Presentación Trabajo Final

Proyecto Inversiones IT

Periodo 2020 - 2021



Integrantes

Carlos Leguizamon Milagros Lizarraga Tannyr

Institución

CoderHouse

Fecha de Entrega

31/03/2022

CONTENIDO

1.Introducción1
2.Descripción de la temática de datos1
3. Alcance
4. Hipótesis1
<u>5. Herramientas</u> 1
6. Datasets
7. Listado de tablas5
8. Listado de columnas por tablas6
9. Diagrama entidad-relación
10. Modelo relacional en Power BI
11. Medidas calculadas:9

1.Introducción

El desarrollo continuo de todos los sistemas informáticos lleva a que cada gobierno de cada nación del mundo tenga que actualizarse día a día para poder trabajar de manera eficiente acorde a dichas actualizaciones.

Lo que conlleva este tipo de actualización, ya sean actualizaciones de hardware o software es una inversión en todo tipo de tecnologías que nos ayuden a poder trabajar de una manera más sofisticada y poder cumplir con objetivos.

En el presente trabajo se abordó esta temática, dado que es un tema muy importante que antes se consideraba el futuro, pero que se está convirtiendo poco a poco en el presente del dia a dia para cada organismo institucional,

Ya que mundialmente existen infinidad de datos sobre inversiones IT decidimos trabajar sobre una base más acotada de inversiones.

En este caso nos enfocamos únicamente en las inversiones IT que realizó el Gobierno de la Nación Argentina, cuyas inversiones fueron realizadas en muchas localidades de las distintas provincias que componen la Nación.

En conclusión en este trabajo observaremos los diferentes proyectos de inversiones en Infraestructura Tecnológica realizada por la Nación Argentina en los diferentes organismos municipales en los años 2020 y 2021, cuales fueron las provincias que más proyectos realizan y en cuales sectores se invirtió más.

2. Descripción de la temática de datos

Para este trabajo hemos usado la información de bases de datos de los proyectos de inversión en IT, obtenidas de la página Datos.gob.ar

Se analizó tanto de la primera inversión propuesta hasta la última, como por ejemplo qué sectores fueron los más desarrollados, cuales provincias fueron las principales en recibir más inversiones, que año fue la convocatoria de cada proyecto y los organismos municipales donde se invirtió.

3. Alcance

El siguiente trabajo está diseñado para un análisis

4. Hipótesis

El objetivo es poder estudiar y cuáles fueron los proyectos realizados por parte del gobierno de la nación Argentina.

Abarca desde la fecha que se realizó la convocatoria del proyecto hasta la realización en el organismo prometido.

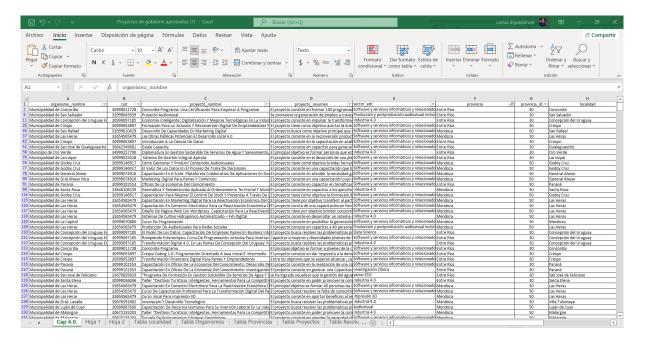
5. Herramientas

Excel: para organización de los Datasets Power point: para el diseño del Mockup

Power BI Desktop: para la creación del tablero de control

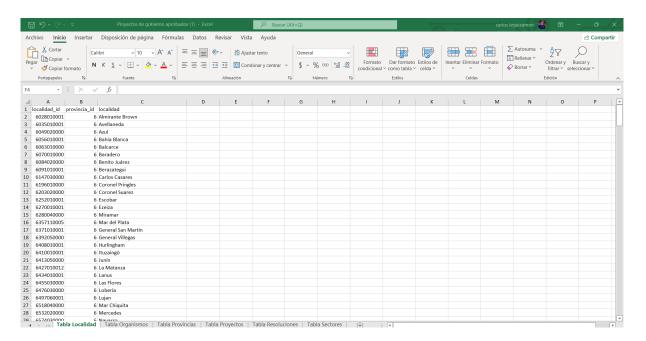
6. Datasets

Al cargar el Dataset teníamos todos los datos muy desorganizados como puede observar en la siguiente imagen



Luego de realizar una limpieza y organización, pudimos crear las siguientes tablas

