



Los países y los Juegos Olímpicos: 1896-2020

Bernabé Aguilar

Braian Vicinelli

Belén Triantafilou

CoderHouse

Profesora: Candela Beyreuther

Comisión: #32485

Fecha: 14/11/2022



Índice

Introducción	3
Dataset	4
Diagrama de Entidad de Relación	4
Tablas	5
Medidas y Columnas	8
Visualizaciones	12
Futuras Líneas de Investigación	20
Diagrama de Entidad de Relación en Power BI	20
Conclusión.....	21



1. Descripción de la temática de los datos

Se extrajeron datos de los resultados históricos de los Juegos Olímpicos. Estos contemplan los datos de todos los atletas que participaron, países, características físicas, medallero, etc.

2. Objetivo:

Comprobar si hay una predisposición de nacionalidad a la hora de obtener medallas en ciertas disciplinas. Analizar si las características físicas de los atletas influyen en el medallero. Generar un top de disciplinas por país y su evolución histórica. Este análisis tiene como objetivo final ayudar en la toma de decisiones económicas de las diferentes organizaciones deportivas en los países participantes.

3. Alcance:

El alcance de este dashboard, es para las diferentes organizaciones deportivas que participan en los Juegos Olímpicos, para que éstas puedan tener al alcance y conocer perfectamente lo que ocurrió en los diferentes eventos a lo largo de la historia, facilitar el manejo de información y la toma de decisiones económicas a la hora de tener que destinar fondos para apoyar a ciertos atletas en las diferentes disciplinas.

4. Tipos de Análisis:

- I. **Descriptivo:** Datos personales de cada uno de los atletas participantes en los Juegos Olímpicos (Fecha de Nacimiento; Nombre, Nacionalidad, Sexo, Altura, Peso, Deporte, Medallero); Países participantes; Juegos Olímpicos (Fechas, Ubicación)
- II. **Diagnóstico:** Mostrar países con mayor rendimiento en las diferentes disciplinas. Evolución histórica de los deportes en el top de sus medalleros. Comprobar características físicas que influyeron o no.
- III. **Predictivo:** Se busca predecir la evolución del medallero de los diferentes países y el porqué de este.
- IV. **Prescriptivo:** Ayudar a los países a proyectar las inversiones en el marco deportivo a atletas en disciplinas con mayores probabilidades de obtener una medalla.



5. Dataset

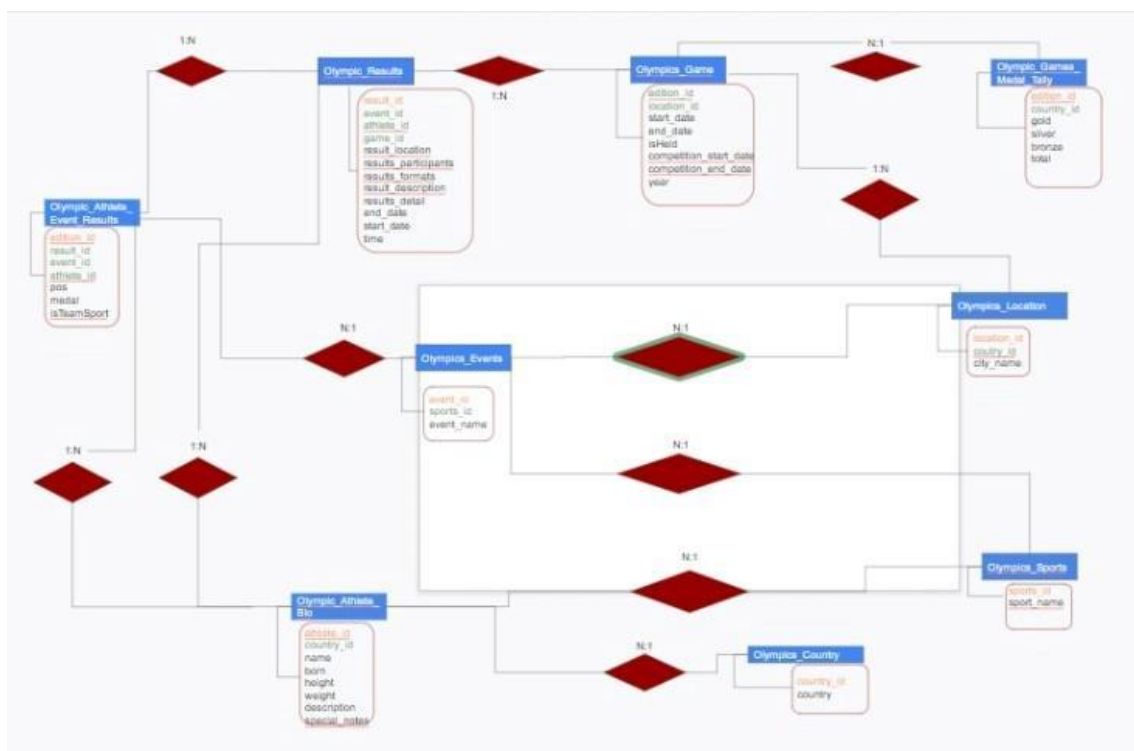
Fuente: <https://www.kaggle.com/datasets/josephcheng123456/olympic-historical-dataset-from-olympediaorg>

