

# CASO PRÁCTICO

GRAVITY PORTAL. MREID

**África Espejo Salinas**



## INDICE

OBJETIVOS .....	2
DESCRIPCIÓN DEL CASO PRÁCTICO .....	2
ETL .....	2
SQL.....	6
PREGUNTAS PLANTEADAS.....	7
POWER BI.....	15
CONCLUSIONES .....	25

## OBJETIVOS

- Análisis y entendimiento de los datos origen.
- Conexión, preparación y carga de datos en una base de datos MySQL (a través de n8n).
- Preparación de estructuras de datos y consultas SQL.
- Responder a las preguntas planteadas.
- Configuración de herramientas de visualización de datos.
- Construcción de al menos un panel o dashboard que permita analizar de forma completa los datos (a través de Power BI Desktop)

## DESCRIPCIÓN DEL CASO PRÁCTICO

Partiendo del Gravity Portal: MREID (Multinational Revenue, Employment, and Investment Database), de la United States International Trade Commission, disponible en <https://www.usitc.gov/data/gravity/mreid.htm> se quiere hacer un análisis visual de los distintos países a nivel mundial.

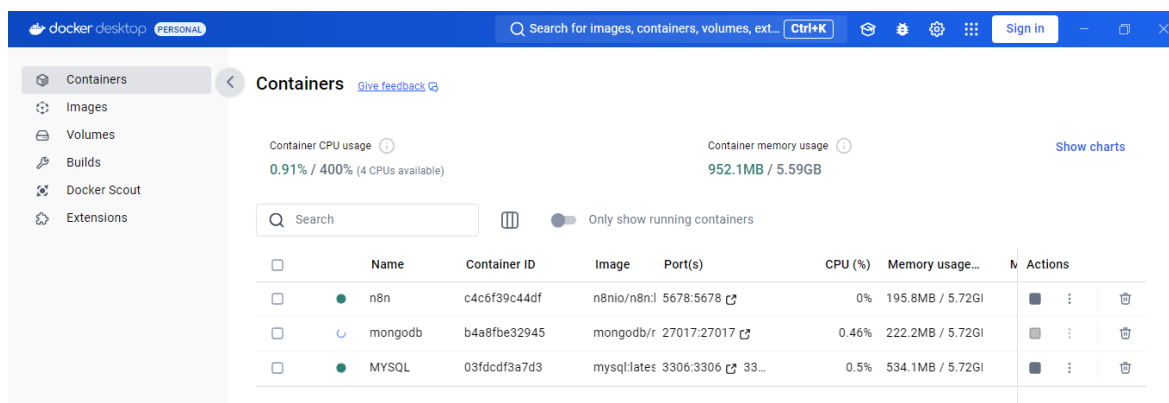
El MREID proporciona información sobre los ingresos, el empleo y la inversión multinacional para un gran número de países e industrias. Las industrias abarcan los sectores de la agricultura, la minería, la energía, la manufactura y los servicios, lo que permite una descripción casi completa de la actividad de inversión extranjera directa (IED) de las empresas multinacionales con filiales transfronterizas en 185 países, 25 industrias y un período de 12 años (2010 - 2021).

Este documento está en formato separado por comas (csv) y se cargara a través de una herramienta ETL, en este caso n8n, a una base de datos relacional (MySQL). Tras ello se realizara un análisis estadístico de los datos y se responderán las preguntas planteadas. Además, se desarrollara un dashboard en Power Bi Desktop para un análisis visual de los datos.

## ETL

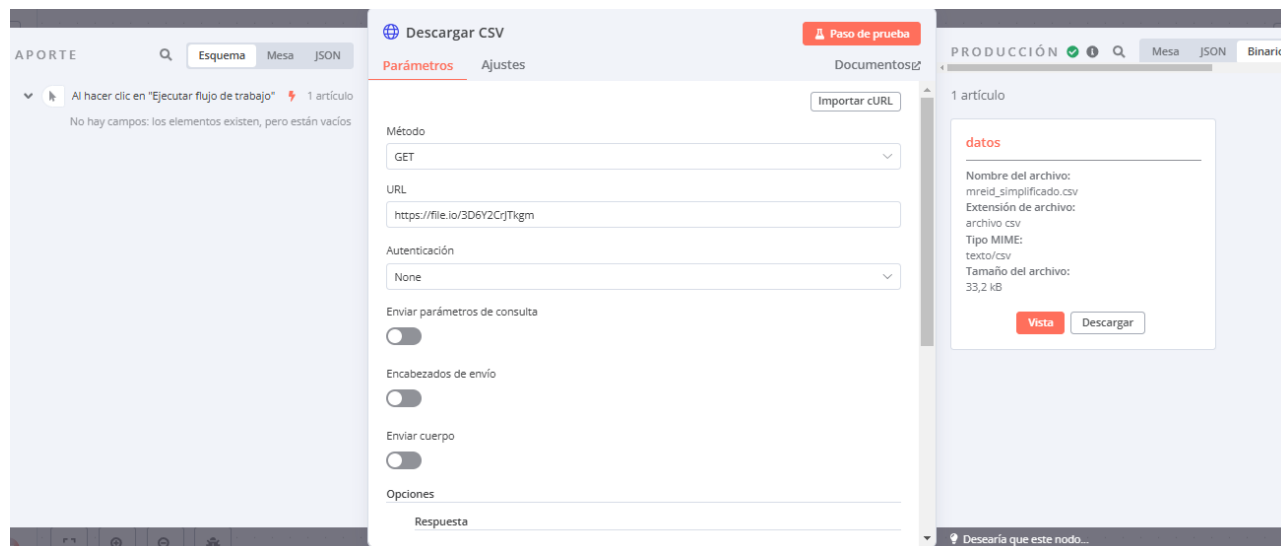
El ETL que se usa en este caso es el n8n, el cual usaremos para descargar un archivo CSV desde la página web Gravity Portal: MREID, que después cargaremos en un DataWarehouse (MySQL) para realizar el análisis de los datos.

En primer lugar, activaremos el n8n y MySQL en Docker para poder usarlos.



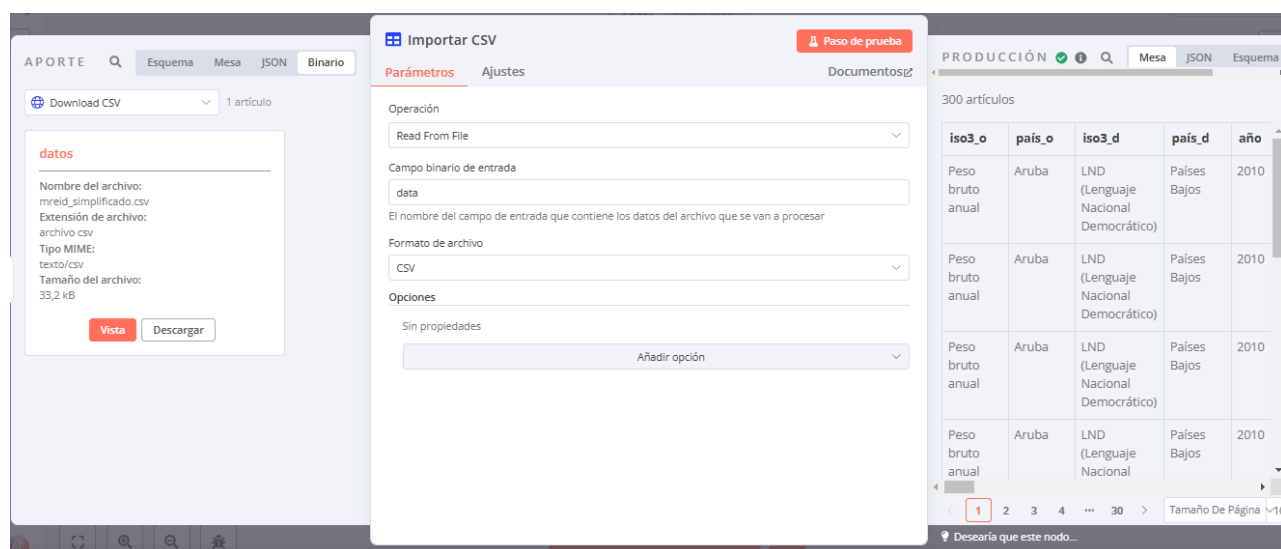
## Pasos a seguir en n8n:

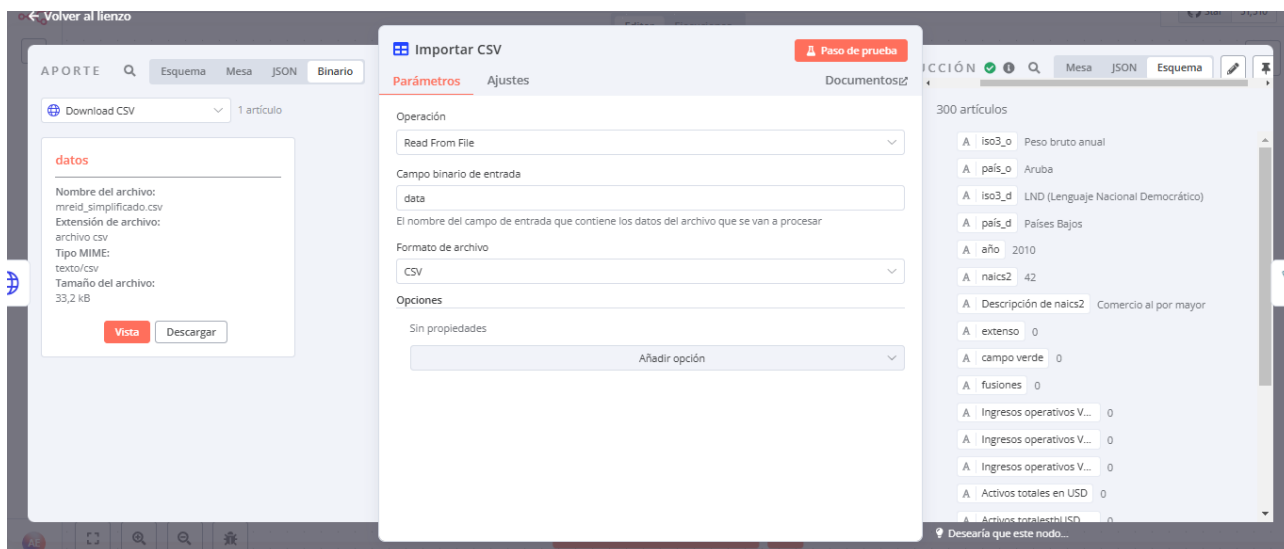
1. Configurar el nodo “HTTP Request” para descargar el archivo CSV desde la página web. En este paso nos encontramos con el problema de que el archivo original es muy grande (40.4 MB), por lo que tendremos que usar file.io para obtener un archivo de menor capacidad.



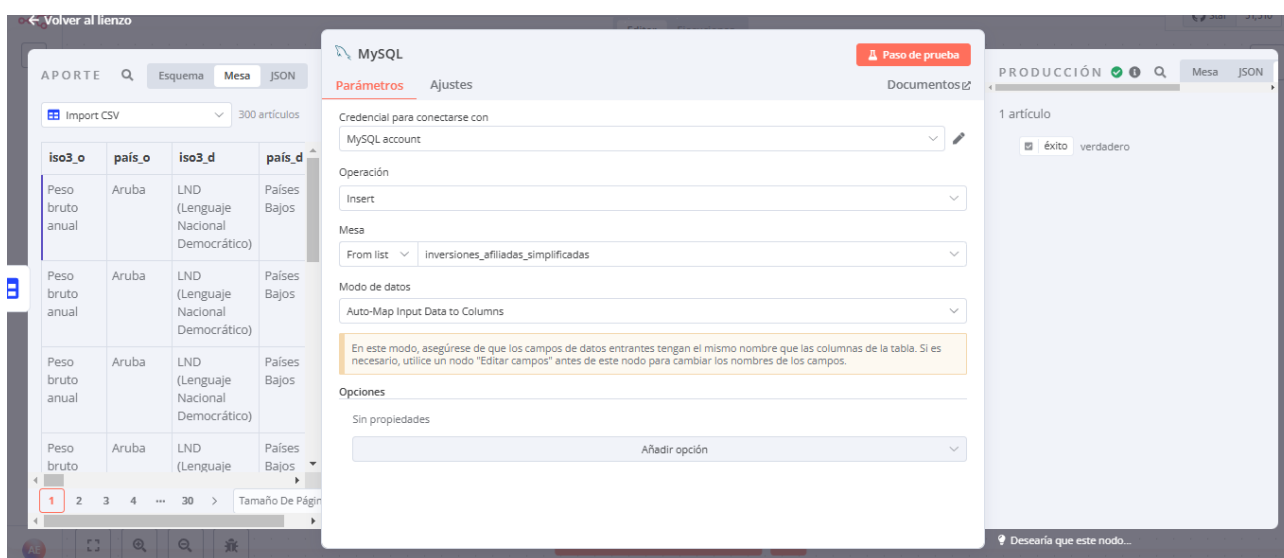
2. Una vez descargado el archivo CSV, usamos el nodo “Read from Binary File” para convertir los datos del archivo CSV en un formato legible.

