# Consigna Clase 19 - Segunda Entrega de Trabajo Final

Juegos Olímpicos: Cien Años de Historia



Para CoderHouse   
presentado el 4/10/2021

Integrantes:

* Hoecker, Felipe
* Miranda, Facundo
* Scidá, Federico

## Índice

[**Consigna Clase 19 - Segunda Entrega de Trabajo Final**](#_heading=h.d64towpf5ad4) **1**

[Índice](#_heading=h.mhof8rrjwxe0) [2](#_heading=h.mhof8rrjwxe0)

[Introducción](#_heading=h.acmsd97iyvma) [3](#_heading=h.acmsd97iyvma)

[Objetivo](#_heading=h.yih73ut47txj) [3](#_heading=h.yih73ut47txj)

[Base de Datos](#_heading=h.cr16wn8kko8) [4](#_heading=h.cr16wn8kko8)

[Diagrama E-R](#_heading=h.gnbjz4n0yov0) [4](#_heading=h.gnbjz4n0yov0)

[Listas de Tablas](#_heading=h.6gtgi2j372m7) [5](#_heading=h.6gtgi2j372m7)

[Columnas por Tablas](#_heading=h.9i895p8w3zmb) [6](#_heading=h.9i895p8w3zmb)

[Visualización](#_heading=h.q69f5e3c716l) [7](#_heading=h.q69f5e3c716l)

[Diagrama E-R en Power BI](#_heading=h.uxgc52x2dcof) [11](#_heading=h.uxgc52x2dcof)

[Medidas Calculadas](#_heading=h.rwjzyrjjwydi) [12](#_heading=h.rwjzyrjjwydi)

[Segmentaciones](#_heading=h.ffkv4ksdy0l4) [13](#_heading=h.ffkv4ksdy0l4)

[Futuras líneas](#_heading=h.sc62tbzd3cto) 13

## Introducción

Se presentará un trabajo acerca de las Olimpiadas en sus ediciones de invierno y verano, desde sus inicios al día de la fecha. La base de datos obtenida permitirá analizar la evolución del evento a lo largo de los años, con el foco puesto específicamente en lo deportivo (y no, por ejemplo, en la audiencia, los presupuestos, o la infraestructura de cada país anfitrión). Evaluaremos así el desempeño de cada equipo y atleta en diferentes disciplinas, obtendremos las tendencias históricas en cuanto a ganadores, analizaremos la evolución en peso y estatura de los participantes en cada deporte y estableceremos un medallero definitivo. Se pondrán en práctica medidas calculadas simples, gráficos vistos en clase y segmentaciones por categorías para facilitar la navegación.