Archivos:

- i. Países-Olimpiadas.xlsx
 1. Olympic_Athlete_Bio.csv
 2. Olympic_Athlete_Event_Results.csv
 3. Olympic_Games_Medal_Tally.csv
 4. Olympic_Results.csv
 5. Olympics_Country.csv
 6. Olympics_Games.csv
 7. Query.csv
- ii. Países-y-Olimpiadas_Aguilar-Vicinelli-Triantafilou.pdf
- iii. Países-y-Olimpiadas-v12.pbix

6. Diagrama de Entidad-Relación

Link: [Diagrama de Entidad-Relación](#)





7. Listado de Tablas:

- I. **Olympic_Results:** Contiene información más detallada de los resultados de todas las participaciones. Cantidad de participantes, de cuántos países, y descripciones sobre la competencia.
 - i. PK: result_id
 - ii. FK: evento_id
 - iii. FK: edition_id
- II. **Olympic_Athlete_Event_Results:** Esta tabla contiene el resultado de cada atleta en cada una de las ediciones, dividido por cada uno de los eventos (disciplinas). Si ganó medalla, si el deporte es grupal o individual y la marca que hayan obtenido.
 - i. CK: edition_id
 - ii. CK: athlete_id
 - iii. FK: event_id
 - iv. FK: result_id
- III. **Olympic_Athlete_Bio:** La biografía de cada uno de los atletas que participaron en las Olimpiadas. Nombre, sexo, fecha de nacimiento, peso y altura del momento de la competición, país y notas para conocer más sobre el atleta.
 - i. PK: athlete_id
 - ii. FK: country_id
 - iii. FK: sports_id
- IV. **Olympic_Games_Medal_Tally:** El resultado final del medallero en cada edición dividido por país.
 - i. CK: edition_id
 - ii. CK: country_id
- V. **Olympics_Country:** Lista de los países que forman parte de las Olimpiadas, como participantes o como sedes.
 - i. PK: country_id
- VI. **Olympics_Game:** El detalle de cada edición de los juegos Olímpicos. Año, fecha, locación y si por alguna razón no se realizó.
 - i. PK: edition_id
 - ii. FK: location_id
- VII. **Olympics_Sports:** Categorías deportivas a las que pertenecen cada uno de los eventos (disciplinas)
 - i. PK: sports_id
 - ii. FK: athlete_id
- VIII. **Olympics_Events:** Todas las disciplinas que tuvieron competencia a lo largo de las Olimpiadas.
 - i. PK: evento_id
 - ii. FK: sports_id
- IX. **Olympics_Location:** Contiene todas las ciudades, con sus respectivos países, en los que fueron celebrados los Juegos Olímpicos.
 - i. PK: location_id



ii. FK: country_id

X. Query: Vista creada desde SQL para tener un listado de todos los resultados con medalla evitando la duplicidad cuando es deporte de equipo y sumando los datos de país y sexo de los atletas.

XI. Primera Medalla: Vista creada para poder obtener la edad en la que cada atleta ganó su primera medalla.

8. Listado de columnas por tablas:

Olympic_Results		
Campo	Tipo de Campo	Tipo de Clave
result_id	int	PK
event_id	int	FK
edition_id	int	FK
result_location	varchar	
result_participants	varchar	
result_format	varchar	
result_detail	varchar	
result_description	varchar	
start_date	datetime	
end_date	datetime	
time	timestamp	

Olympic_Athlete_Event_Results		
Campo	Tipo de Campo	Tipo de Clave
edition_id	int	CK
event_id	int	FK
result_id	int	FK
athlete_id	int	CK
pos	varchar	
medal	varchar	
isTeamSport	varchar	

Olympic_Athlete_Bio		
Campo	Tipo de Campo	Tipo de Clave
athlete_id	int	PK
name	varchar	
sex	varchar	
born	datetime	
height	decimal	
weight	int	
country_id	int	FK
description	varchar	
special_notes	varchar	



