

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

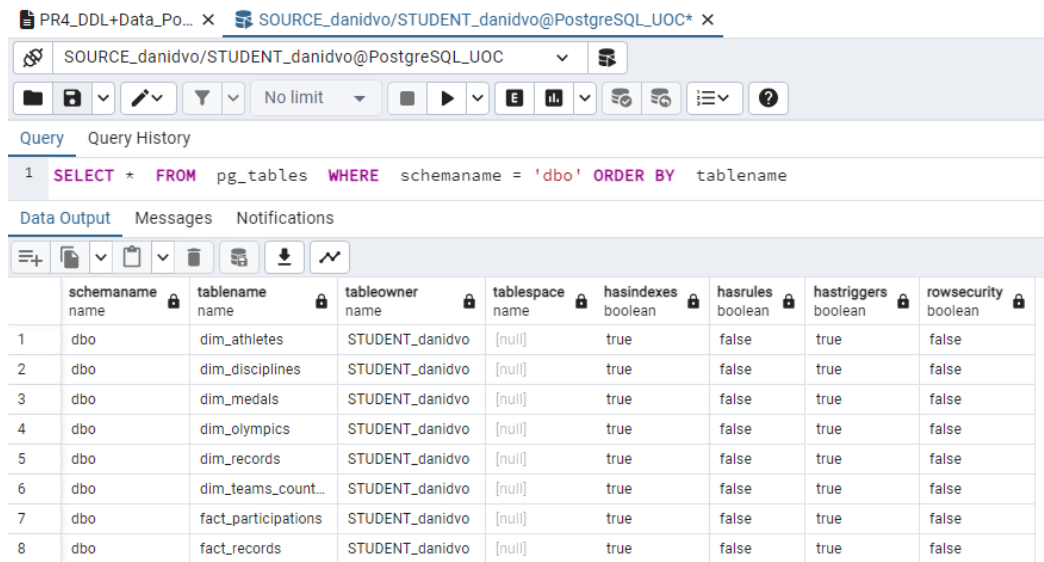
ALUMNO: Daniel Vargas Olivencia

2.2. Ejercicio 1. Análisis de datos mediante autoconsumo (20%)

2.2.1. Autodescubrimiento de la base de datos

Se solicitan las sentencias SQL que den respuesta a las siguientes cuestiones:

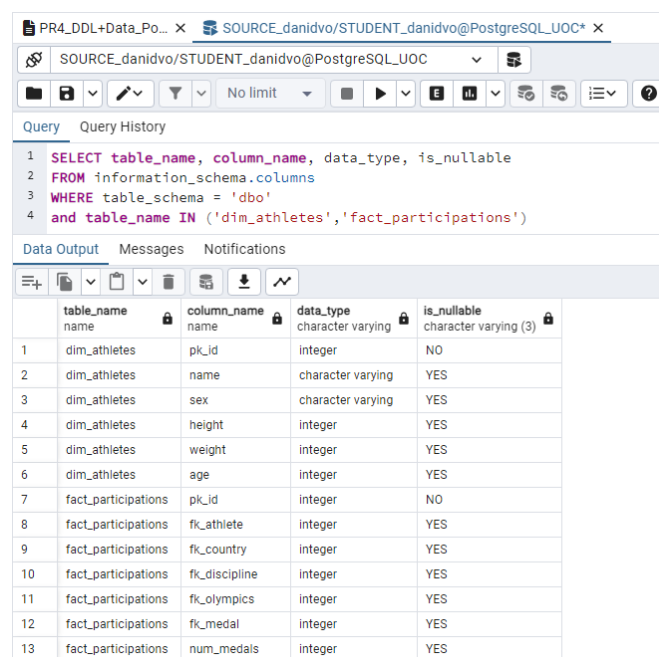
1) Obtener la información existente en el esquema dbo mediante una consulta SQL sobre la vista *pg_tables* (vista que proporciona el catálogo de las tablas almacenadas en la base de datos).



The screenshot shows a PostgreSQL query tool interface. The query executed is: `SELECT * FROM pg_tables WHERE schemaname = 'dbo' ORDER BY tablename`. The results are displayed in a table with 9 columns: schemaname, tablename, tableowner, tablespace, hasindexes, hasrules, hastriggers, and rowsecurity. The results show 8 tables in the dbo schema.

	schemaname	tablename	tableowner	tablespace	hasindexes	hasrules	hastriggers	rowsecurity
1	dbo	dim_athletes	STUDENT_danidvo	[null]	true	false	true	false
2	dbo	dim_disciplines	STUDENT_danidvo	[null]	true	false	true	false
3	dbo	dim_medals	STUDENT_danidvo	[null]	true	false	true	false
4	dbo	dim_olympics	STUDENT_danidvo	[null]	true	false	true	false
5	dbo	dim_records	STUDENT_danidvo	[null]	true	false	true	false
6	dbo	dim_teams_count...	STUDENT_danidvo	[null]	true	false	true	false
7	dbo	fact_participations	STUDENT_danidvo	[null]	true	false	true	false
8	dbo	fact_records	STUDENT_danidvo	[null]	true	false	true	false

2) Utilizando el esquema dbo y la vista *information_schema.columns*, obtener de las tablas *dim_athletes* y *fact_participations* los nombres de las columnas, los tipos de datos de cada columna y si éstas admiten nulos o no.



The screenshot shows a PostgreSQL query tool interface. The query executed is: `SELECT table_name, column_name, data_type, is_nullable FROM information_schema.columns WHERE table_schema = 'dbo' and table_name IN ('dim_athletes', 'fact_participations')`. The results are displayed in a table with 5 columns: table_name, column_name, data_type, and is_nullable. The results show 13 columns for the two specified tables.

	table_name	column_name	data_type	is_nullable
1	dim_athletes	pk_id	integer	NO
2	dim_athletes	name	character varying	YES
3	dim_athletes	sex	character varying	YES
4	dim_athletes	height	integer	YES
5	dim_athletes	weight	integer	YES
6	dim_athletes	age	integer	YES
7	fact_participations	pk_id	integer	NO
8	fact_participations	fk_athlete	integer	YES
9	fact_participations	fk_country	integer	YES
10	fact_participations	fk_discipline	integer	YES
11	fact_participations	fk_olympics	integer	YES
12	fact_participations	fk_medal	integer	YES
13	fact_participations	num_medals	integer	YES

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

3) Ejecutar la siguiente consulta de base de datos y explicar los resultados por cada una de las columnas obtenidas.

SELECT

attname, null_frac, n_distinct, most_common_vals

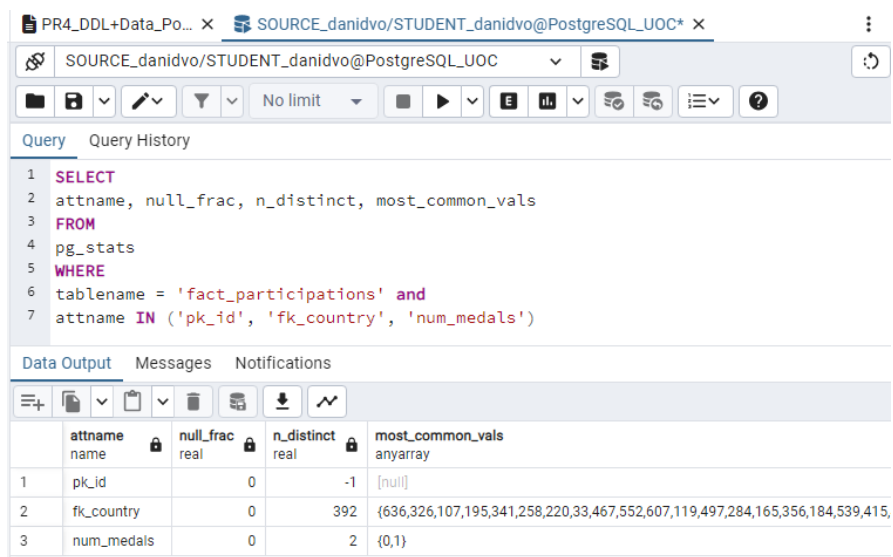
FROM

pg_stats

WHERE

tablename = 'fact_participations' and

attname IN ('pk_id', 'fk_country', 'num_medals')



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

```
1 SELECT
2 attname, null_frac, n_distinct, most_common_vals
3 FROM
4 pg_stats
5 WHERE
6 tablename = 'fact_participations' and
7 attname IN ('pk_id', 'fk_country', 'num_medals')
```

The results are displayed in a table with the following columns: attname, null_frac, n_distinct, and most_common_vals.

	attname	null_frac	n_distinct	most_common_vals
1	pk_id	0	-1	[null]
2	fk_country	0	392	{636,326,107,195,341,258,220,33,467,552,607,119,497,284,165,356,184,539,415,
3	num_medals	0	2	{0,1}

1. **pk_id:**

- **null_frac:** El valor es 0, lo que significa que no existen valores nulos en esta columna. Esto es coherente con la naturaleza de pk_id como clave primaria, ya que por definición no puede contener valores nulos.
- **n_distinct:** El valor es -1. Este resultado indica que todos los valores en la columna son únicos, lo que es consistente con el comportamiento esperado de una clave primaria.
- **most_common_vals:** Este campo está vacío (null), ya que cada valor en la columna es único y no existen valores que se repitan.

2. **fk_country:**

- **null_frac:** El valor es 0, indicando que no hay valores nulos en esta columna. Esto sugiere que todas las participaciones están asociadas a un país válido.

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos
PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

- **n_distinct:** El valor es 392, lo que significa que hay 392 valores únicos en esta columna, representando distintos países o regiones que participaron en los Juegos Olímpicos.
- **most_common_vals:** Contiene una lista de los valores más comunes, como 636, 326, etc. Esto indica que algunos países tienen un número significativamente mayor de participaciones en comparación con otros.

3. num_medals:

- **null_frac:** El valor es 0, lo que significa que todos los registros tienen un valor asociado para esta columna (no hay valores nulos).
- **n_distinct:** El valor es 2, indicando que solo hay dos valores únicos en esta columna, lo que probablemente corresponde a 0 (sin medalla) y 1 (medalla ganada).
- **most_common_vals:** Los valores más comunes son 0 y 1, lo que confirma que las participaciones se dividen entre aquellos que ganaron al menos una medalla y aquellos que no ganaron ninguna.

