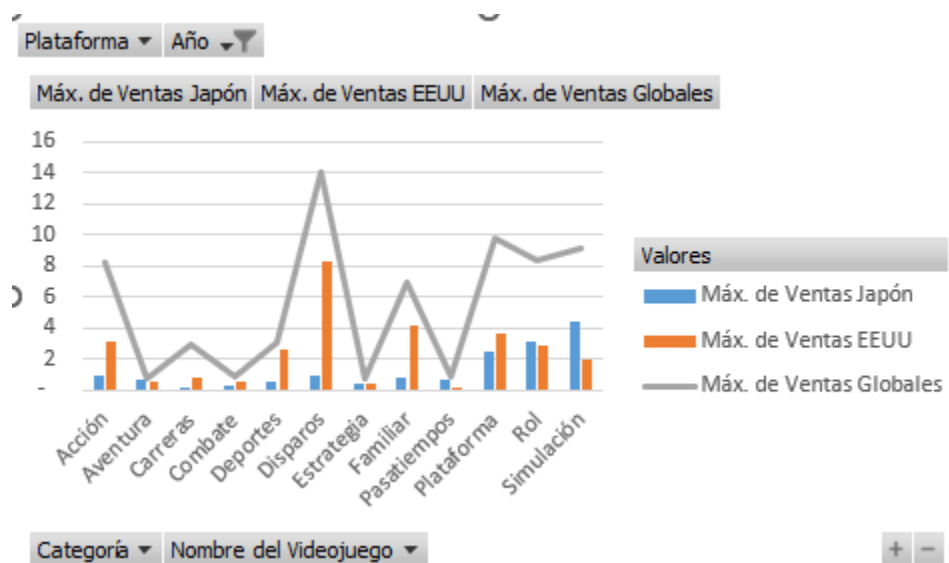
Plataforma	(Todas)			
Año	2012			
Etiquetas de fila	Máx. de Ventas Japón	Máx. de Ventas EEUU	Máx. de Ventas Globales	
⊕ Acción	1	3	8	
⊕ Aventura	1	1	1	
⊕ Carreras	0	1	3	
⊕ Combate	0	0	1	
⊕ Deportes	0	3	3	
⊕ Disparos	1	8	14	
⊕ Estrategia	0	0	1	
⊕ Familiar	1	4	7	
⊕ Pasatiempos	1	0	1	
⊕ Plataforma	2	4	10	
⊕ Rol	3	3	8	
⊕ Simulación	4	2	9	
<b>Total general</b>	<b>4</b>	<b>8</b>	<b>14</b>	

## Establecer practicas de administracion de datos

Aunque para los analistas de datos, tener la informacion en una tabla sea muy explicativa, y facil, para la mayoría de las personas no es asi, por tal razon saber realizar una grafica o imagen explicando los datos encontrados hara que muchas personas entiendan los resultados encontrados en un menor tiempo posible.

Con la hoja de calculo excel podemos realizar Graficas dinamicas que nos ayudan a entender y explicar mejor

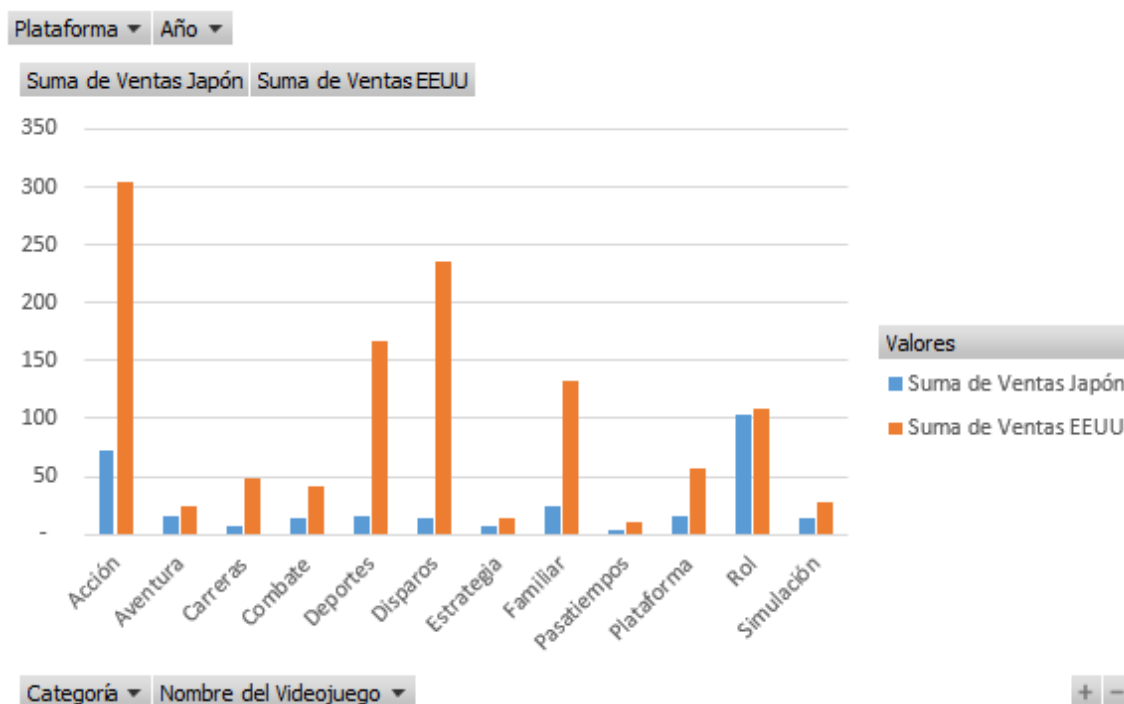
*Imagen 7. Ventas máximas de Japón y Estados Unidos en el año 2012*



En la imagen anterior nos podemos dar cuenta que al igual que la tabla anterior tienen la misma informacion, pero con el grafico podemos destacar que Estados Unidos en comparacion con las venta en Japon vende mas juegos en la categoria de disparos, y en Japon tienen mayor venta los juegos de simulacion en el año 2012.

Y si realizamos un filtro para ver las ventas totales de cada juego, nos damos cuenta que Estados Unidos es por mucho mejor cliente de los juegos a nivel general en comparacion con Japon.

