

# Proyecto Final

## Juegos Olímpicos

Entrega N°3

### Integrantes

Delucchi Pablo Nicolás  
Paez Pamela Gisele

### Profesores

Leandro Abraham  
Rocío Lucchetta  
Marcos Díaz Carmona



## Índice

<b>Tabla de Versionado</b> .....	3
<b>Herramientas implementadas</b> .....	3
<b>1.Descripción de la temática de los datos</b> .....	3
<b>2. Hipótesis</b> .....	3
<b>3. Datasets</b> .....	3
<b>4. Objetivo del proyecto</b> .....	3
<b>5. Alcance</b> .....	4
<b>6. Usuario final y nivel de aplicación del análisis</b> .....	4
<b>7. Diagrama Entidad Relación</b> .....	5
<b>8. Listado de Tablas</b> .....	6
<b>9. Listado de columnas por tablas</b> .....	7
<b>10. Script Creación y filtrado de datos en SQL</b> .....	8
<b>11. Transformaciones realizadas</b> .....	8
<b>12. Medidas calculadas</b> .....	10
<b>13. Columnas Calculadas</b> .....	11
Tabla Competiciones:.....	11
Tabla Competidores: .....	11
Tabla Medallas: .....	11
Tabla Temporadas:.....	11
<b>14. Segmentaciones elegidas</b> .....	11
Deportes:.....	11
Medallas: .....	12
Competidores:.....	15
Juegos:.....	16
<b>15. Conclusiones</b> .....	16

## Tabla de Versionado

Versión	Fecha	Modificaciones
1	14/08/2022	Definición del modelo, creación de consultas y definición de tablas.
2	20/09/2022	Descripción de la hipótesis, objetivo del proyecto, alcance, Transformaciones realizadas, medidas y columnas calculadas y segmentaciones elegidas
3	18/10/2022	Agregado tabla de versionado, herramientas utilizadas y conclusiones finales.

## Herramientas implementadas

Para el desarrollo del presente proyecto se utilizó Excel para la evaluación del data set crudo descargado, SQL Server para realizar la separación y transformación de los datos obtenidos, separándolos e integrándolos según el modelo relacional y Power Bi para realizar las últimas transformaciones y la visual del mismo.

### 1.Descripción de la temática de los datos

Para el presente trabajo se extrajo información acerca de los juegos olímpicos jugados desde 1986 al 2016. El mismo contiene en que países se jugaron, tipo de juego (verano o invierno), categorías jugadas, información acerca de los competidores y sus respectivas medallas.

Además, existe una segunda tabla donde se referencia a como se abrevian los países según tres letras.

### 2. Hipótesis

Identificar qué país ganó más medallas, identificar la asistencia de diferentes países dependiendo del año y tipo de juego. También obtener que país participó en más categorías y realizar un porcentaje de eventos jugados vs medallas obtenidas. Analizar la cantidad de participantes, separándolos por género y deporte en el que compitieron.

### 3. Datasets

Adjunto se encuentra el Excel llamado Juegos Olímpicos Raw Data con el data set que contiene la información descripta anteriormente.

### 4. Objetivo del proyecto

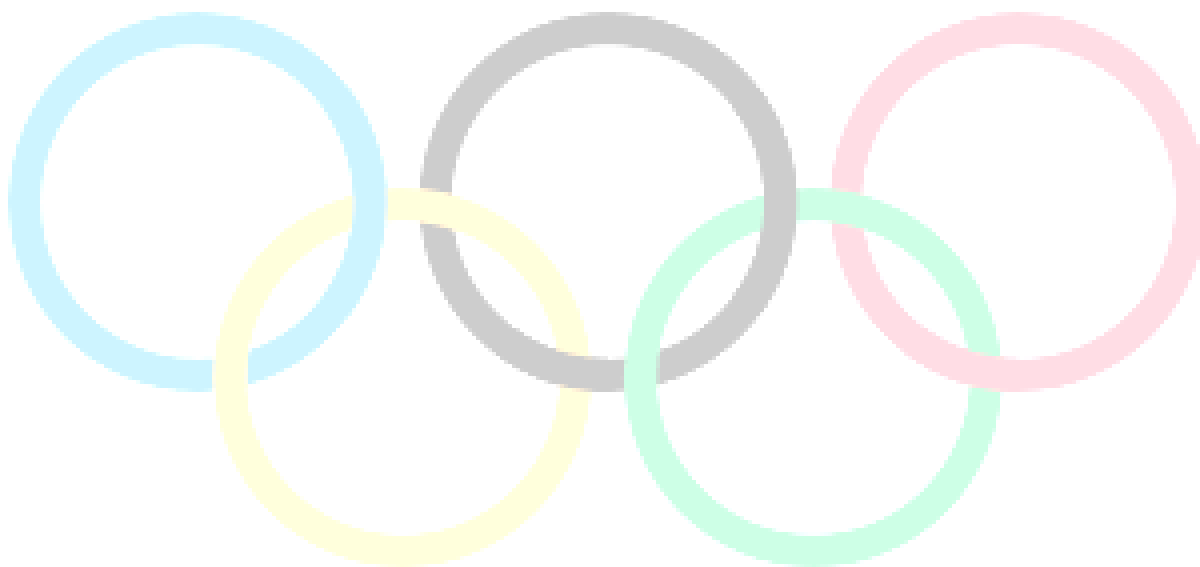
El objetivo del presente proyecto es poder analizar estadísticamente los datos sobre los diferentes juegos olímpicos realizados en un período de tiempo determinado. Un correcto uso del mismo permitirá establecer medidas y prever ciertos datos a futuro en base a las tendencias conseguidas.

