



Universidad Tecnológica de Puebla

Ingeniería en Desarrollo y Gestión de Software

9° Cuatrimestre

Grupo: "D"

PRODUCTO 4

Materia: "Extracción de conocimiento en bases de datos"

Docente:

José Francisco Espinoza Garita

Evaluated:

Salma Cid Morales

Periodo:

Mayo – Agosto 2022

Contenido

Tabla de ilustraciones.....	3
Introducción.....	4
Google Data Studio Dashboard.....	5
Informe 1.....	11
Informe 2.....	12
Informe 3.....	13
Excel Dashboard	14
Informe 4.....	14
Conclusión.....	18

Tabla de ilustraciones

Ilustración 1 Dashboard	4
Ilustración 2 Ruta	5
Ilustración 3 Cambio de datos.....	6
Ilustración 4 Guardar.....	6
Ilustración 5 Google Data Studio.....	7
Ilustración 6 Crear Fuente de Datos.....	8
Ilustración 7 Subir Archivo	8
Ilustración 8 Subir Archivo	9
Ilustración 9 Subir Archivo	9
Ilustración 10 Conexión de Archivo.....	9
Ilustración 11 Cambio de datos.....	10
Ilustración 12 Crear Informe	11
Ilustración 13 Advertencia	11
Ilustración 14 Informe 1	11
Ilustración 15 Filtro	12
Ilustración 16 Informe 2	12
Ilustración 17 Informe 3	13
Ilustración 18 Filtro	13
Ilustración 19 Excel Tabla dinámica	14
Ilustración 20 Excel Tabla dinámica 2	15
Ilustración 21 Excel Grafico dinamico	15
Ilustración 22 Informe 4	16
Ilustración 23 Informe 4	17
Ilustración 24 Logo	18

Introducción

Durante la elaboración del siguiente documento se mostrará la creación de un dashboard utilizando las herramientas de Google Data Studio, el cual nos ayudará paso a paso a subir un archivo .csv con el que se puede trabajar.

También se muestra el uso y manejo de un archivo en Excel para la elaboración de un dashboard.

Los cuales nos ayudaran a analizar y monitorizar diversos datos según nuestro criterio o el resultado que se busca mostrar, además de que se pueden manejar en tiempo real para visualizar el cambio de datos conforme sean necesarios y facilitar el manejo de dicha información.