2.2.2. Análisis de los datos de los atletas en los Juegos Olímpicos

Disponiendo de la FIC (Factoría de Información Corporativa), se solicita dar respuesta a las siguientes cuestiones:

1) Listar el número total de participaciones por Juegos Olímpicos, ordenado por el número de participaciones de forma descendente.

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

PR4_DDL+Data_Po... x SOURCE_danidvo/STUDENT_danidvo@PostgreSQL_UOC* x

SOURCE_danidvo/STUDENT_danidvo@PostgreSQL_UOC

Query Query History

```
1 SELECT
2     fk_olympics,
3     COUNT(*) AS total_participations
4 FROM dbo.fact_participations
5 GROUP BY fk_olympics
6 ORDER BY total_participations DESC;
7
```

Data Output Messages Notifications

	fk_olympics integer	total_participations bigint
1	44	13682
2	52	13443
3	48	13402
4	46	13399
5	50	12524
6	42	11838
7	37	11618
8	39	10326
9	29	9847
10	35	8917
11	27	8402
12	31	8265
13	23	7526
14	25	7368
15	33	6843
16	51	4673
17	49	4378
18	47	4365
19	45	4060

Total rows: 52 of 52 Query complete 00:00:00.214 Ln 6, Col 36

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

PR4_DDL+Data_Po... x SOURCE_danidvo/STUDENT_danidvo@PostgreSQL_UOC* x

SOURCE_danidvo/STUDENT_danidvo@PostgreSQL_UOC

Query Query History

```
1 SELECT
2     o.year,
3     o.city,
4     o.season,
5     COUNT(*) AS total_participations
6 FROM dbo.fact_participations p
7 JOIN dbo.dim_olympics o ON p.fk_olympics = o.pk_id
8 GROUP BY o.year, o.city, o.season
9 ORDER BY total_participations DESC;
10
```

Data Output Messages Notifications

	year character varying (50)	city character varying (30)	season character varying (50)	total_participations bigint
1	2000	Sydney	Summer	13682
2	2016	Rio de Janeiro	Summer	13443
3	2008	Beijing	Summer	13402
4	2004	Athina	Summer	13399
5	2012	London	Summer	12524
6	1996	Atlanta	Summer	11838
7	1988	Seoul	Summer	11618
8	1992	Barcelona	Summer	10326
9	1972	Munich	Summer	9847
10	1984	Los Angeles	Summer	8917
11	1968	Mexico City	Summer	8402
12	1976	Montreal	Summer	8265
13	1960	Roma	Summer	7526
14	1964	Tokyo	Summer	7368
15	1980	Moskva	Summer	6843
16	2014	Sochi	Winter	4673

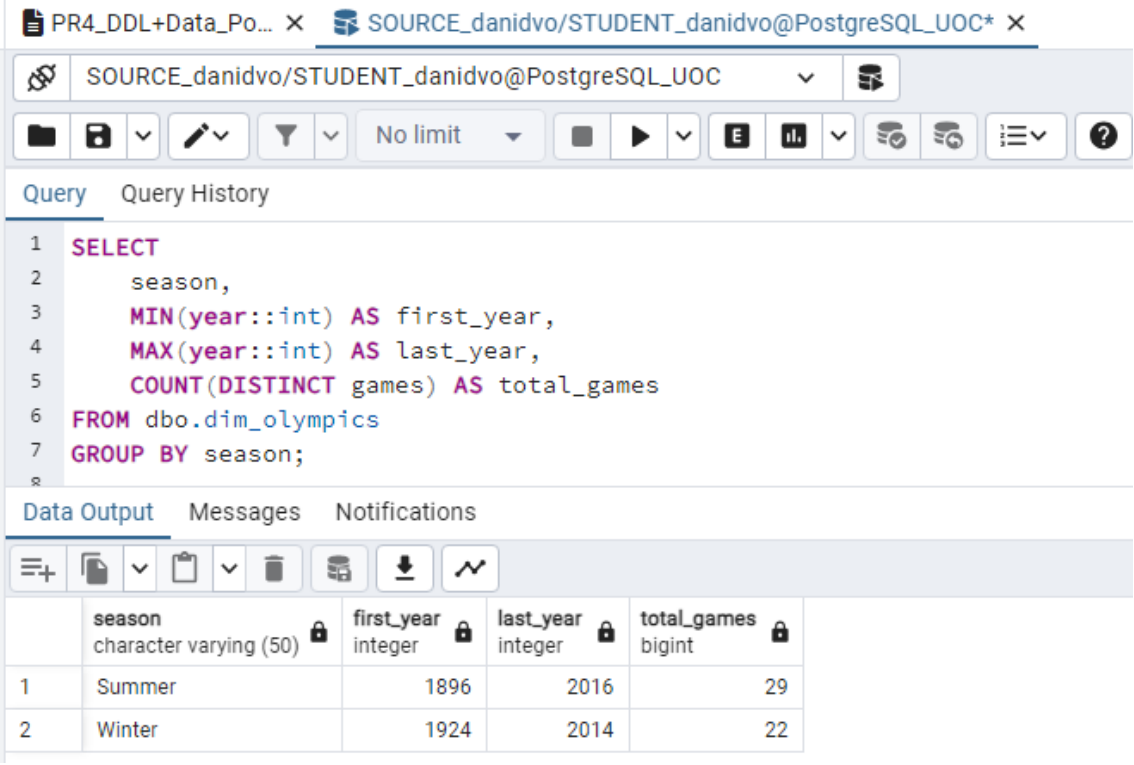
Total rows: 52 of 52 Query complete 00:00:00.164 Ln 10, Col 1

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

2) Listar el primero y el último de los años en que se celebraron unos Juegos Olímpicos, y el número de juegos celebrados, según si son juegos de verano o de invierno.



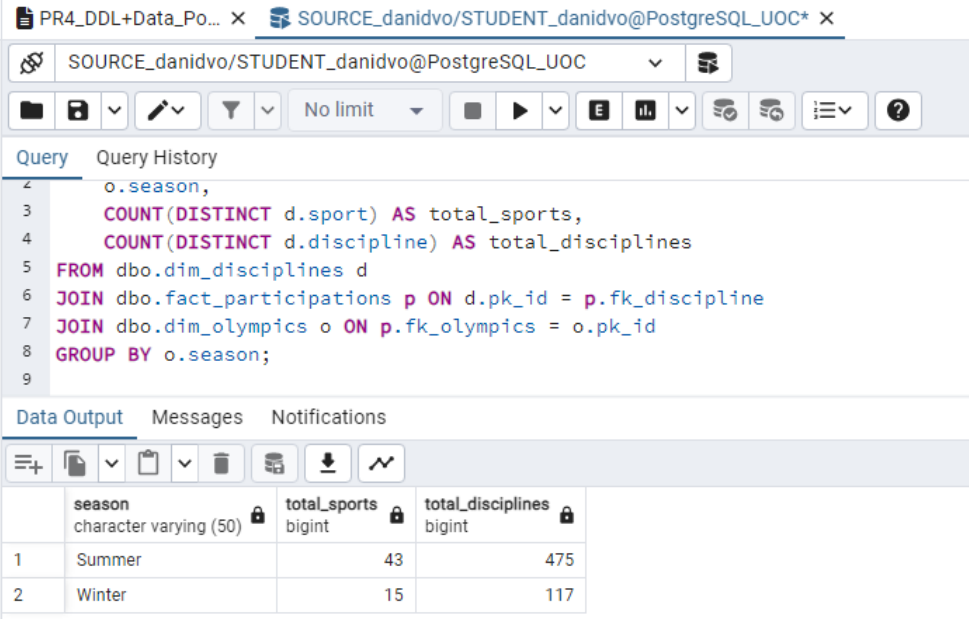
The screenshot shows a PostgreSQL query editor with the following query:

```
1 SELECT
2     season,
3     MIN(year::int) AS first_year,
4     MAX(year::int) AS last_year,
5     COUNT(DISTINCT games) AS total_games
6 FROM dbo.dim_olympics
7 GROUP BY season;
```

The query results are displayed in a table with the following columns: season, first_year, last_year, and total_games.

	season	first_year	last_year	total_games
1	Summer	1896	2016	29
2	Winter	1924	2014	22

3) Listar el número total de deportes y de disciplinas deportivas, según si se han celebrado en juegos de verano o de invierno.



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

```
2 o.season,
3     COUNT(DISTINCT d.sport) AS total_sports,
4     COUNT(DISTINCT d.discipline) AS total_disciplines
5 FROM dbo.dim_disciplines d
6 JOIN dbo.fact_participations p ON d.pk_id = p.fk_discipline
7 JOIN dbo.dim_olympics o ON p.fk_olympics = o.pk_id
8 GROUP BY o.season;
```

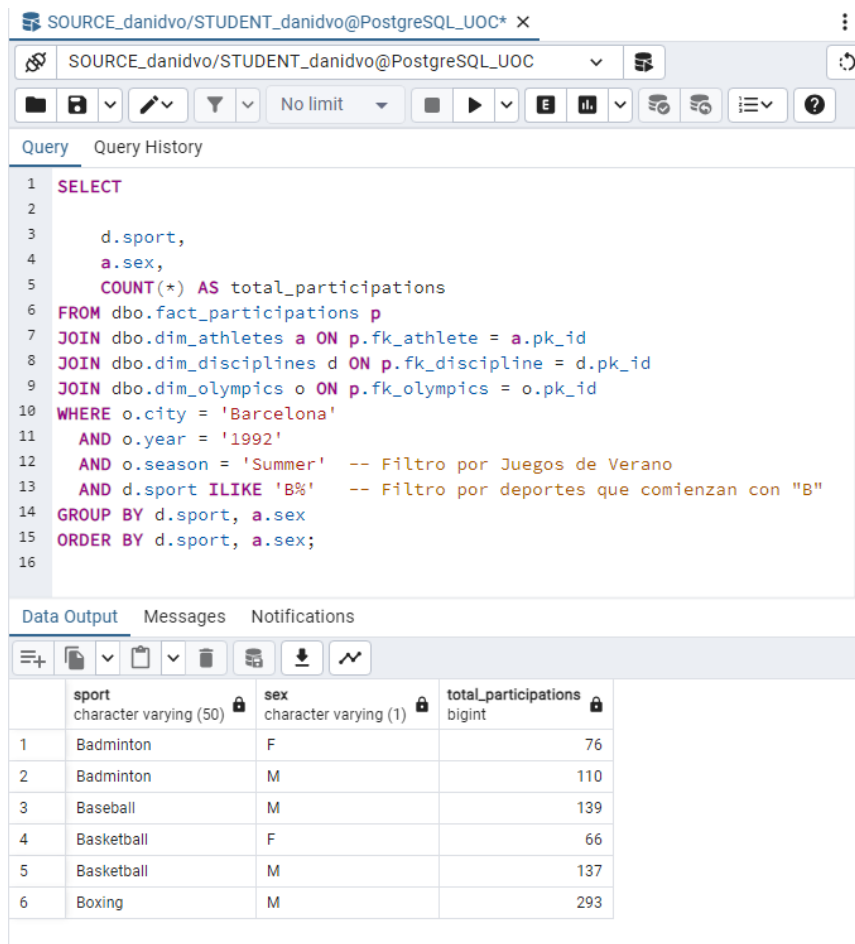
The query results are displayed in a table with the following columns: season, total_sports, and total_disciplines.

	season	total_sports	total_disciplines
1	Summer	43	475
2	Winter	15	117

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos
PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

4) Listar el número total de participaciones de atletas ordenados ascendentemente por el nombre del deporte y por el sexo de los atletas, para los juegos de verano celebrados en Barcelona en 1992, y cuyo deporte empiece por la letra 'B' .