## 5. Alcance

El alcance del siguiente proyecto corresponde al análisis y estadísticas sobre los Juegos Olímpicos comprendidos entre 1986 y 2016

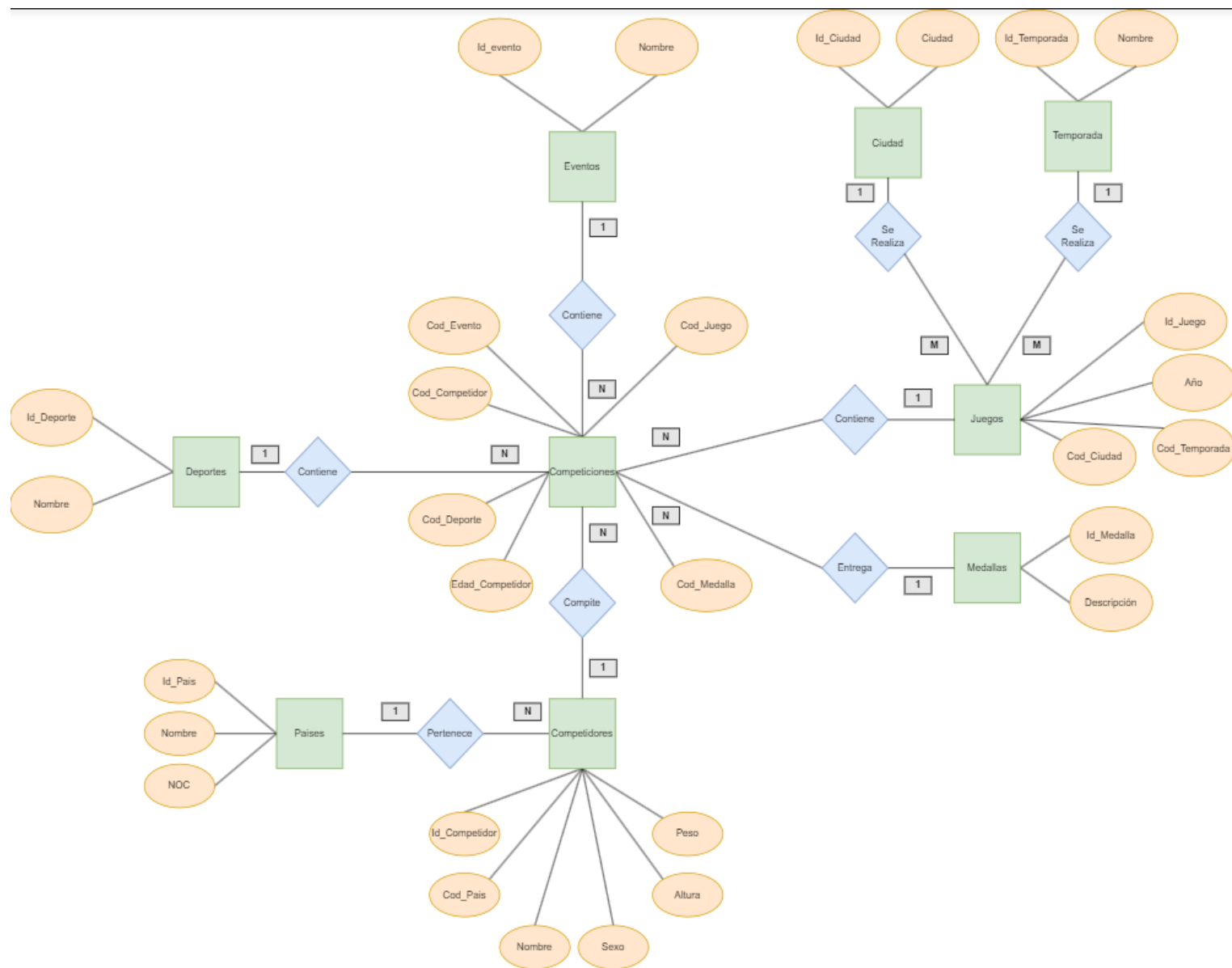
## 6. Usuario final y nivel de aplicación del análisis