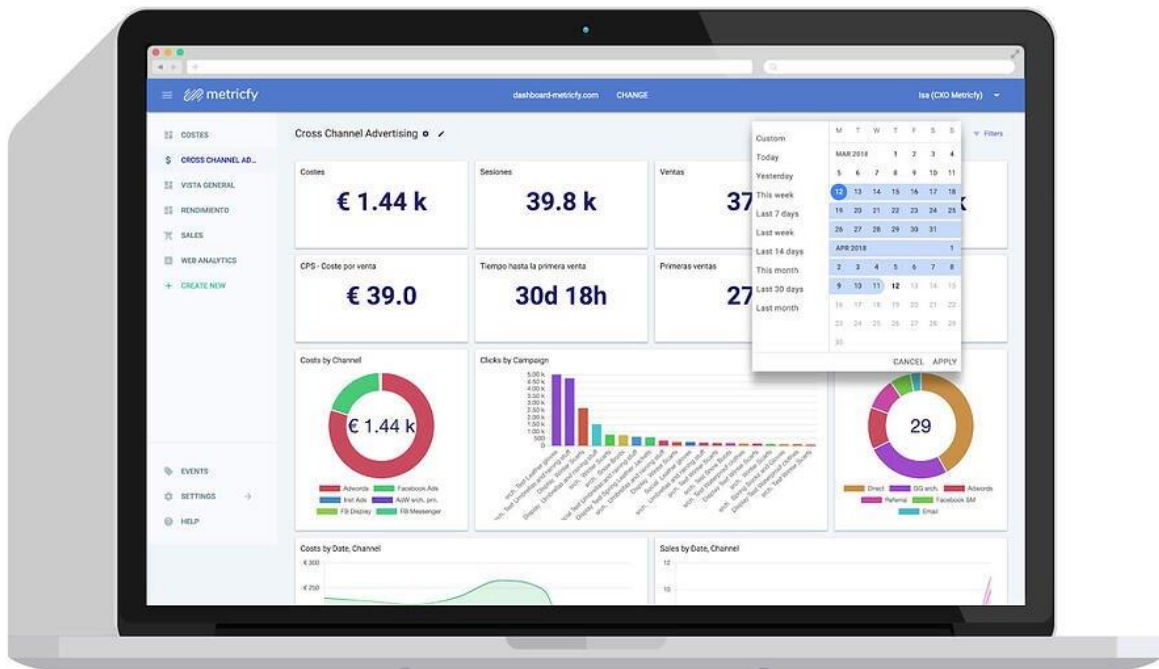


Ilustración 1 Dashboard

Google Data Studio Dashboard

Un dashboard es una herramienta de gestión de la información que monitoriza, analiza y muestra de manera visual los indicadores clave de desempeño (KPI), métricas y datos fundamentales para hacer un seguimiento del estado de una empresa, un departamento, una campaña o un proceso específico.

Para la elaboración del dashboard se utilizará la herramienta de Google Data Studio.

Se hacen los cambios de datos correspondientes, en este caso del archivo ChiapasDB.csv

Se especifica la ruta del archivo, se hacen las importaciones necesarias y se lee el archivo a utilizar

```
path= 'C:/covid2020/'
```

```
import csv  
from datetime import datetime
```

```
import pandas as pd
```

```
dfPuebla = pd.read_csv(path+"ChiapasDB.csv")
```

Ilustración 2 Ruta

Se realizan los cambios necesarios en los datos que se vas a ocupar.

```

dfPuebla['SECTOR'] = dfPuebla['SECTOR'].astype('string')
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["1"], ["CRUZ ROJA"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["2"], ["DIF"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["3"], ["ESTATAL"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["4"], ["IMSS"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["5"], ["IMSS-BIENESTAR"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["6"], ["ISSSTE"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["7"], ["MUNICIPAL"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["8"], ["PEMEX"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["9"], ["PRIVADA"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["10"], ["SEDENA"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["11"], ["SEMAR"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["12"], ["SSA"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["13"], ["UNIVERSITARIO"])
dfPuebla['SECTOR'] = dfPuebla['SECTOR'].replace(["99"], ["NO ESPECIFICADO"])

```

```

dfPuebla['TIPO_PACIENTE'] = dfPuebla['TIPO_PACIENTE'].astype('string')
dfPuebla['TIPO_PACIENTE'] = dfPuebla['TIPO_PACIENTE'].replace(["1"], ["AMBULATORIO"])
dfPuebla['TIPO_PACIENTE'] = dfPuebla['TIPO_PACIENTE'].replace(["2"], ["HOSPITALIZADO"])
dfPuebla['TIPO_PACIENTE'] = dfPuebla['TIPO_PACIENTE'].replace(["99"], ["NO ESPECIFICADO"])

```

```

dfPuebla['NACIONALIDAD'] = dfPuebla['NACIONALIDAD'].astype('string')
dfPuebla['NACIONALIDAD'] = dfPuebla['NACIONALIDAD'].replace(["1"], ["MEXICANA"])
dfPuebla['NACIONALIDAD'] = dfPuebla['NACIONALIDAD'].replace(["2"], ["EXTRANJERA"])
dfPuebla['NACIONALIDAD'] = dfPuebla['NACIONALIDAD'].replace(["99"], ["NO ESPECIFICADO"])

```

Ilustración 3 Cambio de datos

Se guarda el archivo con los cambios.

```

dfPuebla.to_csv(path + "ChiapasTransformado.csv", index=False, encoding='utf-8')

```

Ilustración 4 Guardar

Una vez hechos los cambios se busca en el navegador “Google Data Studio” y se selecciona la opción remarcada.