Olympic_Games_Medal_Tally		
Campo	Tipo de Campo	Tipo de Clave
edition_id	int	CK
country_id	int	CK
gold	int	
silver	int	
bronze	int	
total	int	

Olympics_Country		
Campo	Tipo de Campo	Tipo de Clave
country_id	int	PK
country	varchar	

Olympics_Game		
Campo	Tipo de Campo	Tipo de Clave
edition_id	int	PK
edition	varchar	
year	int	
location_id	int	FK
start_date	datetime	
end_date	datetime	
isHeld	varchar	
competition_start_date	datetime	
competition_end_date	datetime	

Olympics_Sports		
Campo	Tipo de Campo	Tipo de Clave
sports_id	int	PK
sports_name	varchar	
athlete_id	int	FK

Olympics_Events		
Campo	Tipo de Campo	Tipo de Clave
event_id	int	PK
event_name	varchar	
sports_id	int	FK

Olympics_location		
Campo	Tipo de Campo	Tipo de Clave
location_id	int	PK
city_name	varchar	



country_id	int	FK
------------	-----	----

Medidas

Atletas Hombres & Atletas Mujeres:

Estas medidas se utilizan para saber la cantidad de atletas totales, diferenciando hombres y mujeres:

Atletas Hombres =
calculate(DISTINCTCOUNT(Olympic_Athlete_Bio[athlete_id]),Olympic_Athlete_Bio[sex] = "Male")

Atletas Mujeres =
calculate(DISTINCTCOUNT(Olympic_Athlete_Bio[athlete_id]),Olympic_Athlete_Bio[sex] = "Female")

Medallistas Femeninas & Medallistas Masculinos

Estas medidas se utilizan para saber la cantidad de atletas totales que ganaron una medalla, diferenciando hombres y mujeres:

Medallistas Femeninas =
CALCULATE(COUNT(Query[sex]),Query[sex]="Female")

Medallistas Masculinos = CALCULATE(COUNT(Query[sex]),Query[sex]="Male")

AVG_Altura

Esta medida se utiliza para saber la altura promedio de los atletas:

AVG_Altura = IF('Tabla Medidas'[Total_Medallas] >=1, AVERAGE(Olympic_Athlete_Bio[height]))

Hace referencia
AVG_Año = AVERAGE(Calendarario[start_date])

AVG_Peso

Medida utilizada para saber el peso promedio de los atletas:

AVG_Peso = IF('Tabla Medidas'[Total_Medallas] >=1, AVERAGE(Olympic_Athlete_Bio[weight]))

Cant_Atlas

Cuenta la cantidad de atletas:



`Cant_Atletas = DISTINCTCOUNT(Olympic_Athlete_Bio[athlete_id])`

Cant_Events

Cuenta la cantidad de disciplinas que hay/hubo de los Juegos Olímpicos:

`Cant_Events = DISTINCTCOUNT(Olympic_Athlete_Event_Results[event_id])`

Cant_Ediciones

Cuenta la cantidad de ediciones que hubo de los Juegos Olímpicos:

`Cant_Ediciones = DISTINCTCOUNT(Olympics_Games[edition_id])`

Cant_Sports

Cuenta la cantidad de deportes en los que se compite (o competía) en los Juegos Olímpicos:

`Cant_Sports = DISTINCTCOUNT(Olympics_Sports[sports_id])`

Cant_Country

Cuenta la cantidad de países que compiten (o compitieron) en los Juegos Olímpicos:

`Cant_Country = DISTINCTCOUNT(Olympics_Country[country_id])`

pais_disciplina

Esta medida hace referencia a la cantidad de disciplinas por país:

`pais_disciplina = DISTINCTCOUNT(Olympics_Country[country]) * 'Tabla
Medidas'[Cant_Events]`

Summer

Esta medida se utiliza para saber la cantidad de Juegos Olímpicos que hubo en la temporada de verano:

`Summer = calculate(DISTINCTCOUNT('Olympics_Games'[edition]),
'Olympics_Games'[Season] = "Summer")`

Winter

Medida utilizada para saber la cantidad de Juegos Olímpicos que hubo en la temporada de invierno:

`Winter = calculate(DISTINCTCOUNT('Olympics_Games'[edition]),
'Olympics_Games'[Season] = "Winter")`

Equest



Esta medida se utiliza para saber la cantidad de Juegos Olímpicos que hubo en la temporada de “Equest”:

```
Equest = calculate(DISTINCTCOUNT('Olympics_Games'[edition]),  
'Olympics_Games'[Season] = "Equest")
```

Inter

Medida utilizada para saber la cantidad de Juegos Olímpicos “Inter”:

```
Inter = calculate(DISTINCTCOUNT('Olympics_Games'[edition]),  
'Olympics_Games'[Season] = " Inter")
```