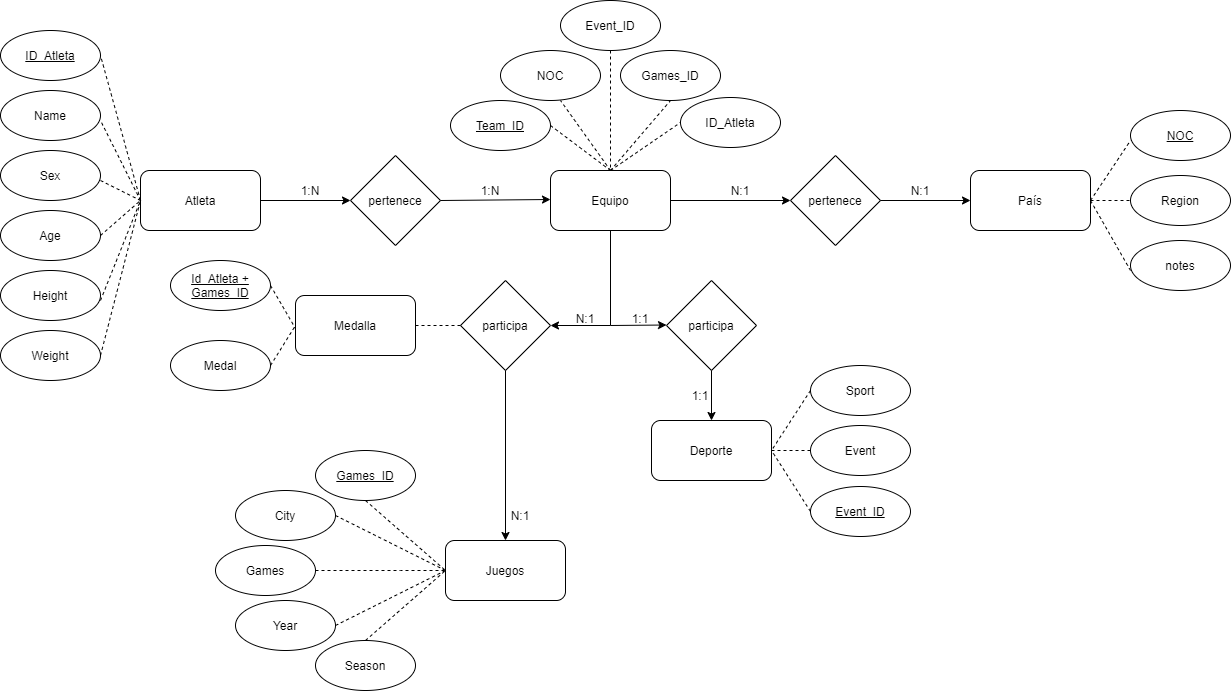
## Objetivo

Este tablero está orientado a un público con un conocimiento superficial de las Olimpiadas. El tablero no ofrece mayor contexto (qué son las Olimpiadas, en qué consisten, cuál es su historia, etc), sino que salta directamente a las características de cada competencia, comparándolas a través de sus ediciones desde lo cuantitativo. Se trata de un tablero de tipo **informativo o didáctico**, dado que difícilmente puedan tomarse decisiones organizacionales partiendo de lo que ofrece.

## Base de Datos

Para darle forma al proyecto, se tuvo que separar la base de datos original en varias tablas, a fin de darle forma al modelo relacional presentado en la sección anterior. Se elaboraron columnas índice para crear las relaciones de correspondencia del modelo y se cambiaron los tipos de datos según se explicita en las secciones siguientes. Por lo demás, no se realizaron mayores transformaciones.

## Diagrama E-R



## 

## Listas de Tablas

| **Tabla** | **PK** | **FK** | **CK** |
| --- | --- | --- | --- |
| Atletas | ID\_Atleta |  |  |
| Equipo | Team\_ID | NOC, Event\_ID, Games\_ID, ID\_Atleta |  |
| País | NOC |  |  |
| Juegos | Games\_ID |  |  |
| Deporte | Event\_ID |  |  |
| Medalla |  | ID\_Atleta, Games\_ID | ID\_Atleta, Games\_ID |

## Columnas por Tablas

Tabla 1: Atletas

| **Tipo de clave** | **Campo** | **Tipo de Campo** |
| --- | --- | --- |
| PK | ID\_Atleta | int |
|  | Name | nvarchar(50) |
|  | Sex | nvarchar(50) |
|  | Age | int |
|  | Height | decimal(5,2) |
|  | Weight | decimal(5,2) |

Tabla 2: Equipo

| **Tipo de clave** | **Campo** | **Tipo de Campo** |
| --- | --- | --- |
| PK | Team\_ID | int |
| FK | NOC | text(3) |
| FK | Event\_ID | int |
| FK | Games\_ID | int |
| FK | ID\_Atleta | Int |

Tabla 3: País

| **Tipo de clave** | **Campo** | **Tipo de Campo** |
| --- | --- | --- |
| PK | NOC | text(3) |
|  | Region | nvarchar(50) |
|  | notes | nvarchar(200) |

Tabla 4: Juegos

| **Tipo de clave** | **Campo** | **Tipo de Campo** |
| --- | --- | --- |
| PK | Games\_ID | int |
|  | City | nvarchar(50) |
|  | Year | int |
|  | Season | nvarchar(50) |
|  | Games | nvarchar(50) |

Tabla 5: Deporte

| **Tipo de clave** | **Campo** | **Tipo de Campo** |
| --- | --- | --- |
|  | Sport | nvarchar(50) |
|  | Event | nvarchar(50) |
| PK | Event\_ID | int |

