

# Redefiniendo el Medallero Olímpico

por Diego Castro



**Data Analytics**

**Comisión 42270**

**25/04/23 al 18/07/23**



# Índice:

---

<b>1 - Introducción a la temática:</b>	<b>3</b>
1.1 - Hipótesis:	3
1.2 - Objetivo:	3
1.3 - Alcance:	4
1.4 - Usuario final y nivel de aplicación del análisis:	4
<b>2 - Dataset:</b>	<b>5</b>
2.1 - Dataset original:	5
2.2 - Transformaciones realizadas:	7
2.3 - Listado de tablas y columnas tras la normalización:	11
2.4 - Diagrama entidad-relación:	14
<b>3 - Power BI:</b>	<b>16</b>
3.1 - Transformaciones llevadas a cabo desde Power BI:	16
3.2 - Modelo Relacional en Power BI:	17
3.3 - Otras transformaciones de datos:	18
3.3.1 - Columnas calculadas:	18
3.3.1 - Medidas calculadas:	18
3.4 - Segmentaciones:	20
<b>4 - Visualizaciones:</b>	<b>21</b>
<b>5 - Conclusión:</b>	<b>24</b>
<b>6 - Futuras líneas:</b>	<b>25</b>

---

# 1 - Introducción a la temática:

---

Si uno revisa el medallero de los Juegos Olímpicos de los últimos 100 años rápidamente notará que está siempre encabezado por Estados Unidos, China y las potencias europeas. Pero ¿Por qué? ¿A qué se debe esta hegemonía?

Si me preguntan a mí, la forma de presentar el medallero siempre me resultó injusta, puesto que no es lo mismo obtener una medalla de oro para EE.UU. con sus 330 millones de habitantes que para un país de apenas 10 millones, ya que a mayor población, mayor probabilidad de que entre ellos nazca alguien con un talento (capacidad genética) mayor para destacarse en determinado deporte. Pero no es sólo que nazca, también hay que encontrar estos talentos y desarrollarlos. Entonces que los países que habitualmente encabezan el medallero sean aquellos con mayor PBI tampoco debe ser casualidad, porque justamente disponen de mayores recursos para destinar a formar deportistas. ¿No?

Entonces, si no es genético, tiene que estar en el ambiente: en este caso estamos hablando de mayores probabilidades de ocurrencia y disponibilidad de recursos que para nada son tenidos en cuenta a la hora de confeccionar el medallero tradicional. Las métricas que hoy en día usamos son, cuanto menos, poco ecuánimes; en favor (oh casualidad) de los que hacen dichas las reglas.

## 1.1 - Hipótesis:

Este trabajo buscará probar si efectivamente es cierto que los países que sistemáticamente dominan el medallero en cada cita olímpica se ubican en esos lugares no tanto por un verdadero mérito deportivo sino más bien por la situación de privilegio económico que atraviesan y cuantificar en qué medida inciden sus tamaños poblacionales y sus PBI per cápita en sus aparentes logros deportivos.

Adicionalmente, aprovecharé para comprobar si la localía también tiene su incidencia en el medallero a través de la comparación de los resultados obtenidos por cada país anfitrión oficiando de local versus el promedio de medallas obtenidas en condición de visitante.

## 1.2 - Objetivo:

El camino que planeo seguir para validar o rechazar la hipótesis consistirá en primera instancia en generar un “medallero ajustado” que compense (o intente compensar) el sesgo producido por el PBI per cápita y la población de cada país, para luego sí, a partir de este indicador, verificar si son las mismas naciones las que mantienen la hegemonía o si por el contrario surgen otras que dentro de su escala estén haciendo mejor las cosas en materia deportiva y que valga la pena considerar.

Adicionalmente, y aprovechando el gran volumen de datos con el que cuento, se intentará realizar un análisis descriptivo de otras tendencias que se pueden observar, como son el aumento a lo largo del tiempo de la cantidad de países participantes, atletas y diversidad de deportes, confirmar que efectivamente cada vez hay mayor paridad entre cantidad de hombres y mujeres participando de la cita olímpica y analizar si existe una disminución en las edades de los medallistas debido a un temprano profesionalismo o si por el contrario los mayores cuidados en la salud han permitido estirar la “vida útil” de los deportistas.

## 1.3 - Alcance:

A partir de la eliminación de los sesgos producidos por el PBI per cápita y el tamaño poblacional los ministerios de deportes de los distintos países tendrán la posibilidad de revisar su gestión a la luz de esta nueva información, al tiempo que les permitirá escoger mejor en qué países fijarse (e imitar) a la hora de seguir una política deportiva.

De igual forma, el Comité Olímpico Internacional (COI) podría utilizar esta misma información para premiar la labor que se hace en materia deportiva desde algunas naciones con recursos mucho más escasos.

Finalmente, los resultados presentados estarán a disposición de todos aquellos que defendemos la meritocracia y queremos ver una forma de presentación de resultados más justa.

## 1.4 - Usuario final y nivel de aplicación del análisis:

Si bien la información que se desprenda del análisis propuesto puede ser utilizada como herramienta por el nivel estratégico -en este caso un ministro de deportes y sus asesores- para tener una noción más objetiva de la situación real en la que se encuentra el país en materia deportiva y trazar metas alcanzables y razonables, el dashboard propuesto, debido a sus características, tendrá su mayor aplicación a nivel Táctico.



Un dashboard orientado a la gerencia debería contener la información aquí presente pero de manera más condensada y adicionar otros datos como podría ser el de presupuesto disponible. Por el contrario, el presente tablero tendrá una naturaleza interactiva, proporcionando a los usuarios la posibilidad de explorar los datos a través de filtros y segmentaciones, pero con mayor visualización de información que los paneles operativos por tratarse de un nivel un poco más alto.

El desarrollo de deportistas es un proceso largo que requiere varios ciclos olímpicos para empezar a dar frutos, por lo que, más allá de que cada 4 años se cuente con nueva información para actualizar los tableros, la misma debería ser utilizada por los líderes para hacer un seguimiento del rendimiento y en todo caso proponer recomendaciones para estrategias futuras a la luz de esta nueva información, pero no para definir un cambio de estrategia desde los altos mandos.

## 2 - Dataset:

### 2.1 - Dataset original:

El Dataset original se descargó de kaggle<sup>1</sup> y consistía en una tabla de 15 columnas y 271.116 registros con datos pertenecientes a todos los deportistas que participaron en algún juego olímpico entre Atenas 1896 y Río de Janeiro 2016.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
ID	Name	Sex	Age	Height [cm]	Weight [Kg]	Team	NOC	Games	Year	Season	City	Sport	Event	Medal
113674	Ivana panovi	F	26	176	65	Serbia	SRB	2016 Summer	2016	Summer	Rio de Janeiro	Athletics	Athletics Women's Long Jump	Bronze
113675	C. Edmond Spapen	M	NA	NA	NA	Belgium	BEL	1928 Summer	1928	Summer	Amsterdam	Wrestling	Wrestling Men's Bantamweight, Freestyle	Silver
113676	Anna Rita Sparaciani	F	21	163	55	Italy	ITA	1980 Summer	1980	Summer	Moskva	Fencing	Fencing Women's Foil, Individual	NA
113676	Anna Rita Sparaciani	F	21	163	55	Italy	ITA	1980 Summer	1980	Summer	Moskva	Fencing	Fencing Women's Foil, Team	NA
113677	Metka Sparavec (-Fain)	F	17	181	60	Slovenia	SLO	1996 Summer	1996	Summer	Atlanta	Swimming	Swimming Women's 50 metres Freestyle	NA
113677	Metka Sparavec (-Fain)	F	17	181	60	Slovenia	SLO	1996 Summer	1996	Summer	Atlanta	Swimming	Swimming Women's 100 metres Freestyle	NA
113678	Ren Sparenberg	M	17	NA	NA	Netherlands	NED	1936 Summer	1936	Summer	Berlin	Hockey	Hockey Men's Hockey	Bronze
113679	Katharina Sporer	F	15	163	52	Italy	ITA	2006 Winter	2006	Winter	Torino	Ice Hockey	Ice Hockey Women's Ice Hockey	NA
113680	John Arthur "Jack" Spargo	M	21	NA	NA	United States	USA	1952 Summer	1952	Summer	Helsinki	Water Polo	Water Polo Men's Water Polo	NA
113681	Arthur Percy Spark	M	29	NA	NA	Great Britain	GBR	1924 Summer	1924	Summer	Paris	Athletics	Athletics Men's Pentathlon	NA

Las columnas correspondían a:

**ID:** Donde se le asignaba un ID distinto a cada deportista, con el detalle de que los deportistas que participaron en varios juegos olímpicos tenían un único ID, por lo que en esta columna había muchos ID repetidos.