The screenshot shows a PostgreSQL query editor interface. The query is as follows:

```
1 SELECT
2
3     d.sport,
4     a.sex,
5     COUNT(*) AS total_participations
6 FROM dbo.fact_participations p
7 JOIN dbo.dim_athletes a ON p.fk_athlete = a.pk_id
8 JOIN dbo.dim_disciplines d ON p.fk_discipline = d.pk_id
9 JOIN dbo.dim_olympics o ON p.fk_olympics = o.pk_id
10 WHERE o.city = 'Barcelona'
11        AND o.year = '1992'
12        AND o.season = 'Summer' -- Filtro por Juegos de Verano
13        AND d.sport ILIKE 'B%' -- Filtro por deportes que comienzan con "B"
14 GROUP BY d.sport, a.sex
15 ORDER BY d.sport, a.sex;
16
```

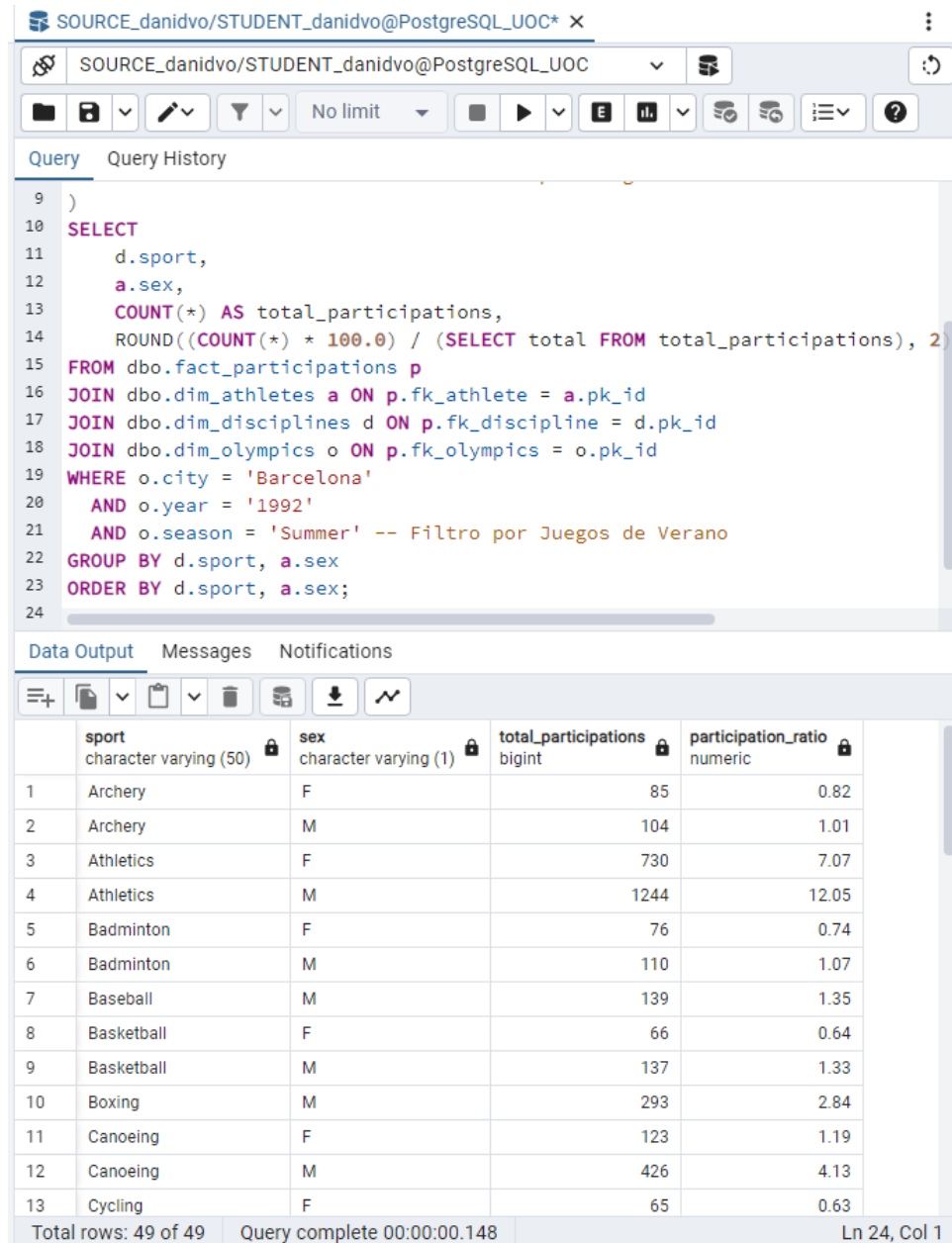
The results are displayed in a table with the following columns: sport, sex, and total_participations.

	sport character varying (50)	sex character varying (1)	total_participations bigint
1	Badminton	F	76
2	Badminton	M	110
3	Baseball	M	139
4	Basketball	F	66
5	Basketball	M	137
6	Boxing	M	293

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos
PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

5) Incorporar a la consulta anterior la ratio de participaciones de atletas ordenados ascendentemente por el nombre del deporte y por el sexo de los atletas, para los juegos de verano celebrados en Barcelona en 1992, considerando todos los deportes, respecto al total de participaciones para los juegos de verano celebrados en Barcelona en 1992.



The screenshot shows a PostgreSQL query editor interface. The query is as follows:

```
9 )
10 SELECT
11     d.sport,
12     a.sex,
13     COUNT(*) AS total_participations,
14     ROUND((COUNT(*) * 100.0) / (SELECT total FROM total_participations), 2)
15 FROM dbo.fact_participations p
16 JOIN dbo.dim_athletes a ON p.fk_athlete = a.pk_id
17 JOIN dbo.dim_disciplines d ON p.fk_discipline = d.pk_id
18 JOIN dbo.dim_olympics o ON p.fk_olympics = o.pk_id
19 WHERE o.city = 'Barcelona'
20 AND o.year = '1992'
21 AND o.season = 'Summer' -- Filtro por Juegos de Verano
22 GROUP BY d.sport, a.sex
23 ORDER BY d.sport, a.sex;
24
```

The results are displayed in a table with the following columns: sport, sex, total_participations, and participation_ratio. The table shows 13 rows of data for various sports and their participation ratios.

	sport character varying (50)	sex character varying (1)	total_participations bigint	participation_ratio numeric
1	Archery	F	85	0.82
2	Archery	M	104	1.01
3	Athletics	F	730	7.07
4	Athletics	M	1244	12.05
5	Badminton	F	76	0.74
6	Badminton	M	110	1.07
7	Baseball	M	139	1.35
8	Basketball	F	66	0.64
9	Basketball	M	137	1.33
10	Boxing	M	293	2.84
11	Canoeing	F	123	1.19
12	Canoeing	M	426	4.13
13	Cycling	F	65	0.63

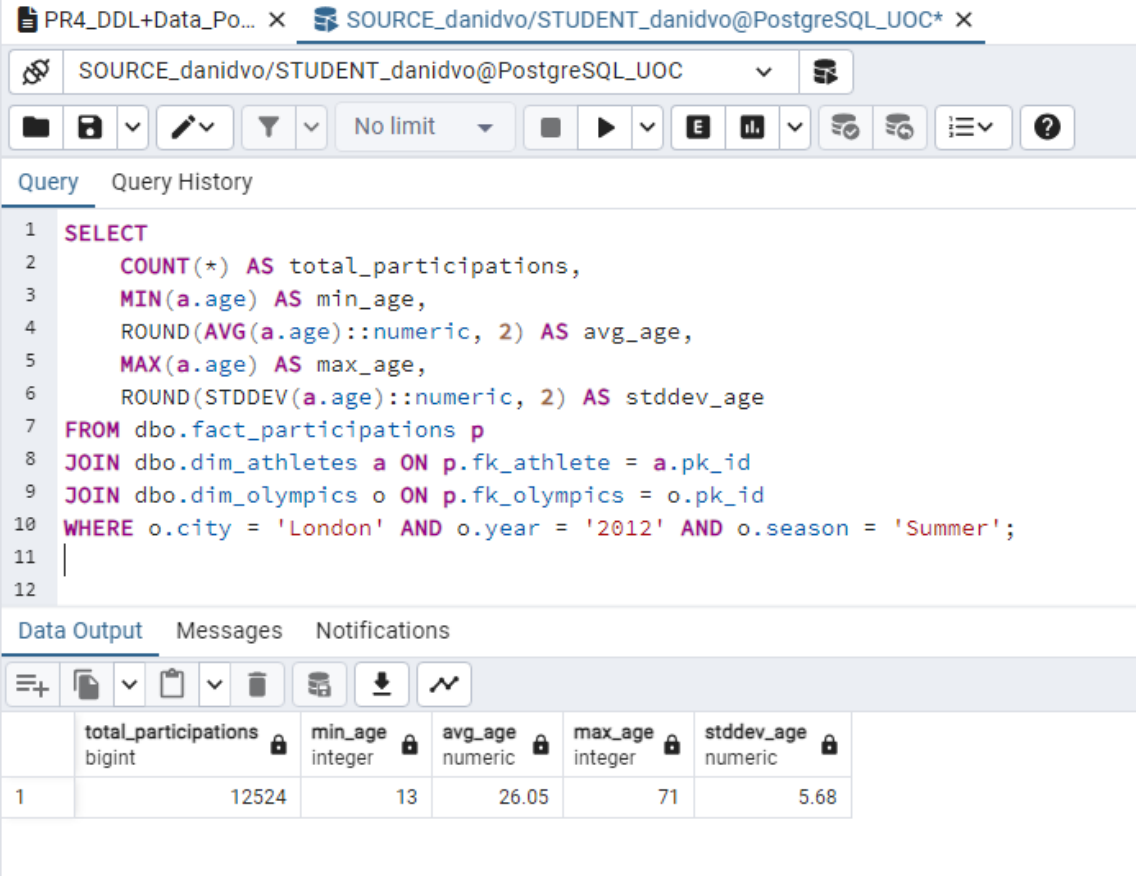
Total rows: 49 of 49 Query complete 00:00:00.148 Ln 24, Col 1