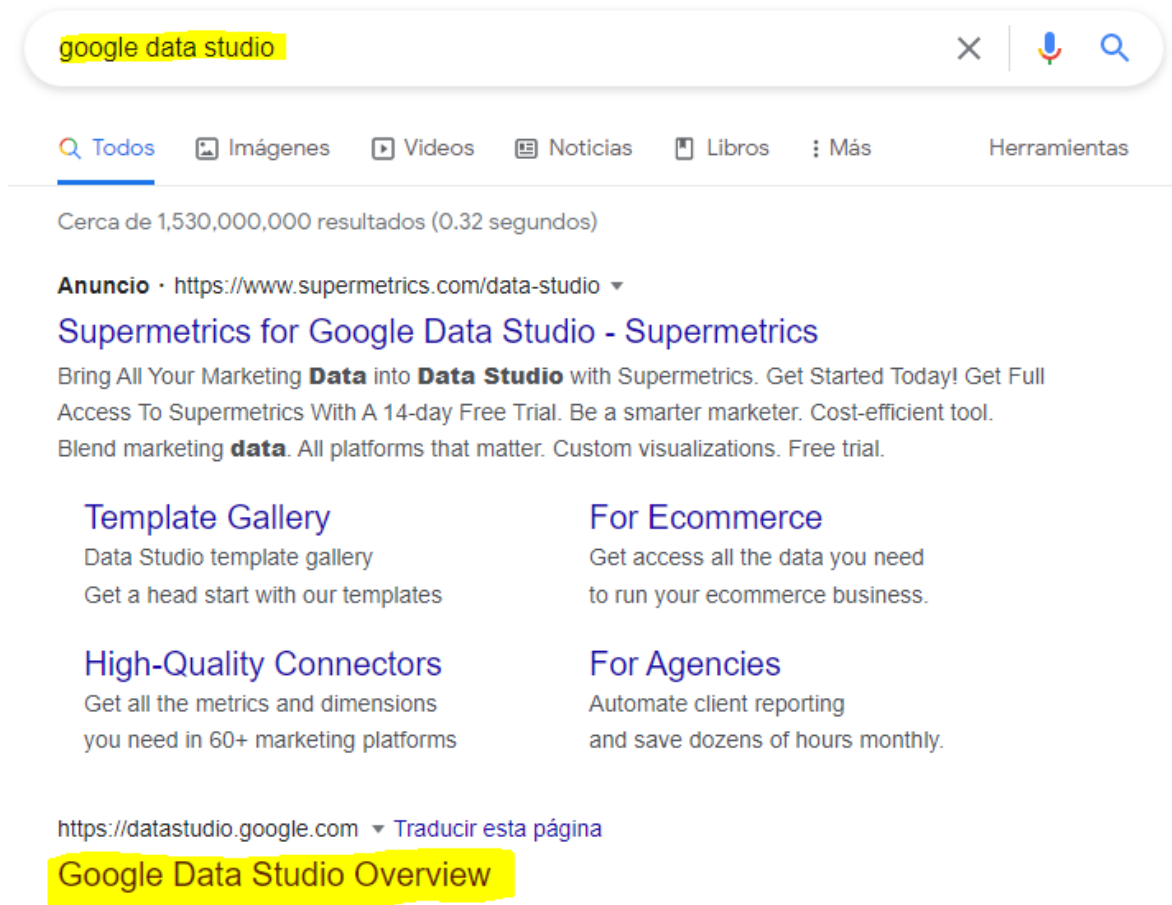


Ilustración 5 Google Data Studio

En la pantalla que aparece se selecciona la opción “crear” y se crea una fuente de datos.

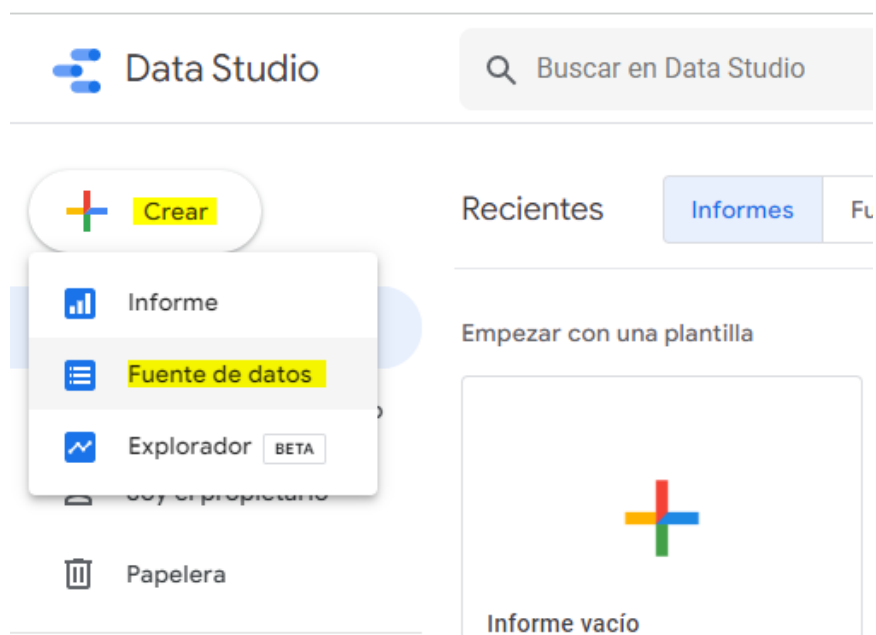


Ilustración 6 Crear Fuente de Datos

Se selecciona la opción subir archivo.