Tabla 6: Medalla

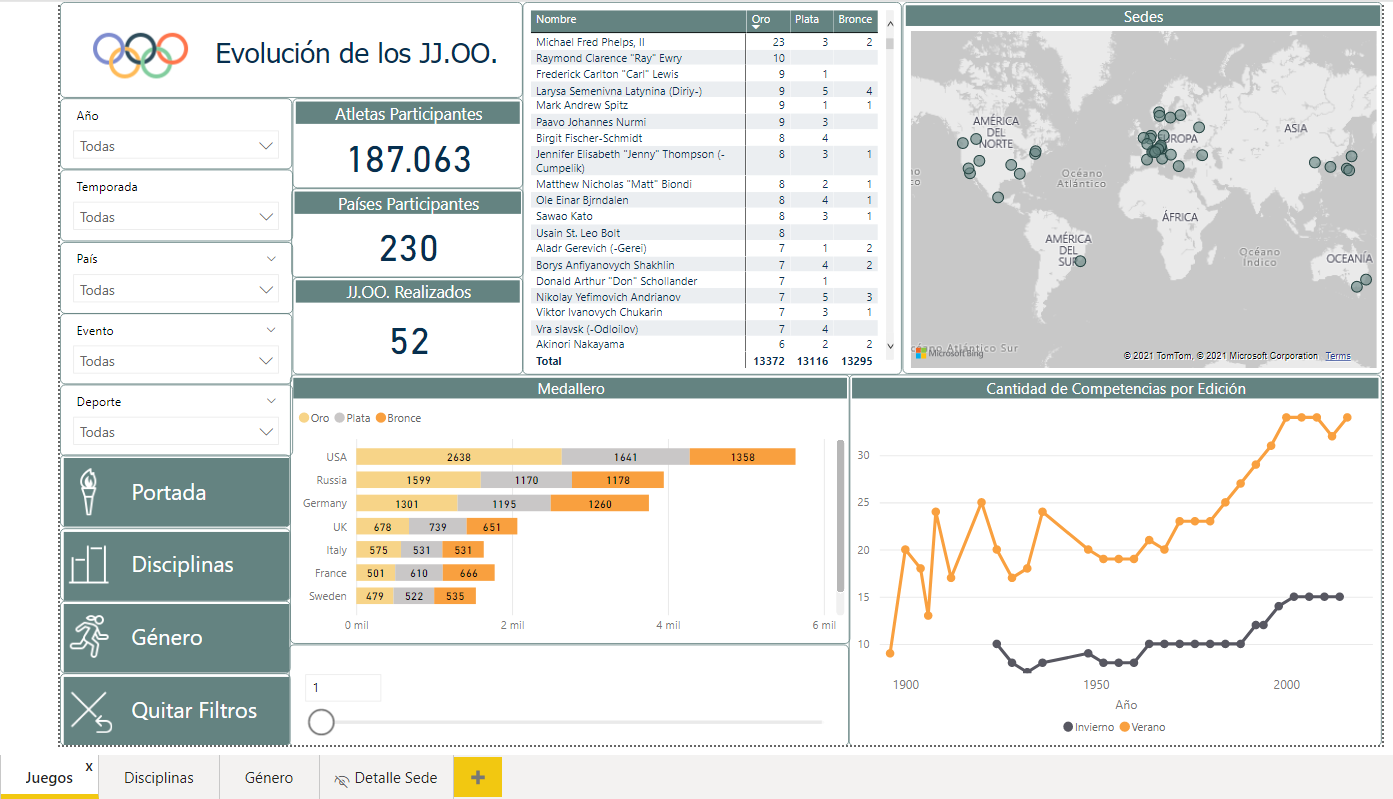
| **Tipo de clave** | **Campo** | **Tipo de Campo** |
| --- | --- | --- |
| FK-CK | ID\_Atleta | int |
| FK-CK | Games\_ID | int |
|  | Medalla | nvarchar(50) |

## Visualización

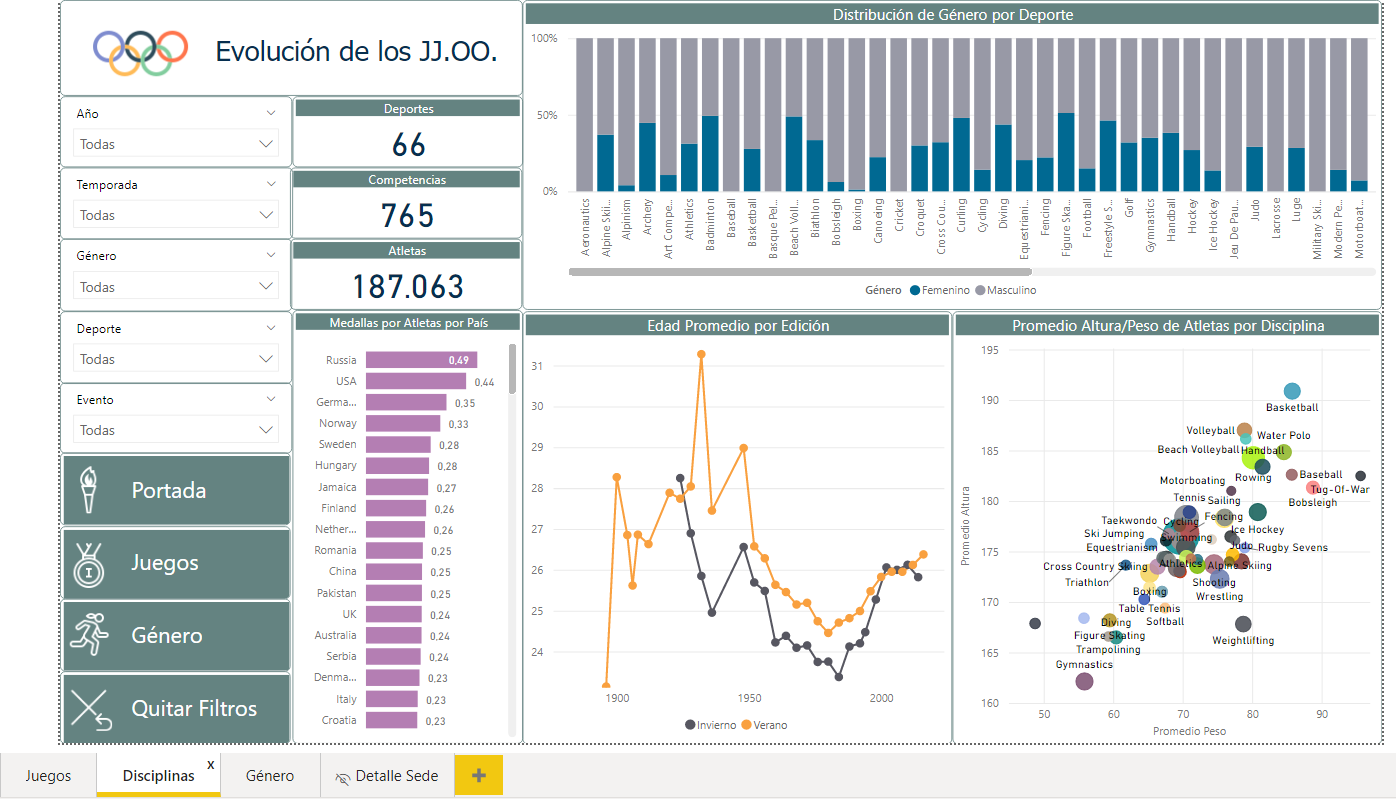
Para la elaboración del tablero tuvo que hacerse un trabajo fino en la confección de las columnas índice, a fin de lograr una adecuada correspondencia en el modelo relacional. Otra adición importante a los datos fue la de la tabla calendario, que permite vincular los datos en un marco temporal coherente.