Para el análisis del proyecto se tienen en cuentas todos los países que participaron de los juegos olímpicos entre los años 1986 y 2016. Y en función a ello, Coderhouse nos solicitó determinar que jugador gano más medallas.



## 7. Diagrama Entidad Relación

[https://drive.google.com/file/d/145D\\_LttZ2sdeb010knX9ryIJKPcDMPuh/view?usp=sharing](https://drive.google.com/file/d/145D_LttZ2sdeb010knX9ryIJKPcDMPuh/view?usp=sharing)



## 8. Listado de Tablas

**Deportes:** Contiene los deportes existentes dentro de los Juegos Olímpicos.

- PK: Id\_Deporte

**Eventos:** Contiene los eventos existentes dentro de los Juegos Olímpicos.

- PK: Id\_Evento

**Juegos:** Contiene información de cuando se realizaron los juegos olímpicos agrupados por Ciudad y Temporada.

- PK: Id\_Juego
- FK: Temporada
- FK: Ciudad

**Medallas:** Contiene los diferentes tipos de medallas entregados en los Juegos.

- PK: Id\_Medalla

**Competidores:** Contiene información acerca de los diferentes competidores, agrupando por Peso, Altura y Sexo.

- PK: Id\_Competidor
- FK: Cod\_Pais

**Países:** Contiene información acerca de los países de los competidores.

- PK: Id\_Pais

**Competiciones:** Contiene información acerca de las diferentes competiciones, la cual referencia a muchas de las tablas ya existentes para así lograr la normalización. Hay que tener en cuenta que la edad de los competidores varía según las diferentes competiciones en las que formó parte, y es por eso que el campo se encuentra en esta tabla y no en Competidores.

- FK: Cod\_Juego
- FK: Cod\_Competidor
- FK: Cod\_Deporte
- FK: Cod\_Juego
- FK: Cod\_Medalla

## 9. Listado de columnas por tablas

Eventos		
Campo	Tipo de Campo	Tipo de Clave
Id_evento	Int	PK – Index
Nombre	varchar(100)	

Paises		
Campo	Tipo de Campo	Tipo de Clave
Id_Pais	Int	PK – Index
Nombre	varchar(50)	
NOC	varchar(5)	

Medallas		
Campo	Tipo de Campo	Tipo de Clave
Id_Medalla	Int	PK – Index
Descripcion	varchar(20)	

Juegos		
Campo	Tipo de Campo	Tipo de Clave
Id_Juego	Int	PK – Index
Cod_Temporada	Int	FK
Cod_Ciudad	Int	FK
Año	Int	

Deportes		
Campo	Tipo de Campo	Tipo de Clave
Id_Deporte	Int	PK – Index
Descripcion	varchar(20)	

Competidores		
Campo	Tipo de Campo	Tipo de Clave
Id	Int	PK – Index
Id_Raw	Int	
Nombre	varchar(150)	
Sexo	varchar(1)	
Altura	Int	
Peso	Int	
Cod_Pais	Int	FK

Competiciones		
Campo	Tipo de Campo	Tipo de Clave
Cod_juego	Int	FK
Cod_Competidor	Int	FK
Edad_Competidor	Int	
Cod_Deporte	Int	FK
Cod_Evento	Int	FK
Cod_Medalla	Int	FK

Temporadas		
Campo	Tipo de Campo	Tipo de Clave
Id_Temporada	Int	PK – Index
Nombre	varchar(50)	

Ciudades		
Campo	Tipo de Campo	Tipo de Clave
Id_Ciudad	Int	PK – Index
Nombre	varchar(50)	

## 10. Script Creación y filtrado de datos en SQL

Adicionalmente, en el siguiente link puede encontrarse el link al archivo de creación de tablas y filtrado de información que se podría utilizar para realizarse en SQL:

<https://drive.google.com/file/d/1PFHHwMsBQKNcqWTrmG3hF2I51Lh3o8wL/view?usp=sharing>

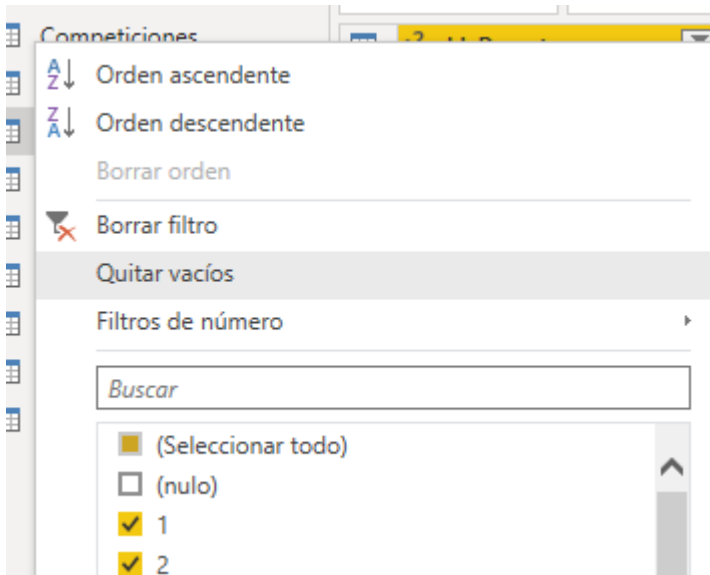
Antes de correrlo, importar la base de datos que contiene la información cruda del data set. Se encuentra en el siguiente link:

[https://drive.google.com/file/d/197gcd7vRV\\_2zdW2DdOfZGm7xOFnhTjjY/view?usp=sharing](https://drive.google.com/file/d/197gcd7vRV_2zdW2DdOfZGm7xOFnhTjjY/view?usp=sharing)

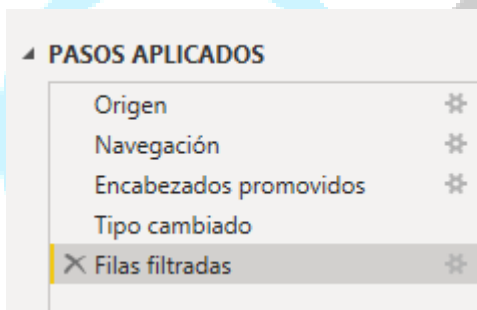
## 11. Transformaciones realizadas

Se utilizó “Power Query” para eliminar los campos nulos adicionales que se obtuvieron al cargar el data set desde “Google Sheets”. Para ellos dentro de Power Query, seleccionamos la columna a eliminar los nulos y tildamos la opción “Filtrar Vacíos”

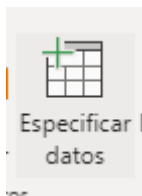




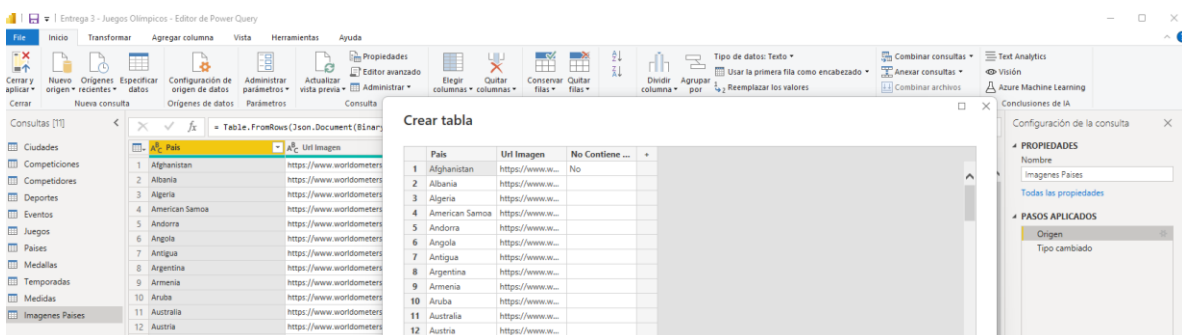
Generando así el paso aplicado de manera secuencial llamado “Filas Filtradas”:



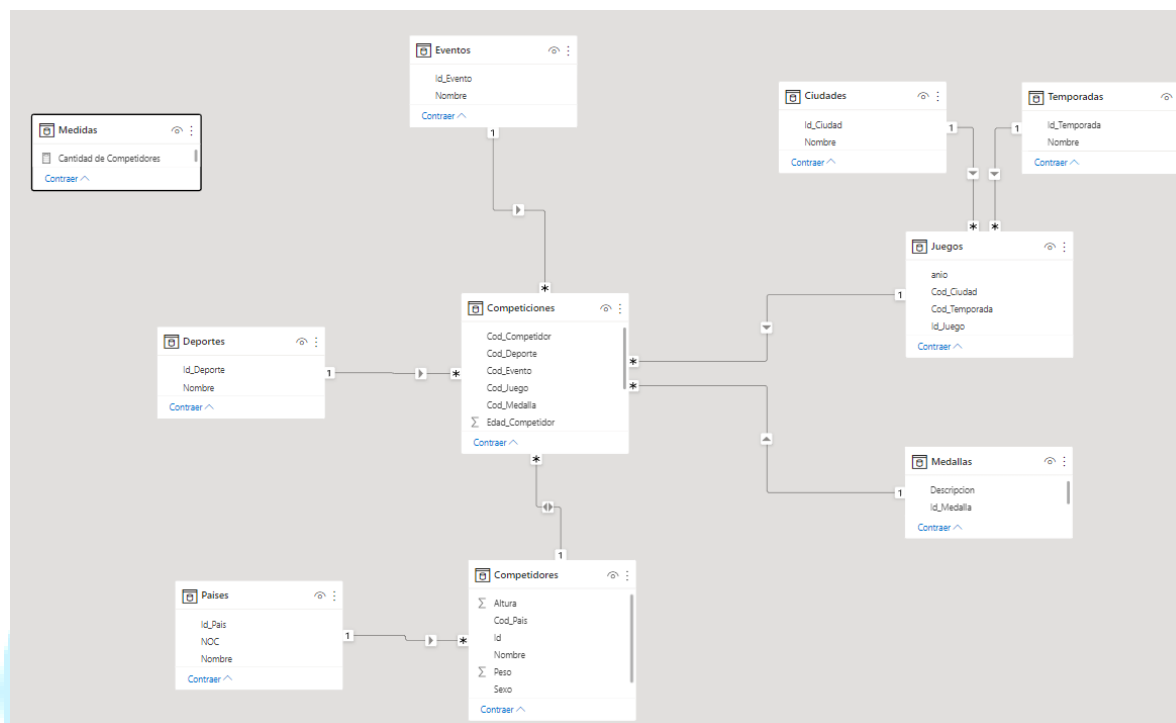
También se insertó una tabla en el modelo con imágenes de los diferentes países mediante la opción de “Especificar Datos”:



Y luego desde power query se incluyó la información necesaria en la tabla modificando el paso “Origen” dentro de “Pasos Aplicados”:



Además, todas las tablas ingestadas fueron relacionadas en Power Bi con el mismo modelado planteado para el DER:



## 12. Medidas calculadas

Se generó una Tabla de Medidas la cual incluirá todas las medidas necesarias para el reporte:

- Cantidad de Competidores = `DISTINCTCOUNT(Competidores[Id])` - Cuenta la cantidad total de competidores de manera distintiva.
- Cantidad de Deportes = `DISTINCTCOUNT(Competiciones[Nombre Deporte])` - Cuenta la cantidad de deportes de manera distintiva.
- Cantidad de Eventos = `DISTINCTCOUNT(Competiciones[Nombre Evento])` - Cuenta la cantidad de eventos de manera distintiva.
- Cantidad de Medallas = `COUNTAX(FILTER(Competiciones,[Cod_Medalla] <> "NULL"),Competiciones[Cod_Medalla])`. Cuenta la cantidad de medallas, excluyendo nulos.
- Cantidad de Países = `DISTINCTCOUNT(Competiciones[Pais Competidor])`. Cuenta la cantidad de países de manera distintiva.
- Cantidad de Países Competidores = `COUNT(Competiciones[Pais Competidor])`. Cuenta la cantidad de Países que compitieron en algún juego.
- Flag\_to\_show = `SELECTEDVALUE('Imágenes Países'[Url Imagen],"https://upload.wikimedia.org/wikipedia/commons/d/d4/World_Flag_%282004%29.svg")`. Muestra la imagen en caso de que la selección sea única. En caso contrario muestra la imagen contenida en la URL especificada.