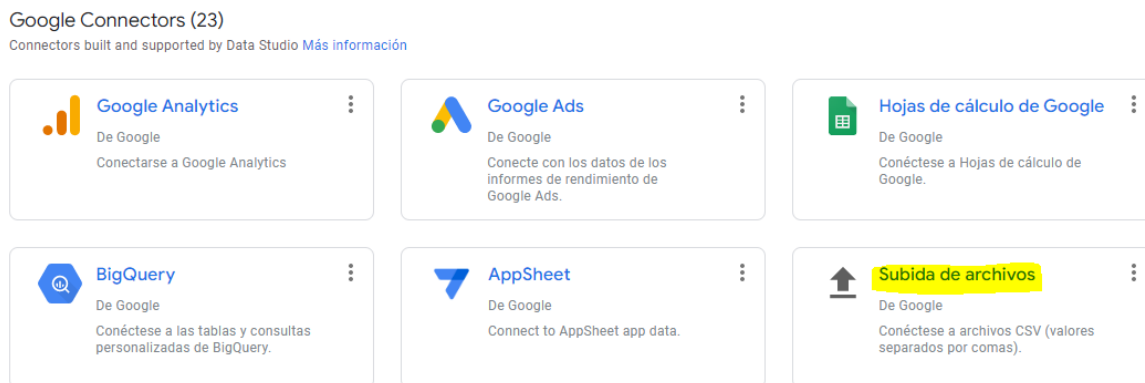


Ilustración 7 Subir Archivo

Se sube el archivo con el que se va a ocupar ChiapasTranformado.csv

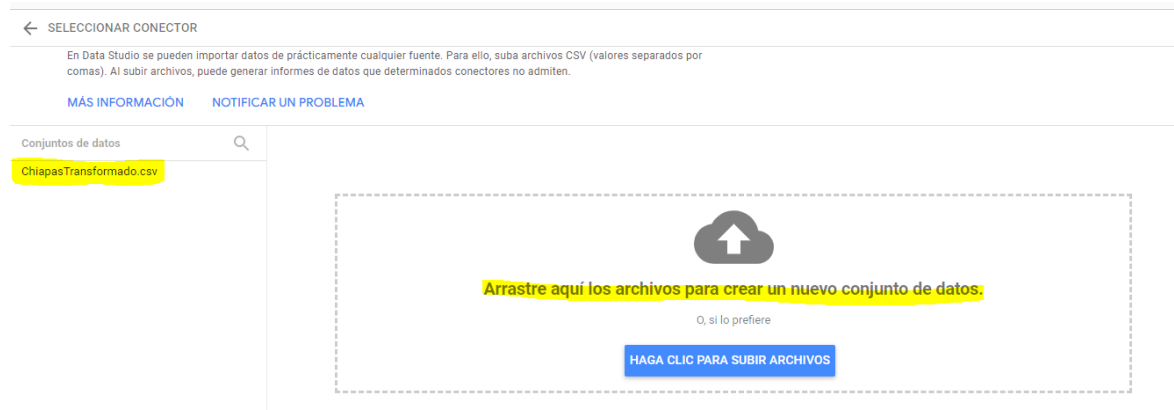


Ilustración 8 Subir Archivo

En este caso el archivo que se va a ocupar ya se subió **ChiapasTranformado.csv**.

Ejemplo:

Se abre la dirección en donde se tiene el archivo y se sube.

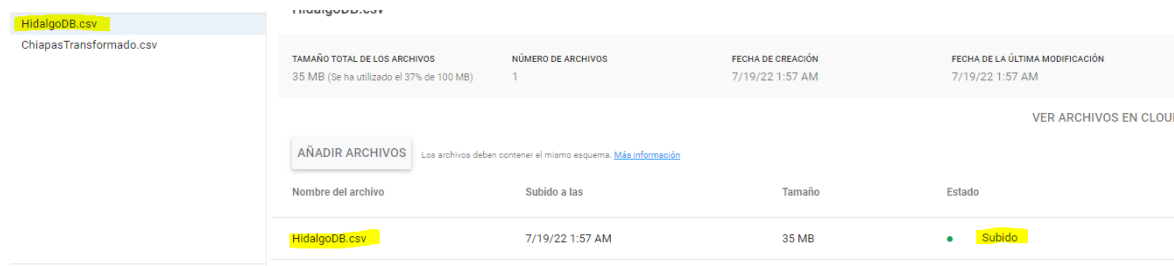



Ilustración 9 Subir Archivo


Abre el archivo a ocupar, en la parte superior derecha se presiona el botón “Conectar”.