En esta imagen se observa el resultado en formato tabla, y en la siguiente en formato lista.

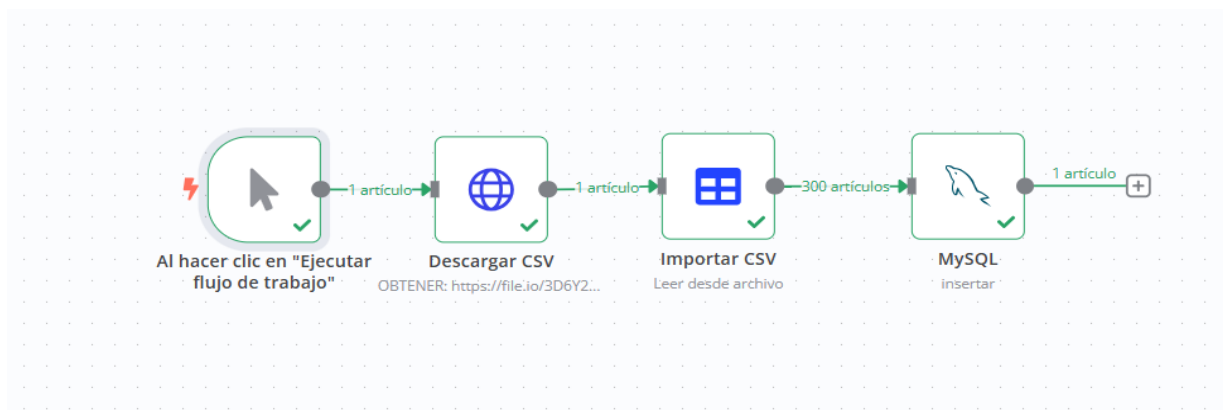




- Después de haber leído el archivo CSV insertaremos estos datos en una base de datos MySQL, a través del nodo "MySQL". Para ello, debemos tener configurada la conexión a MySQL con nuestros datos. Como se muestra en la imagen, el resultado es verdadero, es decir funciona correctamente.



- Ejecutaremos el flujo de trabajo para comprobar que todo funciona correctamente.



Para comprobar que el workflow funciona correctamente generaremos unos datos y los insertaremos en la tabla MySQL.



## 5. Generar los datos a través del nodo "Set"

La interfaz de usuario muestra la configuración del nodo "Añadir campo" en modo "Manual Mapping". Los campos a configurar son:

- country\_o: String
- ejemplo\_n8n: String
- year: String
- 0808: String

El cuadro de resultados a la derecha muestra la siguiente estructura de datos:

A	país_o	ejemplo_n8n
A	año	0808

Un mensaje de estado indica: "Nodo ejecutado exitosamente".

## 6. Insertar los datos en MySQL por el nodo "MySQL insert". Como podemos observar en el cuadro de resultados de la derecha el flow funciona correctamente.

La interfaz de usuario muestra la configuración del nodo "MySQL1" en modo "Insert". Los parámetros configurados son:

- Credencial para conectarse con: MySQL account
- Operación: Insert
- Mesa: From list | inversiones\_afiliadas\_simplificadas
- Modo de datos: Auto-Map Input Data to Columns

El cuadro de resultados a la derecha muestra la siguiente estructura de datos:

país_o	año
ejemplo_n8n	0808

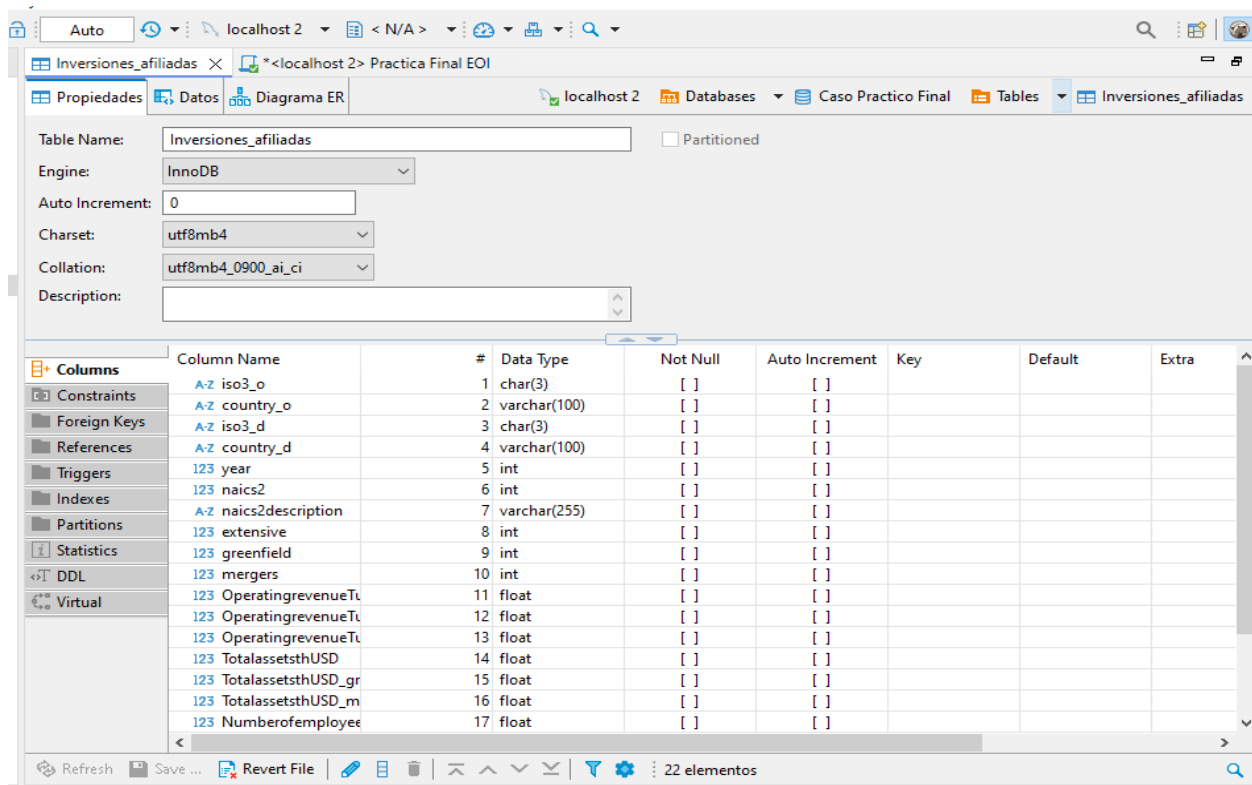
Un mensaje de estado indica: "Nodo ejecutado exitosamente".

## SQL

Para poder realizar el análisis estadístico de los datos y responder a las preguntas, debemos activar MYSQL en Docker. Tras ello, realizaremos el análisis a través de Dveaber.

En primer lugar, crearemos una tabla con la siguiente estructura:

```
CREATE TABLE Inversiones_afiliadas (  
    iso3_o CHAR(3),  
    country_o VARCHAR(100),  
    iso3_d CHAR(3),  
    country_d VARCHAR(100),  
    year INT,  
    naics2 INT,  
    naics2description VARCHAR(255),  
    extensive INT,  
    greenfield INT,  
    mergers INT,  
    OperatingrevenueTurnover FLOAT,  
    OperatingrevenueTurnover_green FLOAT,  
    OperatingrevenueTurnover_mergers FLOAT,  
    TotalassetsthUSD FLOAT,  
    TotalassetsthUSD_green FLOAT,  
    TotalassetsthUSD_mergers FLOAT,  
    Numberofemployees FLOAT,  
    Numberofemployees_green INT,  
    Numberofemployees_mergers INT,  
    FixedassetsthUSD FLOAT,  
    FixedassetsthUSD_green FLOAT,  
    FixedassetsthUSD_mergers FLOAT  
);
```



Después importaremos los datos del MREID a la tabla creada en Dveaber.

	iso3_o	country_o	iso3_d	country_d	year	naics2	naics2descript	extensive	greenfield	m
7188	AUS	Australia	USA	United States	2014	99	Unclassified Establishm	84	1	
7189	AUS	Australia	VGB	British Virgin Islands	2014	21	Oil and Gas Extraction	1	0	
7190	AUS	Australia	VNM	Vietnam	2014	21	Oil and Gas Extraction	1	0	
7191	AUS	Australia	VNM	Vietnam	2014	23	Construction of Buildi	0	0	
7192	AUS	Australia	VNM	Vietnam	2014	32	Materials Manufacturi	1	0	
7193	AUS	Australia	VNM	Vietnam	2014	42	Wholesale Trade	1	0	
7194	AUS	Australia	VNM	Vietnam	2014	52	Finance and Insurance	2	0	
7195	AUS	Australia	VNM	Vietnam	2014	54	Legal Services	0	0	
7196	AUS	Australia	VNM	Vietnam	2014	55	Management of Comp	1	0	
7197	AUS	Australia	VNM	Vietnam	2014	56	Administrative and Suj	2	0	
7198	AUS	Australia	ZAF	South Africa	2014	52	Finance and Insurance	1	0	
7199	AUS	Australia	ARE	United Arab Emirates	2015	42	Wholesale Trade	0	0	
7200	AUS	Australia	ARG	Argentina	2015	52	Finance and Insurance	1	0	
7201	AUS	Australia	AUS	Australia	2015	11	Agriculture, Forestry, F	33	0	
7202	AUS	Australia	AUS	Australia	2015	21	Oil and Gas Extraction	110	5	
7203	AUS	Australia	AUS	Australia	2015	22	Utilities	47	0	
7204	AUS	Australia	AUS	Australia	2015	23	Construction of Buildi	108	3	
7205	AUS	Australia	AUS	Australia	2015	31	Food and Textile Manu	41	1	
7206	AUS	Australia	AUS	Australia	2015	32	Materials Manufacturi	41	0	
7207	AUS	Australia	AUS	Australia	2015	33	Finished Product Mani	59	2	
7208	AUS	Australia	AUS	Australia	2015	42	Wholesale Trade	145	7	
7209	AUS	Australia	AUS	Australia	2015	44	Food and Beverage Stc	35	0	
7210	AUS	Australia	AUS	Australia	2015	45	Miscellaneous Store Ri	10	0	

## PREGUNTAS PLANTEADAS

### 1. ¿Cuál es el top 10 países donde el ratio “Facturación por empleado” es más alto?

```
SELECT
    country_d,
    SUM(OperatingrevenueTurnover) / SUM(Numberofemployees) AS Facturacion_por_empleado
FROM Inversiones_afiliadas ia
GROUP BY country_d
HAVING SUM(Numberofemployees) > 0
ORDER BY Facturacion_por_empleado DESC
LIMIT 10;
```

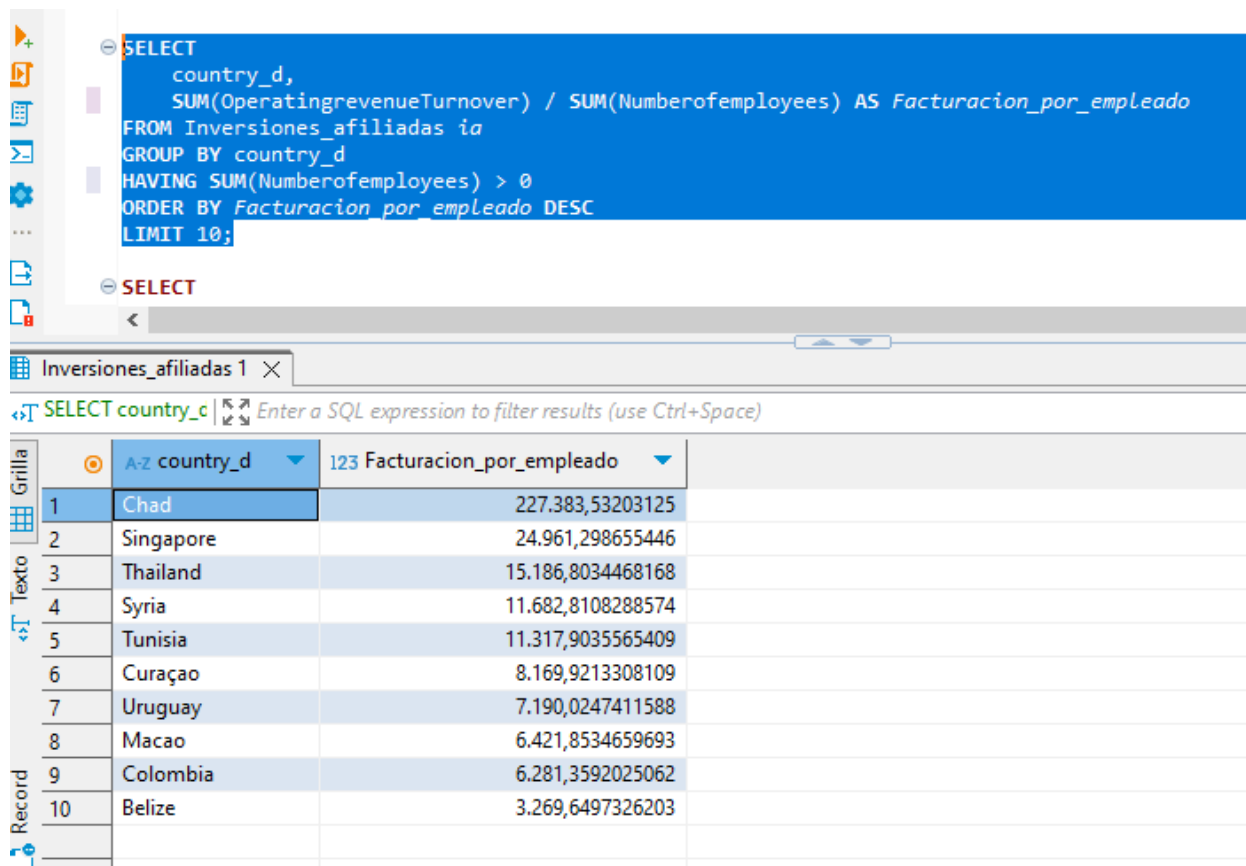
Calculo la ratio de facturación por empleado sumando cada valor por adelantado antes de realizar la división para evitar un error cuando las cifras varían mucho entre diferentes registros para el mismo país, es decir evito hacer la facturación a nivel de cada fila.