A continuación se adjuntan capturas de las pestañas:

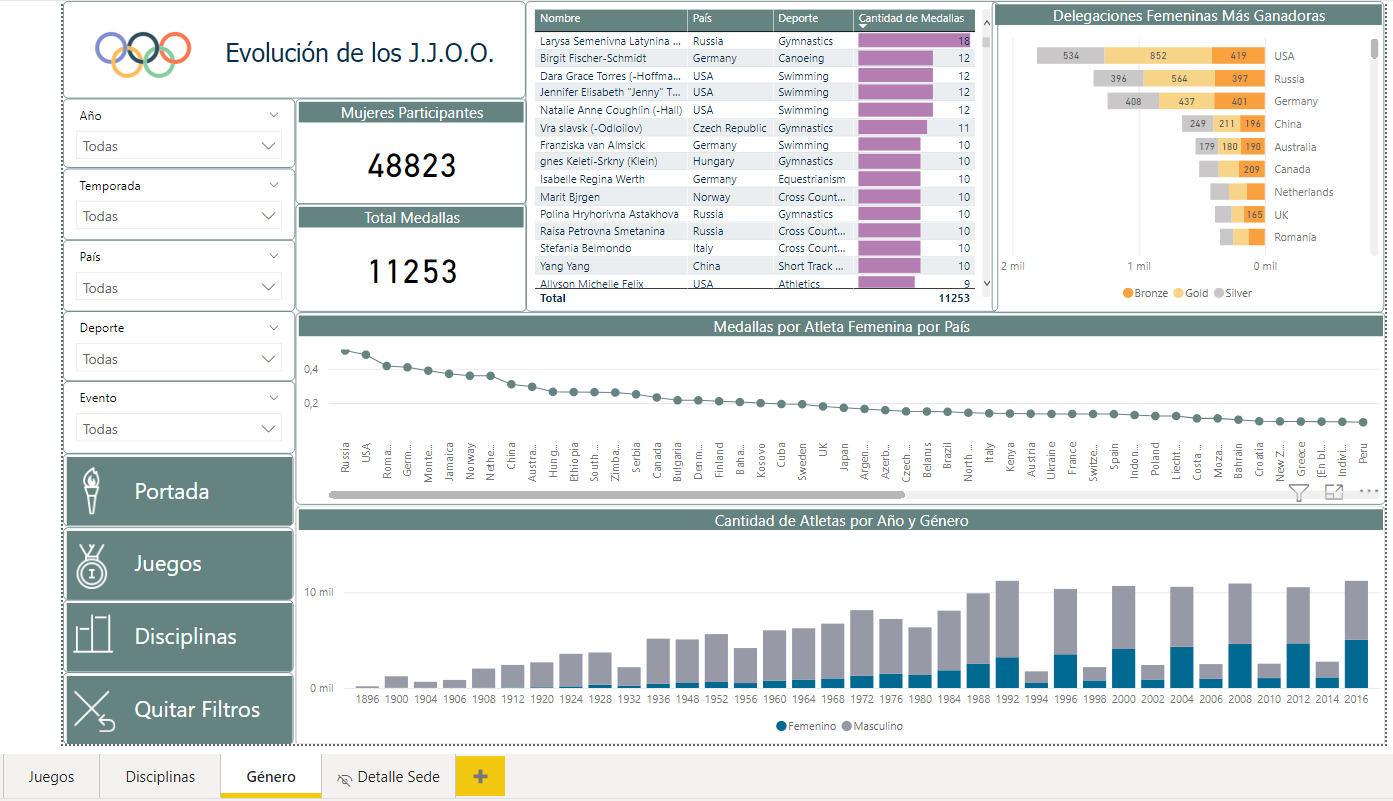
1. La primera pestaña corresponde a la hoja “Juegos” y se analiza la competencia olímpica a lo largo de los años desde una perspectiva general. Se evalúan las delegaciones y atletas más ganadores, las sedes y la evolución en la cantidad de competiciones en cada edición. Un mapa del mundo señala cada sede, y una tooltip permite obtener información sobre cada una con solo flotar el cursor sobre ella. Un gráfico de barras horizontales apiladas señala las delegaciones más ganadoras discriminando por medallas de oro, plata y bronce, y uno de líneas indica cuántas competencias ha habido en cada edición, con una clara tendencia a la alta tanto para los juegos de invierno como para los de verano. Una tabla en la parte superior indica los atletas más ganadores de la historia, y las tarjetas otorgan información resumida sobre los eventos. Un parámetro permite estimar cómo quedaría el medallero si la tendencia se repite en las citas venideras.