**Name:** Nombre y apellido de cada deportista, que en el caso de aquellos que participaron en distintos juegos olímpicos o incluso en varias disciplinas en el mismo juego se encontraban, al igual que el caso anterior, repetidos múltiples veces.

**Sex:** Sexo del atleta: F para female (femenino) y M para male (masculino).

**Age:** Edad de cada deportista en el momento que participó, pero a diferencia de la columna Name y Sex que se repiten invariables en todas las participaciones del mismo atleta, este dato sí cambia de juego a juego, porque los atletas (como todos) también envejecen.

**Height:** La altura de cada deportista en cms, también sujeta a variaciones entre participaciones.

**Weight:** El peso de cada deportista en kgs, que al igual que la edad y la altura podía variar de juego a juego.

**Team:** Es el equipo al que representa cada deportista, en general naciones (aunque no todas existentes como tales en la actualidad). Posteriormente, durante el análisis detallado de los datos, encontré algunas irregularidades en esta columna, como por ejemplo el nombre de la embarcación en algunas competencias de remo o el nombre del caballo en competencias ecuestres en lugar del país por el que participaban, y tuvo que ser corregido.

**NOC:** Es una sigla de 3 letras que identifica de manera unívoca a cada nación que participó en un juego olímpico ante el Comité Olímpico Internacional (COI) por lo que permite saber a qué país representó cada deportista más allá de que la columna Team incluya nombre de embarcaciones o de caballos.

<sup>1</sup> <https://www.kaggle.com/datasets/mysarahmadbhat/120-years-of-olympic-history>

Games: Identifica en qué juego participó cada participante: tanto el año como la identificación de si fue un juego olímpico de verano o de invierno, repetido ininidad de veces, uno para cada registro.

Year: El año de la cita olímpica (información ya incluida en la columna Games)

Season: La temporada en la que se desarrollaron los juegos, con únicamente 2 valores posibles: Summer y Winter, pero de todas formas repetidos a lo largo de los más de 270 mil registros (además de ser redundante con la información presente en la columna Games)

City: La ciudad que resultó anfitriona de los juegos, repetida también múltiples veces en los registros de todos los deportistas que participaron de la misma cita olímpica.

Sport: Deporte en el que participó el atleta (el cual engloba a su vez múltiples disciplinas)

Event: Detalla en qué disciplina específica participó el atleta -dentro del deporte-

Medal: Es el resultado obtenido por el deportista, pudiendo tomar 4 valores: Gold, Silver, Bronze ó NA en caso de no haber obtenido una medalla.

Adicionalmente, en pos de generar este nuevo medallero ajustado pensaba utilizar un coeficiente “c” que tuviera en cuenta el tamaño poblacional de cada país y su situación económica particular, para lo cual iba a trabajar con dos DB adicionales: *PBI* y *Población* de cada país que se vincularían a la DB de trabajo a través de la PK id\_País.

El coeficiente C se expresaría de la siguiente forma:

$$C = \frac{\text{Medallas obtenidas}}{\text{PBI per capita} \times \text{Poblacion}}$$

pero dado que el PBI per cápita no es otra cosa que el PBI por habitante, tenemos:

$$C = \frac{\text{Medallas obtenidas}}{\frac{\text{PBI}}{\text{Poblacion}} \times \text{Poblacion}}$$

por lo que nos queda:

$$C = \frac{\text{Medallas obtenidas}}{\frac{\text{PBI}}{\cancel{\text{Poblacion}}} \times \cancel{\text{Poblacion}}} = \frac{\text{Medallas obtenidas}}{\text{PBI}}$$

Llevando a que únicamente sea preciso contar con información del PBI para generar el coeficiente “c” y por ende el nuevo medallero ajustado.

## 2.2 - Transformaciones realizadas:

Trabajar con esta tabla en Excel era bastante engorroso porque cualquier cambio requería muchísimo tiempo de procesamiento de datos y de espera por lo que empecé por separar el dataset original en tablas más pequeñas, donde pudiera ir deshaciéndome de los datos redundantes.

Primero que nada y bien fácil eliminé la columna Games, puesto que no era más que la concatenación de la información presente en las columnas Year + Season (totalmente redundante) y me resultaba mucho más útil contar con estos datos en columnas separadas, por lo que opté por mantener estas últimas.

Luego decidí pasar toda la información referente a cada atleta en particular (esto es: Nombre, Sexo, Edad, Altura, Peso) a una nueva *tabla Deportistas* vinculada a la original por el *ID\_Deportista* que venía dado. Hacer esto me permitió eliminar de esta última tabla todos los registros duplicados correspondientes a los deportistas que participaron de más de una disciplina en un mismo juego y que por lo tanto en la tabla original contaban con 2 registros casi idénticos, diferenciados únicamente en la columna Event.

Por ejemplo, para el caso del atleta con ID 17 me permitió reemplazar todas las líneas coloreadas por una única ya que todas contenían la misma información – y esto pasaba en muchísimos casos-

A	B	C	D	E	F	G	H	I	J	K	L	M	N
ID	Name	Sex	Age	Height	Weight	Team	NOC	Year	Season	City	Sport	Event	Medal
17	Paavo Johannes Aaltonen	M	28	175	64	Finland	FIN	1948	Summer	London	Gymnastics	Gymnastics Men's Individual All-Around	Gold
17	Paavo Johannes Aaltonen	M	28	175	64	Finland	FIN	1948	Summer	London	Gymnastics	Gymnastics Men's Team All-Around	Bronze
17	Paavo Johannes Aaltonen	M	28	175	64	Finland	FIN	1948	Summer	London	Gymnastics	Gymnastics Men's Floor Exercise	NA
17	Paavo Johannes Aaltonen	M	28	175	64	Finland	FIN	1948	Summer	London	Gymnastics	Gymnastics Men's Horse Vault	Gold
17	Paavo Johannes Aaltonen	M	28	175	64	Finland	FIN	1948	Summer	London	Gymnastics	Gymnastics Men's Parallel Bars	NA
17	Paavo Johannes Aaltonen	M	28	175	64	Finland	FIN	1948	Summer	London	Gymnastics	Gymnastics Men's Horizontal Bar	NA
17	Paavo Johannes Aaltonen	M	28	175	64	Finland	FIN	1948	Summer	London	Gymnastics	Gymnastics Men's Rings	NA
17	Paavo Johannes Aaltonen	M	28	175	64	Finland	FIN	1948	Summer	London	Gymnastics	Gymnastics Men's Pommel Horse	Gold

Hasta ahí la cosa venía sencilla, pero se complejizó cuando caí en la cuenta de que los deportistas que lograron una medalla en deportes grupales, como fútbol, básquet y vóley figuraban todos como ganadores de la medalla correspondiente y esto no me servía a la hora de contabilizar cuántas medallas obtuvo cada país, puesto que en la práctica se consideran como una sola. Por lo tanto hubo que arremangarse:

Lo primero que tuve que hacer es modificar la concepción de mi tabla original que ya no iba a listar a deportistas en forma individual sino a participantes de una competencia específica “x” que bien podían ser deportistas individuales, dúos, tríos o equipos. Fueron cada uno de estos participantes los que obtuvieron una medalla ÚNICA, independientemente de la cantidad que fuesen.