Uso el HAVING para asegurarme que haya empleados y evitar que nos de error por ser imposible dividir entre 0.

Se usa DESC (descendente) en ORDER BY para obtener los resultados con los valores más altos, es decir ordenados de mayor a menor.

La función LIMIT se encarga de limitar la cantidad de resultados que nos devuelve la función.





The screenshot shows a database query editor with a SQL query in a blue box and its results in a table view below. The query is:

```
SELECT
    country_d,
    SUM(OperatingrevenueTurnover) / SUM(Numberofemployees) AS Facturacion_por_empleado
FROM Inversiones_afiliadas ia
GROUP BY country_d
HAVING SUM(Numberofemployees) > 0
ORDER BY Facturacion_por_empleado DESC
LIMIT 10;
```

The table view shows the results of the query, ordered by 'Facturacion\_por\_empleado' in descending order. The table has two columns: 'country\_d' and 'Facturacion\_por\_empleado'.

	country_d	Facturacion_por_empleado
1	Chad	227.383,53203125
2	Singapore	24.961,298655446
3	Thailand	15.186,8034468168
4	Syria	11.682,8108288574
5	Tunisia	11.317,9035565409
6	Curaçao	8.169,9213308109
7	Uruguay	7.190,0247411588
8	Macao	6.421,8534659693
9	Colombia	6.281,3592025062
10	Belize	3.269,6497326203

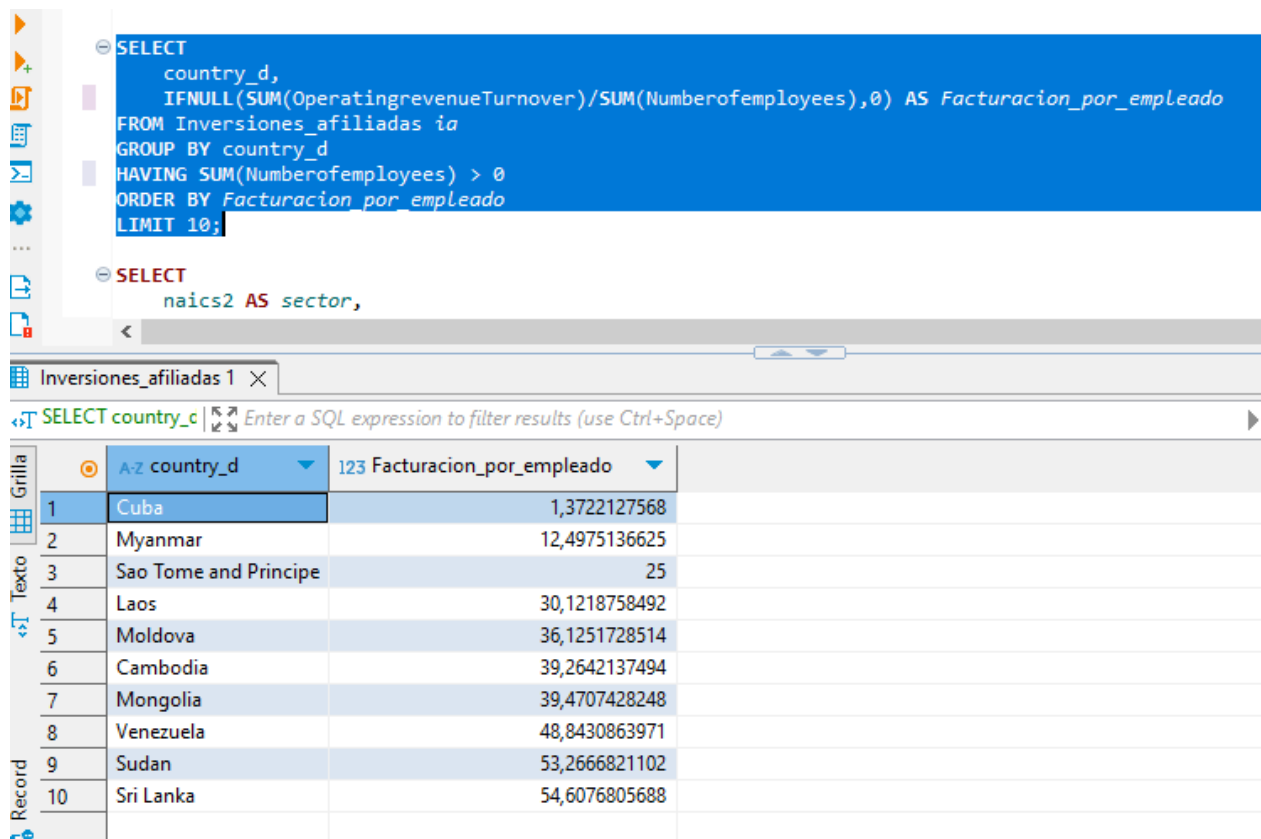
## 2. ¿Cuál es el bottom 10 países donde el ratio “Facturación por empleado” es más bajo?

```
SELECT
    country_d,
    IFNULL(SUM(OperatingrevenueTurnover)/SUM(Numberofemployees),0) AS
    Facturacion_por_empleado
FROM Inversiones_afiliadas ia
GROUP BY country_d
HAVING SUM(Numberofemployees) > 0
ORDER BY Facturacion_por_empleado
LIMIT 10;
```

Se usa la expresión IFNULL para evitar aquellos valores que son nulos y podrían afectar a los cálculos. En este caso si SUM(OperatingrevenueTurnover)/SUM(Numberofemployees) da como resultado NULL, el valor será reemplazado por 0.

Uso el HAVING para asegurarme que haya empleados y evitar que nos de error por ser imposible dividir entre 0.

No es necesario indicar en ORDER BY, el orden ASC (ascendente) porque es el que se realiza por defecto, es decir ordena de menor a mayor directamente.



```

SELECT
    country_d,
    IFNULL(SUM(OperatingrevenueTurnover)/SUM(Numberofemployees),0) AS Facturacion_por_empleado
FROM Inversiones_afiliadas ia
GROUP BY country_d
HAVING SUM(Numberofemployees) > 0
ORDER BY Facturacion_por_empleado
LIMIT 10;

SELECT
    naics2 AS sector,

```

	A-Z country_d	123 Facturacion_por_empleado
1	Cuba	1,3722127568
2	Myanmar	12,4975136625
3	Sao Tome and Principe	25
4	Laos	30,1218758492
5	Moldova	36,1251728514
6	Cambodia	39,2642137494
7	Mongolia	39,4707428248
8	Venezuela	48,8430863971
9	Sudan	53,2666821102
10	Sri Lanka	54,6076805688

3. ¿Cuáles son los 5 sectores con mejor evolución entre 2010 y 2021, en el número de empresas (extensive) en inversión (TotalassetsthUSD) doméstica en España?

```

SELECT
    naics2 AS sector,
    naics2description,
    SUM(CASE WHEN year = 2021 THEN extensive ELSE 0 END)- SUM(CASE WHEN year = 2010
    THEN extensive ELSE 0 END) AS evolucion_2010_2021,
    SUM(CASE WHEN year = 2021 THEN TotalassetsthUSD ELSE 0 END ) AS total_activos_2021
FROM Inversiones_afiliadas ia
WHERE
    country_o = 'Spain'
    AND country_d = 'Spain'
    AND year BETWEEN 2010 AND 2021
GROUP BY naics2, naics2description
ORDER BY evolucion_2010_2021 DESC
LIMIT 5;

```

Como en este caso se pide la evolución entre dos períodos de tiempo uso la cláusula CASE WHEN para comparar los datos, obteniendo la diferencia entre ellos con el fin de observar la evolución, así como para calcular el total de activos (inversión) en 2021.

Para asegurarnos de que sea inversión domestica debemos añadir tanto país de origen como país de destino España.

#3. 5 sectores mejor evolucion 2010-2021, inversion domestica Espana

```

SELECT
  naics2 AS sector,
  naics2description,
  SUM(CASE WHEN year = 2021 THEN extensive ELSE 0 END) - SUM(CASE WHEN year = 2010 THEN extensive ELSE 0 END) AS evolucion_2010_2021,
  SUM(CASE WHEN year = 2021 THEN TotalassetsthUSD ELSE 0 END ) AS total_activos_2021
FROM Inversiones_afiliadas ia
WHERE
  country_o = 'Spain'
  AND country_d = 'Spain'
  AND year BETWEEN 2010 AND 2021
GROUP BY naics2, naics2description

```

Inversiones\_afiliadas 1

SELECT naics2 AS sector

	123 sector	A-Z naics2description	123 evolucion_2010_2021	123 total_activos_2021
1	42	Wholesale Trade	1.043	26.957.842
2	54	Legal Services	1.033	23.547.110
3	23	Construction of Buildings	912	49.770.884
4	53	Real Estate	804	29.624.596
5	22	Utilities	735	137.360.656

En el caso de no querer la evolución temporal sino el número de empresas por sector y año, se lleva a cabo la siguiente función:

```

-- la pregunta anterior sin la evolución
SELECT
  year,
  naics2description,
  SUM(extensive) AS NUMERO_EMPRESAS,
  SUM(TotalassetsthUSD) AS INVERSION
FROM Inversiones_afiliadas ia
WHERE
  country_o = 'Spain'
  AND country_d = 'Spain'
  AND year BETWEEN 2010 AND 2021
GROUP BY year, naics2description
ORDER BY 1, 2;

```

SELECT

```

SELECT
  year,
  naics2description,
  SUM(extensive) AS NUMERO_EMPRESAS,
  SUM(TotalassetsthUSD) AS INVERSION
FROM Inversiones_afiliadas ia
WHERE
  country_o = 'Spain'
  AND country_d = 'Spain'
  AND year BETWEEN 2010 AND 2021
GROUP BY year, naics2description
ORDER BY 1, 2;

```

Inversiones\_afiliadas 1

SELECT year, naics2de

	123 year	A-Z naics2description	123 NUMERO_EMPRESA	123 INVERSION
1	2.010	Accommodation	847	2.424.308,75
2	2.010	Administrative and Support i	884	3.923.221,5
3	2.010	Agriculture, Forestry, Fishing	578	358.059,5625
4	2.010	Air Transportation	913	868.284,375
5	2.010	Arts, Entertainment, and Rec	362	449.466,5625
6	2.010	Construction of Buildings	1.591	3.970.074,25
7	2.010	Educational Services	126	99.766,34375
8	2.010	Executive, Legislative, and O	4	0
9	2.010	Finance and Insurance	724	101.396,272
10	2.010	Finished Product Manufactu	1.042	2.421.689
11	2.010	Food and Beverage Stores	573	334.648,53125
12	2.010	Food and Textile Manufactu	847	2.786.016,5
13	2.010	Health Care and Social Assist	467	336.188,1875