1. La segunda pestaña corresponde a la hoja “Disciplinas”, que hace foco en las distintas competencias y cómo estás han ido cambiando a lo largo de la historia del certamen. Un gráfico de líneas muestra cómo ha ido cambiando la edad promedio de los atletas a lo largo de los años en las ediciones de invierno y verano, y se contrasta peso y estatura promedio de atletas según deporte en un gráfico de burbujas. Se calculó la cantidad de medallas por delegación y la información se puso en un gráfico de barras, y la paridad de género por deporte en la parte superior, en forma de un gráfico de columnas 100% apiladas. Los filtros aplicables permiten dar detalles muy interesantes sobre cada competencia.



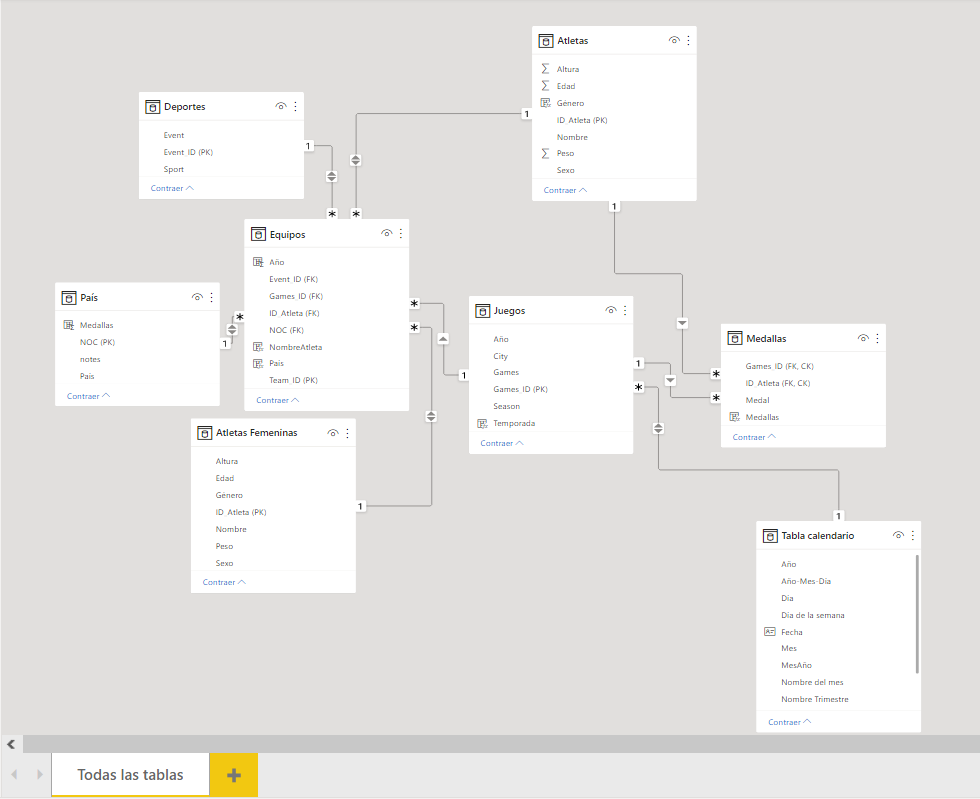
1. La tercera pestaña corresponde a la hoja “Género”, donde se estudia la trayectoria de la competencia desde esa perspectiva. Un gráfico de barras verticales describe cómo ha ido evolucionando la presencia femenina en el certamen, con una clara tendencia a la paridad, y otro de barras horizontales indica cuáles son las delegaciones femeninas más victoriosas. Una tabla con formato condicional da cuenta de las atletas más ganadoras, su delegación y disciplina, y un gráfico de líneas y puntos describe claramente el ratio medallas obtenidas/atletas, demostrando cuáles son las delegaciones femeninas más “eficientes” de la competencia.