Por ejemplo: La selección argentina de básquet obtuvo en el 2004 una medalla de oro y en el dataset original la información se encontraba dispuesta de la siguiente forma:

A	B	C	D	E	F	G	H	I	J	K	L	M	N
ID	Name	Sex	Age	Height	Weight	Team	NOC	Year	Season	City	Sport	Event	Medal
27256	Carlos Francisco Delfino	M	21	196	95	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
34816	Gabriel Diego Fernandez	M	27	204	111	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
40451	Emanuel David "Manu" Ginobili Maccari	M	27	198	91	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
44181	Leonardo Martin Gutierrez	M	26	200	105	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
48185	Walter Hermann	M	25	203	105	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
81471	Alejandro Ariel Montecchia	M	32	182	83	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
86706	Andrés Marcelo Nocioni Roux	M	24	201	102	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
87768	Fabrizio Ral Jess Oberto	M	31	208	115	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
105234	Juan Ignacio "Pope" Sanchez Brown	M	27	193	90	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
107958	Luis Alberto Scialoja Balboa	M	24	206	108	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
107961	Hugo Ariel Sconochini	M	33	193	96	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold
131362	Ruben Oscar Wolkowyski	M	30	208	120	Argentina	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold

En la nueva concepción de la tabla ya no figuran cada uno de los deportistas con su ID correspondiente y su medalla (que dificultaría su contabilización) sino que figura el equipo como una unidad con un único ID y una única medalla:

ID	NOC	Year	Season	City	Sport	Event	Medal
P123916	ARG	2004	Summer	Athina	Basketball	Basketball Men's Basketball	Gold

La idea era luego vincular la tabla Deportistas (donde figuran los datos de nombre, edad, nacionalidad, etc) a esta *tabla Participantes* (donde figura la información de en qué deporte y disciplina participó y qué resultado obtuvo) pero me encontré con un nuevo problema: la relación N:M que se da entre estas tablas:

Por un lado tenemos la situación habitual de que un mismo deportista puede competir en varios eventos (ya sean dentro del mismo juego olímpico o a lo largo de distintas ediciones) por lo que el mismo deportista puede aparecer n-veces en la *tabla Participantes*. Pero por el otro lado, los deportistas que participan en pruebas en equipo tienen un único *ID\_Participante* en la *tabla Participante* que referencia a múltiples registros en la *tabla Deportistas*.

Para salvar esto surge la *tabla Unión*, reemplazando la relación N:M por dos relaciones 1:N, donde las PK *ID\_Deportista* e *ID\_Participante* pasar a ser FK de esta *tabla Unión*, al tiempo que como PK se escogió concatenar ambas columnas para crear el *ID\_P&D*, que es necesariamente única porque un mismo deportista (o equipo) no puede participar 2 veces del mismo evento durante los mismos juegos olímpicos.

A modo de ejemplo podemos ver el caso de Michel Phelps, que en 2004 participó de múltiples competencias de natación de forma individual, para las cuales tiene distintos *ID\_Participante* -uno para cada evento que disputó-, pero un único *ID\_Deportista* que referencia a sus datos particulares: nombre, edad, peso, país. El caso es que también compitió en forma grupal en las postas donde un único *ID\_Participante* se asigna a diferentes *ID\_Deportista* (que son sus compañeros de equipo).

A	B	C
IDP&D (PK)	ID_Participante (FK)	ID_Deportista (FK)
P121667-94406	P121667	94406-2004
P121858-94406	P121858	94406-2004
P121916-94406	P121916	94406-2004
P121966-94406	P121966	94406-2004
P122048-94406	P122048	94406-2004
P124636-94406	P124636	94406-2004
P124636-23868	P124636	23868-2004
P124636-30882	P124636	30882-2004
P124636-45039	P124636	45039-2004
P124636-69288	P124636	69288-2004
P124636-128126	P124636	128126-2004
P124636-131565	P124636	131565-2004

Si quisiéramos saber de qué deportistas se trata (nombre, edad, altura, peso) se puede acceder a esa información usando el *ID\_Deportista*, mientras que para conocer en qué deporte participó y qué resultado obtuvo hay que recurrir a la *tabla Participante* a través del *ID\_Participante*.



Superada esta etapa, el resto de los cambios que siguieron fueron relativamente sencillos y obedecieron a evitar las múltiples repeticiones de los mismos datos en diferentes registros.

Por ejemplo, en la *tabla Participante* figuraba repetida una gran cantidad de veces el nombre del deporte y la disciplina, no solo entre los participantes de esa competencia en esa cita olímpica, sino en distintos juegos (puesto que en su mayoría el evento vuelve a disputarse en cada nueva edición), por lo que se creó una *Tabla Competencia* que permitiera aislar las columnas Sport, Event, Year, Season y City.

Para intentar clarificar lo que se hizo, tomemos el ejemplo de los velocistas jamaquinos que participaron en los JJOO de 2008 y 2012, dónde fácilmente se aprecia que las filas resaltadas tienen información repetida múltiples veces:

A	B	C	D	E	F	G	H
ID_Participante (PK)	Year	Season	City	Sport	Event	Medal	ID_Pais (FK)
P125335	2008	Summer	Beijing	Athletics	Athletics Men's 200 metres	NA	JAM
P142318	2012	Summer	London	Athletics	Athletics Men's 4 x 100 metres Relay	Gold	JAM
P134563	2012	Summer	London	Athletics	Athletics Men's 100 metres	Silver	JAM
P134736	2012	Summer	London	Athletics	Athletics Men's 200 metres	Silver	JAM
P142318	2012	Summer	London	Athletics	Athletics Men's 4 x 100 metres Relay	Gold	JAM
P125177	2008	Summer	Beijing	Athletics	Athletics Men's 100 metres	Gold	JAM
P125336	2008	Summer	Beijing	Athletics	Athletics Men's 200 metres	Gold	JAM
P133246	2008	Summer	Beijing	Athletics	Athletics Men's 4 x 100 metres Relay	NA	JAM
P134562	2012	Summer	London	Athletics	Athletics Men's 100 metres	Gold	JAM
P134734	2012	Summer	London	Athletics	Athletics Men's 200 metres	Gold	JAM
P142318	2012	Summer	London	Athletics	Athletics Men's 4 x 100 metres Relay	Gold	JAM
P133246	2008	Summer	Beijing	Athletics	Athletics Men's 4 x 100 metres Relay	NA	JAM
P142318	2012	Summer	London	Athletics	Athletics Men's 4 x 100 metres Relay	Gold	JAM
P125178	2008	Summer	Beijing	Athletics	Athletics Men's 100 metres	NA	JAM
P133246	2008	Summer	Beijing	Athletics	Athletics Men's 4 x 100 metres Relay	NA	JAM
P142318	2012	Summer	London	Athletics	Athletics Men's 4 x 100 metres Relay	Gold	JAM
P125179	2008	Summer	Beijing	Athletics	Athletics Men's 100 metres	NA	JAM
P133246	2008	Summer	Beijing	Athletics	Athletics Men's 4 x 100 metres Relay	NA	JAM
P134561	2012	Summer	London	Athletics	Athletics Men's 100 metres	NA	JAM
P133246	2008	Summer	Beijing	Athletics	Athletics Men's 4 x 100 metres Relay	NA	JAM
P134735	2012	Summer	London	Athletics	Athletics Men's 200 metres	Bronze	JAM
P125337	2008	Summer	Beijing	Athletics	Athletics Men's 200 metres	NA	JAM

Básicamente se creó un *ID\_Competencia* que funciona como FK en la *tabla Participante* que conduce a una nueva *tabla Competencia* y que permite acceder a la información de temporada y Ciudad anfitriona a través del *ID\_anio* y al evento específico a través del *ID\_Evento*, ambos FK en la *tabla Competencia*, pero PK de sus tablas respectivas.