*Imagen 8. Suma de ventas totales en videojuegos entre EEUU y Japon*



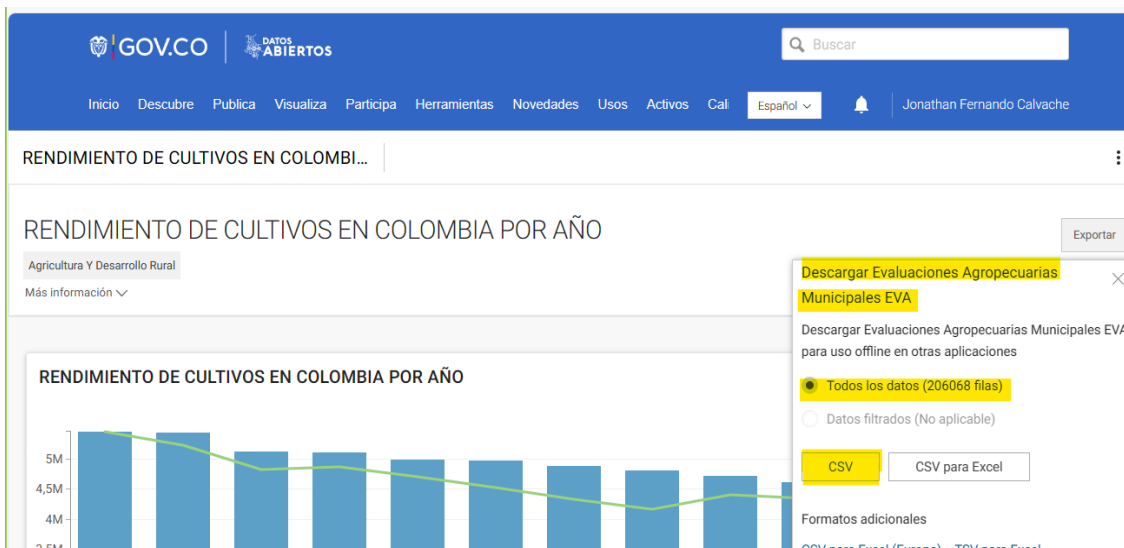
Con la informacion antes encontrada podemos realizar pronosticos de ventas posibles en el futuro, a que juego se debe dedicar una desarrolladora dependiendo del pais en donde quiera sobresalir.

## Ejercicio 2

Para el siguiente ejercicio vamos a analizar datos recolectados de la página del gobierno de Colombia, Datos abiertos, en esta hemos encontrado una base de datos con la siguiente información rendimiento de cultivos en Colombia por año, para este fin lo hemos

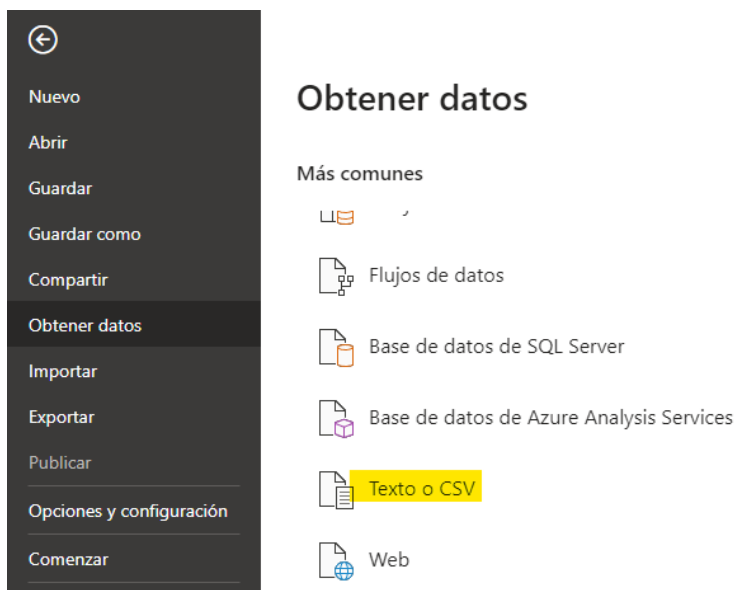
descargado en formato CSV, un formato que es recibido la mayoría de las plataformas de cálculo y análisis de datos.

Imagen 9. Descarga base de datos abiertos página GOV.CO



Para analizar estos datos vamos a pasarlos a la plataforma Power Bi, en este veremos la cantidad de columnas y filas que datos obtendremos importamos los datos directamente.

### Imagen 10. Obtención de datos



Damos la opción de texto CSV y crearemos la base de datos para analizar.

### Imagen 11. Imagen de formato CVS en Excel

**Evaluaciones\_Agropecuarias\_Municipales\_EVA\_20231119.csv**

Origen de archivo: 65001: Unicode (UTF-8) | Delimitador: Coma | Detección del tipo de datos: Basado en las primeras 200 filas

CÓD. DEP.	DEPARTAMENTO	CÓD. MUN.	MUNICIPIO	GRUPO DE CULTIVO	SUBGRUPO DE CULTIVO	CULTIVO	DESAGREGACIÓN REGIONAL Y/O SISTEMA PRODUCTIVO	AÑO	PERIODO	Área Sembrada (ha)	Área Cosechi
15	BOYACA	15114	BUSBANZA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2006	2006B	2	
25	CUNDINAMARCA	25754	SOACHA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2006	2006B	82	
25	CUNDINAMARCA	25214	COTA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2006	2006B	2	
54	NORTE DE SANTANDER	54405	LOS PATIOS	HORTALIZAS	ACELGA	ACELGA	ACELGA	2006	2006B	3	
54	NORTE DE SANTANDER	54518	PAMPLONA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2006	2006B	1	
68	SANTANDER	68377	LA BELLEZA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2006	2006B	1	
25	CUNDINAMARCA	25754	SOACHA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007A	72	
25	CUNDINAMARCA	25214	COTA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007A	2	
54	NORTE DE SANTANDER	54518	PAMPLONA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007A	1	
68	SANTANDER	68377	LA BELLEZA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007A	1	
15	BOYACA	15516	PAIPA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007B	1	
19	CAUCA	19392	LA SIERRA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007B	2	
25	CUNDINAMARCA	25754	SOACHA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007B	62	
25	CUNDINAMARCA	25126	CAJICA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007B	1	
54	NORTE DE SANTANDER	54405	LOS PATIOS	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007B	3	
54	NORTE DE SANTANDER	54518	PAMPLONA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2007	2007B	1	
19	CAUCA	19392	LA SIERRA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2008	2008A	2	
25	CUNDINAMARCA	25754	SOACHA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2008	2008A	62	
25	CUNDINAMARCA	25126	CAJICA	HORTALIZAS	ACELGA	ACELGA	ACELGA	2008	2008A	2	
54	NORTE DE SANTANDER	54405	LOS PATIOS	HORTALIZAS	ACELGA	ACELGA	ACELGA	2008	2008A	3	