Ilustración 10 Conexión de Archivo

Se cambia el tipo de dato de los campos a lo que se les coloco un tipo de dato incorrecto en este caso en el campo “Tabaquismo” tenia el dato “País” y se le cambio por “Texto”.

 ChiapasTransformado.csv

datos:  SALMA CID MORALES


Actualización de datos: 12 horas

Acceso a visualiza

[← EDITAR LA CONEXIÓN](#) | [FILTRAR POR CORREO ELECTRÓNICO](#)


Campo ↓	Tipo ↓
RENAL_CRONICA	RBC Texto
RESULTADO_ANTIGENO	RBC Texto
RESULTADO_LAB	RBC Texto
SECTOR	RBC Texto
SEXO	RBC Texto
TABAQUISMO	RBC Texto
TIPO_PACIENTE	RBC Texto
TOMA_MUESTRA_ANTIG...	RBC Texto
TOMA_MUESTRA_LAB	RBC Texto
UCI	RBC Texto


MÉTRICAS (1)


 ACTUALIZAR LOS CAMPOS

123 Numérico

RBC Texto

 Fecha y hora

 Booleano

 Información geográfica

123 Moneda


 URL

Ilustración 11 Cambio de datos

Se presiona el botón “Crear informe”.



Ilustración 12 Crear Informe

Se muestra una advertencia de informe y se comienza a trabajar para crear los informes.



Ilustración 13 Advertencia

Informe 1

En el primer informe se muestra una tabla con todas las enfermedades, la cual ira cambiando conforme “Una lista de control” que controla los datos de una enfermedad y esta a su vez tiene un filtro de información.

	CARDIO...	DIABETES	ASMA	EPOC	HIPERTENSION	INMUSUPR	NEUMONIA	OBESIDAD	TABAQUISMO	Record ...
1.	NO	NO	NO	NO	NO	NO	NO	NO	NO	197.448
2.	NO	NO	NO	NO	NO	NO	NO	SI	NO	12.192
3.	NO	NO	NO	NO	SI	NO	NO	NO	NO	10.997
4.	NO	SI	NO	NO	NO	NO	NO	NO	NO	7.274
5.	NO	NO	NO	NO	NO	NO	NO	NO	SI	6.229
6.	NO	SI	NO	NO	SI	NO	NO	NO	NO	5.417
7.	NO	NO	SI	NO	NO	NO	NO	NO	NO	3.315
8.	NO	NO	NO	NO	NO	NO	SI	NO	NO	3.265
9.	NO	NO	NO	NO	SI	NO	NO	SI	NO	2.558
10.	NO	SI	NO	NO	SI	NO	NO	SI	NO	1.272
11.	NO	SI	NO	NO	NO	NO	NO	SI	NO	1.185
12.	NO	NO	NO	NO	NO	NO	NO	SI	SI	1.074

Ilustración 14 Informe 1

Nombre

Filtro de FECHA_DEF

ChiapasTransformado.csv

Incluir

FECHA_DEF

Mayor que (>)

17/8/2020

Ilustración 15 Filtro

Informe 2

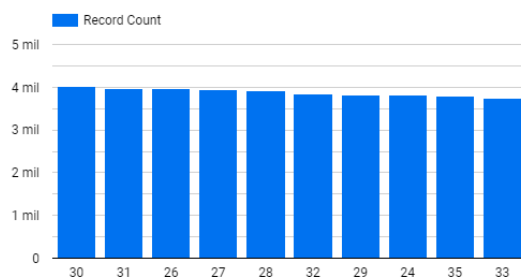
En el segundo informe se muestra dos enfermedades cuyos valores cambian dependiendo de su sexo, además de mostrar una gráfica de barras que muestra las personas cuya edad está registrada junto a la enfermedad que se muestra.