A través de esta normalización, fue posible eliminar todos los registros duplicados, pasando de lo anterior a esto:

ID_Participante (PK)	ID_Competencia (FK)	ID_Resultado (FK)	ID_Pais (FK)
P125177	E61-2008	3	JAM
P125178	E61-2008	0	JAM
P125179	E61-2008	0	JAM
P134563	E61-2012	2	JAM
P134562	E61-2012	3	JAM
P134561	E61-2012	0	JAM
P125335	E66-2008	0	JAM
P125336	E66-2008	3	JAM
P125337	E66-2008	0	JAM
P134736	E66-2012	2	JAM
P134734	E66-2012	3	JAM
P134735	E66-2012	1	JAM
P133246	E75-2008	0	JAM
P142318	E75-2012	3	JAM

Que a partir de *ID\_Competicencia* se vincula a la *tabla Competencia* con la siguiente información:

<b>Id_Competicencia (PK)</b>	<b>Id_Evento (FK)</b>	<b>Id_Año (FK)</b>
E61-2008	E61	2008
E61-2008	E61	2008
E61-2008	E61	2008
E61-2012	E61	2012
E61-2012	E61	2012
E61-2012	E61	2012
E66-2008	E66	2008
E66-2008	E66	2008
E66-2008	E66	2008
E66-2012	E66	2012
E66-2012	E66	2012
E66-2012	E66	2012
E75-2008	E75	2008
E75-2012	E75	2012

Y que a su vez conduce a la *tabla Evento*:

<b>Id_Evento (PK)</b>	<b>Evento</b>	<b>Sport</b>
E61	Athletics Men's 100 metres	Athletics
E66	Athletics Men's 200 metres	Athletics
E75	Athletics Men's 4 x 100 metres Relay	Athletics

Y a la *tabla Año* con información de cada juego<sup>2</sup>:

<b>Id_Año (PK)</b>	<b>Season</b>	<b>Ciudad</b>	<b>País</b>	<b>Continente</b>
2008	Summer	Beijing	China	Asia
2012	Summer	London	Great Britain	Europe

Solo en este ejemplo se logra reducir la tabla de 22x8 (176 celdas) a una tabla de 14x4, una de 14x3, una de 3x3 y otra de 2x5, que totalizan 117 celdas (33% de reducción), donde no sólo no se perdió información sino que por el contrario se adicionó, incorporando ahora datos sobre el país y el continente donde se desarrollaron los juegos, cuando antes solamente se informaba la ciudad.

Siguiendo esta misma idea también se llevó adelante un proceso similar para generar la *tabla Deportes* desde la *tabla Evento*, la *tabla Temporada* desde la *tabla Año* y la *tabla Nombre* desde la *tabla Deportistas*, pero considero que no vale la pena entrar en detalles.

En total, a partir del dataset inicial, se originaron 11 tablas nuevas que permitieron eliminar datos redundantes sin perder información. Las mismas se detallan a continuación:

<sup>2</sup> Nota: se muestran sólo las filas implicadas en el ejemplo propuesto, pero las tablas cuentan con registros para cada participante, para cada evento y para cada año.

## 2.3 - Listado de tablas y columnas tras la normalización:

Tabla Participante: En ella se le asigna una identificación única (ID\_Participante -PK-) a cada deportista o equipo que haya participado en alguna competencia. Esto significa que si un mismo deportista compitió en varios eventos diferentes (ya sea en un mismo Juego Olímpico o en distintos) tendrá un número diferente para cada una de sus participaciones. Adicionalmente, para el caso de los equipos, no hay un registro para cada uno de los deportistas en forma individual sino que se considera al equipo como una unidad y por lo tanto con un único ID. Se decidió realizar esta modificación respecto a la forma original en la que venían presentados los datos porque si bien cada uno de los integrantes del grupo recibe una medalla, a la hora de contabilizarlas se considera como solo una.

Esta tabla cuenta también con otras dos columnas:

- Una con el ID\_Resultado (FK), que permite rastrear el resultado obtenido en la competencia (medalla de oro, plata, bronce o sin medalla)
- Y otra con el ID\_Competencia (FK) que nos conduce a otra tabla de donde se puede obtener la información referente a de qué competencia se trata.

PARTICIPANTES			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Participante	varchar	PK	NOT NULL
Id_Competencia	varchar	FK	NOT NULL
Id_Resultado	int	FK	NOT NULL

Tabla Resultado: Es una tabla sencilla de apenas 4 registros donde para cada ID\_Resultado (PK) se asigna el color de la medalla correspondiente (Gold, Silver, Bronze) o NA en caso de no haber obtenido una. Su objetivo es únicamente evitar la repetición de alguna de estas palabras en todas las líneas de la *Tabla Participante*.

RESULTADOS			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Resultado	int	PK	NOT NULL
Medalla	text		NOT NULL

Tabla Competencia: Se compone de 3 columnas:

- El ID\_Competencia (PK), ya mencionado cuando se habló de la Tabla Participante, que busca identificar cada una de las competencias, solo que aquí funciona como PK, registrándolas de manera única y no aceptando repeticiones.
- El ID\_Año (FK), para poder trackear a que juego olímpico corresponde la competencia en cuestión.
- El ID\_Evento (FK), permite rastrear de qué evento en particular se trata.

La idea es que el ID\_Competencia identifique de manera única a cada una de los eventos realizados en los juegos, aunque se trate del mismo deporte en distintas ediciones. A partir del mismo se debe poder determinar de qué disciplina se trata y en qué año se llevó a cabo.

COMPETENCIA			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Competencia	varchar	PK	NOT NULL
Id_Evento	varchar	FK	NOT NULL
Id_Año	int	FK	NOT NULL

**Tabla Evento:** Enlista la totalidad de los deportes y modalidades practicadas en algún juego olímpico (aunque no haya sido en todos) y les asigna un valor único ID\_Evento (PK), que funcionaba de FK en la *tabla Competencia*. Tiene además otras 2 columnas: Una con el nombre del evento en cuestión y otra con un ID\_Deporte que permite identificar a qué deporte corresponde (ya que hay deportes que tienen varias modalidades distintas).

EVENTO			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Evento	varchar	PK	NOT NULL
Id_Deporte	int	FK	NOT NULL
Evento	varchar		NOT NULL

**Tabla Deporte:** Cuenta únicamente con 2 columnas. El ID\_Deporte (PK) que lo relaciona con la *tabla Evento*, y una con el nombre del deporte al que pertenece la disciplina.

DEPORTE			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Deporte	int	PK	NOT NULL
Deporte	varchar		NOT NULL

**Tabla Juegos:** A cada ID\_Año mencionado en la *tabla Competencia* como FK le corresponde la información de la ciudad donde se llevaron a cabo los Juegos, como así también un ID\_Temporada (FK) que distingue los juegos de verano de los de invierno. En este caso ID\_Año actúa como PK.

JUEGOS			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Año	int	PK	NOT NULL
Id_Temporada	int	FK	NOT NULL
Ciudad	varchar		NOT NULL

**Tabla Temporada:** Es apenas una tabla de 2x2, que tiene por columnas el ID\_Temporada (PK) y la estación en la que se llevaron a cabo los juegos: 1: Verano y 2: Invierno y su finalidad es evitar la sucesiva repetición de estos datos en la tabla anterior.

TEMPORADA			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Temporada	int	PK	NOT NULL
Estacion	text		NOT NULL

Tabla Unión P/D D/P: Esta tabla es la clave, y surge como necesidad de evitar la relación N:M que se daba entre participantes y deportistas. Como ya mencioné en la sección anterior, es habitual que un mismo atleta compita en varios eventos (ya sean dentro del mismo juego olímpico o a lo largo de distintas ediciones) por lo que el mismo deportista puede aparecer n-veces en la *tabla Competencia*. Pero también ocurre que hay deportistas que participan en pruebas en equipo, por lo que se da la situación inversa de que un único ID\_Competencia hace referencia a múltiples registros en la *tabla Deportistas*. Así es como la tabla unión viene a salvar esta relación N:M, reemplazándola por 2 relaciones 1:N. Cuenta básicamente de dos columnas principales: el ID\_Participante (FK) y el ID\_Deportista (FK), cada uno proveniente de la tabla homónima -donde ofician de PK-. Por su parte, como PK de esta tabla se escogió concatenar ambas columnas para crear el ID\_P&D (PK), que como también se mencionó será necesariamente único porque un mismo deportista no puede participar 2 veces del mismo evento durante los mismos juegos olímpicos, imposibilitando que haya más de un Id\_Participante con el mismo Id\_Deportista.

UNION			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_P&D	varchar	PK	NOT NULL
Id_Participante	varchar	FK	NOT NULL
Id_Deportista	varchar	FK	NOT NULL

Tabla Deportistas: Concentra bajo el ID\_Deportista (PK) toda la información personal referida a los atletas, como la edad, el sexo, la altura, el peso y el país al que representan – a través del ID\_Pais (FK) -.

Debe aclararse que un deportista que se presenta a 2 juegos olímpicos distintos tendrá 2 ID\_Deportista diferentes, puesto que la información de edad y talla habrá cambiado en los 4 años que separan las citas.

En principio, la única información personal que no está sujeta a este cambio es el nombre, por lo que para evitar repeticiones sucesivas, la tabla deportistas no cuenta con el nombre del atleta sino con un ID\_Nombre (FK) igual para todos los registros que hacen referencia a la misma persona, (más allá de que tenga varios ID\_Deportista por haber participado en diferentes ediciones de los JJ.OO.) en una relación 1:N.