Con una tabla de 17 columnas y 206646 271 valores diferentes, seguiremos el siguiente método para analizar los datos:

#### 11. Determinar dirección

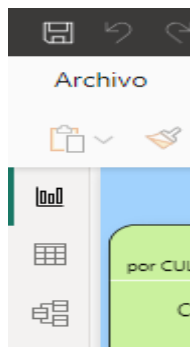
12. Establecer parámetros
13. Datos agregados
14. Seguimiento del progreso
15. Eliminar datos irrelevantes
16. Ejecutar análisis estadístico
17. Establecer prácticas de administración de dato
18. Involucrar la tecnología
19. Emplear visualización de datos
20. construcción de narrativas

Antes de iniciar con la explicación del proceso de análisis vamos a dar una breve explicación que es Power bi; *“Creado en 2014 por Microsoft para mejorar la productividad corporativa, Power BI es un software de inteligencia de negocios que permite recolectar, analizar y visualizar datos dentro de una empresa. Su interfaz es intuitiva y amigable para los usuarios que necesitan combinar, modelar y compartir datos provenientes de diversas fuentes, como CRM, Excel y Google Analytics, mediante paneles, gráficos y tablas.”* (EBAC, s.f.)

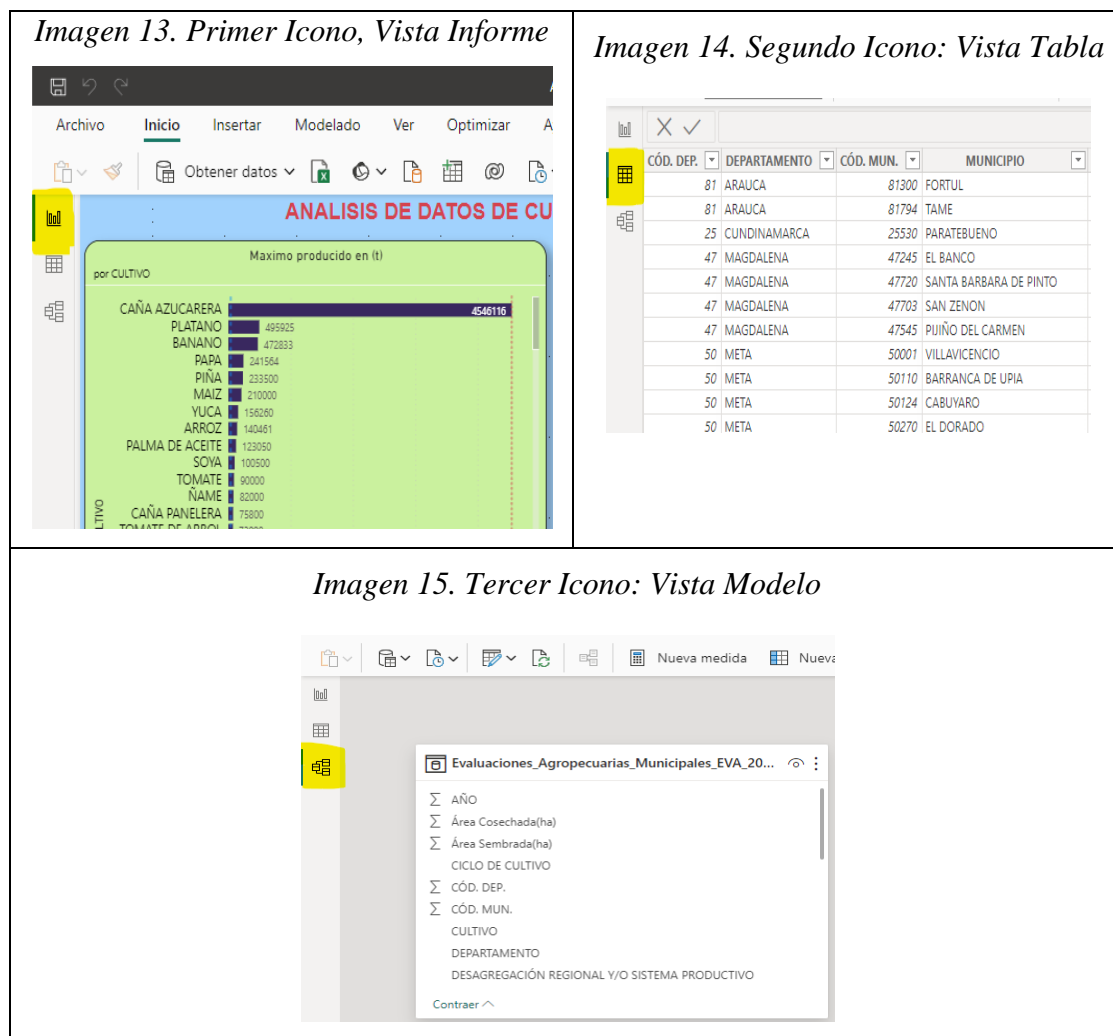
Para completar la explicación anterior podemos decir que Power bi, está enfocado en mostrar los datos que vamos a analizar en imágenes simples y fáciles de crear y utilizar con los datos que tenemos.

Vamos a explicar un poco cómo funciona la herramienta:

Imagen 12. Iconos



En la imagen anterior podemos observar 3 iconos, el primer icono de arriba hacia abajo corresponde a la configuración de cómo se ven las visualizaciones, el segundo icono nos muestra los datos en forma de tabla a los cuales la mayoría de personas estamos enseñados, y el tercer icono nos permite crear enlaces entre diferentes tablas.