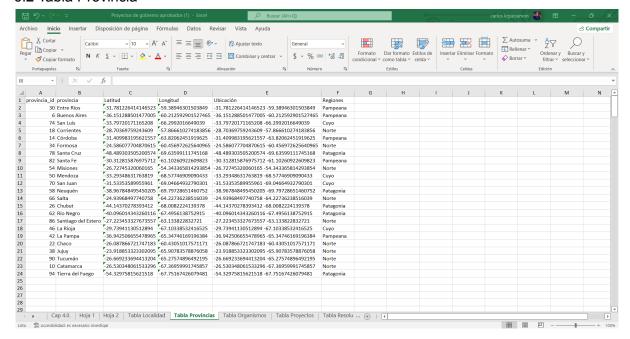
6.1 Tabla Localidades



Que nos muestra la columna Localidad_id como Primary Key de la tabla, en donde podemos encontrar el nombre de las localidades y también el provincia id como Foreign Key de la tabla

También realizamos una limpieza de datos duplicados, ya que el dataset tenía muchos datos duplicados, para mayor comodidad hemos diferenciado los datos de las localidades, tanto nombre, id y lo hemos filtrado borrando los duplicados para crear la tabla para poder crear fácilmente la tabla entidad-relación

6.2 Tabla Provincia

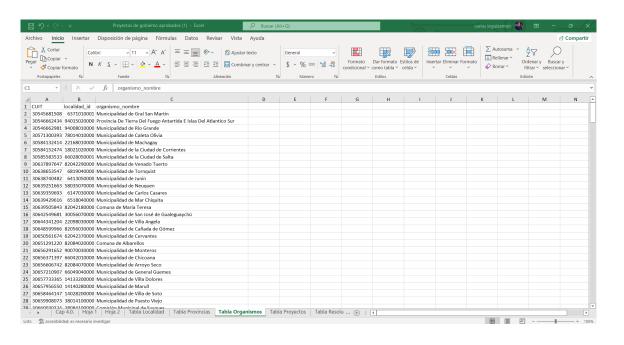


Para realizar la tabla provincia hemos también realizado un filtrado de datos duplicados ya que en la dataset se repetían los datos en la columna provincia id y provincia

El cual elegimos a provincia id como Primary Key

También hemos agregado los datos de longitud, latitud, ubicación y región de cada provincia para un análisis geográfico

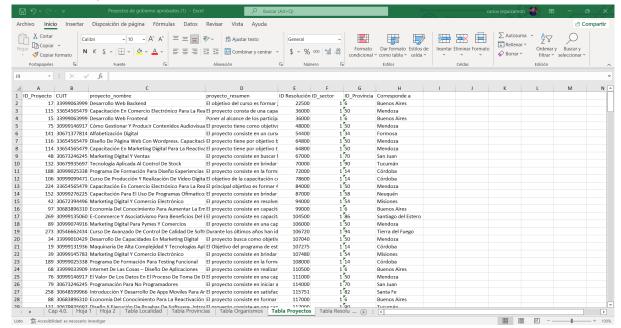
6.3 Tabla Organismos



En la tabla organismos trabajamos también filtrando y limpiando datos duplicados, en este caso se trata de los organismos municipales donde se realizan las inversiones en IT

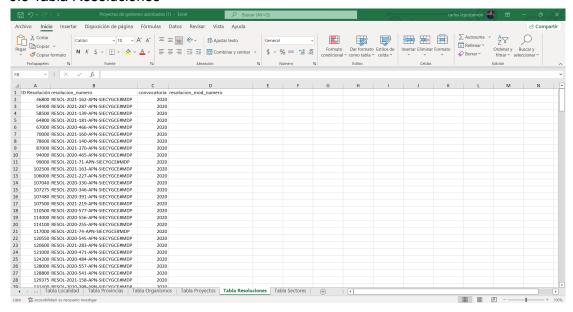
Tiene una Primary Key que vendría a ser el CUIT de cada organismo municipal, el nombre del organismo y agregando a la localidad_id como Foreign Key

6.4 Tabla Proyectos



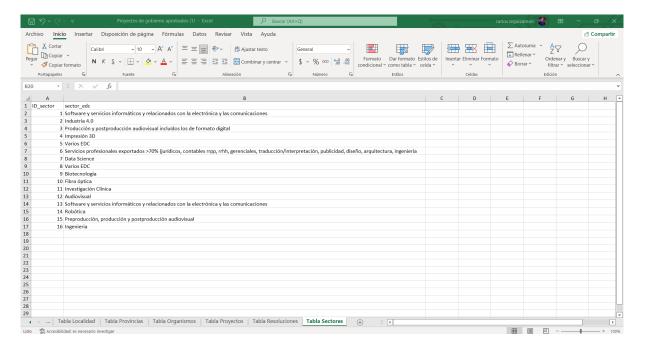
La tabla proyectos es la que más datos presenta, esta formada por las columnas proyecto_id que identifica cada proyecto, es la Primary Key, la tabla también contiene información del nombre del proyecto, un breve resumen de que se trata tal proyecto, un CUIT que trabaja como Foreign Key, id_sector que también trabaja como Foreign Key, id_resolucion que es el número que identificamos la resolución del proyecto y un id_provincia para identificar de qué provincia se trata.