Puede verse que la *Tabla Nombres* tiene 43.903 registros menos que la *Tabla Deportistas*, es decir que con esta sencilla modificación logramos deshacernos de 43.903 campos que tendría información con nombres innecesariamente duplicados

DEPORTISTAS			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Deportista	varchar	PK	NOT NULL
Id_Nombre	varchar	FK	NOT NULL
Id_Pais	text	FK	NOT NULL
Age	int		
Sex	text		NOT NULL
Height	int		
Weight	int		

**Tabla Nombres:** Se compone únicamente de 2 columnas: una con el ID\_Nombre (PK), y otra con el nombre del deportista propiamente dicho y este ID\_Nombre asociado al atleta será siempre el mismo independientemente de en cuántos Juegos se presente.

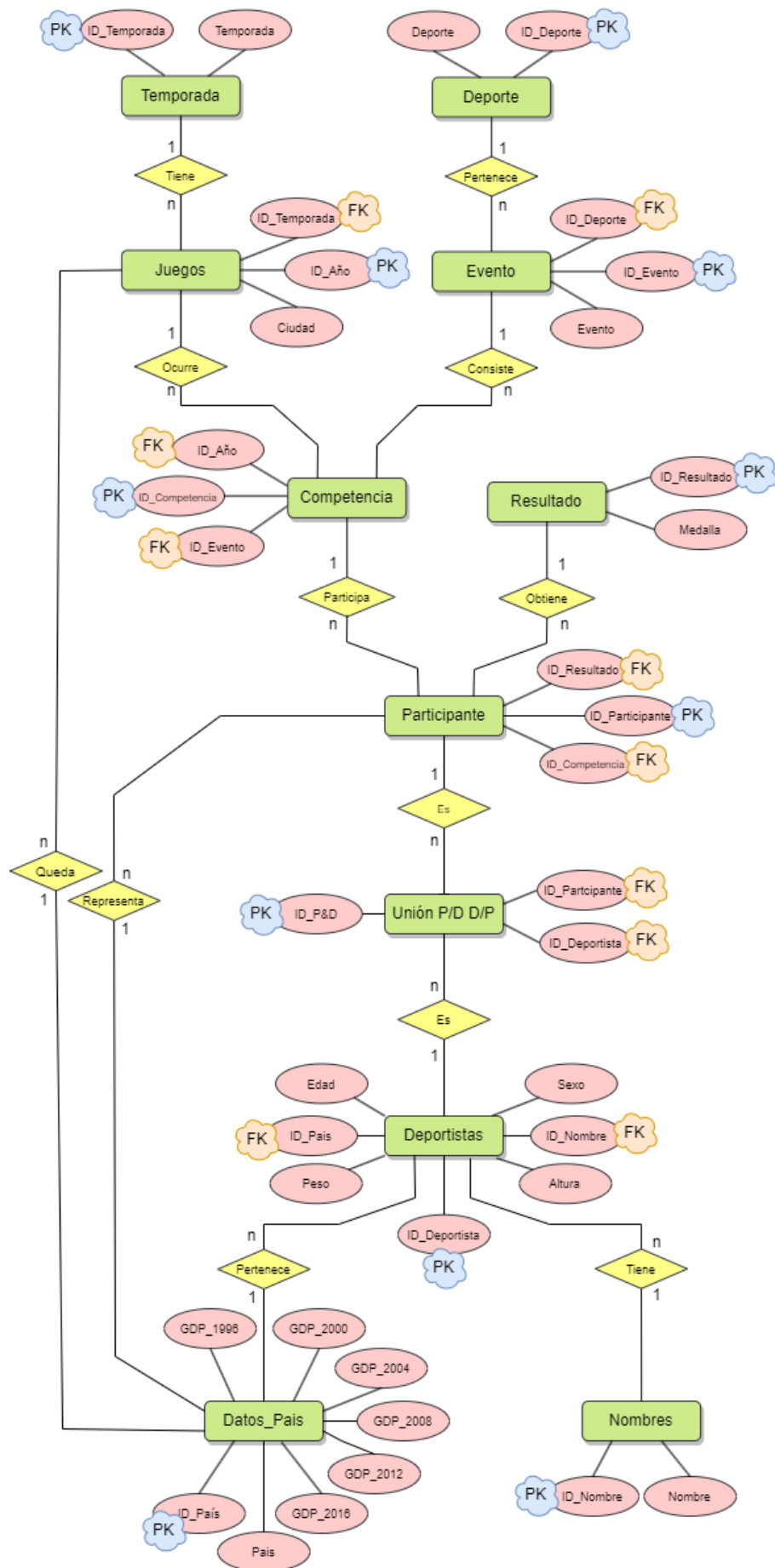
NOMBRES			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Nombre	varchar	PK	NOT NULL
Nombre	varchar	IK	NOT NULL

**Tabla País:** Como ID\_Pais se utilizó el NOC, que es la sigla de 3 letras que identifica a cada país en el Comité Olímpico Internacional (COI) y ya venía dado en el dataset original, pero adicionalmente esta tabla cuenta también con el nombre del país y con los datos del PBI (GDP) en años olímpicos, que serán necesarios para el análisis posterior.

PAISES			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Id_Pais	text(3)	PK	NOT NULL
País	varchar	IK	NOT NULL
GDP1996	bigint		
GDP2000	bigint		
GDP2004	bigint		
GDP2008	bigint		
GDP2012	bigint		
GDP2016	bigint		

## 2.4 - Diagrama entidad-relación:

Tras realizar estas modificaciones y aquella propuesta por la tutora de vincular las tablas *Juegos* y *Participantes* a la *Tabla País*, el diagrama ER quedó como se muestra a continuación:



## 3 - Power BI:

### 3.1 - Transformaciones llevadas a cabo desde Power BI:

Los cambios más grandes que tuve que realizar en Power BI respecto al modelo ER anterior fueron la inclusión de tablas con datos que necesitaba a la hora de generar los gráficos y que me resultó mucho más sencillo trabajar en Excel y luego incorporar al modelo.

La *tabla Juegos\_Datos* tiene la información agrupada de cuantos hombres, cuantas mujeres y cuantos países participaron en cada uno de los juegos olímpicos y cuantos deportes distintos se disputaron en cada cita y se vincula en una relación 1:1 con la *tabla Juegos* a través del *ID\_Año*.

JUEGOS_DATOS			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
Anio	Whole number	PK	NOT NULL
Cant Hombres	Whole number		
Cant Mujeres	Whole number		
Cant Eventos	Whole number		
Cant Deportes	Whole number		
Cant Países	Whole number		
Cant Países Medallistas	Whole number		
Cant Países Campeones	Whole number		

De igual forma incorporé nuevas dimensiones a la *tabla País* ya existente, como ser el continente al que pertenece, la cantidad de participaciones olímpicas de cada país y el total de medallas obtenidas en condición de local y visitante por cada uno.

PAÍS			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
ID_Pais	Text	PK	NOT NULL
País	Text		
Continente	Text		
Continente_Corregido	Text		
Cant_Particip_Local	Whole number		
Cant_Particip_Visitante	Whole number		
Medallas local	Whole number		
Medallas visitante	Whole number		
Prom_Medal_Tradicional	Decimal number		
Prom_Medal_Ajustado	Decimal number		



Además, originalmente pensaba incorporar los datos correspondientes al GDP como dimensiones, con un registro por país y una columna por año. El problema con el que me topé con este planteo es que me obligaba a trabajar con demasiadas columnas (y a su vez varias columnas calculadas para cada una de ellas) al tiempo que, al ser justamente los años los nombres de las dimensiones no encontraba la forma de que Power BI los considerara como medidas e interactuaran con los filtros.