TopPaíses

Esta medida se utiliza para hacer referencia al top 3 de países con mayor cantidad de medallas:

```
TopPaíses = RANKX(ALLSELECTED(Olympics_Country[country]), 'Tabla  
Medidas'[Total_Medallas])
```

Total_Medallas

Esta medida se utiliza para tener la cuenta de medallas totales:

```
Total_Medallas = sum (Olympic_Games_Medal_Tally[gold]) + sum  
(Olympic_Games_Medal_Tally[silver]) +  
sum(Olympic_Games_Medal_Tally[bronze])
```

Bronce

Esta medida se utiliza para calcular la cantidad de medallas de Bronce:

```
Bronce = CALCULATE(COUNT('Query'[medal]), 'Query'[medal]="Bronze")
```

Plata

Esta medida se utiliza para calcular la cantidad de medallas de Plata:

```
Plata = CALCULATE(COUNT('Query'[medal]), 'Query'[medal]="Silver")
```

Oro

Esta medida se utiliza para calcular la cantidad de medallas de Oro:

```
Oro = CALCULATE(COUNT('Query'[medal]), 'Query'[medal]="Gold")
```

Participaciones

Esta medida se utiliza para calcular la cantidad de veces que se participó en una disciplina:



Participaciones = `DISTINCTCOUNT(Olympic_Athlete_Event_Results[result_id])`

Tasa de Exito

Esta medida se utiliza para calcular la efectividad de un país en una disciplina en base a las participaciones que tuvo y la cantidad de medallas que ganó:

Tasa de Exito = `DIVIDE([Total Medallas],[Participaciones])`

Bronce por Atleta

Esta medida se utiliza para calcular cantidad de medallas de Bronce que ganó cada atleta:

Bronce por Atleta =
`CALCULATE(COUNT(Olympic_Athlete_Event_Results[athlete_id]),
Olympic_Athlete_Event_Results[medal]="bronze")`

Plata por Atleta

Esta medida se utiliza para calcular cantidad de medallas de Plata que ganó cada atleta:

Plata por Atleta =
`CALCULATE(COUNT(Olympic_Athlete_Event_Results[athlete_id]),
Olympic_Athlete_Event_Results[medal]="silver")`

Oro por Atleta

Esta medida se utiliza para calcular cantidad de medallas de Oro que ganó cada atleta:

Oro por Atleta =
`CALCULATE(COUNT(Olympic_Athlete_Event_Results[athlete_id]),
Olympic_Athlete_Event_Results[medal]="gold")`

Medallero por Atleta

Esta medida se utiliza para calcular cantidad de medallas que ganó cada atleta:

Medallero por Atleta =
`CALCULATE(COUNT(Olympic_Athlete_Event_Results[athlete_id]),
Olympic_Athlete_Event_Results[medal]="bronze") +
CALCULATE(COUNT(Olympic_Athlete_Event_Results[athlete_id]),
Olympic_Athlete_Event_Results[medal]="silver")+CALCULATE(COUNT(Olympic
_Athlete_Event_Results[athlete_id]),
Olympic_Athlete_Event_Results[medal]="gold")`



Columnas

Edad Primera Medalla

Esta columna se creó para calcular la edad en la que los atletas ganaron su primera medalla:

```
Edad Primera Medalla = Datediff('Primera  
Medalla'[Olympic_Athlete_Bio.born], 'Primera  
Medalla'[Olympic_Results.start_date], year)
```

Rango Etario

Esta columna se creó para poder clasificar dentro de un rango etario la edad en la que los atletas ganan su primera medalla:

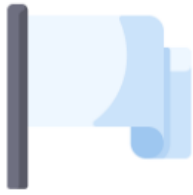
```
Rango Etario = IF ('Primera Medalla'[Edad Primera Medalla] <18, "Menor de  
18", IF('Primera Medalla'[Edad Primera Medalla] <=24, "18 a 24", IF('Primera  
Medalla'[Edad Primera Medalla] <=29, "25 a 29", IF('Primera Medalla'[Edad  
Primera Medalla] <=34, "30 a 34", "+35"))))
```

Visualizaciones

Contexto Olímpico

La primera página del reporte nos da un contexto de lo que estamos analizando. Conocemos la cantidad de Países, Atletas, Ediciones y Disciplinas. Además, obtenemos un primer vistazo del medallero olímpico. En esta página tenemos dos filtros para aplicar: El tipo de evento (verano, invierno, equest o inter) y el año de la edición.