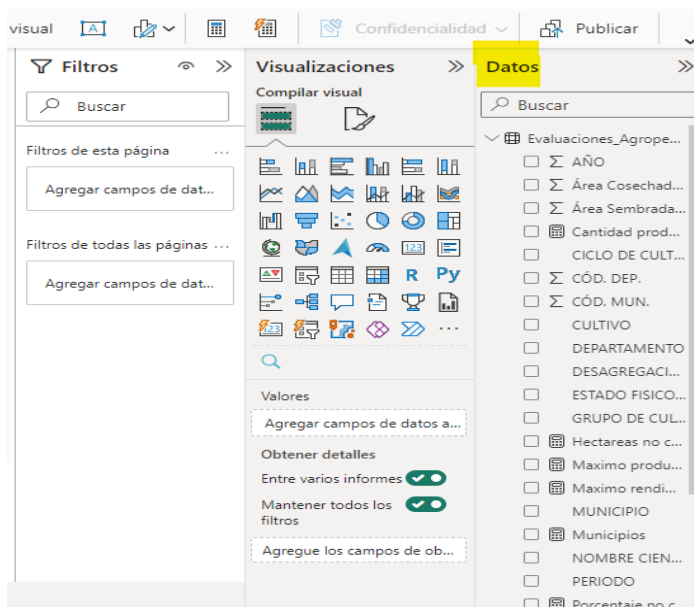
Ahora que sabemos brevemente cómo funciona la herramienta, nos vamos a enfocar solamente en los dos primeros iconos la visual y los datos en tabla, esta segunda para revisar como podemos adaptar nuestros datos.

Vamos a seguir el método de análisis:

1. **Determinar dirección:** vamos a enfocarnos en encontrar los cultivos con mayor producción dependiendo de su año, departamento y municipio.

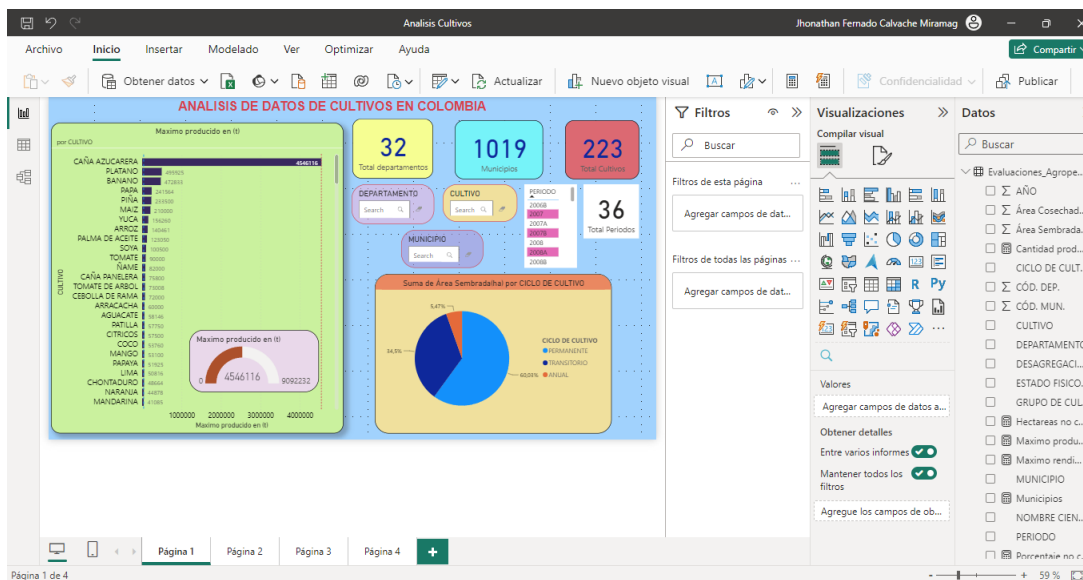
En power BI, en el primer icono sale los menús despegables con la información que tenemos en nuestra tabla.

*Imagen 16. Menús despegables*



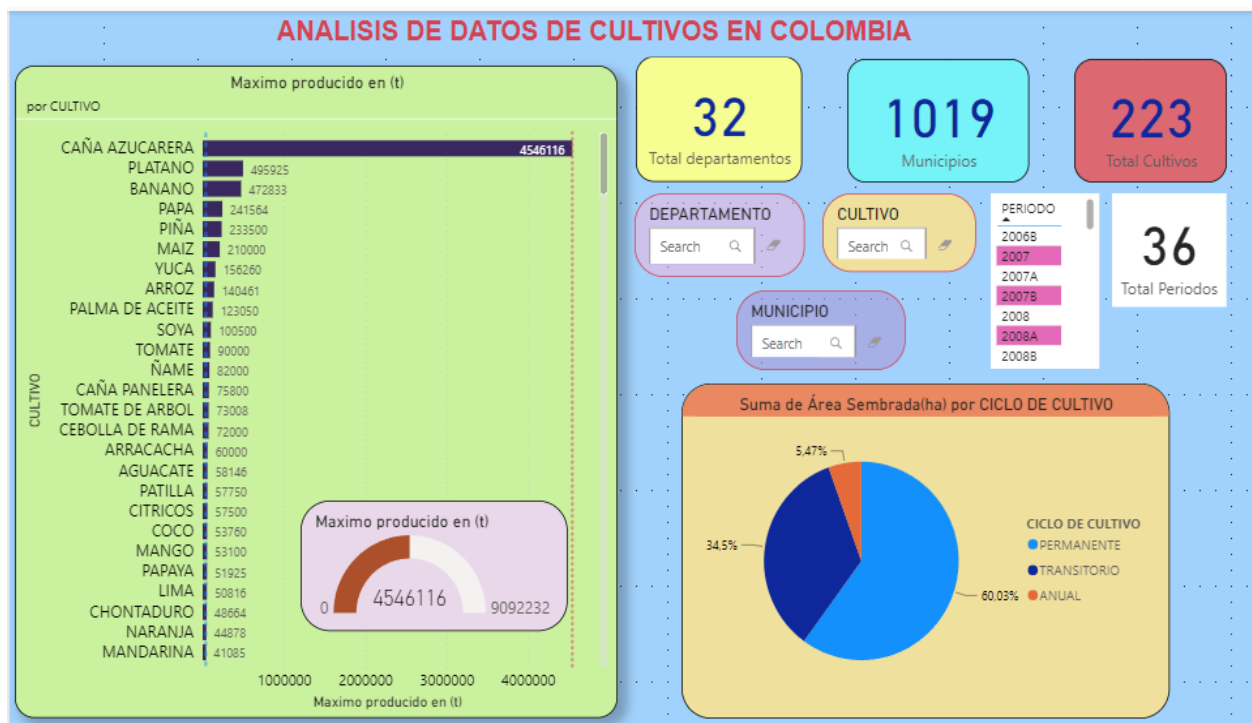
Al seleccionar uno de los cuadros con un visto y escogiendo las visualizaciones correspondientes el programa inmediatamente nos genera las gráficas correspondientes, para efectos del análisis hemos querido representar un primer análisis de que es lo que vamos a analizar.