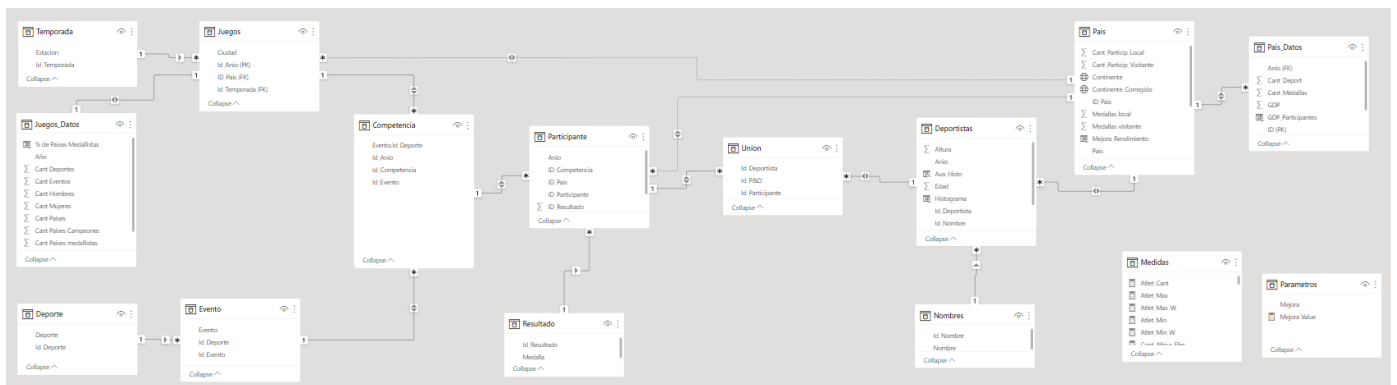
La mejor solución que encontré fue cargar mis datos de forma diferente: Armé una nueva *tabla Pais\_Datos* en la que hay un registro distinto para cada país en cada año y una única columna GDP [como una anulación de dinamización]. Esta tabla se vincula a la *tabla País* en una relación 1:N a través del Id\_Pais (FK en *Pais\_Datos* y PK en *País*) y me permite no sólo filtrar por país y año sino que debo generar muchas menos columnas calculadas puesto que ahora todos mis datos están en una única, simplificando la tarea y reduciendo los errores, al tiempo que me permite realizar gráficos de una forma más sencilla.

También aproveché para incorporar como dimensiones la cantidad de deportistas y la cantidad de medallas obtenidas por cada país en cada año de competencia, que al estar todos en la misma columna me simplificó mucho la tarea a la hora de generar los gráficos.

PAÍS DATOS			
CAMPO	TIPO DE CAMPO	TIPO DE CLAVE	CARACTERISTICAS
ID (País+Año)	Text	PK	NOT NULL
ID_Pais	Text	FK	NOT NULL
Año	Whole number	FK	NOT NULL
GDP	Whole number		
Cant_Deportistas	Whole number		
Cant_Medallas	Whole number		

## 3.2 - Modelo Relacional en Power BI:

Tras las modificaciones mencionadas, el diagrama ER quedó planteado en PBI de la siguiente manera:



## 3.3 - Otras transformaciones de datos:

### 3.3.1 - Columnas calculadas:

Se generaron varias columnas calculadas en distintas tablas:

- En la *Tabla Deportistas* fue necesario crear dos columnas para poder graficar un histograma: una que segmenta las edades de los deportistas en subconjuntos de 4 años y una auxiliar con los nombres de los intervalos;
- En la *Tabla Juegos\_Datos* se utilizaron para calcular la cantidad de países sin medallas y el porcentaje de países medallistas,
- En la *Tabla País* se calcularon el promedio de medallas como local y como visitante tanto así como la mejora en el rendimiento,
- Mientras que en la *Tabla Pais\_Datos* hay columnas calculadas que determinan si un país participó o no de cierta cita olímpica, el GDP total de únicamente los países participantes en el juego en cuestión y una última que calcula cómo quedaría el medallero con el ajuste por PBI propuesto

Adicionalmente en la *Tabla Paises\_Datos* también se concatenaron columnas para generar la Primary Key a partir de otras 2 columnas, y se usaron index columns para los rankings de posiciones en la *Tabla País*.

### 3.3.1 - Medidas calculadas:

Al momento de redactar esta documentación el modelo cuenta con varias medidas calculadas por lo que solo haré referencia a un par de ellas.

Por ejemplo: Se contabilizaron las medallas ganadas agrupadas por continente de la siguiente manera:

```
Cont_Africa_Med =  
CALCULATE(SUM(Participante[ID_Resultado]),  
Participante[Pais.Continente_Corregido]="Africa")
```

Y a partir del mismo se pudo calcular la efectividad de los deportistas de cada región dividiendo el valor obtenido por la cantidad de atletas que participaron de cada continente:

```
Cont_Africa_Efec =  
[Cont_Africa_Med]/CALCULATE(COUNT(Participante[ID_Participante]),  
Participante[Pais.Continente_Corregido]="Africa")
```

Otro: Para determinar en qué juegos olímpicos se registró el máximo número de deportistas participantes se utilizó la siguiente expresión:

```
Atlet_Max_W = LOOKUPVALUE(Juegos_Datos[Año], Juegos_Datos[Total  
Deportistas], [Atlet_Max])
```

donde:

```
Atlet_Max = MAX(Juegos_Datos[Total Deportistas])
```

Por último, y entrando más en el objetivo que perseguía el trabajo se utilizó la siguiente medida calculada para determinar la cantidad de medallas adicionales que debería ganar la delegación argentina para igualar la media mundial:

```
Medallas_Mejora_Necesaria =  
([Medallero_Ajustado]-[Medallero_Ajustado_Arg])*  
CALCULATE([Medallas_GDP_Partici]/1000000000000,  
Pais_Datos[ID_Pais (FK)]="Arg")
```

donde:

```
Medallero_Ajustado =  
[Medallas_Cantidad]*1000000000000/[Medallas_GDP_Partici]
```

```
Medallero_Ajustado_Arg =  
CALCULATE([Medallas_Cantidad]*1000000000000/[Medallas_GDP_Partici]  
, Pais_Datos[ID_Pais (FK)]="Arg")
```

```
Medallas_GDP_Partici = SUM(Pais_Datos[GDP_Participantes])
```

```
Medallas_Cantidad = SUM(Pais_Datos[Cant_Medallas])
```

y donde sólo cabría agregar que se divide a los GDP por un billón debido a la magnitud de estos números y por lo tanto el medallero ajustado termina siendo en realidad **la cantidad de medallas equivalentes ganadas por un país por cada billón de dólares de GDP**.

### 3.4 - Segmentaciones:

Debido a que gran parte del trabajo se basa en visualizar la evolución de parámetros tales como la cantidad de deportistas, la cantidad de países participantes y las medallas obtenidas por los mismos a lo largo del tiempo no es de extrañar que la segmentación más común en el dashboard sea la referida a años. De hecho aparece en 4 solapas distintas, pero siempre con la idea de brindar mayor profundidad al análisis en un juego olímpico en particular tras haber visto la foto completa.<sup>3</sup>

Por ejemplo, en la solapa en la que se muestra la evolución en el tiempo de la cantidad de deportistas participando agrupados por sexo, el filtro por año nos permite acceder a un gráfico de tortas específico donde es más sencillo apreciar la repartición por género en cada juego. De esta forma podemos visualizar rápidamente que las representaciones de unos y otros es cada vez más justa, siendo en 2016 54,97% hombres y 45,03% mujeres -muy alejado del 100% hombres que participaron en los JJOO de Atenas 1896-

Por otro lado, en la solapa referente a la evolución etaria, el segmentador por año viene acompañado también por filtros que nos permiten seleccionar el sexo y la medalla obtenida, de modo tal de poder ir jugando con los 3 de forma alternada e ir notando los cambios en el tiempo. De esta interacción podemos obtener por ejemplo que la edad promedio de las deportistas mujeres que obtienen una medalla de oro es de 24,49 años, mientras que los hombres alcanzan su pico de rendimiento 2 años después, a los 26,53. Y si bien es cierto que esta tendencia viene estable no dejan de sorprender los picos, con un promedio de edad de 55 años entre las deportistas campeonas en los juegos de 1904, un participante de 93 años en 1928 y uno de apenas 10 en los juegos de 1896.