### 13. Columnas Calculadas

#### Tabla Competiciones:

- Nombre Competidor = `RELATED(Competidores[Nombre])` - Trae el campo "Nombre" relacionado de la tabla Competidores a la tabla Competiciones.
- Nombre Deporte = `RELATED(Deportes[Nombre])` - Trae el campo "Nombre" relacionado de la tabla Deportes a la tabla Competiciones.
- Nombre Evento = `RELATED(Eventos[Nombre])` - Trae el campo "Nombre" relacionado de la tabla Eventos a la tabla Competiciones.
- Pais Competidor = `RELATED(Paises[Nombre])` - Trae el campo "Nombre" relacionado de la tabla Paises a la tabla Competiciones

#### Tabla Competidores:

- Sexo\_Completo = `IF(Competidores[Sexo] == "F", "Femenino", "Masculino")` - Devuelve "Femenino" o "Masculino" dependiendo del valor del campo "Sexo" de la tabla "Competidores"

#### Tabla Medallas:

```
Traducción Medallas = IF(Medallas[Descripcion] == "Gold", "Oro",  
                          IF(Medallas[Descripcion] == "Silver", "Plata",  
                              IF(Medallas[Descripcion] == "Bronze", "Bronce", "SinMedalla"))  
) - Devuelve la traducción al español del nombre de las medallas dependiendo del valor encontrado en el campo "Descripcion" de la tabla "Medallas"
```

#### Tabla Temporadas:

Traducción Temporada = `IF(Temporadas[Nombre] == "Summer", "Verano", "Invierno")` - Devuelve la traducción al español del nombre de las temporadas dependiendo del valor encontrado en el campo "Nombre" de la tabla "Temporadas"

### 14. Segmentaciones elegidas

Se eligieron diferentes segmentaciones dependiendo de cada página, por lo que se analizará cada página por separado, explicando brevemente cada una y señalando el tipo de Segmentaciones seleccionadas.

#### Deportes:

La misma tiene el objetivo de mostrar la cantidad de deportes y la cantidad de eventos encontrados durante los Juegos Olímpicos. Además, se analiza la cantidad de países que participaron en los diferentes deportes.

Las segmentaciones utilizadas fueron:

- Año: Permitiendo filtrar el intervalo en el cual se realizaron los juegos:

Año

1896

2016



- Deporte: Permite elegir los deportes que se quieren visualizar. En dicho caso, se añadió para permitir la búsqueda con el objetivo de facilitar el filtrado:

Deporte

Todas

Buscar

- ☐ Aeronautics
- ☐ Alpine Skiing
- ☐ Alpinism
- ☐ Archery
- ☐ Art Competitions
- ☐ Athletics

- Evento: Permite elegir los eventos que se quieren visualizar. En dicho caso, se añadió para permitir la búsqueda con el objetivo de facilitar el filtrado:

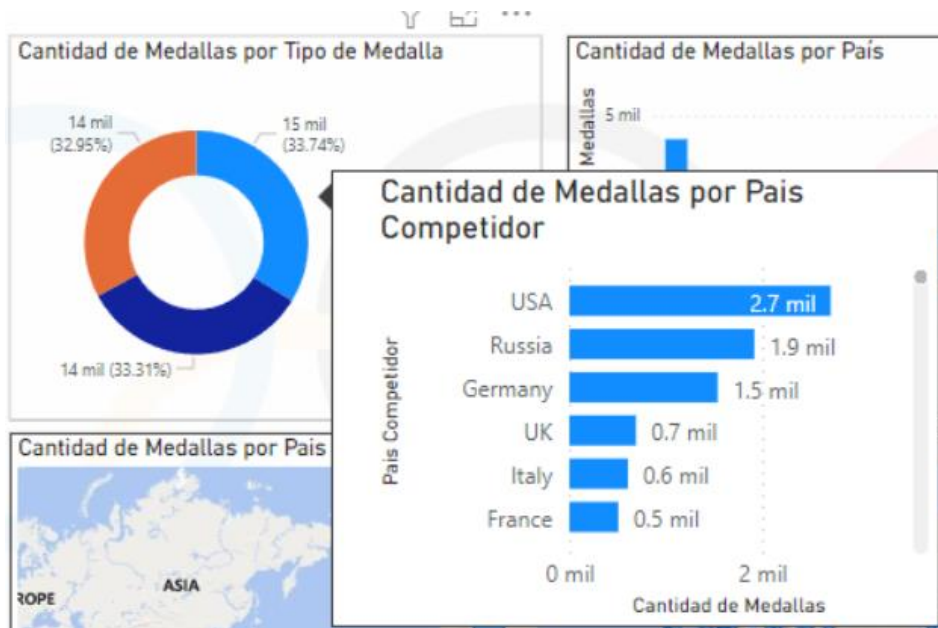
Buscar

- ☐ Aeronautics Mixed Aeronautics
- ☐ Alpine Skiing Men's Combined
- ☐ Alpine Skiing Men's Downhill
- ☐ Alpine Skiing Men's Giant Slal...
- ☐ Alpine Skiing Men's Slalom
- ☐ Alpine Skiing Men's Super G