Imagen 17. Ventana de configuración e ingreso de datos, icono visto informe



Para seguir con el método vamos a explicar la dirección que vamos a tomar para encontrar el análisis y la información más importante de la base de datos a analizar, primero debemos entender que vamos a analizar, para tal fin vamos a enfocarnos en explicar la imagen que tenemos nuestra primera información:

Imagen 18. Principal Análisis cultivos en Colombia



2. **Establecer Parámetros:** Vamos a analizar el máximo rendimiento producción, cosecha y sembrado en el país Colombia, para tal fin debemos tener claro que cantidad de elementos vamos a analizar, como lo demuestra la imagen.

Vamos a analizar 223 cultivos, 1019 municipios o población de muestra, 32

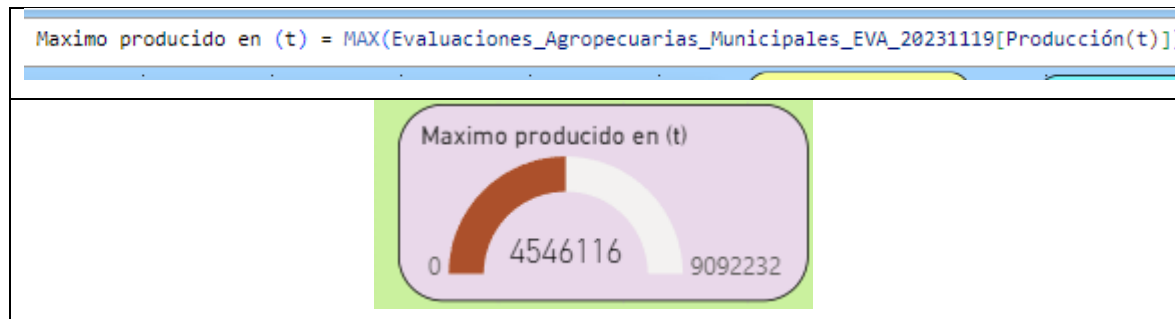
departamentos, 3 tipos de ciclo de cultivo como permanente transitorio, Anual.

Para tal fin hemos creado filtros para enfocarnos por búsqueda en departamento, cultivo y municipio.

3. **Datos Agregados:** Los datos que descargamos en la página carecen de información importante que debemos agregar, como sumatorias de datos, diferencias entre uno y otro elemento que debemos hacer, en Power BI, esto es muy fácil.

Se crea una nueva medida y creamos nuestro calculo para que nos de los datos que necesitamos.

Imagen 19. Fórmula para encontrar la máxima producción e imagen como muestra el resultado



4. **Seguimiento del proceso:** Power bi, nos permite crear botones y seleccionar directamente los datos que necesitamos realizando un filtro detallado de las cosas que necesitemos.

Inicialmente tenemos datos globales con sus máximos, pero en este primer informe podemos buscar por departamento, municipio cultivo para que la información sea más puntual.

Imagen 20. Tipos de filtro



**Eliminar datos irrelevantes**

**Ejecutar análisis estadísticos**

**Establecer prácticas de administración de datos**

**Involucrar la Tecnología**

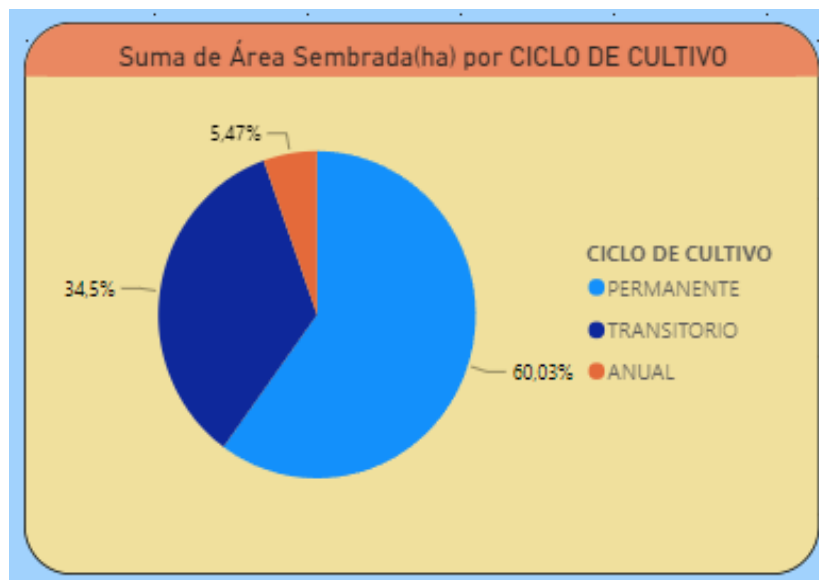
**Emplear visualización de datos**

Los requisitos del método a seguir; antes mencionados se deben tener en cuenta en Power bi, antes de empezar a seleccionar las gráficas los datos estadísticos como máximo mínimo promedio, varianza se deben sacar con las fórmulas que el programa nos permite, los filtros que creamos nos permiten administrar los datos, Power Bi es la tecnología la cual nos agiliza el análisis de forma sorprendente.

5. **Construcción de narrativas:** Power Bi, nos permite crear varios informes dinámicos como lo mostraremos a continuación con las mismas tablas descargadas generaremos diferentes informes como hojas en Excel, que nos ayudan a entender los datos a analizar.

Narrativas creadas con los datos que escogimos para analizar:

Imagen 21. Tipos de ciclos de cultivos

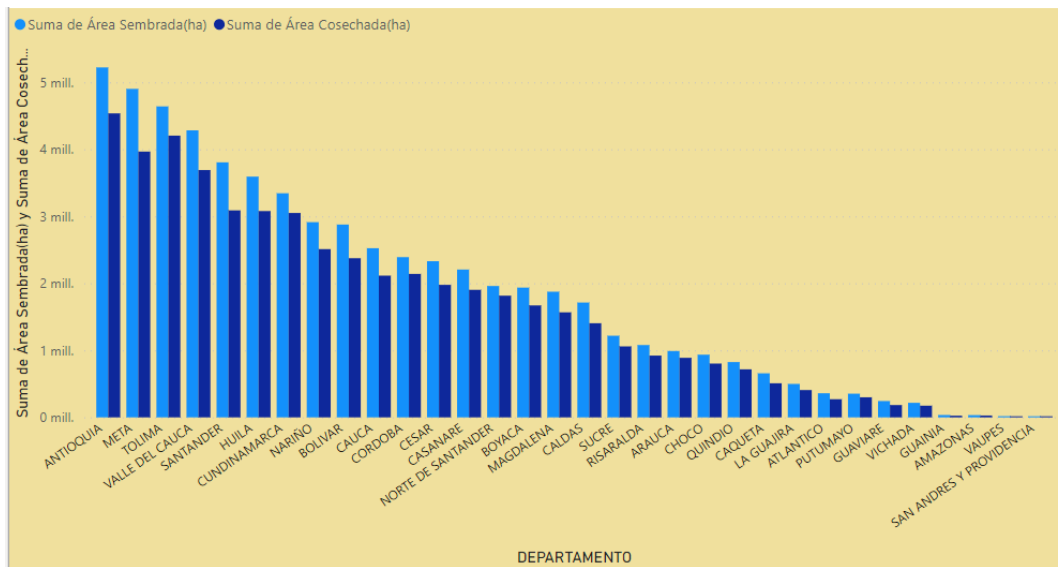


Hay 3 diferentes tipos de ciclo de cultivos que nos indica los datos descargados, los permanentes, los transitorios y los anuales, cada uno de estas representa la demora en cosecha de cada uno de ellos, si vemos la imagen podemos deducir lo siguiente:

El 60.03% de los cultivos sembrados en Colombia son de tipo permanente, el 34.5% de los cultivos son transitorios, y solo el 5.47% son de tipo Anual.

Podemos deducir o describir que los cultivos que siembra el área agrícola en Colombia son los cultivos que duran más de una cosecha.

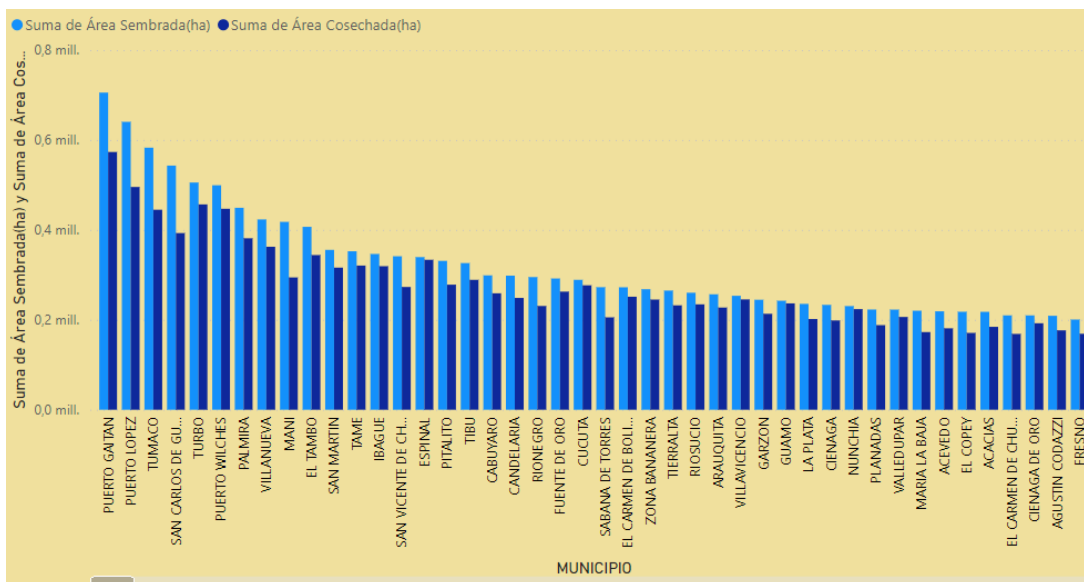
Imagen 22. Comparativa área sembrada y cosechada



La imagen antes descrita nos muestra la comparativa del área sembrada y cosechada en los departamentos de Colombia, podemos observar que el mayor número de siembra y cosecha se realiza en el departamento de Antioquia, que es mucho más el área que se siembra y hay un porcentaje de perdidas.

Podemos encontrar un parámetro visual y es que cuando hablamos de cultivos a nivel general, siempre será mayor el área sembrada que cosechada, y a medida que mas aumenta la siembra de cultivos mayor será el área que no será posible cosecharla.

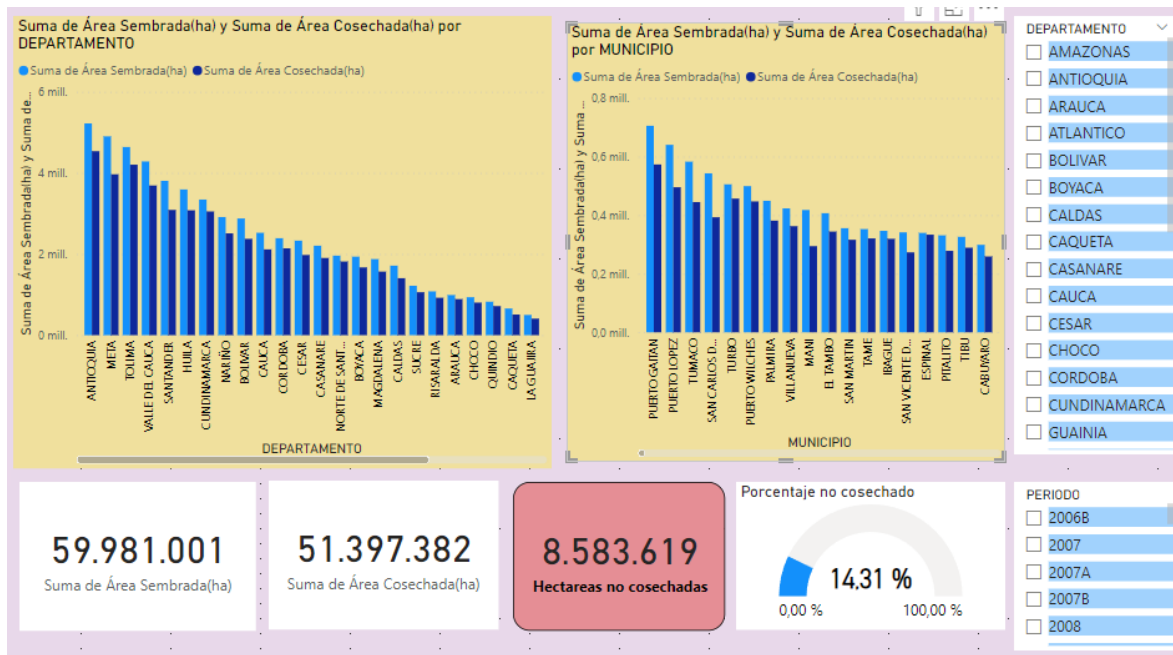
Imagen 23. Municipios con más área sembrada y cosechada



En la imagen podemos observar que, aunque en la métrica por departamento el área con mayor siembra es Antioquia, el municipio con más área sembrada es puerto Gaitán del departamento del Meta, el informe nos detalla la información.