Por su parte, en la solapa en la que se agrupan los resultados por continente si bien se utilizó al tiempo como la forma de interactuar con los gráficos, en lugar de utilizar un segmentador dropdown opté por un selector de periodo. Me pareció que para este caso, por tratarse de valores acumulados, tenía más sentido un filtro que permitiera segmentar por espacios de tiempo más largos para analizar períodos históricos y tendencias. Manipulando el selector podemos notar, por ejemplo, que durante los primeros 50 años de los JJOO (1896-1946) el 75% de los deportistas era europeo y obtenían el 66,78% de las medallas. Luego de la segunda guerra mundial, de la mano primero de la URSS y luego de China Asia pasó del 3,5% del total de deportistas al 20,39% (el segundo continente con más atletas luego de Europa) y a recolectar el 26,92 % del total de medallas en juego, cuando en el período pre-guerra se hacía solo con el 1,78%.

Finalmente, toca mencionar también que en la solapa donde se analiza la ventaja o no de ser local hay filtro por país organizador y no por año que permite acceder a medidas calculadas que se visualizan a través de tarjetas y tarjetas multi-líneas.

---

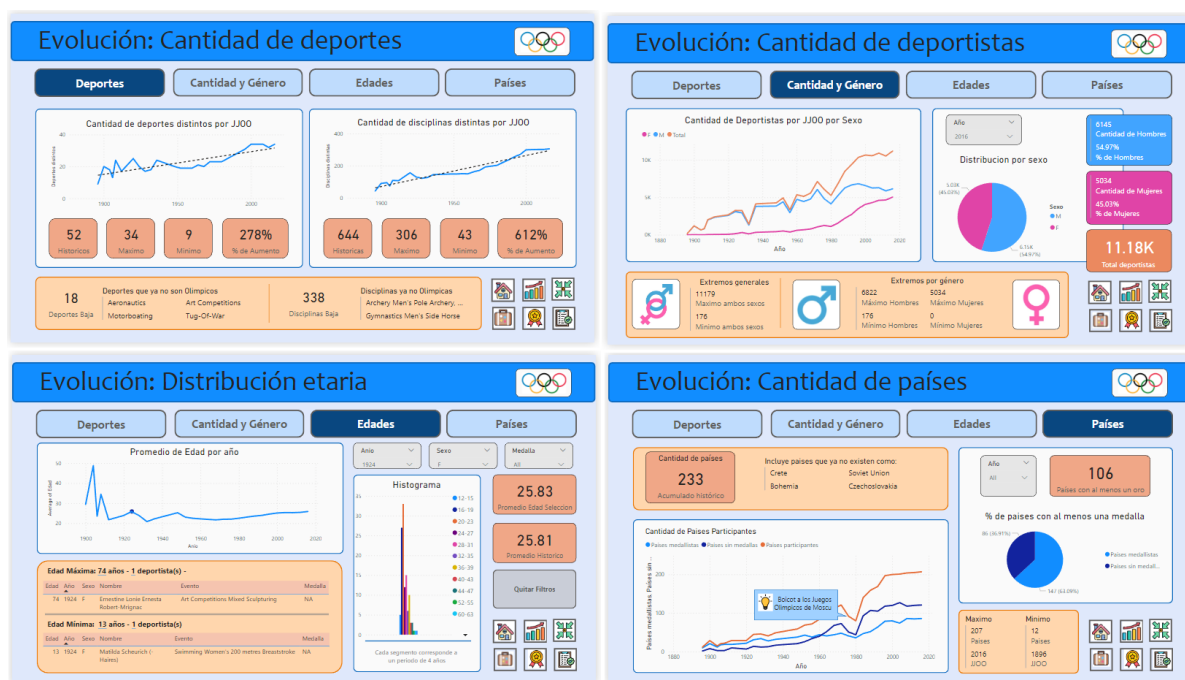
<sup>3</sup> La interfaz gráfica de todos los segmentadores pueden verse en la siguiente sección: "Visualizaciones"

## 4 - Visualizaciones:

Para ir de lo más general a lo más específico, y aprovechando los datos con los que contaba, decidí comenzar con toda la información referente a la evolución en el tiempo de distintos parámetros que hacen a los juegos olímpicos, donde quedara de manifiesto la importancia que fueron adquiriendo a lo largo de los años.

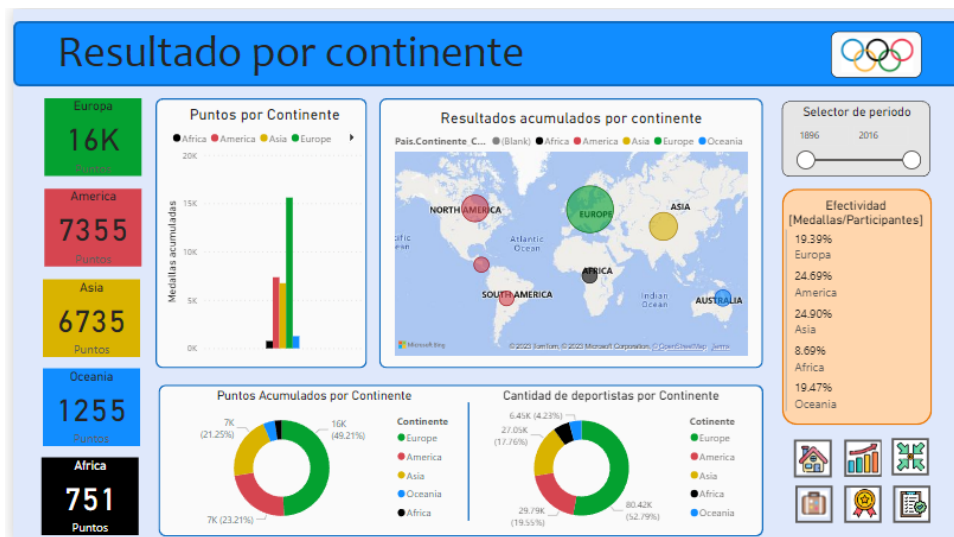
Así es como pueden visualizarse:

- gráficos y KPIs referidos al aumento en la cantidad de deportes y la cantidad de disciplinas junto con ejemplos de deportes y disciplinas que ya no son olímpicos.
- la distribución etaria de los atletas (que a su vez puede filtrarse por año, sexo y resultado obtenido + un botón que permite limpiar los filtros), un histograma que segmenta las edades en grupos de 4 años y tablas que muestran los deportistas más jóvenes y más longevos dependiendo de la selección realizada.
- la cantidad de atletas participantes y su representación por sexo junto a un gráfico de torta que permite apreciar las proporciones de unos y otros según el año escogido para filtrar, y
- el aumento en la cantidad de países que participan, con el detalle de qué porcentaje de ellos obtiene al menos una medalla, también segmentable por año.



Originalmente toda esta información se encontraba distribuida en 4 solapas distintas con botones que permitían la navegabilidad entre ellas, pero en pos de no incumplir con la consigna de cantidad total de solapas y aprovechando que todas referían a evolución de parámetros en el tiempo decidí unificarlas y utilizar Bookmarks y Selections para permitir la visualización de la información que solicite el usuario reutilizando los botones originales.

A continuación, en orden de empezar a evidenciar las diferencias que existen entre países y utilizar otras herramientas, decidí agrupar la información por continente. Queda plasmada la gran diferencia que saca Europa respecto al resto, seguido por América (empujada principalmente por Estados Unidos). Además, como ya se mencionó, a través del uso del selector de período, puede advertirse también el crecimiento de Asia de los últimos años.



El gráfico de columnas y el de burbujas contienen prácticamente la misma información, con la salvedad que en el de burbujas América se dividió en Sur, Central y Norte para que quede en evidencia el peso de Norteamérica en el total, pero decidí incluir el de columnas (que no estaba en el mockup) porque me parece una representación más clara.

La siguiente solapa analiza la hipótesis secundaria de en qué tanto incide la localía en los resultados obtenidos por el país anfitrión, comparando en un gráfico de columnas el resultado promedio obtenido de visitante vs el resultado promedio obtenido de local y permite a su vez filtrar por cada uno de los países que organizaron un juego para un análisis más detallado.