Todas

### Medallas:

La página de medallas tiene el objetivo de mostrar la Cantidad total de medallas entregadas separadas por categoría y poder corroborar que país obtuvo más medallas. Además, se añadió un Tooltip sobre el gráfico de dona que permitirá ver la cantidad de países que obtuvieron una medalla en particular:



Y también se añadió una visualización del estilo “Mapa” para mostrar la cantidad de Medallas por países:

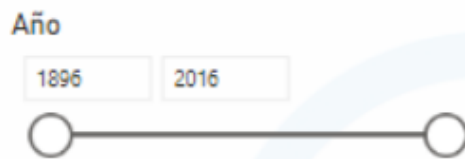


Y se agregó la imagen de cada bandera que cambia en base a la selección de cada país:



Las segmentaciones utilizadas fueron:

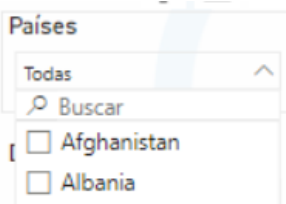
- Año: Permitiendo filtrar el intervalo en el cual se realizaron los juegos:



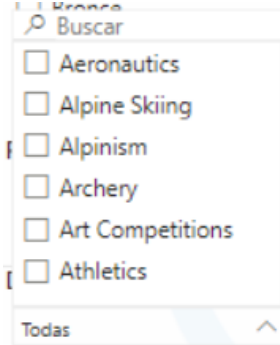
- Tipo de Medalla: Permite filtrar que medalla analizar en el tablero:



- Países: Permite filtrar a nivel de país para ver que medallas obtuvo y en que deporte se desempeña mejor.



- Deportes: Permite filtrar a nivel de deporte, en cuál se obtuvieron más medallas y que país es mejor en uno u otro deporte



Buscar

- ☐ Aeronautics
- ☐ Alpine Skiing
- ☐ Alpinism
- ☐ Archery
- ☐ Art Competitions
- ☐ Athletics

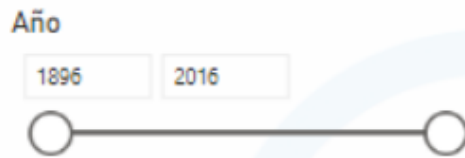
Todas

#### Competidores:

La siguiente página tiene el objetivo de analizar la cantidad total de competidores, la cantidad total de competidores por deporte y separar los mismos por sexo.

Las Segmentaciones utilizadas fueron:

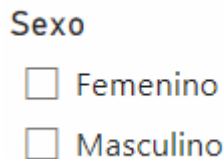
- Año: Permitiendo filtrar el intervalo en el cuál se realizaron los juegos:



Año

1896 2016

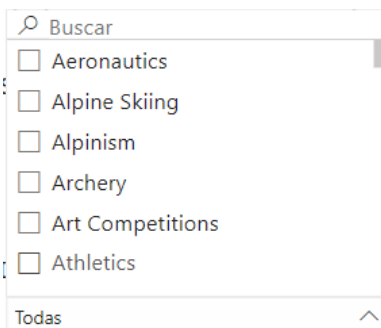
- Sexo: Permitiendo separar por Femenino o Masculino según corresponda.



Sexo

- ☐ Femenino
- ☐ Masculino

- Deporte: Permitiendo elegir y comparar a un nivel de detalle que deportes tienen mas participación de un sexo u otro



Buscar

- ☐ Aeronautics
- ☐ Alpine Skiing
- ☐ Alpinism
- ☐ Archery
- ☐ Art Competitions
- ☐ Athletics


Todas

### Juegos:

- Año: Permitiendo filtrar el intervalo en el cuál se realizaron los juegos:

Año

1896 2016



- Ciudad: Permite filtrar por el lugar donde ocurrieron los Juegos

Ciudad

Todas

Buscar

☐ Albertville

☐ Amsterdam

- Temporada: Permite filtrar por el tipo de Juego, ya que varían los deportes de los mismos