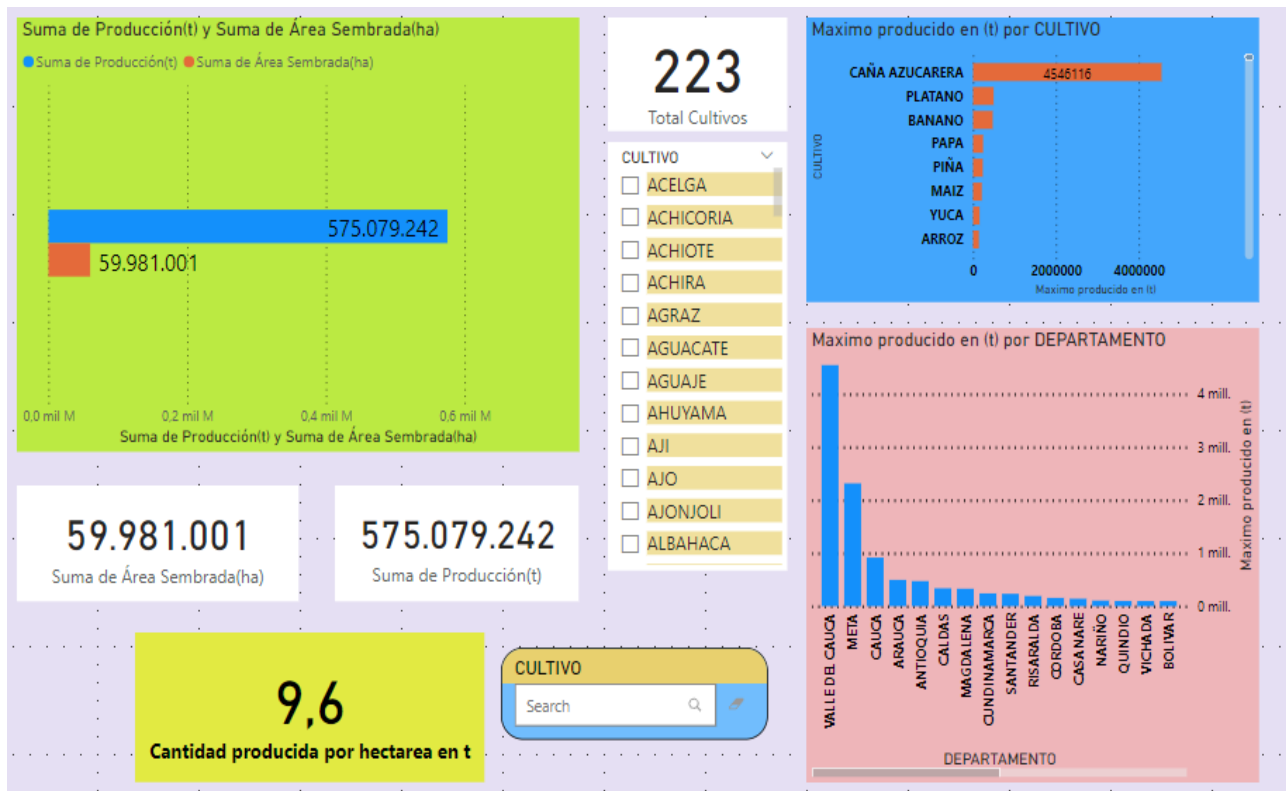
Al igual que la información departamental observamos que siempre se va a sembrar más área y se cosechara un porcentaje menor. Cabe anotar que esto es muy diferente a la producción y rendimiento de una cosecha.

Imagen 24. Comparativas área sembrada y cosechada filtros por departamento y periodo de cultivo



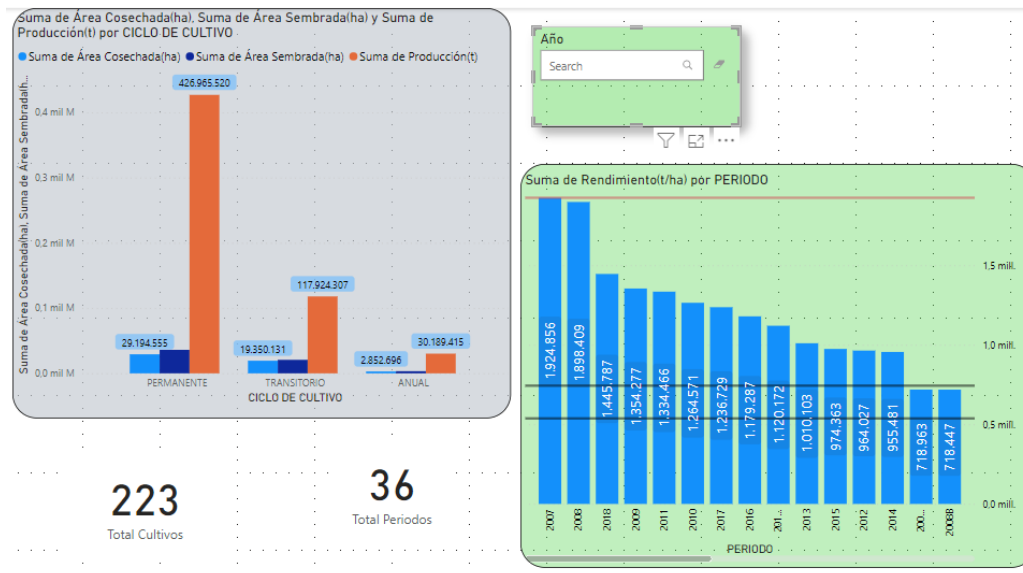
Si nos ponemos a detallar la imagen anterior producida por Power bi, podemos encontrar de forma rápida, que el área no cosechada en hectáreas corresponde al 14.31% del área sembrada en total, y si queremos detallar mas a fondo en departamentos con datos máximos de áreas no cosechadas, podemos filtrar con los filtros del área izquierda.

Imagen 25. Cantidad de producción que y cuál es su diferencia



En este informe podemos observar la cantidad de área sembrada y la cantidad de producción en toneladas, tenemos un promedio de la cantidad producida por hectárea y podemos filtrar por cultivo y nos muestra el departamento, los datos se ven fáciles de mirar y si queremos ampliar más la información tiene sus filtros característicos.