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos
PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

6) Listar en dos consultas distintas:

a. el cálculo del número total de participaciones, del valor mínimo, del valor medio, del valor máximo y de la desviación estándar de la edad de los atletas participantes, en los juegos de verano celebrados en Londres el año 2012.



The screenshot shows a PostgreSQL query editor interface. The top bar indicates the connection is 'SOURCE_danidvo/STUDENT_danidvo@PostgreSQL_UOC'. The query editor displays the following SQL query:

```
1 SELECT
2     COUNT(*) AS total_participations,
3     MIN(a.age) AS min_age,
4     ROUND(AVG(a.age)::numeric, 2) AS avg_age,
5     MAX(a.age) AS max_age,
6     ROUND(STDDEV(a.age)::numeric, 2) AS stddev_age
7 FROM dbo.fact_participations p
8 JOIN dbo.dim_athletes a ON p.fk_athlete = a.pk_id
9 JOIN dbo.dim_olympics o ON p.fk_olympics = o.pk_id
10 WHERE o.city = 'London' AND o.year = '2012' AND o.season = 'Summer';
11
12
```

Below the query editor, the 'Data Output' tab is active, showing the results of the query in a table format:

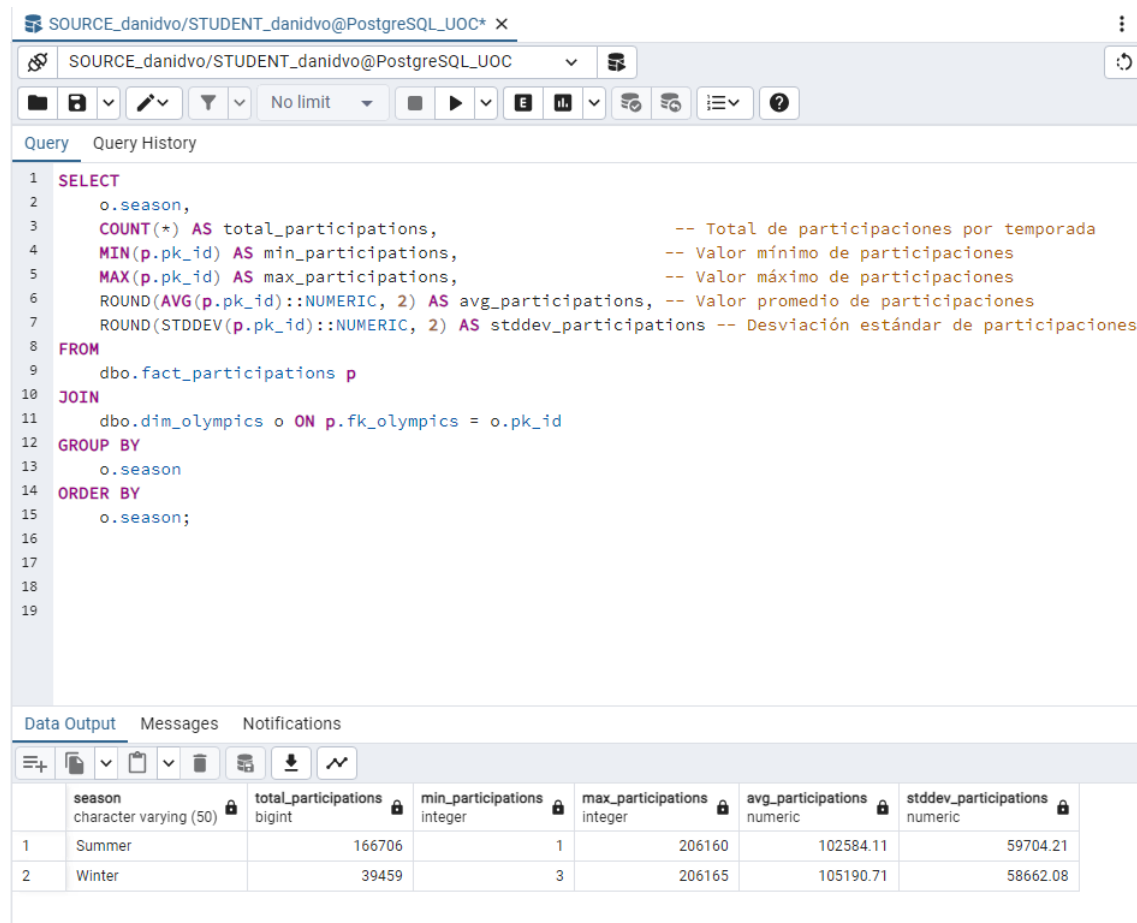
	total_participations bigint	min_age integer	avg_age numeric	max_age integer	stddev_age numeric
1	12524	13	26.05	71	5.68

b. el cálculo del número total, del valor mínimo, del valor medio, del valor máximo y de la desviación estándar de participaciones, según si se han celebrado en juegos de verano o de invierno.

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia



The screenshot shows a PostgreSQL query editor interface. The query is as follows:

```
1 SELECT
2     o.season,
3     COUNT(*) AS total_participations,           -- Total de participaciones por temporada
4     MIN(p.pk_id) AS min_participations,         -- Valor mínimo de participaciones
5     MAX(p.pk_id) AS max_participations,         -- Valor máximo de participaciones
6     ROUND(AVG(p.pk_id)::NUMERIC, 2) AS avg_participations, -- Valor promedio de participaciones
7     ROUND(STDDEV(p.pk_id)::NUMERIC, 2) AS stddev_participations -- Desviación estándar de participaciones
8 FROM
9     dbo.fact_participations p
10 JOIN
11     dbo.dim_olympics o ON p.fk_olympics = o.pk_id
12 GROUP BY
13     o.season
14 ORDER BY
15     o.season;
```

The results are displayed in a table with the following columns: season, total_participations, min_participations, max_participations, avg_participations, and stddev_participations.

	season	total_participations	min_participations	max_participations	avg_participations	stddev_participations
1	Summer	166706	1	206160	102584.11	59704.21
2	Winter	39459	3	206165	105190.71	58662.08

2.3. Ejercicio 2. Análisis de datos multidimensional (15%)

Objetivo:

Conocer la aplicación para diseñar cubos multidimensionales (OLAP) con el objetivo de incorporar información procedente de un almacén y analizarla mediante cubos.

Indicaciones para el desarrollo del ejercicio:

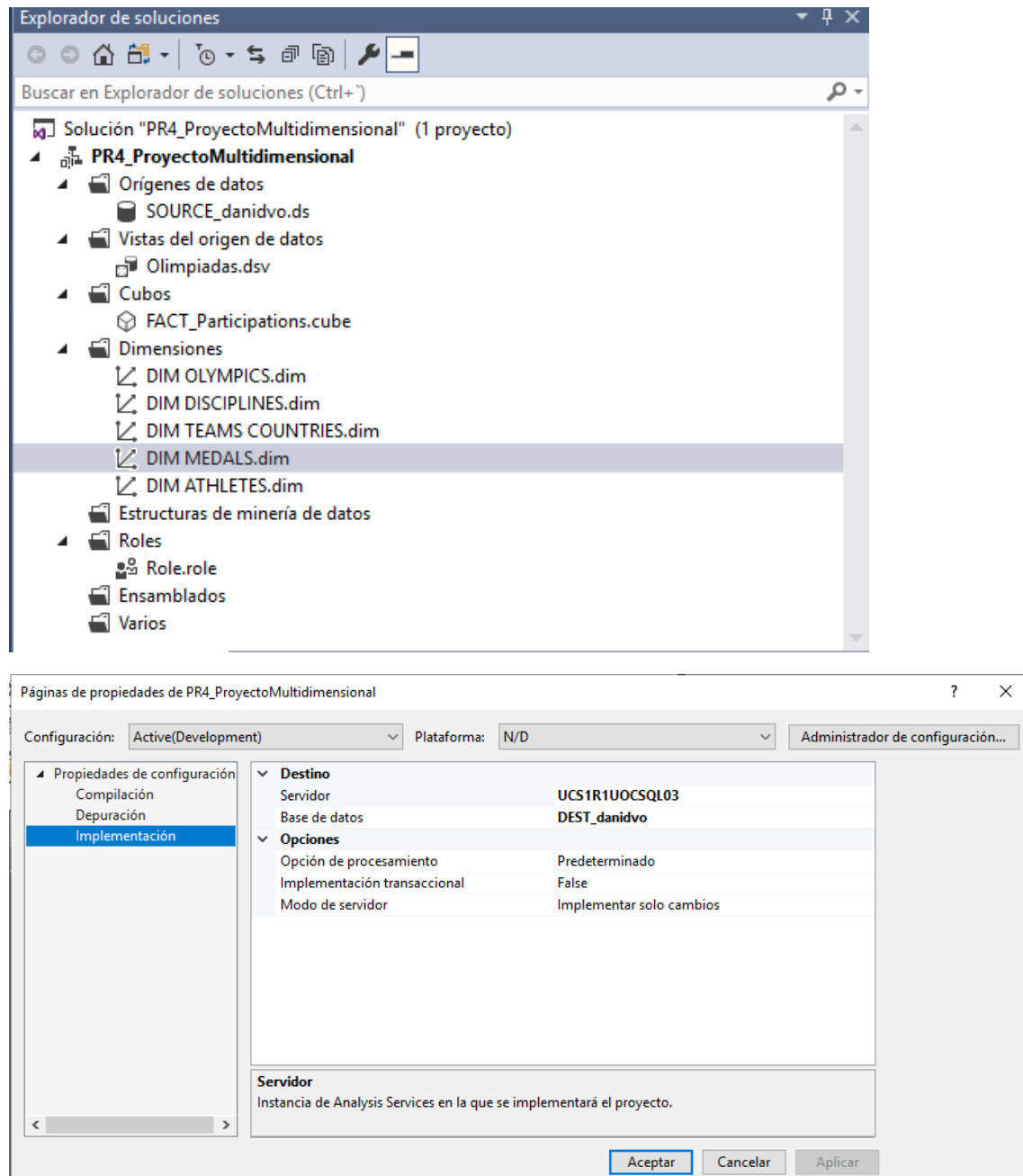
A continuación, se detallarán los diferentes pasos necesarios para la creación de un cubo que almacene la información de la tabla de hechos *FACT_participations* y sus dimensiones vinculadas, con el objetivo de poder realizar un análisis mediante las herramientas de análisis multidimensional facilitadas en la VDI.