DIABETES		Record Count
1.	NO	129.066
2.	SI	10.508
3.	SE IGNORA	94

1 - 3 / 3 < >

HIPERTENSION		Record Count
1.	NO	125.292
2.	SI	14.279
3.	SE IGNORA	97

1 - 3 / 3 < >



SEXO (1)		Record Count
Escriba el término de búsqueda		
<input checked="" type="checkbox"/>	FEMENINO	139,7 mil
<input type="checkbox"/>	MASCULINO	125,9 mil

Ilustración 16 Informe 2

Informe 3

En este informe se mostrará un control de los embarazos los cuales se fueron a realizar un estudio médico, dependiendo de su nacionalidad y aplicando un filtro.

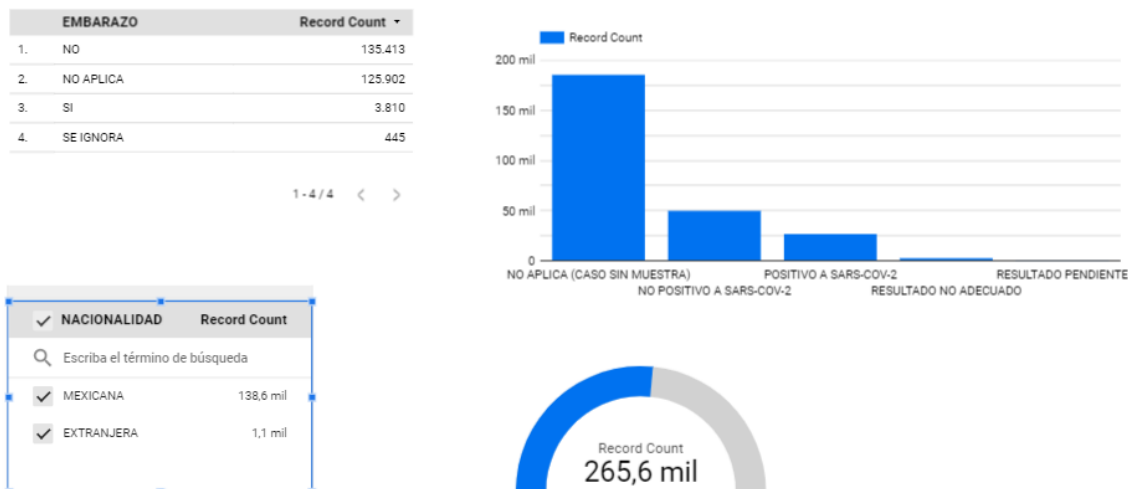


Ilustración 17 Informe 3

Nombre:

Filtro de SEXO:

ChiapasTransformado.csv

Incluir: ABC SEXO Igual que (=) FEMENINO

Ilustración 18 Filtro

Excel Dashboard

Para la elaboración de un dashboard en Excel se realizan los siguientes pasos.

Informe 4

Se desea mostrar los resultados de la tabla de embarazos cuya edad este entre los 18 y 30 años, los que sean menores de edad y mayores a los 14 años, además de que su sexo sea “Femenino”

Se abre el archivo a ocupar, en este caso ChiapasTransformado.csv, en la parte superior se abre el apartado de “Insertar/Tabla dinámica/De una tabla o rango” se selecciona “Nueva hoja de cálculo” y aceptar.