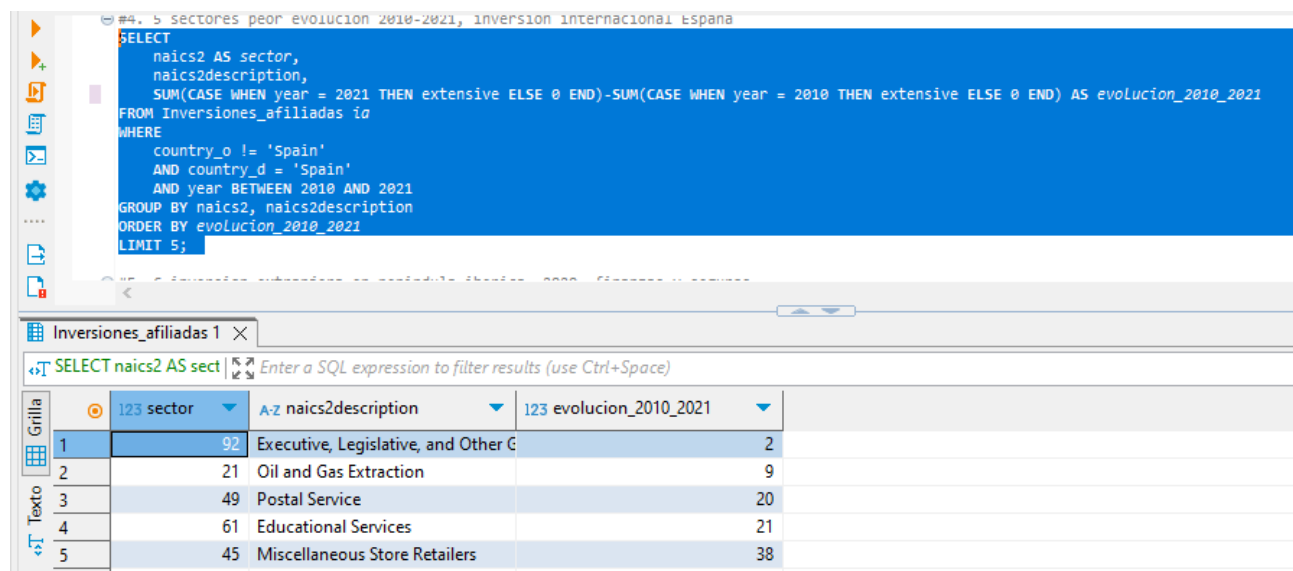
Refresh Save Cancel

Exportar datos ...

4. ¿Cuáles son los 5 sectores con peor evolución entre 2010 y 2021, en el número de empresas (extensive) en inversión (TotalassetsthUSD) internacional en España? (empresas con origen internacional y con destino España)

```
SELECT
    naics2 AS sector,
    naics2description,
    SUM(CASE WHEN year = 2021 THEN extensive ELSE 0 END)-SUM(CASE WHEN year = 2010
    THEN extensive ELSE 0 END) AS evolucion_2010_2021
FROM Inversiones_afiliadas ia
WHERE
    country_o != 'Spain'
    AND country_d = 'Spain'
    AND year BETWEEN 2010 AND 2021
GROUP BY naics2, naics2description
ORDER BY evolucion_2010_2021
LIMIT 5;
```

En este caso como se hace referencia a que es inversión internacional, definimos que el país de origen es diferente a España con la expresión != que significa no es igual.



	123 sector	A-z naics2description	123 evolucion_2010_2021
1	92	Executive, Legislative, and Other G	2
2	21	Oil and Gas Extraction	9
3	49	Postal Service	20
4	61	Educational Services	21
5	45	Miscellaneous Store Retailers	38

En el caso de no querer la evolución temporal sino el número de empresas por sector y año, se lleva a cabo la siguiente función:

```
-- la pregunta anterior sin la evolución
SELECT
    year,
    naics2description,
    SUM(extensive) AS NUMERO_EMPRESAS,
    SUM(TotalassetsthUSD) AS INVERSION
FROM Inversiones_afiliadas ia
WHERE
    country_o != 'Spain'
    AND country_d = 'Spain'
    AND year BETWEEN 2010 AND 2021
GROUP BY year, naics2description
ORDER BY 1, 2;
```

-- la pregunta anterior sin la evolución

```

SELECT
    year,
    naics2description,
    SUM(extensive) AS NUMERO_EMPRESAS,
    SUM(TotalassetsthUSD) AS INVERSION
FROM Inversiones_afiliadas ia
WHERE
    country_o != 'Spain'
    AND country_d = 'Spain'
    AND year BETWEEN 2010 AND 2021
GROUP BY year, naics2description
ORDER BY 1, 2;

```

Inversiones\_afiliadas 1

SELECT year, naics2de | Enter a SQL expression to filter results (use Ctrl+Space)

	123 year	A-Z naics2description	123 NUMERO_EMPRESA	123 INVERSION
1	2.010	Accommodation	180	5.117.841,538970947
2	2.010	Administrative and Support	330	814.034,8950042725
3	2.010	Agriculture, Forestry, Fishing	67	194.703,6028900146
4	2.010	Air Transportation	318	1.638.768,6779785156
5	2.010	Arts, Entertainment, and Rec	185	183.072,593963623
6	2.010	Construction of Buildings	265	1.032.502,4758529663
7	2.010	Educational Services	64	66.556,8346252441
8	2.010	Executive, Legislative, and O	1	0
9	2.010	Finance and Insurance	222	30.456.801,00793457
10	2.010	Finished Product Manufactu	571	10.194.275,601020813
11	2.010	Food and Beverage Stores	225	658.438,59375
12	2.010	Food and Textile Manufactu	211	1.217.969,2949047089
13	2.010	Health Care and Social Assist	376	1.297.206,845703125
14	2.010	Information	330	2.436.373,0036270307

Refresh Save Cancel Exportar datos

##### 5. Importe en € (EUROS), en el 2020 de la inversión extranjera en la Península Ibérica, en el sector "Finanzas y Seguros".

-- calculo cada país que forma la península ibérica en inversión extranjera

```

SELECT
    country_d,
    year,
    naics2description,
    SUM(OperatingrevenueTurnover) * 0.85 AS Importe_en_Euros
FROM Inversiones_afiliadas ia
WHERE
    country_d IN ('Spain', 'Portugal', 'Andorra')
    AND year=2020
    AND TRIM(naics2description) = 'Finance and Insurance'
GROUP BY country_d, year, naics2description
ORDER BY Importe_en_Euros ;

```

-- calculo el total para toda la península ibérica

```

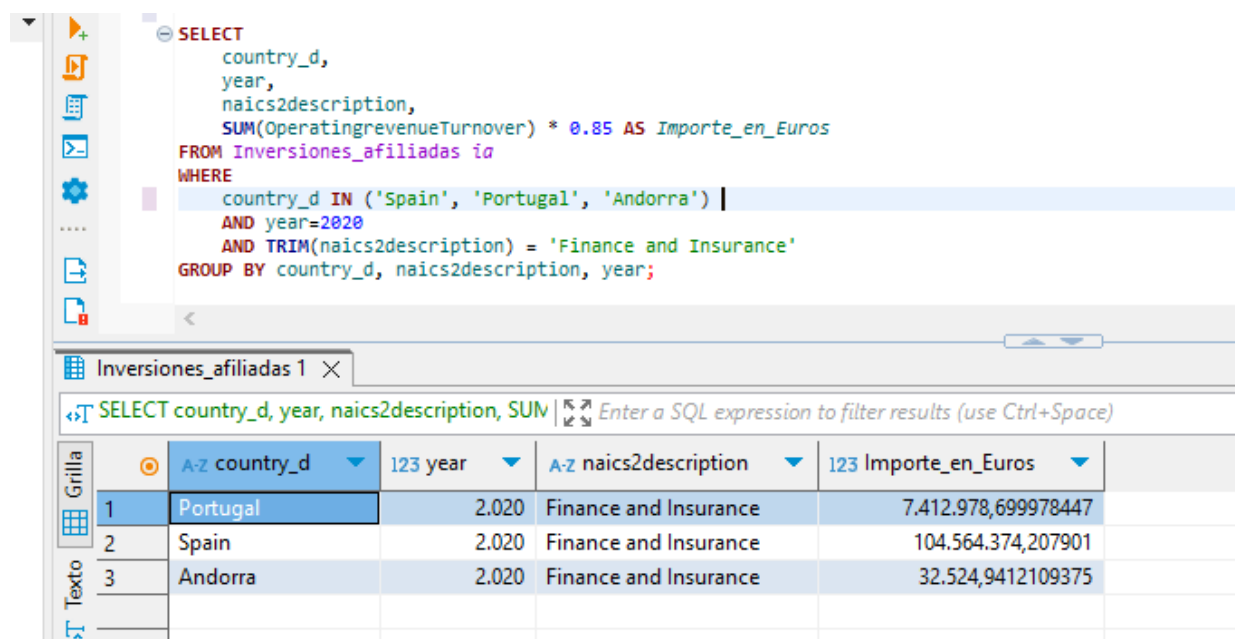
SELECT SUM(OperatingrevenueTurnover) * 0.85 AS Importe_en_Euros
FROM Inversiones_afiliadas ia
WHERE
    country_d IN ('Spain', 'Portugal', 'Andorra')
    AND year=2020
    AND TRIM(naics2description) = 'Finance and Insurance';

```

Al referirse a la Península Ibérica cogemos los datos de España, Portugal y Andorra en el sector de las finanzas y seguros.

Para cambiar de dólares a euros revisamos el valor del cambio actual que es 1USD=0.85EUR, para hacer la conversión usamos la expresión: OperatingrevenueTurnover \* 0.85. Además, se usa la función SUM para averiguar el importe total en euros.

Uso la función TRIM para eliminar los espacios adicionales al principio y al final de la cadena 'Finance and Insurance' en naics2description, porque al realizar la consulta sin él no nos devolvía ningún valor, sin embargo al comprobar individualmente los datos si encontrábamos resultados, por lo que tenía que ser un problema en la manera de almacenamiento de la información.



```

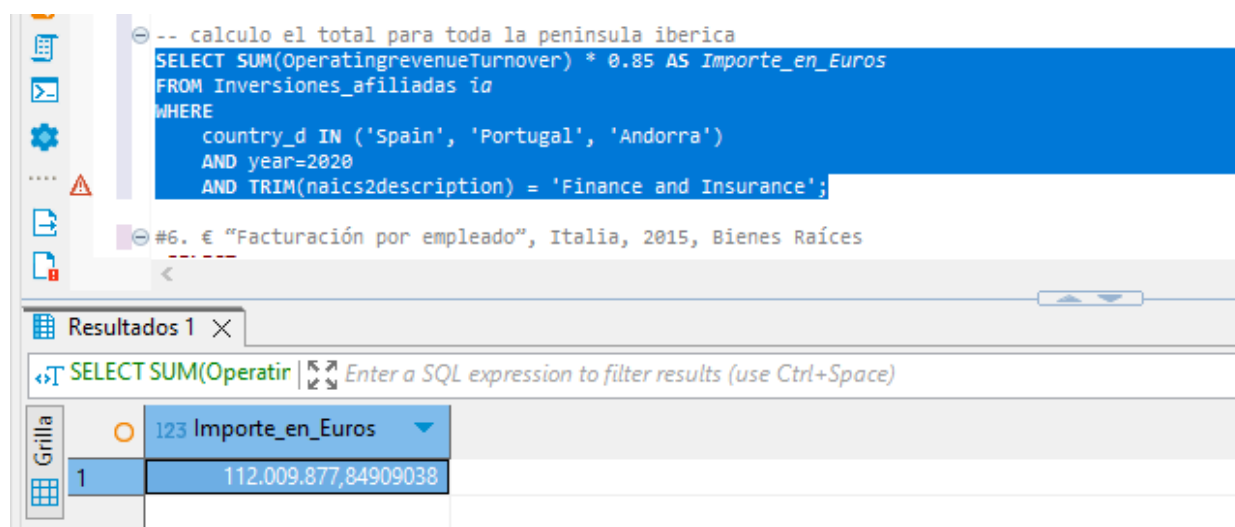
SELECT
    country_d,
    year,
    naics2description,
    SUM(OperatingrevenueTurnover) * 0.85 AS Importe_en_Euros
FROM Inversiones_afiliadas ia
WHERE
    country_d IN ('Spain', 'Portugal', 'Andorra')
    AND year=2020
    AND TRIM(naics2description) = 'Finance and Insurance'
GROUP BY country_d, naics2description, year;

```

Inversiones\_afiliadas 1 X

SELECT country\_d, year, naics2description, SUM | Enter a SQL expression to filter results (use Ctrl+Space)

	A-Z country_d	123 year	A-Z naics2description	123 Importe_en_Euros
1	Portugal	2.020	Finance and Insurance	7.412.978,699978447
2	Spain	2.020	Finance and Insurance	104.564.374,207901
3	Andorra	2.020	Finance and Insurance	32.524,9412109375



```

-- calculo el total para toda la peninsula iberica
SELECT SUM(OperatingrevenueTurnover) * 0.85 AS Importe_en_Euros
FROM Inversiones_afiliadas ia
WHERE
    country_d IN ('Spain', 'Portugal', 'Andorra')
    AND year=2020
    AND TRIM(naics2description) = 'Finance and Insurance';