PAÍSES

234

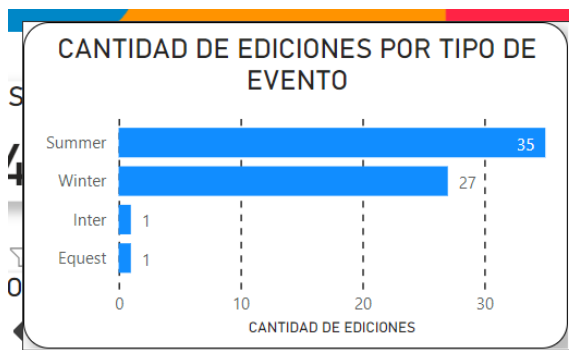
Cantidad de países que participan en los Juegos Olímpicos según el tipo de Evento seleccionado y el Año.



EVENTOS

64

Cantidad de ediciones que se realizaron de los Juegos Olímpicos según el tipo de Evento seleccionado y el Año.



Tooltip que aparece cuando nos situamos sobre Eventos y nos muestra la cantidad de ediciones que se realizaron de cada evento de los Juegos Olímpicos.



DISCIPLINAS

916

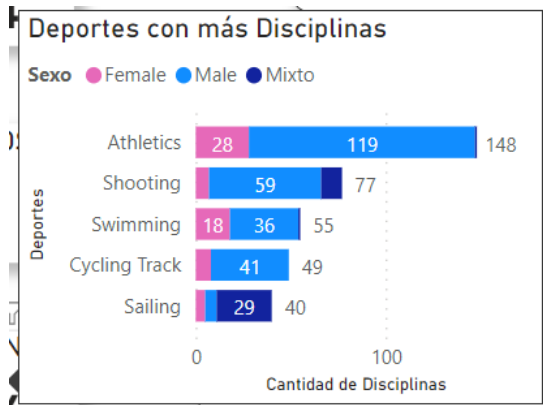
Cantidad de disciplinas que se juegan/jugaron en los Juegos Olímpicos según el tipo de Evento seleccionado y el Año.



DEPORTES

97

Cantidad de deportes que engloban las diferentes disciplinas en los Juegos Olímpicos según el tipo de Evento seleccionado y el Año.

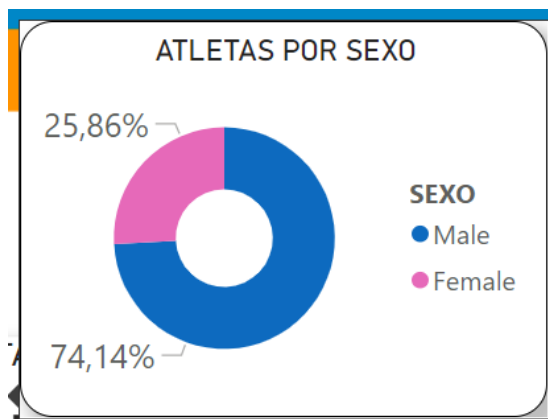


Tooltip que aparece cuando nos situamos sobre Deportes y nos muestra un Top 5 con los deportes que tienen más disciplinas, discriminado por género.

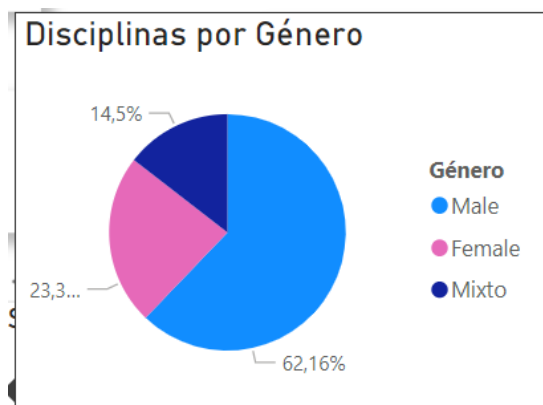


ATLETAS
155 mil

Cantidad de atletas que jugaron en los Juegos Olímpicos según el tipo de Evento seleccionado y el Año.



Tooltip que aparece cuando nos situamos sobre Atletas y nos muestra el porcentaje de atletas de cada género que participaron de los Juegos Olímpicos.



Tooltip que aparece cuando nos situamos sobre Disciplinas y nos muestra un Top gráfico de torta con la cantidad de disciplinas disponibles para competir para cada género.



TOP 3 MEDALLERO POR PAÍS				
country	Suma de gold	Suma de silver	Suma de bronze	Total_Medallas
United States	1179	961	839	2979
Soviet Union	473	376	355	1204
Germany	349	368	359	1076

Tabla con el top 3 de países con mayor cantidad de medallas según el tipo de Evento seleccionado y el Año.