6.5 Tabla Resoluciones



También trabajada y filtrando los duplicados, obtenemos una tabla con los datos de las resoluciones de cada proyecto, con id_resolucion como Primary Key, el número de la resolución y la fecha de convocatoria de cada proyecto realizado

6.6 Tabla Sectores



La tabla sectores representa los datos de los distintos sectores a los que se les atribuyen las inversiones, esta trabajada ya que filtramos para no tener duplicados, y asignamos un valor id para cada uno de los sectores, haciéndola una Primary Key Id_Sector, y también está formado por una columna con el nombre de cada sector a la que hace referencia la inversión.

7. Listado de tablas

En este apartado se hará mención de las tablas junto con una breve descripción de lo que representan las mismas y definición de claves primarias y foráneas

Provincia: Contiene el ID de la provincia que realiza el proyecto y el nombre de la misma, la ubicación geográfica y la región a la cual pertenece

PK: Provincia_id

Nombre_provincia

Longitud

Latitud

Ubicación

Región

Localidad: Contiene el ID de la localidad que realiza el proyecto, el nombre y también tiene el ID provincia como Foreign key

PK: Localidad_id Nombre_localidad FK: Provincia_id Organismo: Contiene el CUIT del organismo que realiza el proyecto, el nombre del mismo y el id de la localidad a la cual pertenece el organismo como Foreign Key

PK: CUIT

Nombre_organismo FK: Localidad_id

Proyecto: Contiene el id proyecto al cual se le atribuye la inversión, el nombre , un breve resumen de lo que se basa el proyecto y luego tiene tres foreign key que hace relacionar a la tabla proyecto con las tablas organismos, resoluciones y sector

PK: id_proyecto Nombre_Proyecto Proyecto_Resumen FK: CUIT

FK: id_resolucion FK: id_sector

Resolución: Contiene el ID de la resolución que tuvo el proyecto, la fecha de la convocatoria y el número de dicha resolución

PK: ID_Resolucion Numero_resolución Convocatoria resolucion_mod_num

Sectores: Esta tabla contiene el sector a la cual se le atribuye la inversion del proyecto, contiene el id_sector y los nombres de los sectores

PK: id_sector
sector_edc

8. Listado de columnas por tablas

Localidad		
Campo	Tipo de campo	Tipo de clave
Localidad_id	INT	PK
Provincia_id	INT	
Localidad	Varchar(50)	

Provincia			
Campo	Tipo de campo	Tipo de clave	
Provincia_id	INT	PK	
Provincia	INT		
Latitud	Varchar(50)		
Longitud	varchar50)		
Ubicación	Varchar(50)		
Región	varchar(20)		

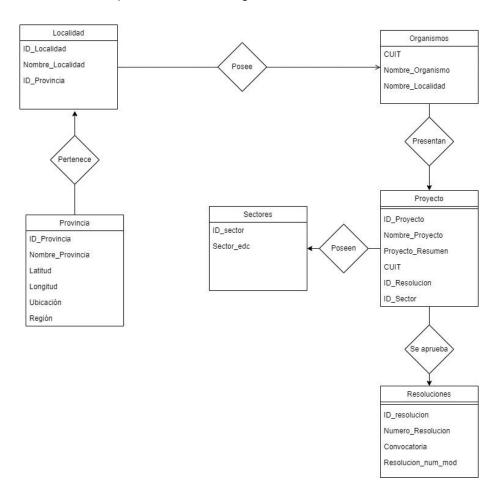
Proyectos			
Campo	Tipo de camp	o Tipo de clav	/e
ID_Proyecto	INT	FK	
CUIT	INT		
Proyecto_nombre	Varchar(500)	
Proyecto_resumen	Varchar(500)	
ID_Resolucion	INT	PK	
ID_Sector	INT	PK	
ID_Provincia	INT	PK	
Organismo			
Campo	Tipo de campo	Γipo de clave	
CUIT	INT	PK	
Localidad_id	INT	FK	
organismo_nombre	Varchar(100)		