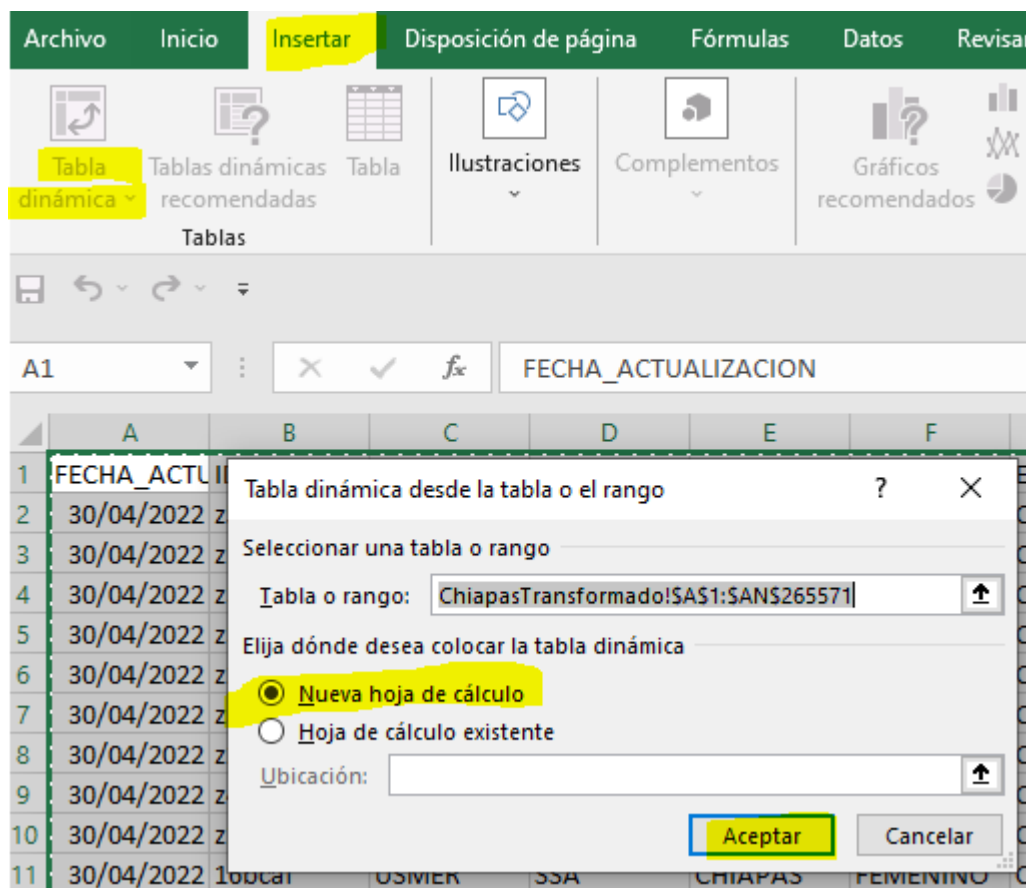


Ilustración 19 Excel Tabla dinámica

Se mostrar la siguiente tabla, con la cual se podrá comenzar a trabajar.

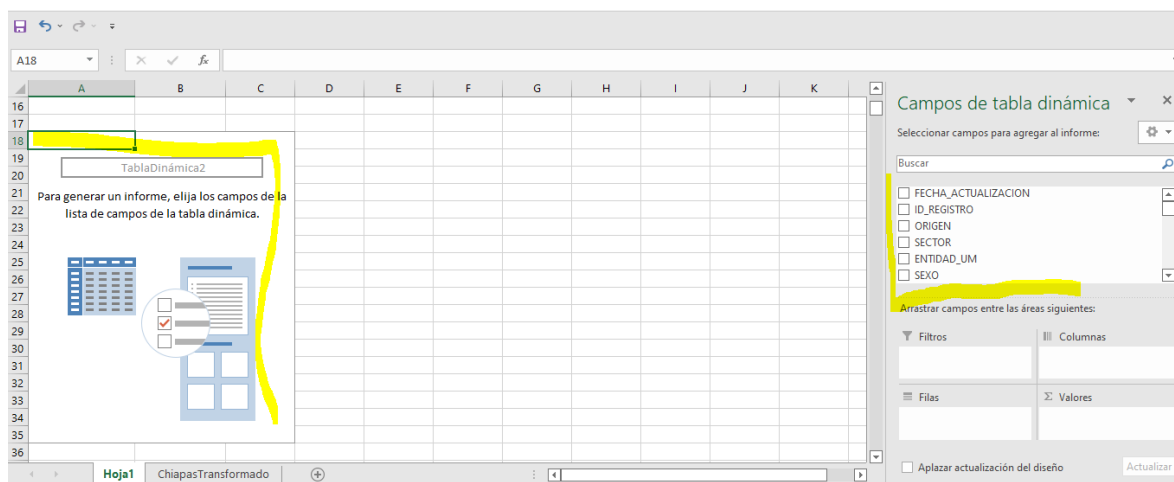


Ilustración 20 Excel Tabla dinámica 2

Para la elaboración del dashboard se comienza por manipular los datos de la parte derecha los cuales se verán reflejados sobre la tabla dinámica.

Se seleccionan los datos a ocupar “Sexo, Edad, Embarazo”, en la parte inferior se van arrastrando los valores, el campo a tomar en cuenta es “Embarazo”, los valores a tomar en cuenta son “Sexo” y los filtros a aplicar son que el Sexo sea Femenino de entre 18 y 30 años.

Para crear el grafico, se va al apartado de grafico dinamico y se selecciona el grafico a ocupar.