*Imagen 26. Comparativa área sembrada cosechada producción y rendimiento en tipo de cultivos por periodo*



En esta comparativa observamos 2 graficas con los datos de producción por cultivo y el mayor rendimiento en el año.

## Imágenes

<i>Imagen 1. Mapa conceptual.....</i>	<i>6</i>
<i>Imagen 2. Primera tabla por analizar.....</i>	<i>16</i>
<i>Imagen 3. Opciones de tablas dinámicas.....</i>	<i>17</i>
<i>Imagen 4 Tabla dinamica sencilla.....</i>	<i>18</i>
<i>Imagen 5. Tabla completa suma de ventas en videojuegos.....</i>	<i>19</i>
<i>Imagen 6. Tabla de ventas Maximos en Japon Estados Unidod y globlales.....</i>	<i>19</i>
<i>Imagen 7. Ventas máximas de Japón y Estados Unidos en el año 2012.....</i>	<i>20</i>
<i>Imagen 8. Suma de ventas totales en videojuegos entre EEUU y Japon.....</i>	<i>21</i>

	37
<i>Imagen 9. Descarga base de datos abiertos página GOV.CO .....</i>	<i>22</i>
<i>Imagen 10. Obtención de datos .....</i>	<i>23</i>
<i>Imagen 11. Imagen de formato CVS en Excel .....</i>	<i>23</i>
<i>Imagen 12. Iconos.....</i>	<i>24</i>
<i>Imagen 13. Primer Icono, Vista Informe .....</i>	<i>25</i>
<i>Imagen 14. Segundo Icono: Vista Tabla.....</i>	<i>25</i>
<i>Imagen 15. Tercer Icono: Vista Modelo.....</i>	<i>25</i>
<i>Imagen 16. Menús despegables .....</i>	<i>26</i>
<i>Imagen 17. Ventana de configuración e ingreso de datos, icono visto informe.....</i>	<i>27</i>
<i>Imagen 18. Principal Análisis cultivos en Colombia.....</i>	<i>28</i>
<i>Imagen 19. Fórmula para encontrar la máxima producción e imagen como muestra el resultado .....</i>	<i>29</i>
<i>Imagen 20. Tipos de filtro.....</i>	<i>29</i>
<i>Imagen 21. Tipos de ciclos de cultivos .....</i>	<i>31</i>
<i>Imagen 22. Comparativa área sembrada y cosechada.....</i>	<i>32</i>
<i>Imagen 23. Municipios con más área sembrada y cosechada.....</i>	<i>33</i>
<i>Imagen 24. Comparativas área sembrada y cosechada filtros por departamento y periodo de cultivo.....</i>	<i>34</i>
<i>Imagen 25. Cantidad de producción que y cuál es su diferencia .....</i>	<i>35</i>
<i>Imagen 26. Comparativa área sembrada cosechada producción y rendimiento en tipo de cultivos por periodo .....</i>	<i>36</i>

## Conclusiones

Los ejercicios resueltos nos muestran que analizar datos no es difícil, cuando se siguen los métodos recomendados las herramientas recomendadas. Antes de iniciar con cualquier análisis debemos proyectarnos que tenemos que encontrar inicialmente, cuáles son esos datos primarios que nos ayudaran a avanzar el segundo paso, entender las variables y características para hacer un buen análisis, nos permite descubrir que hay muchas herramientas incluso conocidas de donde podemos sacar todo el potencial como Excel, una aplicación con historia en el mundo tecnológico y empresarial y que lleva mucho tiempo en el juego, esto nos enseña que hay herramientas fuertes y que están al alcance de todos para tener una información veraz.

Tanto analizar y encontrar información es importante. No debe quedar en segundo plano saber transmitir la información a las personas; aquellas encargadas de la empresa o objetivos que tomaran esa información para realizar o tomar las acciones, por tal razón debe predominar en los informes la información más exacta y fácil.

## Bibliografía

- crehana*. (s.f.). Obtenido de crehana: <https://www.crehana.com/>
- Diferenciando. (s.f.). <https://diferenciando.com/tipos-de-datos/>. Obtenido de <https://diferenciando.com/tipos-de-datos/>.
- EBAC. (s.f.). *Escuela britanica de artes creativas y tecnologia*. Obtenido de EBAC: <https://ebac.mx/>
- ebac.mx*. (s.f.). Obtenido de ebac.mx: <https://ebac.mx/blog/que-es-el-analisis-de-datos>
- economipedia.com*. (s.f.). Obtenido de economipedia.com: [https://economipedia.com/definiciones/metodo.html#google\\_vignette](https://economipedia.com/definiciones/metodo.html#google_vignette)
- gcfglobal*. (s.f.). Obtenido de gcfglobal: <https://edu.gcfglobal.org/es/estadistica-basica/que-es-la-estadistica/1/#>
- geekflare.com*. (s.f.). Obtenido de geekflare.com: <https://geekflare.com/es/ai-data-analysis-tools/>
- Gestiopolis*. (s.f.). Obtenido de Gestiopolis: <https://www.gestiopolis.com/que-es-un-pronostico-empresarial/>
- IMF blog de Tecnologia*. (2023). Obtenido de IMF blog de Tecnologia: [https://blogs.imf-formacion.com/blog/tecnologia/las-10-mejores-herramientas-de-analisis-de-datos-202006/#Formacion\\_Relacionada](https://blogs.imf-formacion.com/blog/tecnologia/las-10-mejores-herramientas-de-analisis-de-datos-202006/#Formacion_Relacionada)
- rockcontent.com*. (2023). <https://rockcontent.com/es/blog/tipos-de-almacenamiento-de-datos/>. Obtenido de <https://rockcontent.com/es/blog/tipos-de-almacenamiento-de-datos/>.
- Significados*. (s.f.). Obtenido de Significados: <https://www.significados.com/cuantitativa/>
- unir.net*. (s.f.). Obtenido de unir.net: <https://www.unir.net/ingenieria/revista/analista-de-datos-cual-es-el-perfil-de-estos-profesionales/>
- Wikipedia. (23 de 10 de 18). <https://es.wikipedia.org/wiki/Dato>. Obtenido de <https://es.wikipedia.org/wiki/Dato>.
- zipreporting*. (2023). Obtenido de zipreporting: [zipreporting.com/es/data-analysis-method/types-of-data-analysis.html](https://zipreporting.com/es/data-analysis-method/types-of-data-analysis.html)