Resoluciones			
Campo	Tipo de campo	Tipo de clave	
ID_Resolucion	INT	FK	
Resolucion_numero	Varchar(100)		
Convocatoria	Varchar(500)		
Resolucion_mod_numero	Varchar(500)		

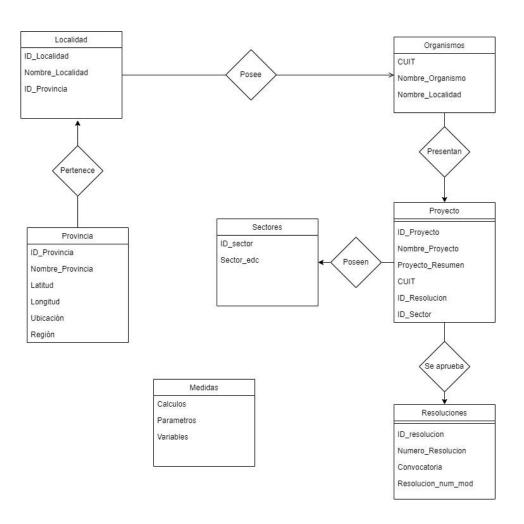
Sectores			
Campo	Tipo de campo	Tipo de clave	
ID_Sector	INT	FK	
Sector edc	Varchar(100)		

9. Diagrama entidad-relación

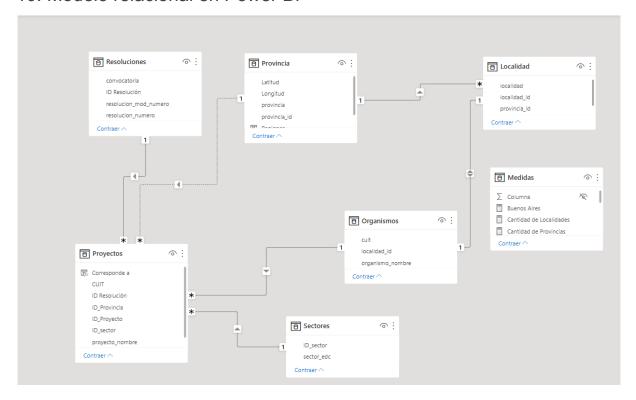
A continuación presentamos el diagrama entidad-relacion, asi fue al comienzo del proyecto



Y asi quedo al final del proyecto habiéndose agregado la tabla medidas, conformada por cálculos, parámetros y variables.



10. Modelo relacional en Power BI



Una vez cargado los datos

Hemos realizado los siguientes cambios

Tabla Provincia: Hemos agregado las columnas Latitud, Longitud, Ubicación y Región para poder segregar de mejor manera en un mapa donde esta cada provincia

Tabla Sectores: La hemos realizado un filtrado de datos para eliminar los duplicados y poder trabajar mejor a la hora de realizar gráficos

Tabla Localidad: También al haber filtrado datos para eliminar los duplicados, pudimos trabajar de mejor manera desarrollando filtros

11. Medidas calculadas:

Creamos las medidas calculadas

- 1.Cantidad de Localidades : Cantidad de Localidades = COUNT(Localidad[localidad_id])
- 2.Cantidad de Provincias : Cantidad de Provincias = COUNT(Provincia[provincia id])
- 3.Cantidad de Proyectos : Cantidad de proyectos = COUNT(Proyectos[ID Proyecto])
- 4.Cantidad de Sectores : Cantidad de Sectores = COUNT(Sectores[sector edc])
- 5 Para crear un gráfico de torta con el TOP 5 de provincias con más proyectos creamos las siguientes medidas:
- 5.1 Buenos Aires: Buenos Aires = VAR Total_Proyectos = COUNT(Proyectos[ID_Provincia])

```
VAR Proyectos_Buenos_Aires = CALCULATE(COUNT(Proyectos[ID_Provincia]), Proyectos[ID_Provincia] = "6")
```

RETURN DIVIDE(Proyectos_Buenos_Aires,Total_Proyectos)

5.2 Córdoba : Córdoba = VAR Total Proyectos = COUNT(Proyectos[ID Provincia])

VAR Proyectos_Cordoba = CALCULATE(COUNT(Proyectos[ID_Provincia]),Proyectos[ID_Provincia] = "14")

RETURN DIVIDE(Proyectos_Cordoba,Total_Proyectos)

5.3 Entre Ríos: Entre Ríos = VAR Total Proyectos = COUNT(Proyectos[ID Provincia])

VAR Proyectos Entre Rios = CALCULATE(COUNT(Proyectos[ID Provincia]), Proyectos[ID Provincia] = "30")