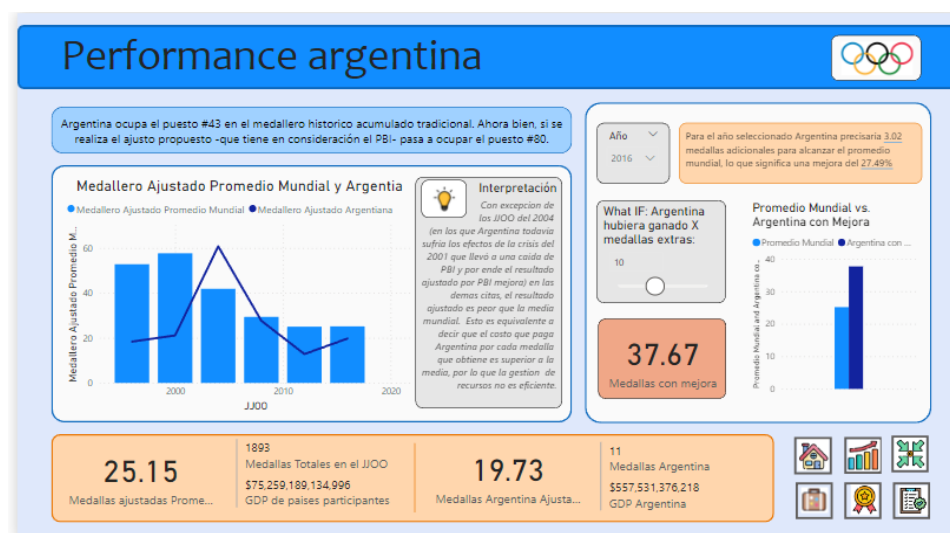
En principio, con jueces imparciales y canchas con las medidas reglamentarias, ser local no debería generar una diferencia en el rendimiento, pero los números indican otra cosa.

La última solapa que incluía el Mockup apuntaba a darle respuesta a la hipótesis original del trabajo: cuánto incide el PBI en los resultado deportivos. Para esto se realizaron 2 gráficos de barras con el TOP 7 de países que dominan el medallero –promedio de los últimos 20 años- calculados de la forma tradicional y con el método de ajuste propuesto.



De esta solapa, lo que salta a la vista es que ninguno de los países que dominan un medallero siquiera figuran en el otro. De hecho, pueden irse recorriendo los gráficos a través del tooltip y comparar qué lugar ocupa cada país en cada una de las versiones.

Por último, (y por ser la que mayor grado de detalle aporta), la solapa que cierra el dashboard no estaba en el mockup, pero la incorporé como resultado de determinar que el usuario final sería alguien perteneciente al ministerio de deportes nacional. En función de eso evalué la situación particular de la Argentina a la luz de este medallero ajustado para determinar si la gestión estaba siendo exitosa o no.



En esta hoja se incluye un parámetro que permite jugar con las medallas adicionales que tendría que haber sacado Argentina e ir comparando esta mejora en el resultado contra el promedio mundial. Su inclusión es sin dudas forzada, pero se incorporó para poder usar la herramienta.



## 5 - Conclusión:

---

Tal como era de esperar, los resultados deportivos obtenidos están vinculados a factores externos que escapan de los límites del campo de juego, pero lo que llama mi atención es lo fuerte que es este vínculo y la claridad con lo que esto queda reflejado en los números:

Para empezar la incidencia de la localía es marcada: Salvo Canadá, todos los países anfitriones consiguieron más medallas en casa que el promedio de medallas obtenidas de visitante, pero el número que sorprende es que esta mejora sea en promedio del 272,3%, siendo Bélgica el país donde se marca más este efecto, con 876,4% de mejora en condición de local. Nota al margen (y no tanto): El de Río 2016 fue el primer JJOO realizado en Sudamérica y a la fecha nunca hubo un país que oficie de organizador perteneciente al continente africano.

En cuanto a la influencia del PBI, caía de maduro que esto iba a ser así porque las economías grandes tienen, por lo general, las necesidades básicas de sus habitantes resueltas, lo que les permite disponer de una mayor cantidad de fondos para dedicar a planes deportivos y de esparcimiento. Nuevamente lo que impacta de los resultados obtenidos no es la confirmación de que esto ocurra sino su magnitud. Ni uno sólo de los países que dominan el medallero tradicional aparecen en el otro, siendo Rusia el que mejor se posiciona con su 3er puesto en el tradicional y 25 en el ajustado.

Más allá de eso, la comparación directa de los medalleros carece un poco de sentido puesto que cuentan dos historias distintas. El medallero ajustado calcula cuántas medallas se obtienen por cada billón de USD de PBI, lo no es otra cosa que la representación de la inversa del costo de obtener una medalla. Dicho de otra forma, los países que encabezan el medallero ajustado son aquellos para los que el costo de obtener una medalla es menor, ó, por qué no, los que consiguen gestionar sus recursos de una manera más eficiente. Es por esto que Estados Unidos pasa de ocupar el puesto 1 en el medallero tradicional a caer hasta el 93 en el ajustado; porque si bien consigue una cantidad increíble de medallas en cada cita olímpica, la cantidad de recursos con que dispone hacen que el costo de obtener cada medalla sea altísimo.

Pero, ¿qué ocurre con la situación de la Argentina? La primera lectura que uno puede hacer es que la delegación nacional pasa del puesto 43 en el medallero tradicional al puesto 80 en el ajustado, (retroceso considerable) y esto ocurre porque hay muchos países que sacan algunas pocas medallas menos que Argentina, pero que tienen un PBI sensiblemente menor, por lo que este ratio les da mejor.

Ahora bien, alejándonos de la comparación directa país vs país, podemos simplemente intentar verificar en qué situación se encuentra Argentina respecto al promedio de costos mundial de obtener una medalla y veremos que salvo en los juegos olímpicos del 2004 –coletazo de la caída de nuestro PBI debido a la crisis del 2001- en los últimos 20 años el costo por medalla para Argentina fue superior a la media mundial. Este análisis nos permite tener una referencia de que nuestro país no está siendo eficiente en cuanto al uso de recursos y hay camino por andar y cosas por mejorar, aunque no debe dejar de mencionarse que en las últimas ediciones las actuaciones no fueron malas, acercándose a la media -siempre por debajo-.



## 6 - Futuras líneas:

---

Habiendo concluido que Argentina está haciendo un uso poco eficiente de recursos -puesto que los costos medios de conseguir una medalla para la delegación nacional son superiores al promedio mundial-, ¿en qué podemos mejorar y cuál es el camino a seguir ahora?

Considero que lo primero sería centrarse en los países que mejor se posicionan en el ranking ajustado (que son los que están haciendo mejor uso de sus recursos) y analizar la razón por la que se colocan allí y qué políticas están llevando adelante.

Por ejemplo Jamaica lidera el medallero ajustado, pero la mayoría de sus medallas las obtiene por la arrolladora diferencia que sacan en las pruebas de velocidad de atletismo (100 m, 200m y 4x100m tanto en hombres como en mujeres), sin destacarse tanto en otros deportes. ¿Será acaso una ventaja genética? ¿O solo invierte recursos en desarrollar esta especialidad? Por lo que sea, descartado.

Tonga ocupa el segundo lugar en el medallero, pero únicamente porque su PBI es realmente muy bajo. Lo cierto es que las últimas medallas que consiguió fueron en Atlanta 1996. Descartado también.

Los Georgianos se caracterizan por contar con deportistas con mucha fuerza, por lo que no es extraño que la mayoría de sus victorias se dan en lucha, judo y levantamiento de pesas.

Siguiendo el orden aparece Cuba, que de todos los países participantes es el que logra el mejor equilibrio entre ambos medalleros, ocupando el puesto 11 en el tradicional y ascendiendo hasta ubicarse en el puesto 4 en el ajustado. Pero no solo esto, sino que más allá de contar con una gran tradición en boxeo amateur, en los últimos 3 juegos olímpicos ha obtenido medallas en deportes tan variados como decatlón, salto en largo, salto con garrocha, 110 metros con vallas, lanzamiento de martillo y de disco, judo, taekwondo, levantamiento de pesas, lucha, shooting, ciclismo y baseball.

Por lo que si me preguntan a mí, tomaría a Cuba como ejemplo y vería en detalle qué políticas está llevando a cabo este país insular en materia deportiva y cuales son extrapolables a nuestra situación. De no poder llevarse adelante ninguna, será cuestión de seguir recorriendo el medallero ajustado: hay 75 países más que están haciendo una gestión de recursos mejor que la nuestra.

Contamos con el potencial humano. Falta acompañarlo con políticas que le brinden un contexto propicio para poder desarrollarse. *Si no es genético, entonces es el ambiente.*