Para poder disponer de un set de datos destinados al análisis, es necesario ejecutar el script de creación de tablas y carga de registros en la base de datos *SQLServer*

PR4_DDL+Data_SQLS.sql facilitado junto al enunciado.

2.3.1. Creación del proyecto, configuración del destino y origen de los datos

Mediante la herramienta Visual Studio se solicita crear un proyecto nuevo de tipo “Proyecto multidimensional y de minería de datos”, que permita construir cubos analíticos u analizar los datos a través de ellos.



Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

Administrador de conexiones

Proveedor: OLE DB nativo\SQL Server Native Client 11.0

Nombre del servidor: UCS1R1UOCSQL03 Actualizar

Conexión con el servidor

Autenticación: Autenticación de SQL Server

Nombre de usuario: STUDENT_danidvo

Contraseña: ●●●●●●●●

☒ Guardar mi contraseña

Establecer conexión con una base de datos

☒ Seleccionar o escribir el nombre de la base de datos:

SOURCE_danidvo

☐ Adjuntar un archivo de base de datos:

Examinar...

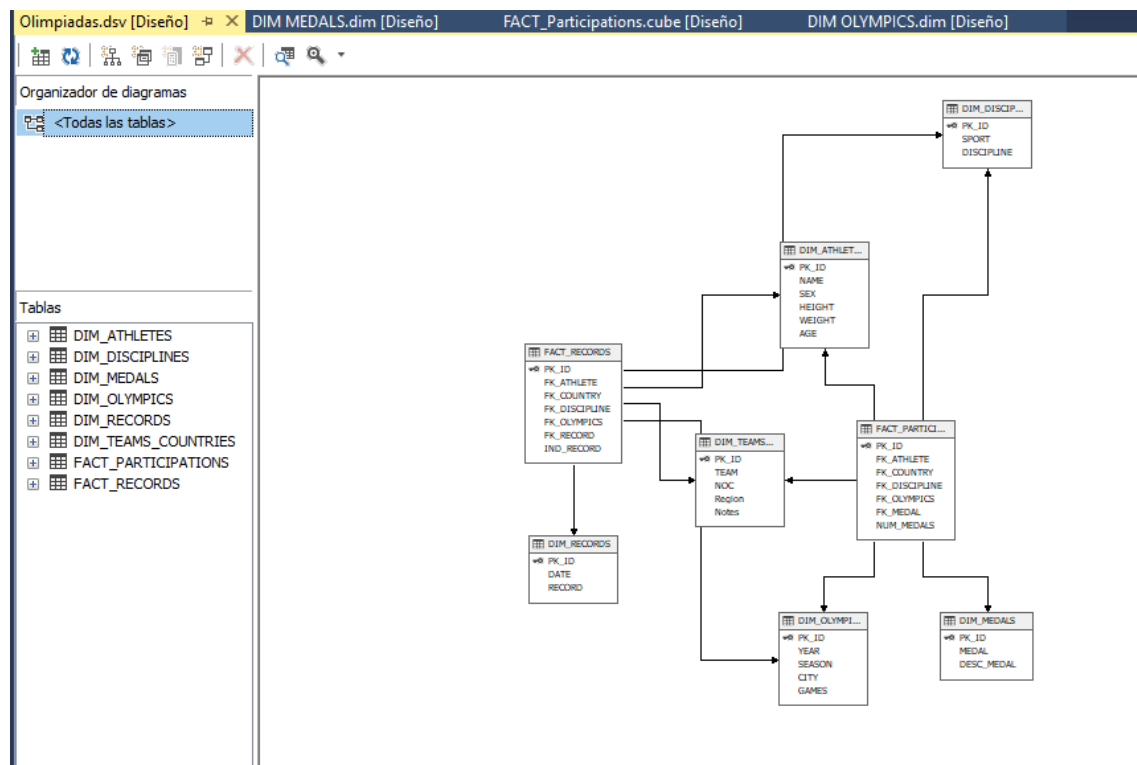
Nombre lógico:

Probar conexión Aceptar Cancelar Ayuda

2.3.2. Crear una vista de origen de datos para el modelo de estrella de la

FACT_PARTICIPATIONS, así como crear y procesar el cubo para el modelo de estrella de la FACT PARTICIPATIONS

Se solicita la creación de la vista FACT_PARTICIPATIONS destinada a incorporar los datos del modelo estrella, así como la creación y procesado del cubo destinado a la explotación de los datos.



ALUMNO: Daniel Vargas Olivencia



<Todos>

Recuento FACT PARTICIPATIONS

- KPI
 - DIM ATHLETES
 - AGE
 - HEIGHT
 - NAME
 - PK ID
 - SEX
 - WEIGHT
 - DIM DISCIPLINES
 - DISCIPLINE
 - PK ID
 - SPORT
 - DIM MEDALS
 - DESC MEDAL
 - MEDAL
 - PK ID
 - DIM OLYMPICS
 - CITY
 - GAMES
 - PK ID
 - SEASON
 - YEAR

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

2.3.4. Explotación del modelo OLAP

Mediante el uso de la FIC (Factoría de Información Corporativa) se solicita dar respuesta a las siguientes cuestiones:

1) Listar el número total de participaciones por cada uno de los Juegos Olímpicos celebrados y su ciudad anfitriona.

YEAR	CITY	GAMES	Recuento FACT PARTI...
1896	Athina	1896 Summer	40
1900	Paris	1900 Summer	71
1904	St. Louis	1904 Summer	119
1906	Athina	1906 Summer	201
1908	London	1908 Summer	366
1912	Stockholm	1912 Summer	551
1920	Antwerpen	1920 Summer	428
1924	Chamonix	1924 Winter	39
1924	Paris	1924 Summer	543
1928	Amsterdam	1928 Summer	622
1928	Sankt Moritz	1928 Winter	44
1932	Lake Placid	1932 Winter	43
1932	Los Angeles	1932 Summer	449
1936	Berlin	1936 Summer	827
1936	Garmisch-Partenkirchen	1936 Winter	69
1948	London	1948 Summer	914
1948	Sankt Moritz	1948 Winter	102
1952	Helsinki	1952 Summer	1914
1952	Oslo	1952 Winter	144
1956	Cortina d'Ampezzo	1956 Winter	334
1956	Melbourne	1956 Summer	2155
1956	Stockholm	1956 Summer	106

2) Listar el número de medallas de oro conseguidas por cada equipo olímpico que participó en los Juegos Olímpicos del verano del año 2008, por atletas masculinos, en el deporte de natación.

NOC	Region	SEX	GAMES	SPORT	MEDAL	NUM MEDALS	Recuento FACT PART...
BRA	Brazil	M	2008 Summer	Swimming	Gold	1	1
FRA	France	M	2008 Summer	Swimming	Gold	1	1
JPN	Japan	M	2008 Summer	Swimming	Gold	2	2
KOR	South Korea	M	2008 Summer	Swimming	Gold	1	1
NED	Netherlands	M	2008 Summer	Swimming	Gold	1	1
TUN	Tunisia	M	2008 Summer	Swimming	Gold	1	1
USA	USA	M	2008 Summer	Swimming	Gold	29	29

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

3) De la consulta anterior, se desea profundizar en conocer cuáles fueron los atletas que ganaron las medallas de oro del equipo procedente de Estados Unidos y cuantas medallas de oro ganaron.

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetro
DIM MEDALS	MEDAL	Igual	{ Gold }	
DIM ATHLETES	SEX	Igual	{ M }	
DIM DISCIPLINES	SPORT	Igual	{ Swimming }	
DIM OLYMPICS	SEASON	Igual	{ Summer }	
DIM OLYMPICS	YEAR	Igual	{ 2008 }	
DIM TEAMS COUNTRIES	NOC	Igual	{ USA }	

NOC	Region	SEX	GAMES	SPORT	MEDAL	NAME	NUM MEDALS	Recuento FACT P...
USA	USA	M	2008 Summer	Swimming	Gold	Aaron Wells Peirsol	2	2
USA	USA	M	2008 Summer	Swimming	Gold	Benjamin Marshall "Ben" Wildman...	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Brendan Joseph Hansen	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Cullen Andrew Jones	1	1
USA	USA	M	2008 Summer	Swimming	Gold	David "Dave" Walters	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Erik K. Vendt	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Garrett Weber-Gale	2	2
USA	USA	M	2008 Summer	Swimming	Gold	Jan Lowell Crocker	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Jason Edward Lezak	2	2
USA	USA	M	2008 Summer	Swimming	Gold	Klete D. Keller	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Mark Daniel Gangloff	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Matthew "Matt" Grevers	2	2
USA	USA	M	2008 Summer	Swimming	Gold	Michael Fred Phelps, II	8	8
USA	USA	M	2008 Summer	Swimming	Gold	Nathan Ghar-Jun Adrian	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Peter William Vanderkaay	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Richard "Ricky" Berens	1	1
USA	USA	M	2008 Summer	Swimming	Gold	Ryan Steven Lochte	2	2

4) Listar el número total de participaciones y de medallas concedidas en los Juegos Olímpicos celebrados en los años 2000, 2004 y 2008, por año, temporada, y ciudad anfitriona.