Estas tres pestañas tienen filtros según lo especificado en la sección 9.Segmentaciones, y todas tienen también tarjetas con agregados clave que ayudan a resumir la información, según se indica en los títulos de cada una. Para facilitar la navegación, se incluye abajo a la izquierda de cada hoja una serie de botones que permite pasar de una pestaña a otra, o bien volver a la portada.

Existe también una cuarta pestaña que permanece oculta, cuya única función es brindar los datos al tooltip del mapa en la primera pestaña.

## Diagrama E-R en Power BI



Se incluyó una tabla calendario a fin de activar las relaciones entre las tablas en torno a las fechas. No están incluidas en la imagen la Tabla Métricas ni la Tabla Parámetro, dado que no tienen ninguna relación con el modelo.

## Columnas Calculadas

Para las tablas empleadas se generaron columnas calculadas adicionales, con el propósito, principalmente, de traducir del inglés al castellano los contenidos de la tabla, o bien de evidenciar las relaciones con otras tablas, como ocurre con las columnas calculadas de la tabla “Equipos”:

Tabla Atletas:

* Género = IF(Atletas[Sexo] = "M", "Masculino", "Femenino")

Tabla Equipos:

* Año = RELATED(Juegos[Año])
* NombreAtleta = RELATED(Atletas[Nombre])
* País = RELATED('País'[País])

Tabla Juegos:

* Temporada = IF(Juegos[Season]="Summer","Verano","Invierno")

Tabla Medallas:

* Medallas = IF(Medallas[Medal]="Gold","Oro",IF(Medallas[Medal]="Silver","Plata",IF(Medallas[Medal]="Bronze","Bronce","NA")))

Tabla País

* Medallas = [Medallas]

También se incluyó una tabla calendario.

## Medidas Calculadas

Todas las medidas calculadas se encuentran en la tabla “Tabla Métricas”, y las que se emplean en el modelo son las siguientes:

* Atleta\_fem\_mas\_medallas = CALCULATE([Medallas], Atletas[Sexo]="F")
  + Se filtra, en la tabla Medallas, todas las filas que correspondan a una atleta femenina, por medio del campo Sexo de la tabla Atletas. Sin una correcta interacción del modelo relacional, este cálculo no sería posible.
* Cantidad De Juegos = COUNT(Juegos[Games\_ID (PK)])
  + Cuenta la cantidad de ediciones a partir del campo Games\_ID, PK de la tabla Juegos.
* Cantidad De Atletas = DISTINCTCOUNT(Equipos[ID\_Atleta (FK)])
  + Cuenta la cantidad de valores únicos en el campo ID\_Atleta de la tabla Equipos, que es FK en esa tabla y PK en la tabla Atletas. Es equivalente a hacer COUNT(Atletas[ID\_Atleta])
* Cantidad\_Países = COUNT('País'[País])
  + Cuenta valores del campo País, PK de la tabla homónima
* Cantidad de Deportes = DISTINCTCOUNT(Equipos[Event\_ID (FK)])
  + Cuenta valores únicos del campo Event\_ID de la tabla Equipos, que es FK de esa tabla y PK de la tabla Deportes. Es equivalente a hacer COUNT(Deportes[Event\_ID])
* Cantidad de Competencias = DISTINCTCOUNT(Deportes[Event])
  + Cuenta la cantidad de valores únicos en el campo Event de la tabla Deportes
* Bronce = calculate(count(Medallas[Medallas]), Medallas[Medallas]="Bronce")
  + Se filtra, en la tabla Medallas, todas las filas donde el valor de la columna homónima es Bronce. Esto permite calcular el medallero en la primera pestaña.
* Oro = calculate(count(Medallas[Medallas]), Medallas[Medallas]="Oro")
  + Se filtra, en la tabla Medallas, todas las filas donde el valor de la columna homónima es Oro. Esto permite calcular el medallero en la primera pestaña.
* Plata = calculate(count(Medallas[Medallas]), Medallas[Medallas]="Plata")
  + Se filtra, en la tabla Medallas, todas las filas donde el valor de la columna homónima es Plata. Esto permite calcular el medallero en la primera pestaña.
* Mujeres = CALCULATE([Cantidad De Atletas], Atletas[Sexo]= "F")
  + Actúa sobre la medida calculada “Cantidad de Atletas”, descripta anteriormente, filtrando solamente por atletas femeninas, según el campo Sexo de la tabla Atletas. Sin una correcta interacción del modelo relacional este cálculo no sería posible.
* Medallas por Atletas = ([Oro]+[Plata]+[Bronce])/COUNT(Atletas[ID\_Atleta (PK)])
  + Esta medida permite conocer la relación entre cantidad de medallas (indistintamente del tipo) obtenidas por una delegación y la cantidad de atletas que llevó al certamente. Es una forma de “eficiencia”. Se calcula como la suma de los campos Oro, Plata y Bronce, anteriormente descriptos, y el conteo de valores del campo ID\_Atleta de la tabla Atletas, de la cual es PK.
* Medallas por Atleta Femenina = [Atleta\_fem\_mas\_medallas]/[Mujeres]
  + Medida de la “eficiencia” de las delegaciones, según lo descripto en la viñeta anterior, aplicado solamente a atletas femeninas. Se emplea la medida calculada Atleta\_fem\_mas\_medallas y se divide en la medida Mujeres, ambas ya detalladas con anterioridad.