TIPO DE EVENTO

Todas

Filtro que aplica a toda la página y permite seleccionar el tipo de Evento del que quiero obtener información.

AÑO

Todas

Filtro que aplica a toda la página y permite seleccionar el año del que quiero obtener información.

Disciplinas en el Podio

La segunda página del reporte nos permite conocer en qué disciplinas le ha ido mejor históricamente a cada país, podemos separar la evolución en 4 periodos diferentes. Además, vemos la evolución del medallero a lo largo de los años diferenciando entre hombres y mujeres y el tipo de medalla. En esta página tenemos cuatro filtros para aplicar: La disciplina (por si queremos info más detallada sobre una en particular), el país a analizar, el tipo de evento (verano, invierno, equest o inter) y el sexo.





Disciplinas Destacadas

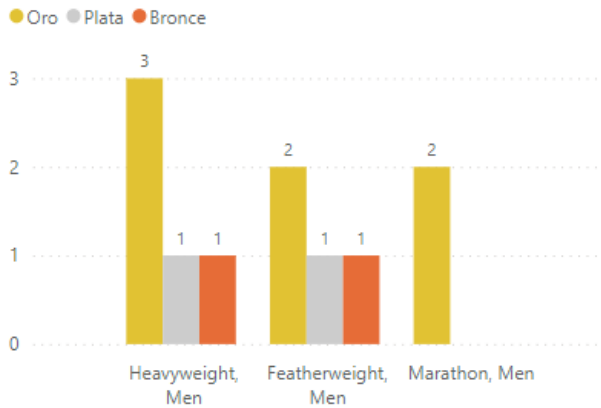


Gráfico de columnas ordenado descendentemente que muestra las disciplinas en las que el país seleccionado tiene más medallas.

Evolución del Medallero

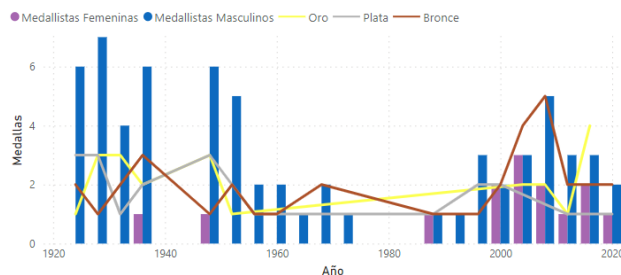


Gráfico de columnas y líneas que nos muestra la evolución del medallero según los filtros aplicados destacando las medallas ganadas según su tipo y la cantidad de medallistas masculinos y femeninas.

Seleccione el país a analizar:

Argentina

Filtro que aplica a toda la página y permite seleccionar el país del que quiero obtener información.

TIPO DE EVENTO

Todas

Filtro que aplica a toda la página y permite seleccionar el tipo de Evento del que quiero obtener información.

SEXO

Todas

Filtro que aplica a toda la página y permite seleccionar el sexo del que quiero obtener información.

DISCIPLINAS

Todas

Filtro que aplica a toda la página y permite seleccionar la disciplina de la que quiero obtener información.