☐ Invierno

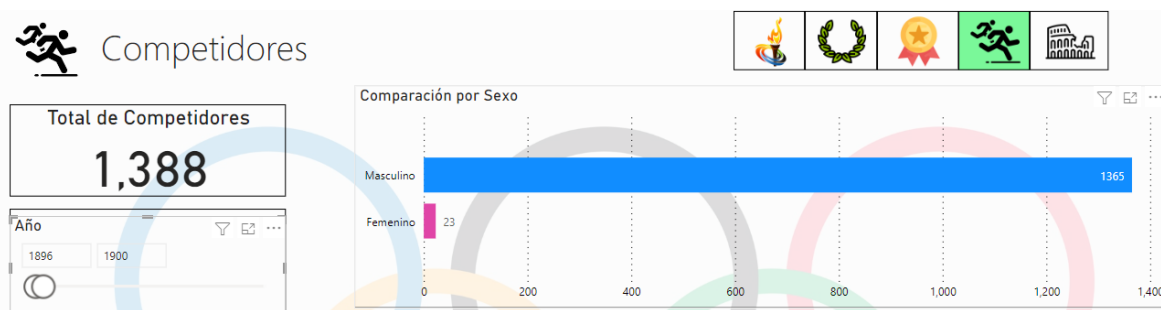
☐ Verano

Todas

## 15. Conclusiones

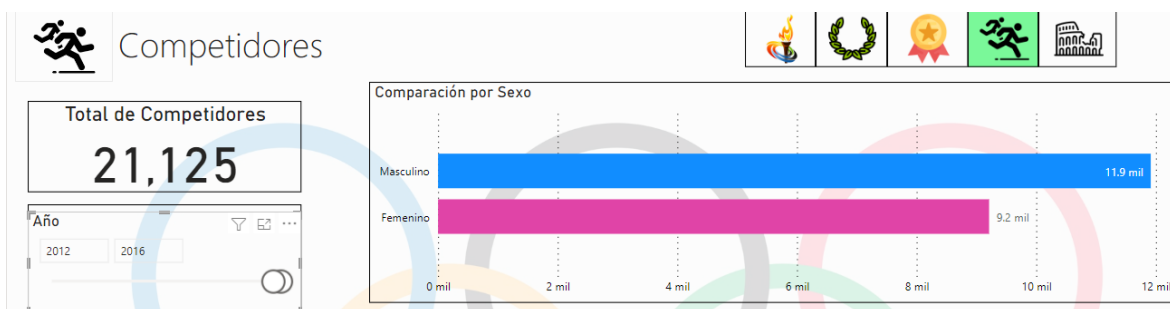
Luego de realizar el reporte basándonos en las hipótesis planteadas, pudimos notar ciertos datos al analizarlos en detenimiento. Por ejemplo, si miramos la solapa de Competidores, podremos ver como la participación del genero Femenino fue aumentando conforme nos acercamos a la actualidad:

**1986 – 1990:**

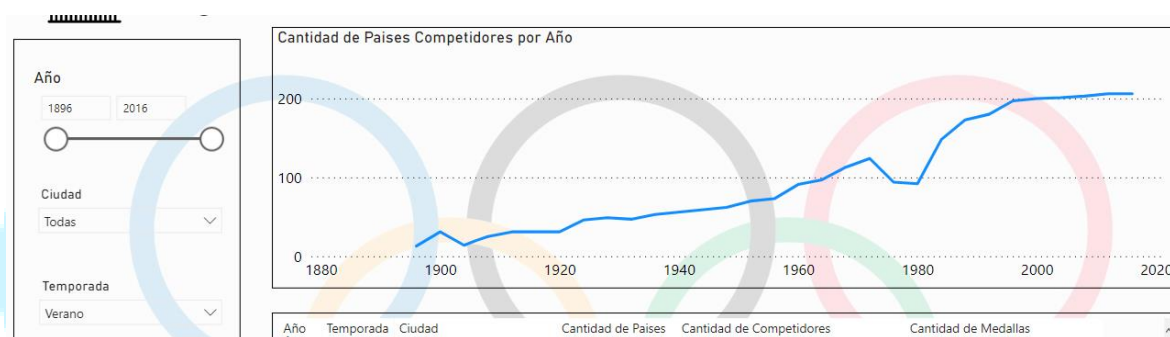




**2012-2016:**



También puede observarse por ejemplo, si analizamos la solapa de “Juegos”, si seleccionamos una de las temporadas, veremos como el aumento de los países a lo largo del tiempo fue en crecimiento, habiendo muchos mas participantes que en un principio:



Podemos concluir entonces que el análisis descriptivo mediante la utilización del modelo relacional es muy útil para comprender tendencias y establecer conclusiones acerca de los diferentes datos.