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetro
DIM OLYMPICS	YEAR	Igual	{ 2000, 2004, 2008 }	

GAMES	MEDAL	CITY	Recuento FACT PARTICIPACIONES	NUM MEDALS
2000 Summer	Bronze	Sydney	676	676
2000 Summer	Gold	Sydney	659	659
2000 Summer	NA	Sydney	11689	0
2000 Summer	Silver	Sydney	658	658
2004 Summer	Bronze	Athina	676	676
2004 Summer	Gold	Athina	664	664
2004 Summer	NA	Athina	11399	0
2004 Summer	Silver	Athina	660	660
2008 Summer	Bronze	Beijing	706	706
2008 Summer	Gold	Beijing	664	664
2008 Summer	NA	Beijing	11367	0
2008 Summer	Silver	Beijing	665	665

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

5) Determinar por equipo olímpico (NOC y región), el número de participaciones, número de medallas conseguidas y la ratio de medallas conseguidas sobre el número de participaciones.

Dimensión	Jerarquía	Operador	Expresión de filtro	Pará...
DIM OLYMPICS	YEAR	Igual	{ 2000, 2004, 2008 }	

NOC	Region	Recuento FACT PARTICIPATIONS	NUM MEDALS	Ratio
AFG	Afghanistan	8	1	12,5
AHO	Curacao	15	0	0
ALB	Albania	24	0	0
ALG	Algeria	176	7	3,97727272727273
AND	Andorra	16	0	0
ANG	Angola	91	0	0
ANT	Antigua	14	0	0
ARG	Argentina	488	120	24,5901639344262
ARM	Armenia	70	6	8,57142857142857
ARU	Aruba	11	0	0
ASA	American Samoa	8	0	0
AUS	Australia	1957	489	24,9872253449157
AUT	Austria	299	15	5,01672240802676
AZE	Azerbaijan	112	15	13,3928571428571
BAH	Bahamas	100	20	20
BAN	Bangladesh	15	0	0
BAR	Barbados	53	1	1,88679245283019
BDI	Burundi	16	0	0
BEL	Belgium	276	15	5,43478260869565
BEN	Benin	13	0	0

Nombre: Ratio

Jerarquía primaria: Measures

Miembro primario:

Expresión: ([Measures].[NUM MEDALS]/[Measures].[Recuento FACT PARTICIPATIONS])*100

Metadatos:

- FACT_Participations
- Measures
- KPI
- DIM ATHLETES
- DIM DISCIPLINES
- DIM MEDALS
- DIM OLYMPICS
- DIM TEAMS COUNTRIES

Funciones:

- (Todos)
- KPI
- Metadata
- Navigation
- Other
- Set
- Statistical
- String
- Time
- UI
- Value

Aceptar Cancelar

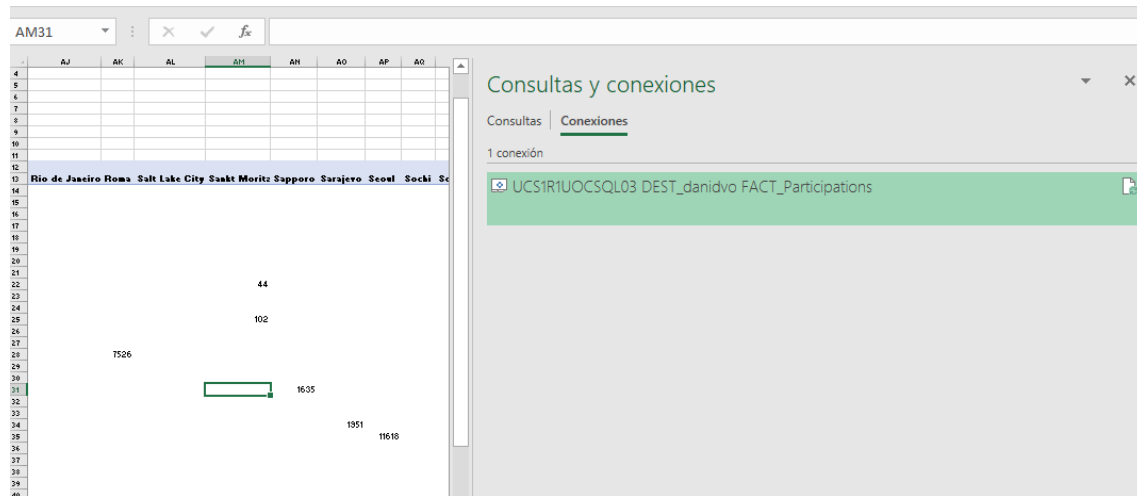
6) Adicionalmente a la realización de consultas mediante el explorador de cubos, a continuación, se solicita reproducir las dos primeras consultas 1) y 2) pero desde la herramienta ofimática Microsoft Excel.

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

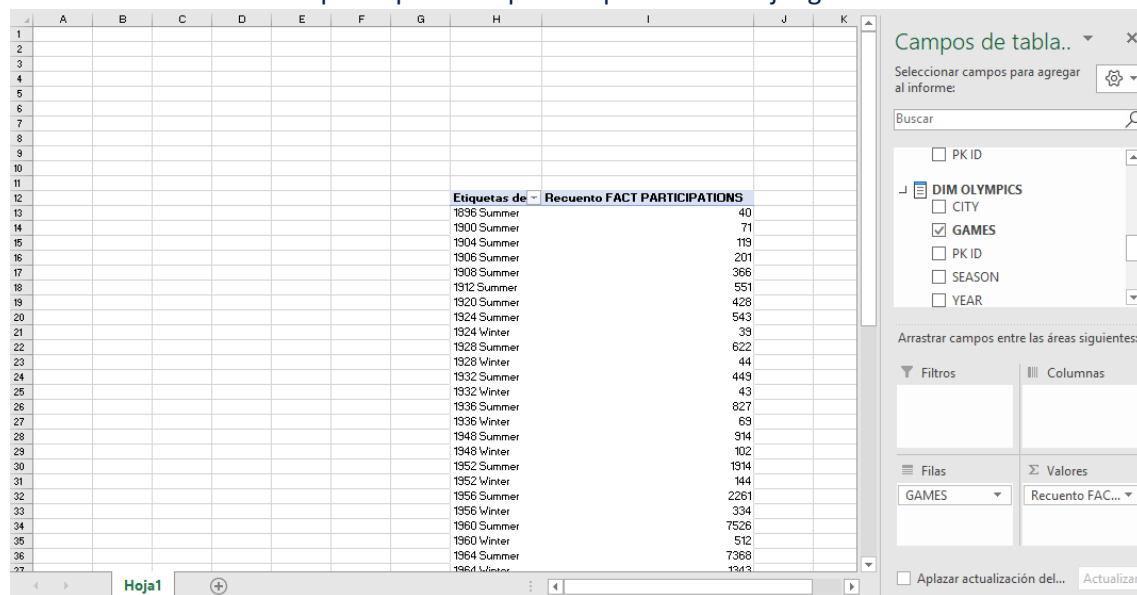
PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

Conexión con Visual Studio



d. A partir de este punto ya se pueden crear informes o gráficos dinámicos, como, por ejemplo, contabilizar el número de participaciones por temporada de los juegos.



Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

e. Construcción de las dos siguientes consultas

1. Listar el número total de participaciones por deporte.

The screenshot shows a data table with two columns: 'Etiquetas de fila' (Row Labels) and 'Recuento FACT PARTICIPATIONS' (Count of FACT PARTICIPATIONS). The table lists various sports and their corresponding participation counts.

Etiquetas de fila	Recuento FACT PARTICIPATIONS
Alpine Skiing	6322
Archery	1883
Art Competitions	28
Athletics	32374
Badminton	1362
Baseball	846
Basketball	3668
Beach Volleyball	538
Biathlon	4615
Bobsleigh	2205
Boxing	4363
Canoeing	5550
Cross Country Skiing	7529
Curling	410
Cycling	7775
Diving	2011
Equestrianism	4553
Fencing	6537
Figure Skating	1512
Football	4443
Freestyle Skiing	913
Golf	108
Gymnastics	18271
Handball	3319
Hockey	4264
Ice Hockey	4573
Judo	3366
Lacrosse	2
Luge	1377
Modern Pentathlon	1259

The 'Campos de tabla' pane on the right shows the 'SPORT' field selected under 'DIM DISCIPLINES' and 'DIM MEDALS'. The 'Filtros' area is empty, and the 'Columnas' area shows 'Recuento FAC...'. The 'Filas' area shows 'SPORT'.

2. Crear una matriz con número total de participaciones distribuido por la ciudad huésped en el eje y, y por el año de celebración de los juegos en el eje x.

The screenshot shows a PivotTable with 'CITY' on the vertical axis and 'YEAR' on the horizontal axis. The table displays participation counts for various cities across different years.