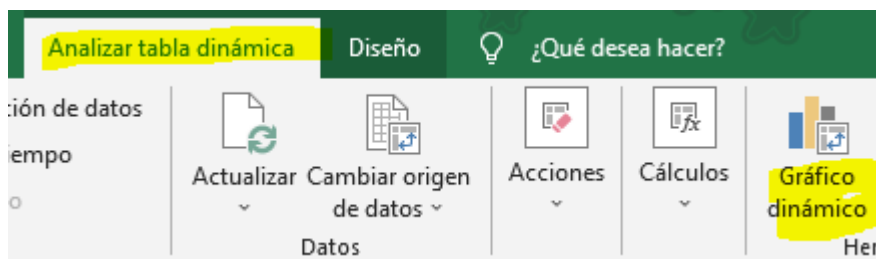


Ilustración 21 Excel Grafico dinamico

El informe queda de esta manera para los datos establecidos.

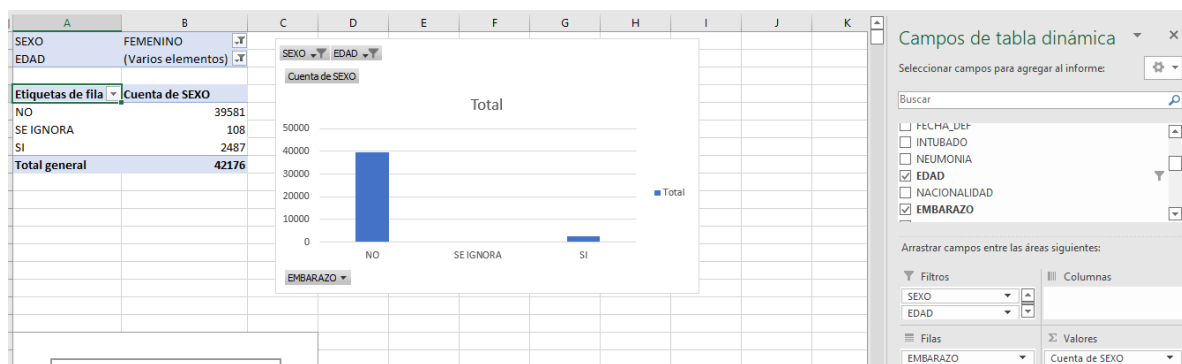


Ilustración 22 Informe 4

El segundo filtro a aplicar es que sean de entre 15 y 17 años de sexo Femenino.

Se siguen los pasos anteriores y el informe queda de la siguiente manera.

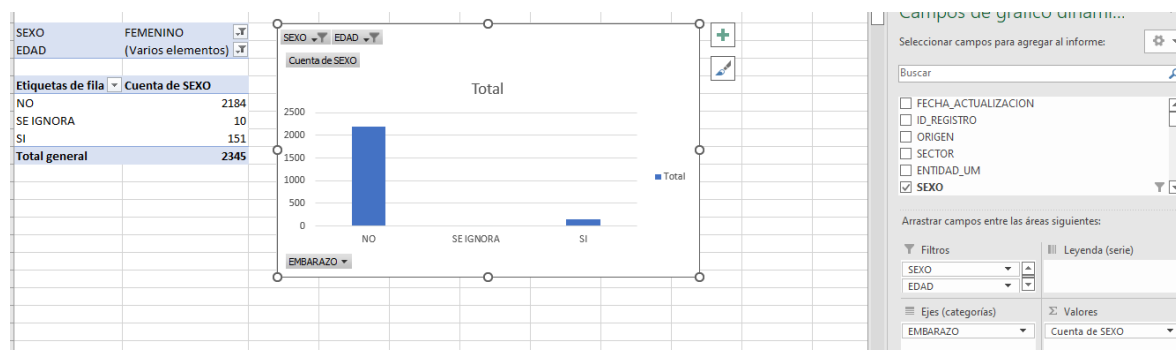


Ilustración 23 Informe 4

Conclusión

Durante la elaboración de este documento, me di cuenta que un Dashbohard es de alguna manera sencilla siempre y cuando tengas un objetivo al cual quieres llegar y mostrar además de tener una limpieza de datos correcta.

Las herramientas que se utilizaron, aunque muy distintas, se trabajan casi de la misma forma ya que puedo seleccionar los datos a ocupar, colocar filtros y valores necesarios para llegar a mi objetivo.

A mi parecer me agrado mas el uso del Excel ya que es más fácil de comprender y su interfaz en un poco más amigable para trabajar. Aun que Google Data Studio tiene la ventaja de ser en línea y puedo hacer los informes que necesite en cualquier dispositivo.



Ilustración 24 Logo