```

#6. € "Facturación por empleado", Italia, 2015, Bienes Raíces

Resultados 1 X

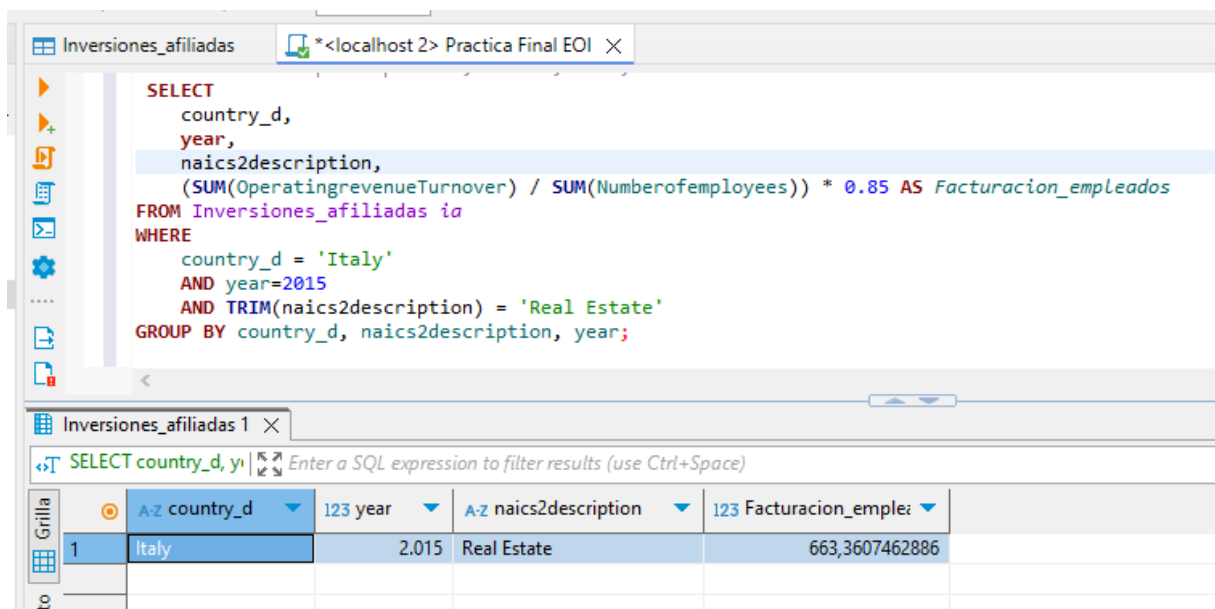
SELECT SUM(Operativ | Enter a SQL expression to filter results (use Ctrl+Space)

	123 Importe_en_Euros
1	112.009.877,84909038

6. Importe en € (EUROS), en el 2015 del ratio “Facturación por empleado” en Italia, para el sector de “Bienes Raíces”

```
SELECT
    country_d,
    year,
    naics2description,
    (SUM(OperatingrevenueTurnover) / SUM(Numberofemployees)) * 0.85 AS Facturacion_empleados
FROM Inversiones_afiliadas ia
WHERE
    country_d = 'Italy'
    AND year=2015
    AND TRIM(naics2description) = 'Real Estate'
GROUP BY country_d, naics2description, year;
```

Para resolver esta pregunta se utilizaran las formulas usadas previamente, combinando la formula de facturación por empleado con la conversión de dolares a euros.



The screenshot shows a SQL query editor with the following query:

```
SELECT
    country_d,
    year,
    naics2description,
    (SUM(OperatingrevenueTurnover) / SUM(Numberofemployees)) * 0.85 AS Facturacion_empleados
FROM Inversiones_afiliadas ia
WHERE
    country_d = 'Italy'
    AND year=2015
    AND TRIM(naics2description) = 'Real Estate'
GROUP BY country_d, naics2description, year;
```

Below the query editor, the results are displayed in a grid. The grid has columns for country\_d, year, naics2description, and Facturacion\_empleados. The results show one row for Italy in 2015 in the Real Estate sector with a value of 663,360,746,2886.

	A-Z country_d	123 year	A-Z naics2description	123 Facturacion_empleados
1	Italy	2015	Real Estate	663,360,746,2886

## POWER BI

Para poder usar los datos en Power Bi primero debemos cargarlos, para ello usaremos el comando obtener datos de una base de datos MySQL.

The screenshot shows the Power BI Desktop interface. The 'Obtener datos' (Get Data) dialog box is open, displaying a list of connectors. 'Base de datos MySQL' is selected. The 'Datos' (Data) pane on the right shows a list of tables, including 'greenfield' and 'mergers'. The 'country\_o' table is selected in the main view.

iso3_o	country_o	iso3_d	country_d
AGO	Angola	CHE	Switzerland
AGO	Angola	PRT	Portugal
AIA	Anguilla	MKD	Macedonia
ARE	United Arab Emirates	AUS	Australia
ARE	United Arab Emirates	AUT	Austria
ARE	United Arab Emirates	BIH	Bosnia and Herzegovina
ARE	United Arab Emirates	DNK	Denmark
ARE	United Arab Emirates	ITA	Italy
ARE	United Arab Emirates	JOR	Jordan
ARE	United Arab Emirates	KWT	Kuwait
ARE	United Arab Emirates	LUX	Luxembourg
ARE	United Arab Emirates	QAT	Qatar
ARE	United Arab Emirates	SRB	Serbia
ARE	United Arab Emirates	TZA	Tanzania
ARE	United Arab Emirates	USA	United States
ARG	Argentina	FRA	France
AUS	Australia	CZE	Czech Republic
AUS	Australia	IND	India
AUS	Australia	IRL	Ireland
AUS	Australia	MAR	Morocco
AUS	Australia	MOZ	Mozambique
AUS	Australia	RUS	Russia

The screenshot shows the Power BI Desktop interface with the 'greenfield' table loaded in the main view. The 'country\_o' table is also visible in the 'Datos' pane. The 'greenfield' table has columns: iso3\_o, country\_o, iso3\_d, country\_d, year, naics2, naics2description, extensive, greenfield, mergers, OperatingrevenueTurnover, and year.

iso3_o	country_o	iso3_d	country_d	year	naics2	naics2description	extensive	greenfield	mergers	OperatingrevenueTurnover	year
AGO	Angola	CHE	Switzerland	2010	42	Wholesale Trade	1	0	0	0	2010
AGO	Angola	PRT	Portugal	2010	42	Wholesale Trade	1	0	0	0	2010
AIA	Anguilla	MKD	Macedonia	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	AUS	Australia	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	AUT	Austria	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	BIH	Bosnia and Herzegovina	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	DNK	Denmark	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	ITA	Italy	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	JOR	Jordan	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	KWT	Kuwait	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	LUX	Luxembourg	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	QAT	Qatar	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	SRB	Serbia	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	TZA	Tanzania	2010	42	Wholesale Trade	1	0	0	0	2010
ARE	United Arab Emirates	USA	United States	2010	42	Wholesale Trade	1	0	0	0	2010
ARG	Argentina	FRA	France	2010	42	Wholesale Trade	1	0	0	0	2010
AUS	Australia	CZE	Czech Republic	2010	42	Wholesale Trade	1	0	0	0	2010
AUS	Australia	IND	India	2010	42	Wholesale Trade	1	0	0	0	2010
AUS	Australia	IRL	Ireland	2010	42	Wholesale Trade	1	0	0	0	2010
AUS	Australia	MAR	Morocco	2010	42	Wholesale Trade	1	0	0	0	2010
AUS	Australia	MOZ	Mozambique	2010	42	Wholesale Trade	1	0	0	0	2010
AUS	Australia	RUS	Russia	2010	42	Wholesale Trade	1	0	0	0	2010

Una vez insertados en Power BI, empezamos a crear los dashboards.

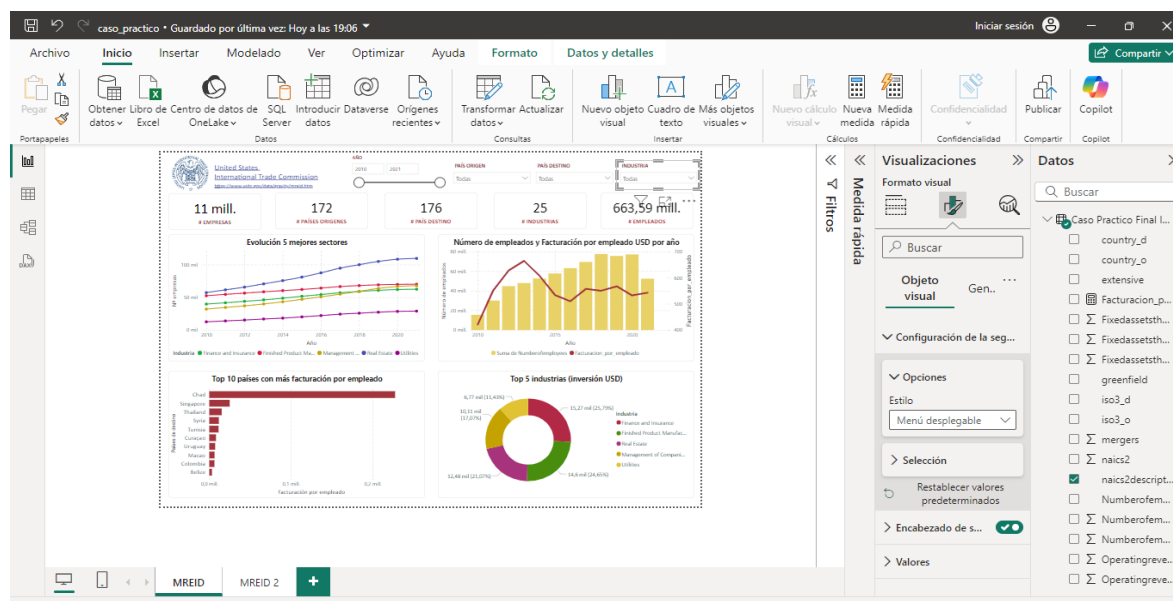
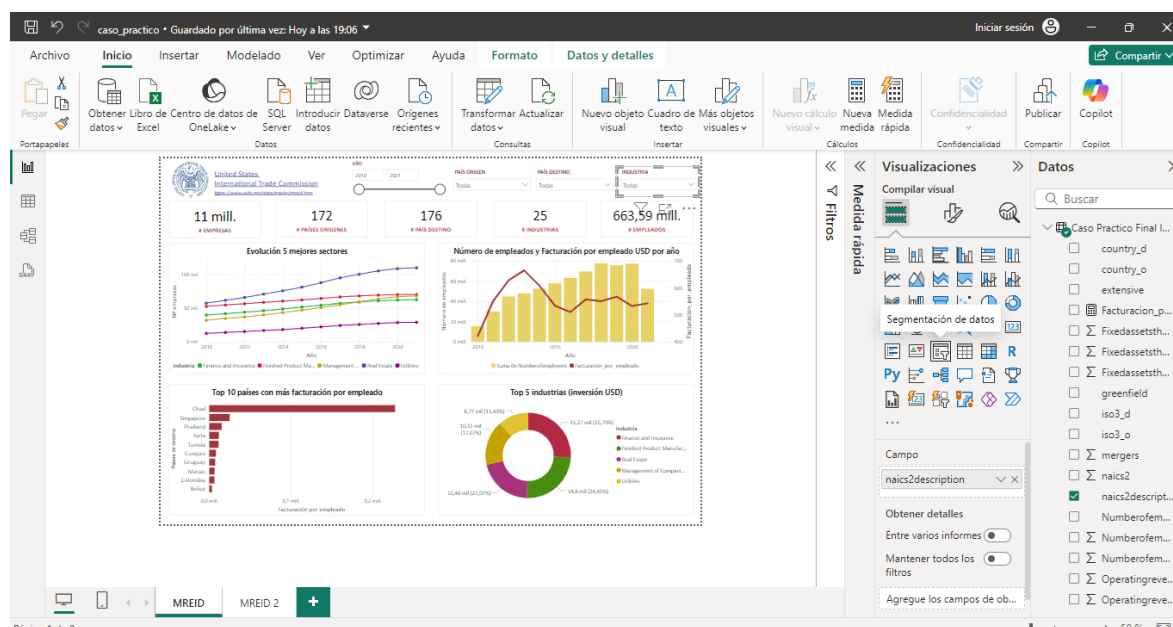


En primer lugar le daremos un formato general para usarlo en todos los paneles de MREID, de esta manera obtenemos la siguiente visión:

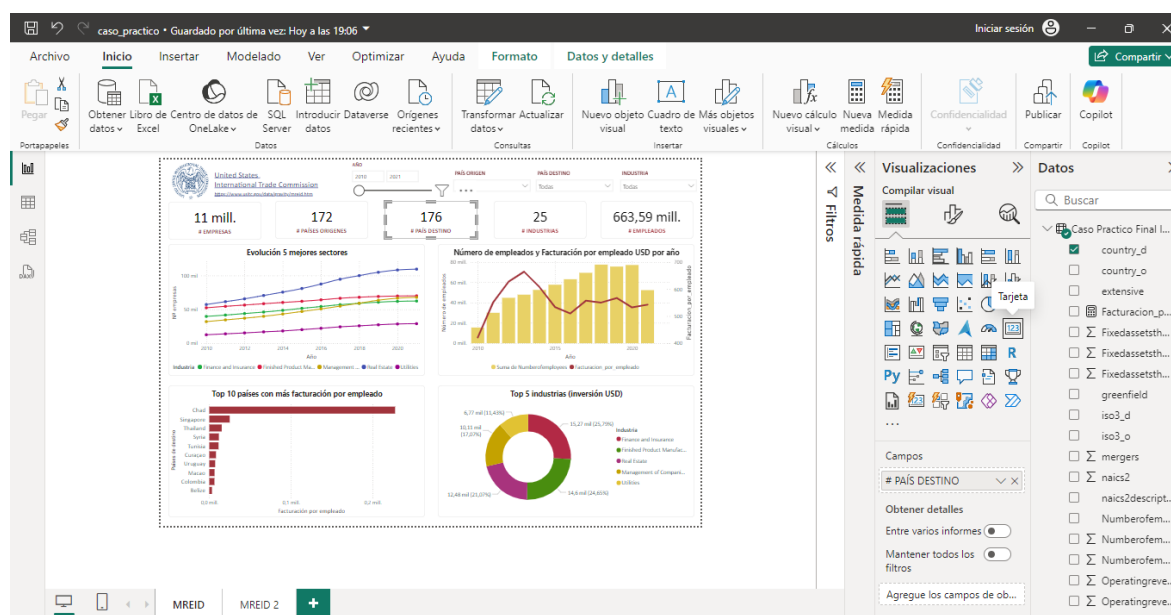


Inserte como imagen el logo de United States International Trade Commission, y añadí un cuadro de texto con el nombre de Centro de la página así como el enlace a la web.

Usando la opción de segmentación de datos creo la visualización de año con la configuración entre período de tiempo, mientras que en el caso del país de origen, de destino y de industria uso la configuración de menú desplegable.



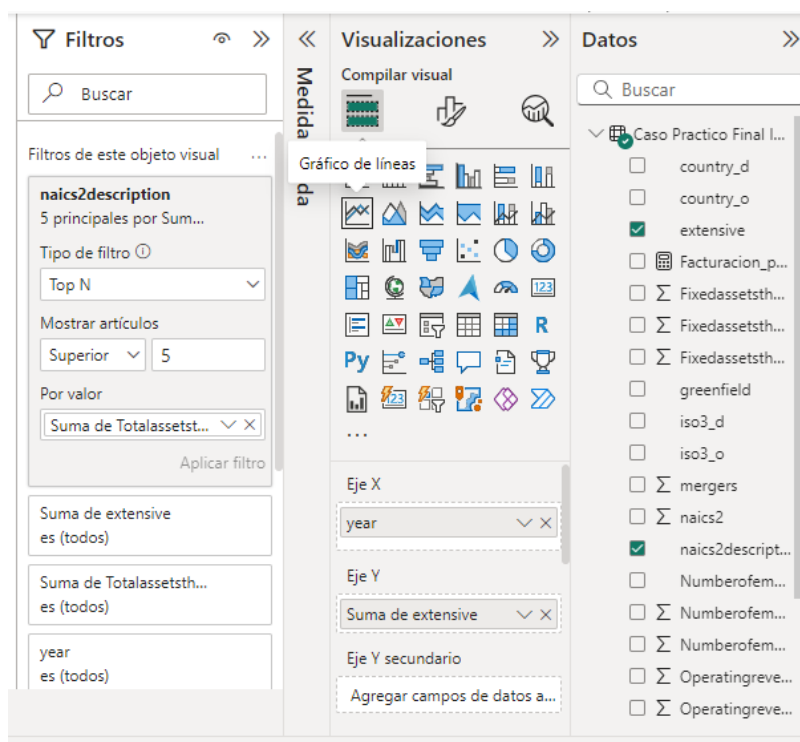
Selecciono la opción de tarjeta para crear las visualizaciones del número de empresas, países de origen, país de destino, industrias y empleados.

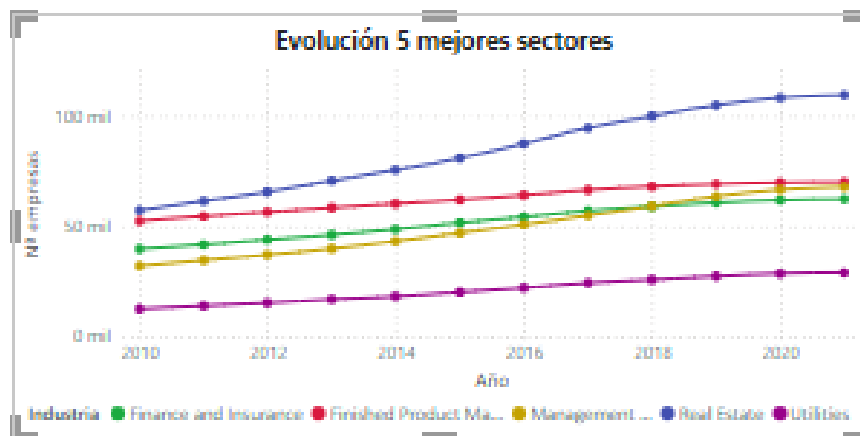


Una vez que le he dado el formato deseado a todos los objetos comienzo a elaborar los diversos gráficos. En esta ocasión dividí mi presentación de Power BI en dos paneles, MREID Y MREID 2.

En el caso de MREID:

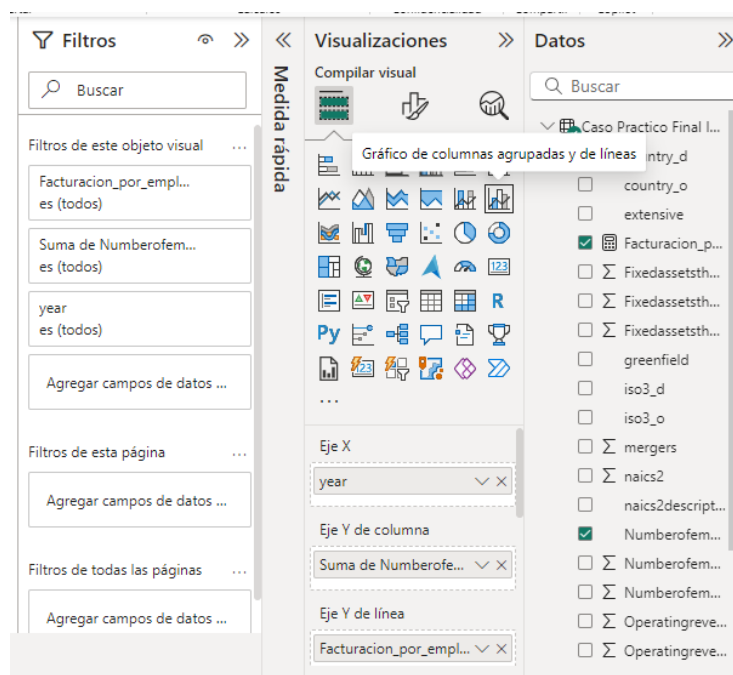
El primer gráfico muestra la evolución de los 5 mejores sectores industriales. Para ello, elegí el gráfico de líneas puesto que se utiliza para visualizar una tendencia en los datos a lo largo de intervalos de tiempo (una serie temporal), por lo que la línea se dibuja cronológicamente. Para establecer los 5 mejores es necesario instaurar un filtro en el tipo de sectores Top N, superior 5 por valor de la inversión.





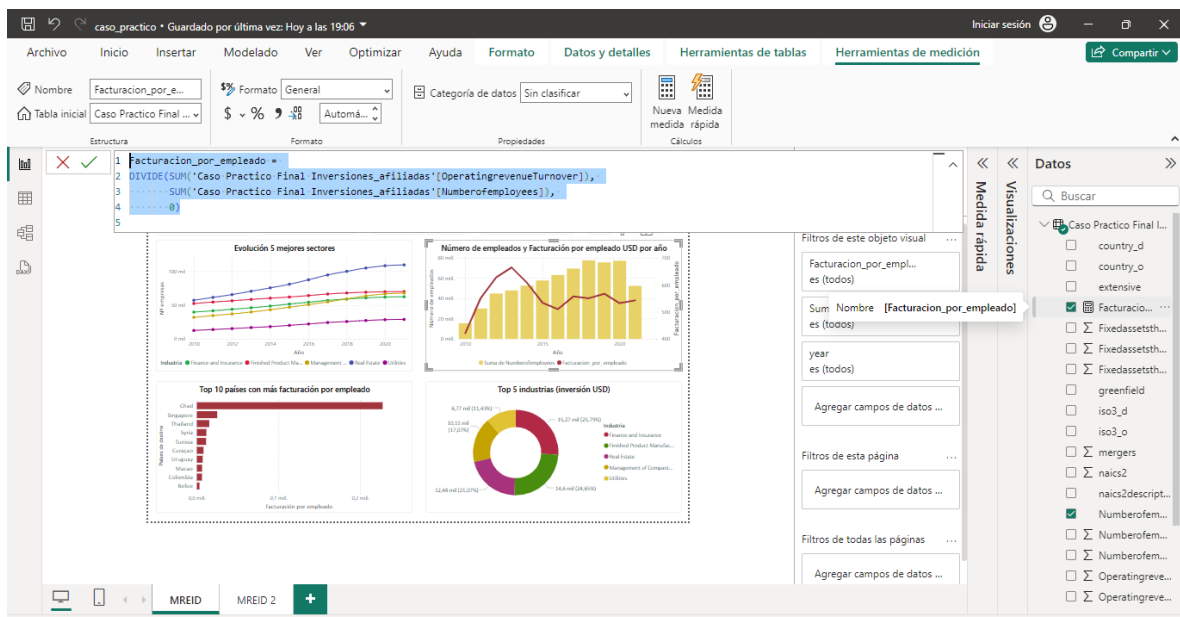
Como se puede observar añadí en la leyenda el tipo de sector.

En el segunda gráfico muestro el número de empleados y su facturación en dólares por año. Para realizarlo elegí la visualización de gráfico de columnas y líneas, es decir un gráfico combinado con el fin de mostrar la evolución de distintos tipos de datos a lo largo de un mismo periodo de tiempo.

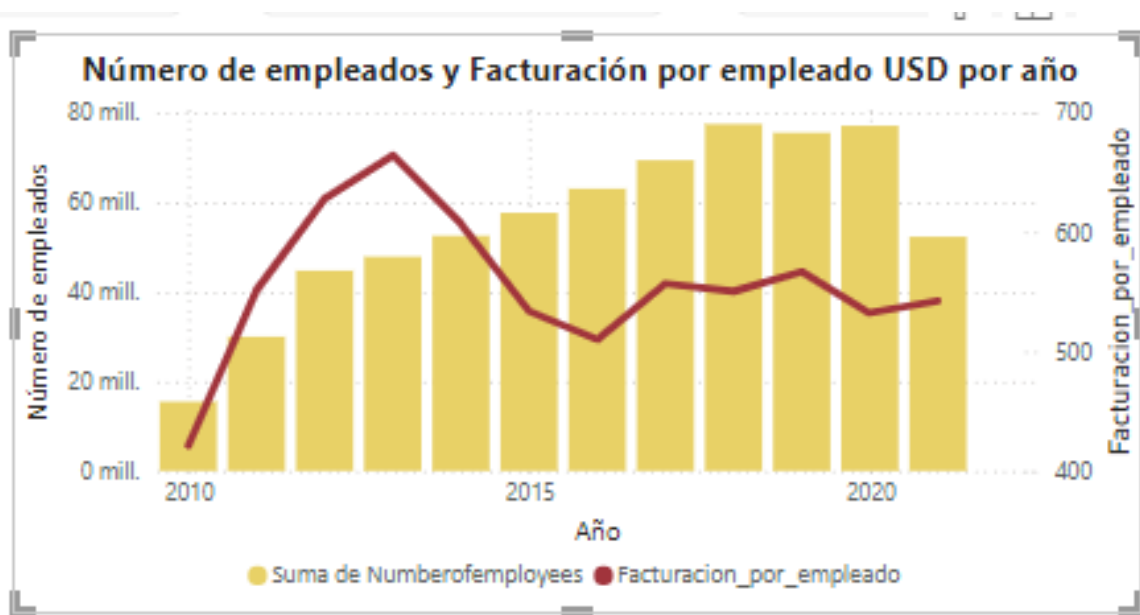


Para calcular la facturación por empleado se tiene que crear una nueva medida e insertar esta fórmula para que calcule los valores:

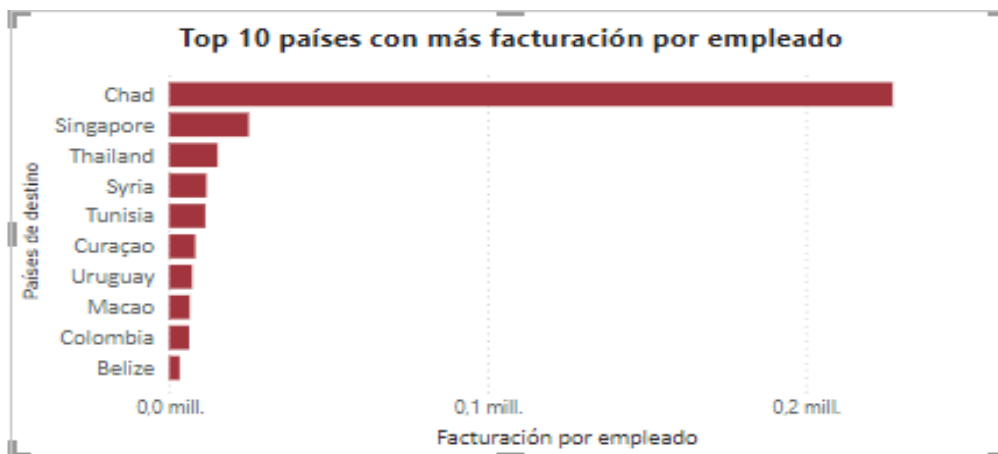
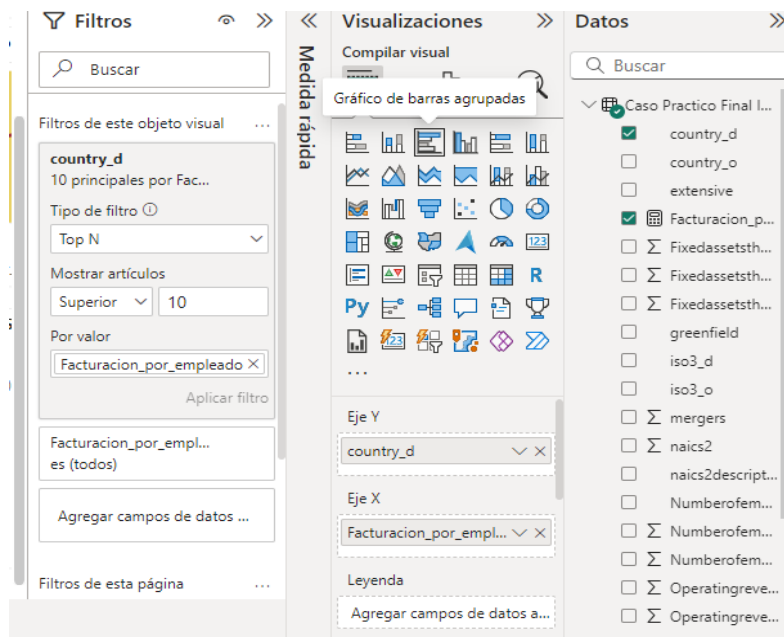
```
Facturacion_por_empleado =
DIVIDE(SUM('Caso Practico Final Inversiones_afiliadas'[OperatingrevenueTurnover]),
SUM('Caso Practico Final Inversiones_afiliadas'[Numberofemployees]),
0)
```



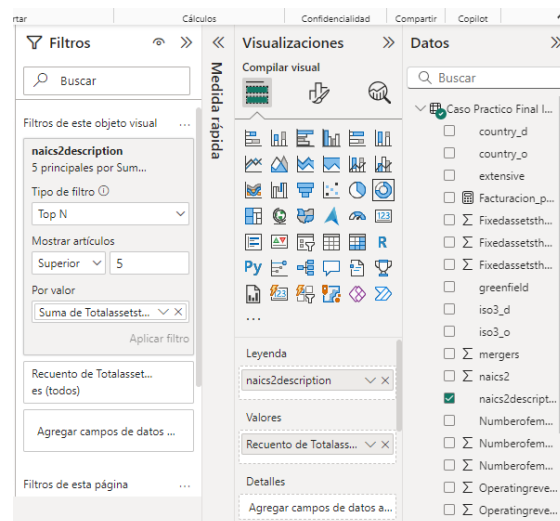
Obteniendo como resultado:

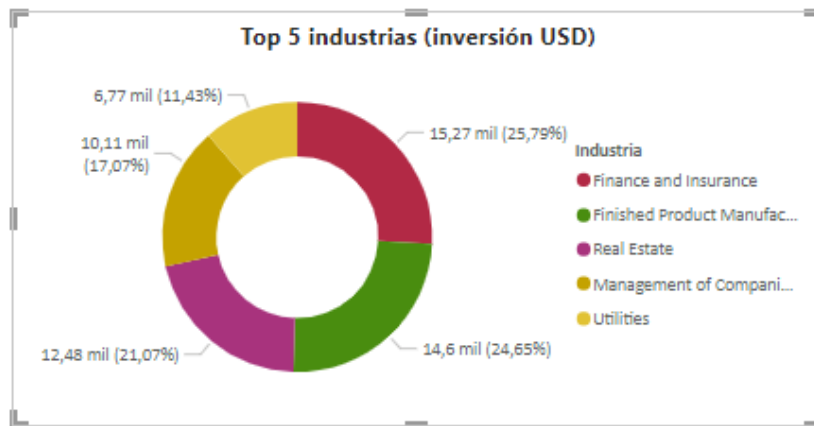


El tercer gráfico consiste en una visualización de columnas puesto que pretendo mostrar una comparación entre los 10 países con más facturación por empleado. Para ello tuve que usar un filtro en el país de destino por el tipo Top N, superior 10 por la facturación por empleado calculada previamente para el gráfico anterior.

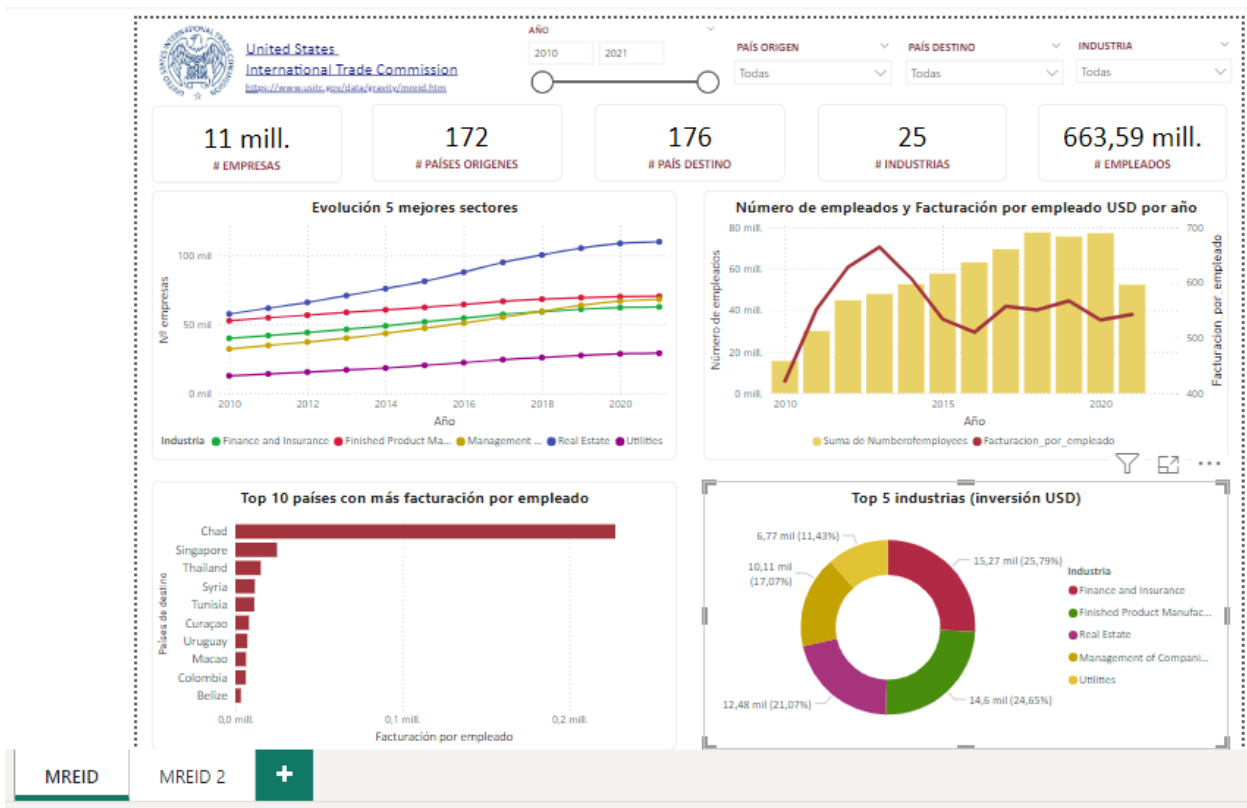


El cuarto gráfico es de anillos para representar las proporciones de datos categóricos, en este caso para visualizar las 5 mejores industrias según la inversión en dólares. Tuve que establecer un filtro en el tipo de sector como Top N superior 5 con el valor de las inversiones.





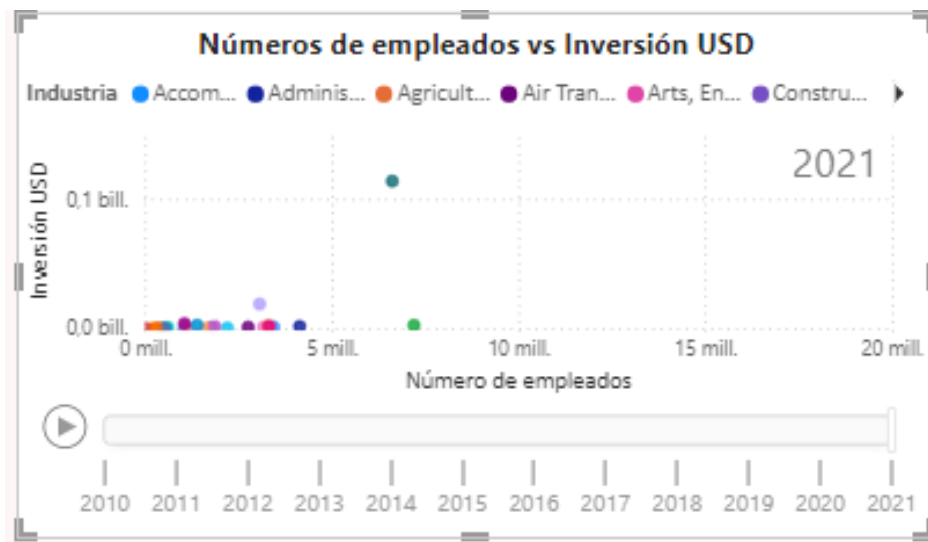
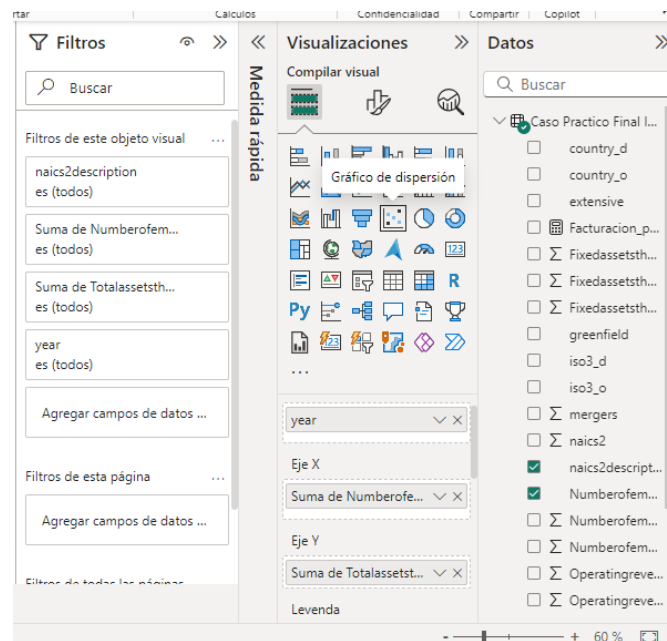
De esta forma el panel final se podrá ver en este formato:



En el segundo panel, MREID 2:

Establezco el formato general explicado previamente y elaboro los siguientes gráficos.