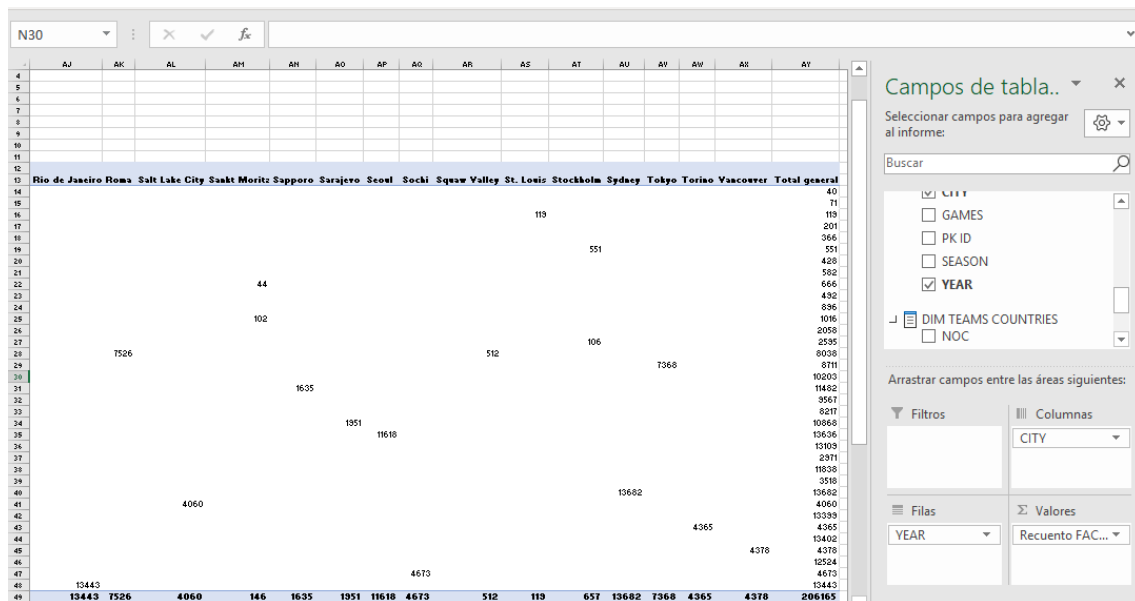
CITY	1996	2000	2004	2008	2012	2016
Amsterdam	40					
Antwerpen						
Athina						
Atlanta						
Barcelona						
Beijing						
Berlin						
Calgary						
Chamonix						
Cortina d'Ampezzo						
1996						
1998						
2000						
2002						
2004						
2006						
2008						
2010						
2012						
2014						
2016						
Total general	2783	622	428	13640	11838	10326

The 'Campos de tabla' pane on the right shows the 'CITY' field selected under 'DIM TEAMS COUNTRIES' and 'YEAR' field selected under 'DIM DISCIPLINES'. The 'Filtros' area is empty, and the 'Columnas' area shows 'CITY'. The 'Filas' area shows 'YEAR'.

Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

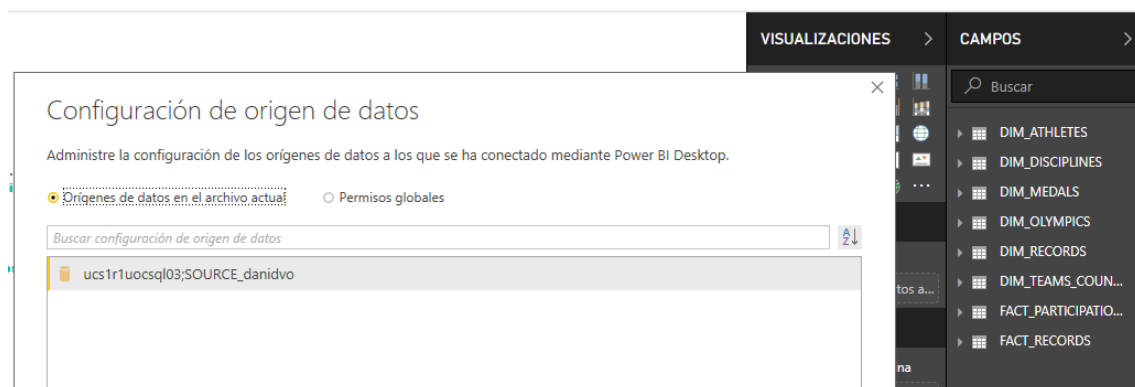
PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia



2.4. Ejercicio 3. Análisis de datos mediante Dashboards (35%)

2.4.1. Configurar un origen de datos (DataSource) conectado a SOURCE_loginuoc



Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

PowerBI-PR4 - Power BI Desktop

InicioModeladoAyuda

Archivos

Obtener datos

Orígenes recientes

Especificar datos

Editar consultas

Actualizar

Nueva página

Nuevo objeto visual

Hacer una pregunta

Botones

Formas

Cuadro de texto

Imagen

Objetos visuales personalizados

Desde Marketplace

De archivo

Cambiar tema

Administrar relaciones

Nueva medida

Nueva columna

Nueva medida rápida

Publicar

Inicio sesión

Portapapeles

Datos externos

Insertar

Objetos visuales personalizados

Temas

Relaciones

Cálculos

Compartir

PK_ID	NAME	SEX	HEIGHT	WEIGHT	AGE
1713	Franck Adisson	M	180	70	23
2723	Kamel Ait Daoud	M	180	70	23
3267	Hamdan Awadh Al-Bishi	M	180	70	23
5860	Prince Amara	M	180	70	23
8049	Martin Ansa Ortiz	M	180	70	23
8112	Jnos Antal	M	180	70	23
9380	Ainsley Edward Armstrong	M	180	70	23
9381	Ainsley Edward Armstrong	M	180	70	23
18206	Charles Wade Bell	M	180	70	23
20115	Eckart Berkes	M	180	70	23
27296	KC Boutlette	M	180	70	23
27901	Kenneth Braaten	M	180	70	23
29538	Nejc Brodar	M	180	70	23
29539	Nejc Brodar	M	180	70	23
29540	Nejc Brodar	M	180	70	23
29541	Nejc Brodar	M	180	70	23
32208	David George Brownlow Cecil Burghley	M	180	70	23
32209	David George Brownlow Cecil Burghley	M	180	70	23
32246	Leonid Iosypovych Buriak	M	180	70	23
34065	Juan ngel Calzado de Castro	M	180	70	23
34292	Pablo Jos Campana Saenz	M	180	70	23
35032	Giulio Capitano	M	180	70	23
35033	Giulio Capitano	M	180	70	23
35034	Giulio Capitano	M	180	70	23
37727	Daniel Cerisey	M	180	70	23
37728	Daniel Cerisey	M	180	70	23
43304	Casey Colby	M	180	70	23
43305	Casey Colby	M	180	70	23
43306	Casey Colby	M	180	70	23
44244	Walter Convents	M	180	70	23
44245	Walter Convents	M	180	70	23
49172	Emmerich Danzer	M	180	70	23
49794	Alejandro Davila McAllister	M	180	70	23
55856	Euripidis Dimosthenous	M	180	70	23
74985	Florentin Mihai Galiceanu	M	180	70	23
74986	Florentin Mihai Galiceanu	M	180	70	23