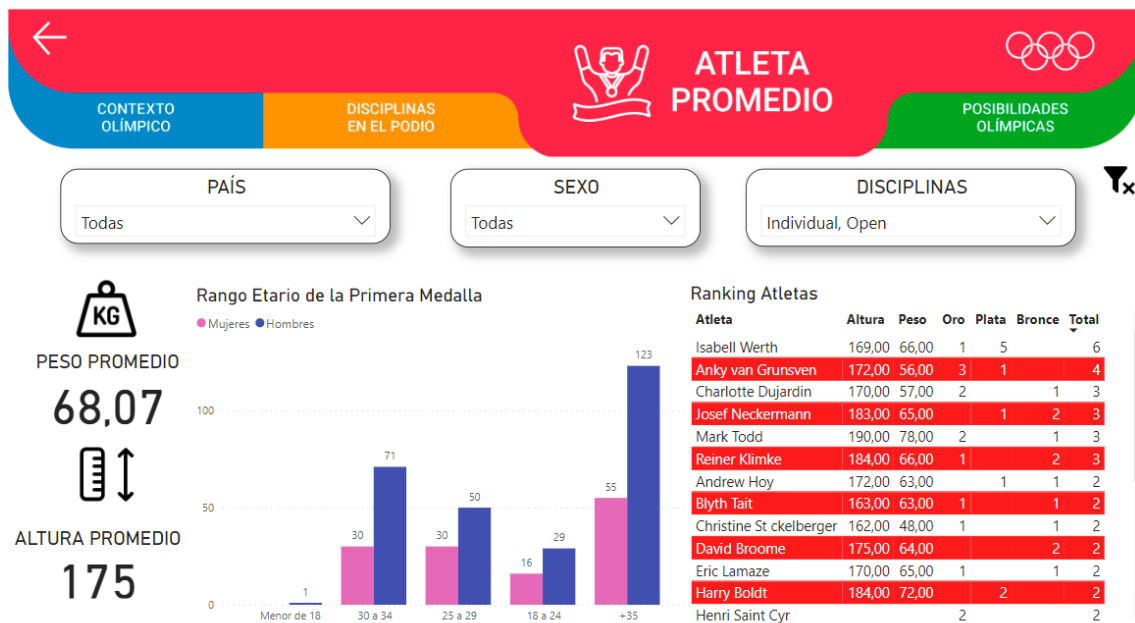


1896 - 1950

Botones que nos permiten obtener la información del periodo histórico seleccionado.

Atleta Promedio

Esta página del reporte nos permite conocer las características físicas promedio de los atletas ganadores de una medalla según cada disciplina. Además, vemos el histograma que nos demuestra en qué rango etario hay más atletas ganando su primera medalla según la disciplina y el sexo. Completando la experiencia, sumamos un ranking de los 10 atletas con más medallas según los filtros aplicados, en los que podemos ver además la altura y peso para poder compararla con el promedio.



PESO PROMEDIO
71,87

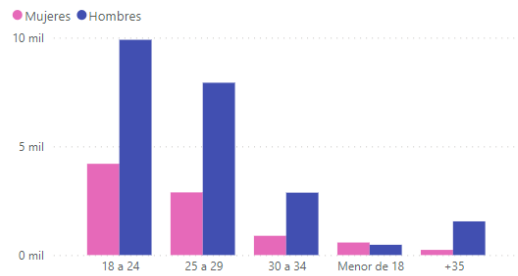
Tarjeta que muestra el peso promedio de los atletas ganadores de por lo menos una medalla según la disciplina y el sexo seleccionado.

ALTURA PROMEDIO
176

Tarjeta que muestra la altura promedio de los atletas ganadores de por lo menos una medalla según la disciplina y el sexo seleccionado.



Rango Etario de la Primera Medalla



Histograma que muestra la cantidad de atletas que ganan su primera medalla dentro de un rango etario según la disciplina.

Ranking Atletas

Atleta	Altura	Peso	Oro	Plata	Bronce	Total
Isabell Werth	169,00	66,00	1	5		6
Anky van Grunsvan	172,00	56,00	3	1		4
Charlotte Dujardin	170,00	57,00	2		1	3
Josef Neckermann	183,00	65,00		1	2	3
Mark Todd	190,00	78,00	2		1	3
Reiner Klimke	184,00	66,00	1		2	3
Andrew Hoy	172,00	63,00		1	1	2
Blyth Tait	163,00	63,00	1		1	2
Christine Stichelberger	162,00	48,00	1		1	2
David Broome	175,00	64,00			2	2
Eric Lamaze	170,00	65,00	1		1	2
Harry Boldt	184,00	72,00		2		2
Henri Saint Cyr			2			2

Tabla que muestra el ranking de atletas con mayor cantidad de medallas según los filtros aplicados. Además, nos muestra la información de altura y peso para poder compararlo con el promedio.

DISCIPLINAS

Individual, Open



Filtro que aplica a toda la página y permite seleccionar la disciplina de la que quiero obtener información.

SEXO

Todas



Filtro que aplica a toda la página y permite seleccionar el sexo del que quiero obtener información.

PAÍS

Todas



Filtro que aplica a toda la página y permite seleccionar el país del que quiero obtener información.



Posibilidades Olímpicas

La última página del reporte nos permite conocer la efectividad de cada país en la disciplina seleccionada. En base a la cantidad de participaciones y las medallas obtenidas calculamos la tasa de éxito lo que podría significar una mayor probabilidad de ganar nuevamente en el futuro. En esta página aplican cuatro filtros: El país, el tipo de evento, el sexo y la disciplina.



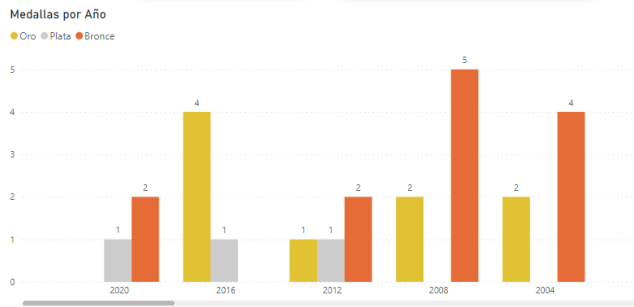


Gráfico de columnas que muestra la cantidad de medallas ganadas por año diferenciando entre Oro, Plata y Bronce.