RETURN DIVIDE(Proyectos_Entre_Rios,Total_Proyectos)

5.4 Mendoza : Mendoza = VAR Total_Proyectos = COUNT(Proyectos[ID_Provincia])

miArgentina

```
VAR Proyectos_Mendoza = CALCULATE(COUNT(Proyectos[ID_Provincia]),Proyectos[ID_Provincia] = "50")
RETURN DIVIDE(Proyectos Mendoza, Total Proyectos)
 5.5 Santa Fe: Santa Fe = VAR Total_Proyectos = COUNT(Proyectos[ID_Provincia])
VAR Proyectos_Santa_Fe = CALCULATE(COUNT(Proyectos[ID_Provincia]), Proyectos[ID_Provincia] = "82")
RETURN DIVIDE(Proyectos_Santa_Fe,Total_Proyectos)
6. Promedio Proyectos Provincia de Buenos Aires: Promedio Proyectos Provincia de Buenos Aires = VAR Total Proyectos =
COUNT(Proyectos[ID_Provincia])
VAR Proyectos_BuenosAires = CALCULATE(COUNT(Proyectos[ID_Provincia]), Proyectos[ID_Provincia] = "6")
RETURN DIVIDE (Proyectos_BuenosAires,Total_Proyectos)
7. Proyectos 2020 : Proyectos 2020 = VAR Total_proyectos_2020 = CALCULATE(COUNT(Resoluciones[convocatoria]), Resoluciones[convocatoria] = 2020)
RETURN Total_proyectos_2020
8. Proyectos 2021 : Proyectos 2021 = VAR Total_proyectos_2021 = CALCULATE(COUNT(Resoluciones[convocatoria]),Resoluciones[convocatoria] = 2021)
RETURN Total_proyectos_2021
9. Proyectos Totales (Parámetro): Proyectos total = CALCULATE([Cantidad de proyectos]) * '% del total'[Valor % del total]
```

FILMINA 1

Agregación de filtros: Agregamos 4 filtros para poder realizar segregaciones en base a provincia, año de convocatoria, regiones y sectores donde se realizan las inversiones



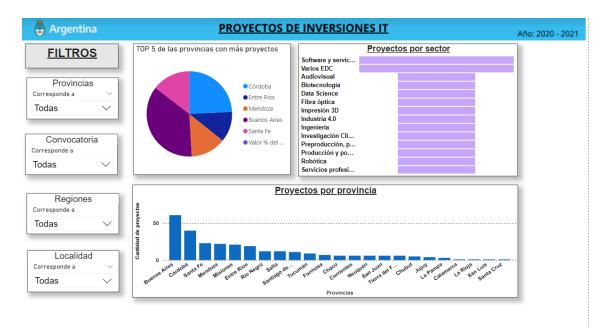
Gráfico Torta: Nos muestra un TOP 5 de las provincias que más proyectos desarrollaron



Gráfico de tablas : nos muestra cuantos proyectos tiene cada provincia, desde la que más proyectos tiene hasta la que menos proyectos tiene



Todo junto nos muestra en una solapa de gráficos: las provincias por proyectos, convocatoria, regiones y localidad.



FILMINA 2

Mapa argentino: donde mostramos la ubicación geográfica de cada provincia. Este gráfico está sujeto a un filtro de provincias



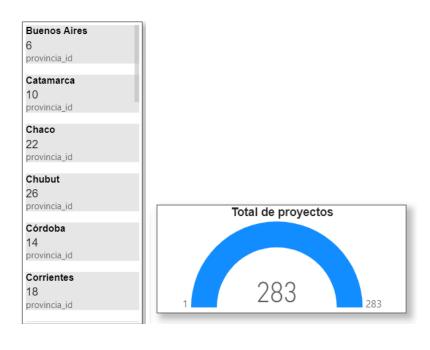
Adicionamos los siguientes filtrados:

-Filtro por regiones, donde el interactúa con los gráficos delimitando las provincias que se ubican allí

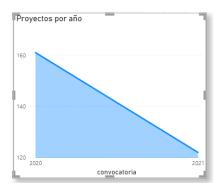
Regiones

Cuyo
Norte
Pampeana
Patagonia

-Filtro por provincias con su respectivo id_provncia que nos calcula el KPI total de proyectos que tiene cada provincia



Por último, agregamos un gráfico donde nos muestra la variación interanual de la cantidad de proyectos que fueron creados durante el periodo de 2020 - 2021



Aquí mostramos como queda todo junto en la última solapa donde mostramos el mapa del país, donde podemos ver como se ve trabajado y filtrado en este caso con la provincia de Buenos Aires