El primer gráfico que usamos es el de dispersión para mostrar como dos variables se relacionan entre sí, en este caso el número de empleados y la inversión en USD. Añadiremos en el campo de valor los datos de año para poder analizar el cambio en el tiempo y en la leyenda el tipo de sectores.

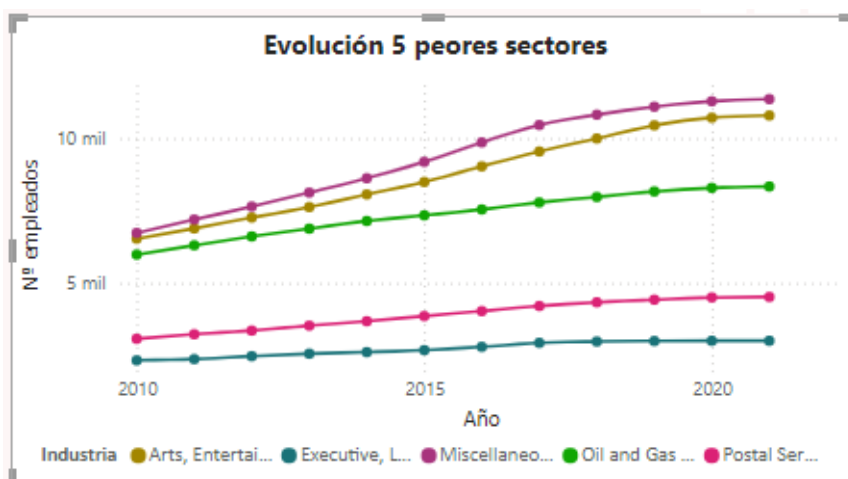
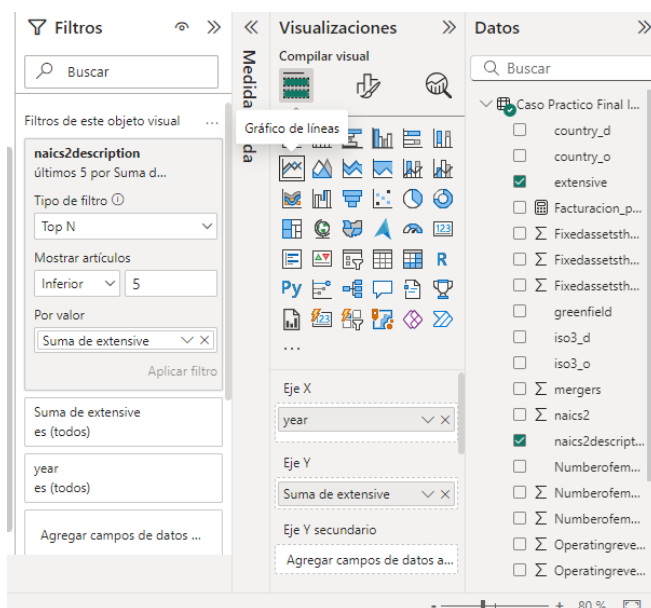


El segundo gráfico es un mapa que muestra la cantidad de empresas por países. Para ello establecimos en el campo de ubicación los datos de países y en el tamaño de las burbujas los datos del número de empleados. La personalizamos para que se pueda desplazar por el mapa y seleccionar el país o región que necesita analizar.





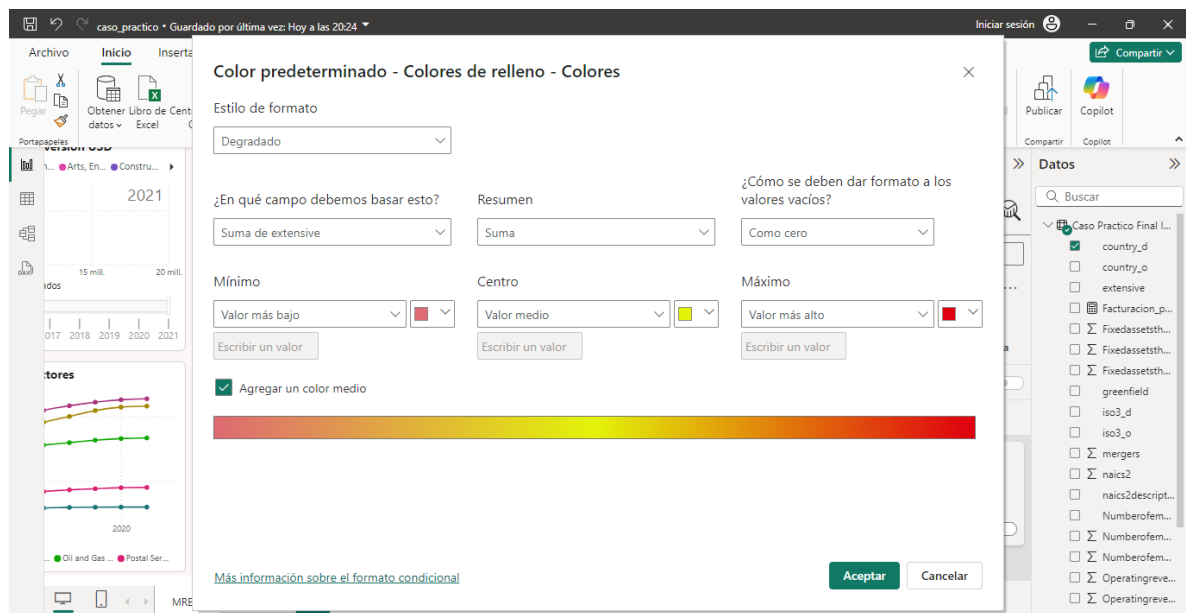
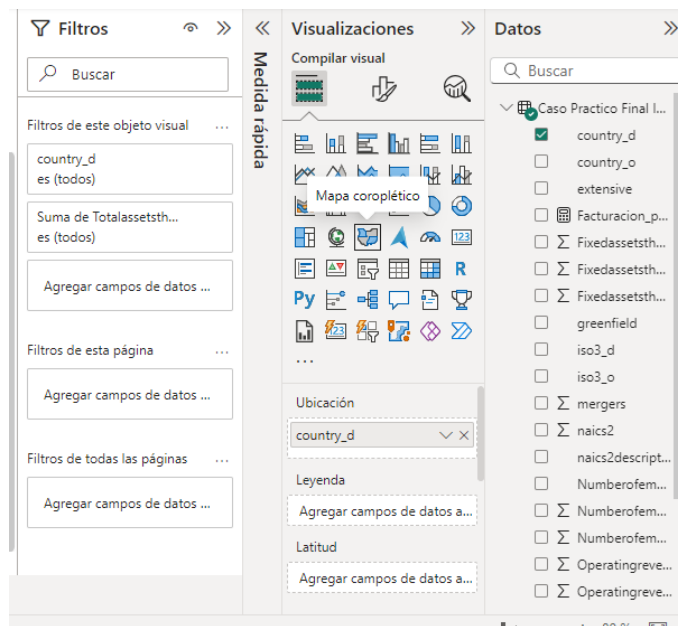
El tercer gráfico muestra la evolución de los 5 peores sectores industriales a través de un gráfico de líneas, puesto que buscamos visualizar una tendencia en los datos a lo largo de un intervalo de tiempo. Para establecer los 5 peores es necesario instaurar un filtro en el tipo de sectores Top N, inferior 5 por valor de la inversión.



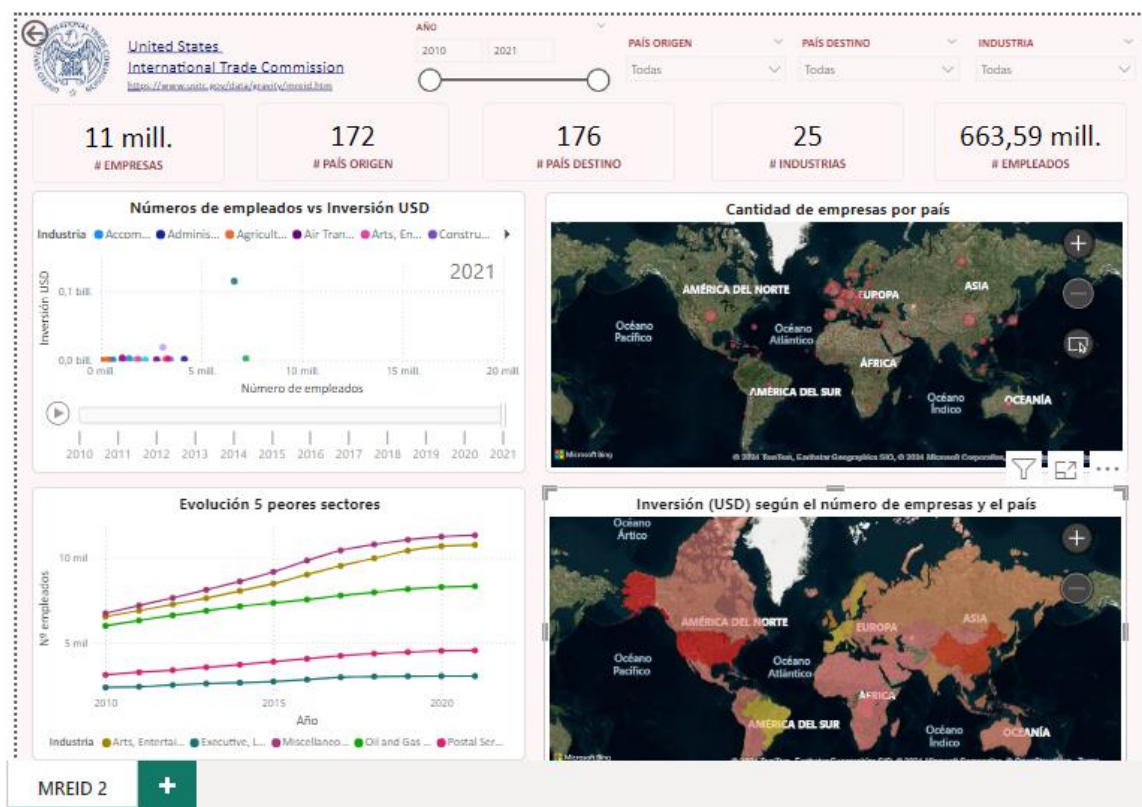
El último gráfico es un mapa coroplético en el cual mostramos la inversión en dólares teniendo en cuenta el número de empresas y el país. Aunque no se establece ningún filtro en esta ocasión si se personaliza a través de una regla de degradado según el número de empresas.







De esta manera el panel final para MREID 2 se podrá ver en este formato:



## CONCLUSIONES

En este proyecto hemos elaborado un workflow que nos ha permitido cargar un fichero CSV de una web , en una aplicación de gestión de datos que usa MySQL , a través del cual respondimos a las preguntas planteadas, lo que nos sirvió para comprender los datos de origen y prepararlos para usarlos en la creación de dashboards de Power BI.

A través de las consultas SQL y los dashboards he podido observar que los países con más ratio por facturación de empleado son Chad, Singapur, Tailandia y Siria, esto se puede deber a que en los últimos años se han convertido en focos de inversión y expansión empresarial, en el caso de Singapur es conocido por su rica economía y su sociedad avanzada, lo que justifica un alto ratio de facturación por empleado. Sin embargo, los países con menos ratio son Cuba, Myanmar, Sao Tome and Príncipe y Laos, los cuales son países con grandes dificultades como una menor productividad, economías más cerradas o falta de inversión lo que impacta negativamente en la facturación de los empleados.

Por otro lado los sectores con mejor evolución en los últimos años en España respecto del número de empresas y de inversión son el comercio mayorista, los servicios legales y la construcción. El crecimiento del comercio mayorista sugiere una demanda elevada de bienes, mientras que la construcción de edificios apunta a una posible recuperación o expansión del sector inmobiliario. Los servicios legales pueden estar creciendo debido a la mayor necesidad de regulación y cumplimiento en diversos sectores. Si atendemos a los datos sobre los peores sectores son la administración pública, los servicios postales y la extracción de petróleo y gas. Los sectores como la administración pública y los servicios postales podrían requerir innovación tecnológica para mejorar su competitividad. La extracción de petróleo y gas podría estar enfrentando una reducción de inversión por los cambios globales hacia energías más limpias y renovables, lo

que impacta la actividad en este sector, así como en la escasez de España como proveedora de estos productos.

Estas observaciones pueden guiar decisiones estratégicas sobre inversiones, expansión geográfica o innovación sectorial.

Hay que tener en cuenta que la amplitud de la cobertura de los datos nos ha permitido analizar diversos parámetros, pero el escaso tiempo de elaboración del proyecto también ha limitado poder profundizar más en ellos. No obstante, ha permitido que nos familiaricemos más profundamente con Power BI y adquirir muchas habilidades necesarias para el trabajo de un analista de datos.