DISCIPLINA

Todas

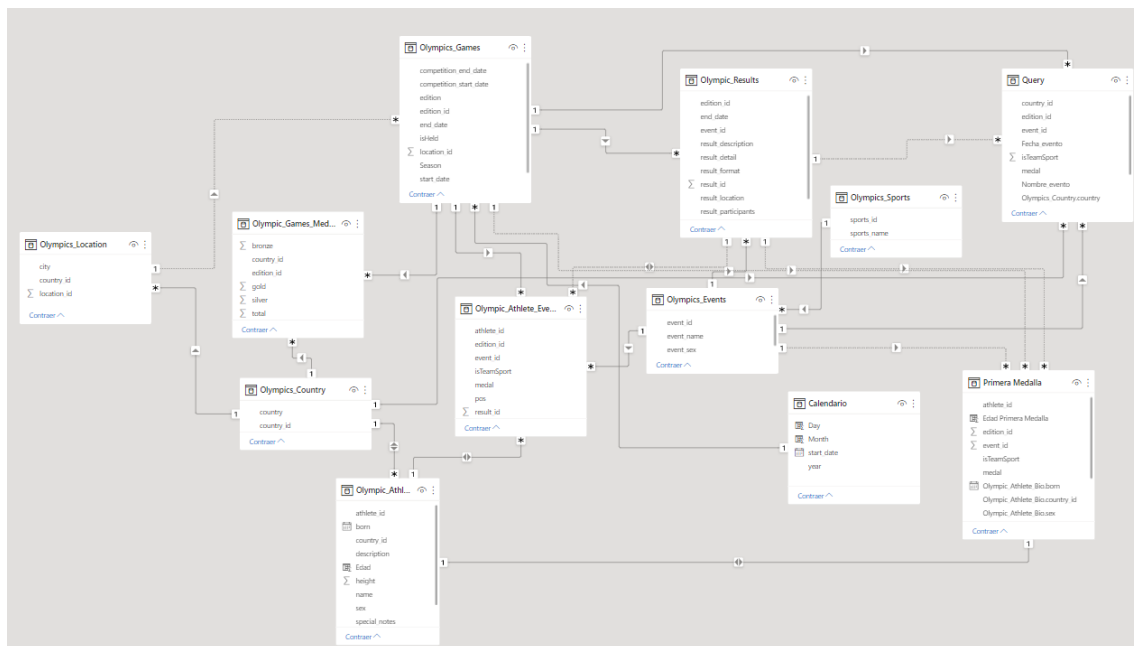
Filtro que aplica a toda la página y permite seleccionar la disciplina de la que quiero obtener información.

PAÍS

Argentina

Filtro que aplica a toda la página y permite seleccionar el país del que quiero obtener información.

Diagrama de Entidad de Relación en Power BI



Futuras Líneas de Investigación

Para poder complementar la investigación, sería ideal en un futuro poder contar con información de las inversiones económicas que realizaron las organizaciones en los diferentes años, para ver si un ROI determinado en base a la inversión y la cantidad de



medallas ganadas por edición. Además, obtener detalle de en qué deportes y géneros se hace la inversión más grande, se podría de esta manera crear una política de inclusión que busque la equidad de género ya que de por sí, la cantidad de disciplinas en las que pueden participar las atletas femeninas es mucho menor que la de los hombres.

Para aumentar las predicciones, también se podrían sumar los resultados de las participaciones previas en eventos clasificatorios para entender las oportunidades que tiene cada atleta, por lo que ya no se basaría solamente en probabilidades históricas basadas en nacionalidades o características promedio, sino que sumaríamos la performance a la ecuación.

Conclusión

Partimos de un dataset que nos brindaba la posibilidad de elegir entre diferentes objetivos. En nuestro caso decidimos encarar por ayudar en la búsqueda de respuestas a la hora de tener que decidir en qué destinar el presupuesto de deporte de ciertas organizaciones, pensando más que nada en el contexto tercermundista en el que quizás los fondos son escasos y hay que hacer una elección más estratégica. Ahora, si bien este dataset nos brinda muchísima información, creemos que para hacer un análisis realmente exhaustivo al respecto, aún nos está faltando información. Nuestro dashboard es una buena introducción al tema, brindando el contexto e información relevante sobre los atletas y las disciplinas, pero aún no termina de ser concluyente para la toma de una decisión.

Lo que proponemos en las futuras líneas de investigación creemos que podría ayudar a completar el propósito de nuestro trabajo.