## Segmentaciones

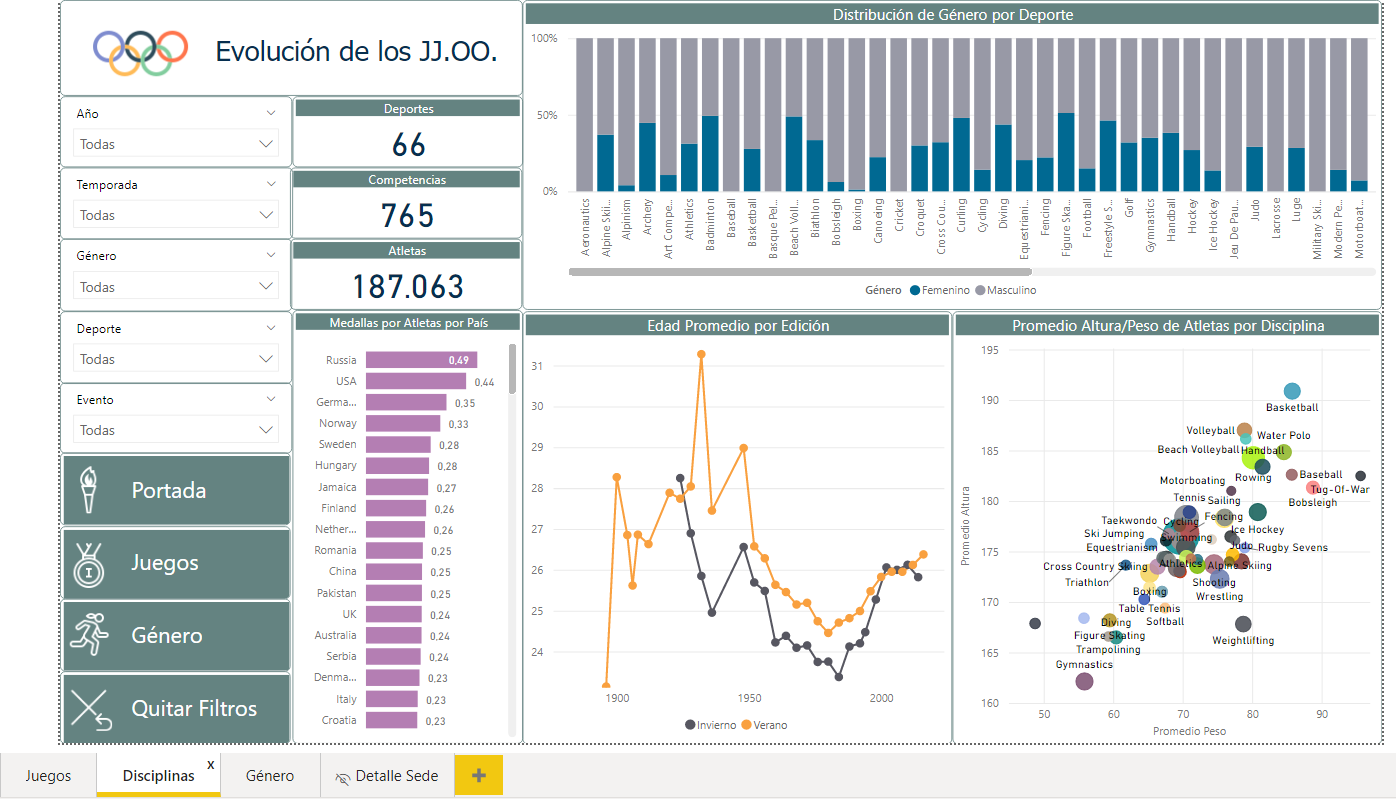
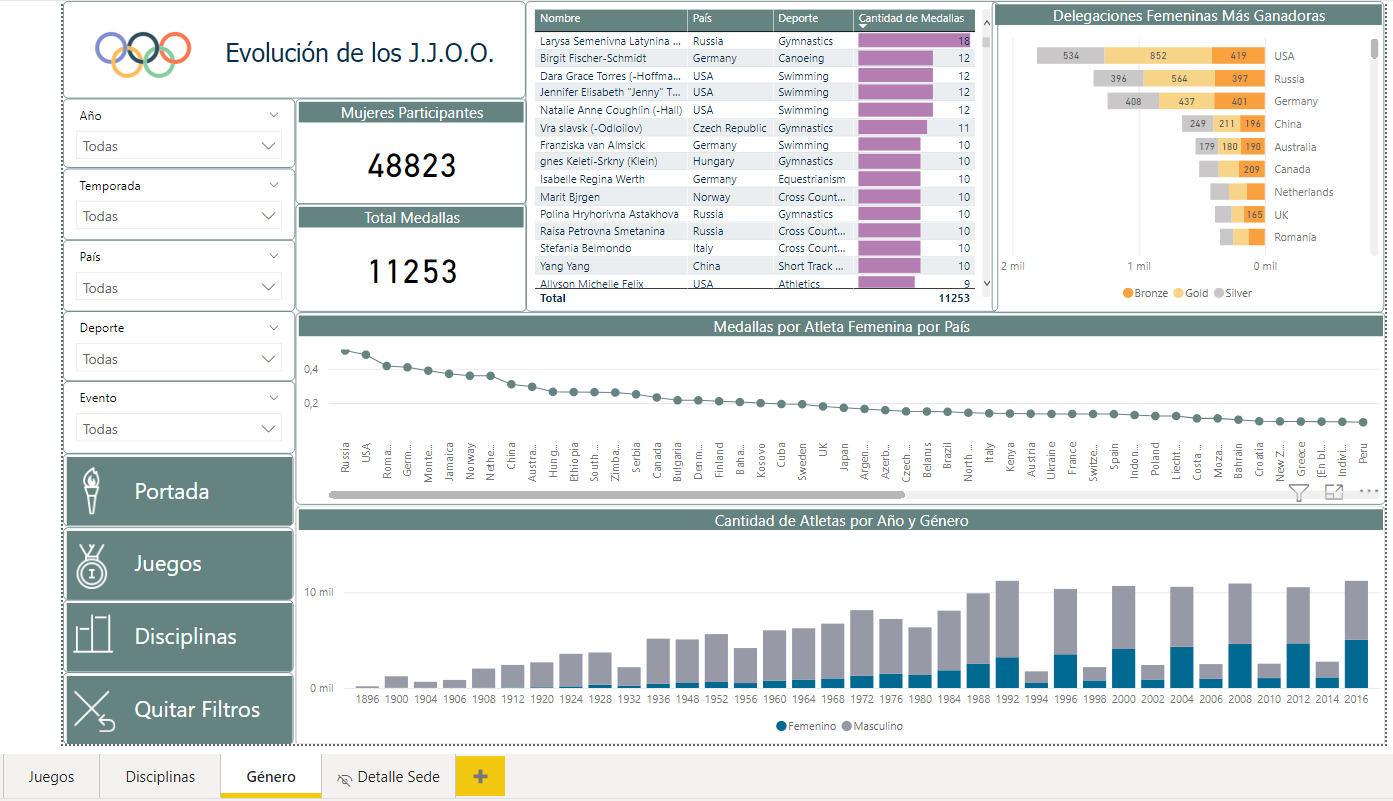
Se elaboraron las siguientes segmentaciones, según sirven para cada una de las solapas:

* Por año
* Por edición
* Por deporte
* Por disciplina
* Por género
* Por país

## 

## 12. Tabla Versionado

| **Versión** | **Descripción** |
| --- | --- |
| 0,5 | Modelo de datos |
| 1,0 | Primeras medidas |
| 1,0 | Creación de botones |
| 1,0 | Creación de logos |
| 1,0 | Creación de gráficos |
| 1,0 | Creación de filtros |
| 1,1 | Modificación de botones |
| 1,2 | Modificación de estructura y orden de la solapa |
| 1,3 | Creación de solapas: Portada, juegos, disciplinas y genero |
| 1,4 | Edición y interacción de filtros |
| 1,5 | Edición y creación de botones con acciones |
| 1,6 | Creación de medidas calculadas |
| 1,7 | Corrección de gráficos de barra en base a la devolución de la 1er entrega final |



## 13. Tecnología Utilizada

Para la elaboración de este tablero se emplearon una serie de programas y servicios web, a saber:

* Microsoft Excel, para el estudio de la primera base de datos obtenida,
* Microsoft SQL Server Management Studio, para la primera segmentación de esta base en diferentes tablas, por medio de comandos SQL,
* Power BI, para la elaboración de gráficos a partir de las tablas obtenidas,
* Lenguaje M, para la estructuración de ciertas partes de las tablas,
* Lenguaje DAX, para las medidas calculadas,
* El sitio coolors.com, para la obtención de paletas de colores,
* El sitio icons8.com, para la obtención de íconos

## 14. Líneas Futuras

Una mejora interesante que podría hacérsele al tablero es la incorporación de datos no ligados a lo deportivo, sino más bien a lo económico, por ejemplo: presupuesto e ingresos por edición, presupuesto por comitiva, población y PBI de países para cada certamen, etc. Esto va a permitir comparar el desempeño de las delegaciones en relación a las condiciones macroeconómicas de sus países, y estudiar, entre otras cosas, cómo ha ido cambiando la política presupuestaria del certamen. Otro aspecto que podría brindar observaciones interesantes es el de las Olimpiadas como espectáculo: audiencia, patrocinadores, difusión, etc.