CAMPOS

Buscar

DIM_ATHLETES

DIM_DISCIPLINES

DIM_MEDALS

DIM_OLYMPICS

DIM_RECORDS

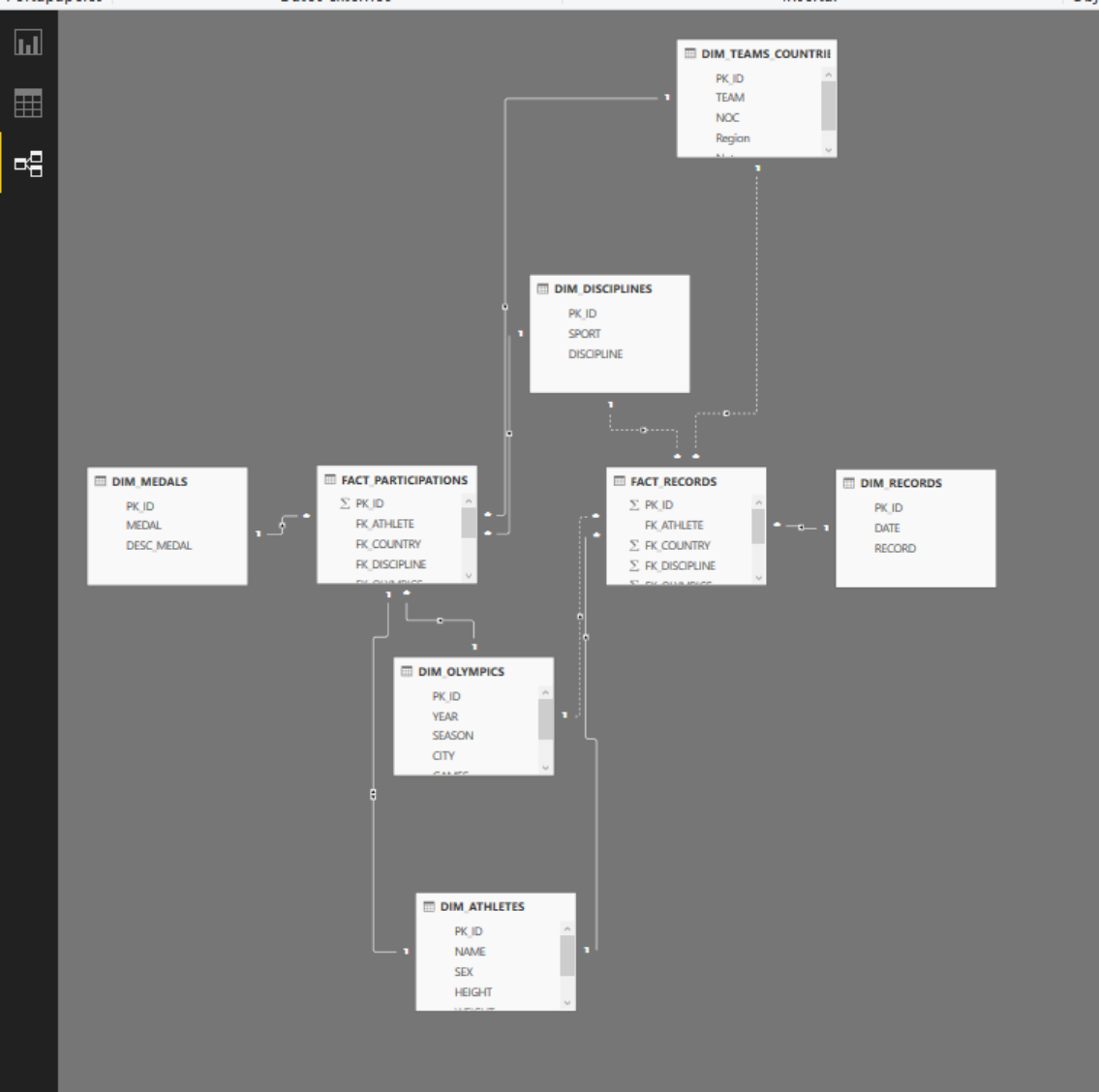
DIM_TEAMS_COUNTRIES

FACT_PARTICIPATIONS

FACT_RECORDS

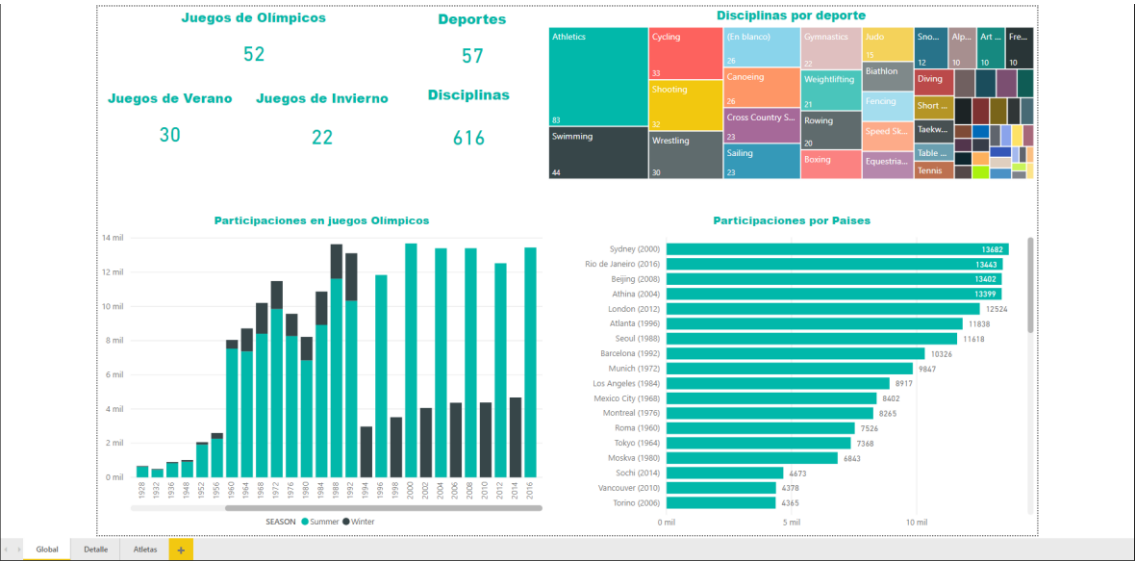
TABLA: DIM_ATHLETES (206.165 filas)

ACTUALIZACIÓN DISPONIBLE (CLIC PARA DESCARGARLA)

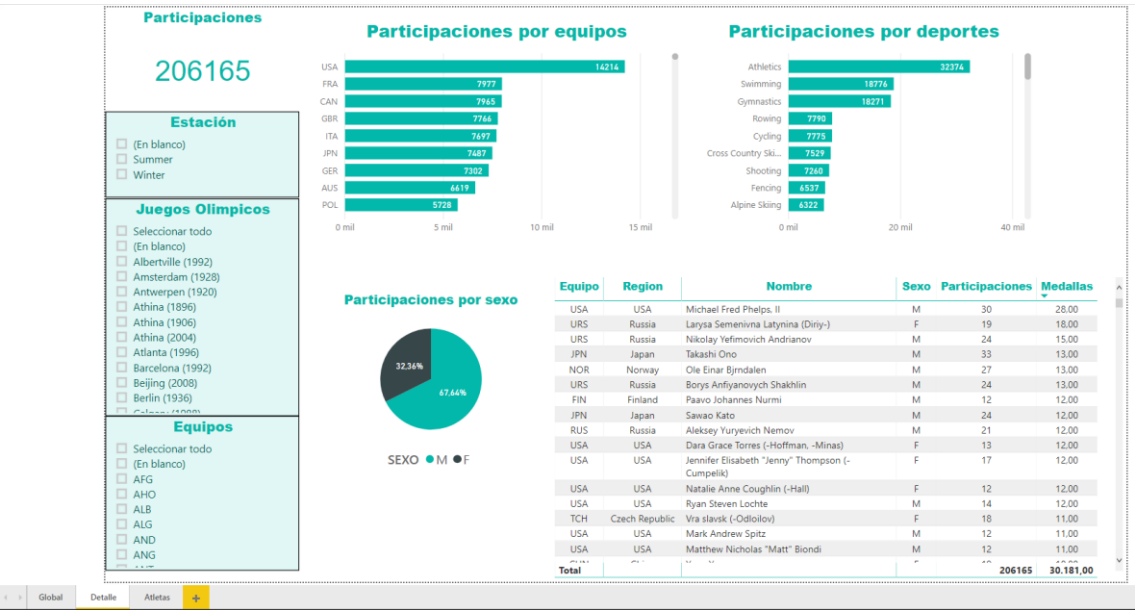


CAMPOS	
<div>Buscar</div>	
▲	DIM_ATHLETES
■	AGE
■	HEIGHT
■	NAME
■	PK_ID
■	SEX
■	WEIGHT
▲	DIM_DISCIPLINES
■	DISCIPLINE
■	PK_ID
■	SPORT
▲	DIM_MEDALS
■	DESC_MEDAL
■	MEDAL
■	PK_ID
▲	DIM_OLYMPICS
■	CITY
■	ConcatenatedColumn
■	GAMES
■	PK_ID
■	SEASON
■	YEAR
▲	DIM_RECORDS
■	DATE
■	PK_ID
■	RECORD
▲	DIM_TEAMS_COUNTRIES
■	NOC
■	Notes
■	PK_ID
■	Region
■	TEAM
▲	FACT PARTICIPATIONS

2.4.2. Construcción de la pestaña Global



2.4.3. Construcción de la pestaña Detalle



Caso práctico: Almacén de datos para el análisis de los resultados de los Juegos Olímpicos

PR4 – Análisis de datos

ALUMNO: Daniel Vargas Olivencia

2.4.4. Construcción de la pestaña Atletas

Estación	Sexo	Id. Inscripción	Equipo	Region	nombre	sexo	edad	altura	peso	medalla	disciplina
<input type="checkbox"/> (En blanco)	<input type="checkbox"/> (En blanco)	1	CHN	China	A. Djiang	M	24.00	180.00	80.00	No se ha conseguido medalla	Basketball Men's Basketball
<input type="checkbox"/> Summer	<input type="checkbox"/> F	2	CHN	China	A. Lamusi	M	23.00	170.00	60.00	No se ha conseguido medalla	Judo Men's Extra-Lightweight
<input type="checkbox"/> Winter	<input type="checkbox"/> M	3	NED	Netherlands	Christine Jacobs Aaftink	F	21.00	185.00	82.00	No se ha conseguido medalla	Speed Skating Women's 500 metres
Juegos Olímpicos		4	NED	Netherlands	Christine Jacobs Aaftink	F	21.00	185.00	82.00	No se ha conseguido medalla	Speed Skating Women's 1,000 metres
<input type="checkbox"/> Seleccionar todo		5	NED	Netherlands	Christine Jacobs Aaftink	F	25.00	182.00	82.00	No se ha conseguido medalla	Speed Skating Women's 500 metres
<input type="checkbox"/> (En blanco)		6	NED	Netherlands	Christine Jacobs Aaftink	F	25.00	185.00	82.00	No se ha conseguido medalla	Speed Skating Women's 1,000 metres
<input type="checkbox"/> Albertville (1992)		7	NED	Netherlands	Christine Jacobs Aaftink	F	27.00	183.00	82.00	No se ha conseguido medalla	Speed Skating Women's 500 metres
<input type="checkbox"/> Amsterdam (1928)		8	NED	Netherlands	Christine Jacobs Aaftink	F	27.00	185.00	82.00	No se ha conseguido medalla	Speed Skating Women's 1,000 metres
<input type="checkbox"/> Antwerpen (1920)		9	USA	USA	Per Knut Aaland	M	31.00	188.00	75.00	No se ha conseguido medalla	Cross Country Skiing Men's 10 kilometres
<input type="checkbox"/> Athina (1896)		10	USA	USA	Per Knut Aaland	M	31.00	188.00	75.00	No se ha conseguido medalla	Cross Country Skiing Men's 50 kilometres
<input type="checkbox"/> Athina (1906)		11	USA	USA	Per Knut Aaland	M	31.00	188.00	75.00	No se ha conseguido medalla	Cross Country Skiing Men's 10/15 kilometres Pursuit
Equipos		12	USA	USA	Per Knut Aaland	M	31.00	188.00	75.00	No se ha conseguido medalla	Cross Country Skiing Men's 4 x 10 kilometres Relay
<input type="checkbox"/> Seleccionar todo		13	USA	USA	Per Knut Aaland	M	33.00	188.00	75.00	No se ha conseguido medalla	Cross Country Skiing Men's 10 kilometres
<input type="checkbox"/> (En blanco)		14	USA	USA	Per Knut Aaland	M	33.00	188.00	75.00	No se ha conseguido medalla	Cross Country Skiing Men's 30 kilometres
<input type="checkbox"/> AFG		15	USA	USA	Per Knut Aaland	M	33.00	188.00	75.00	No se ha conseguido medalla	Cross Country Skiing Men's 10/15 kilometres Pursuit
<input type="checkbox"/> AHO		16	USA	USA	Per Knut Aaland	M	33.00	188.00	75.00	No se ha conseguido medalla	Cross Country Skiing Men's 4 x 10 kilometres Relay
<input type="checkbox"/> ALB		17	USA	USA	John Aalberg	M	31.00	183.00	72.00	No se ha conseguido medalla	Cross Country Skiing Men's 10 kilometres
<input type="checkbox"/> AUG		18	USA	USA	John Aalberg	M	31.00	183.00	72.00	No se ha conseguido medalla	Cross Country Skiing Men's 50 kilometres
Deporte		19	USA	USA	John Aalberg	M	31.00	183.00	72.00	No se ha conseguido medalla	Cross Country Skiing Men's 10/15 kilometres Pursuit
<input type="checkbox"/> Seleccionar todo		20	USA	USA	John Aalberg	M	31.00	183.00	72.00	No se ha conseguido medalla	Cross Country Skiing Men's 4 x 10 kilometres Relay
<input type="checkbox"/> (En blanco)		21	USA	USA	John Aalberg	M	33.00	183.00	72.00	No se ha conseguido medalla	Cross Country Skiing Men's 10 kilometres
<input type="checkbox"/> Alpine Skiing		22	USA	USA	John Aalberg	M	33.00	183.00	72.00	No se ha conseguido medalla	Cross Country Skiing Men's 30 kilometres
<input type="checkbox"/> Archery		23	USA	USA	John Aalberg	M	33.00	183.00	72.00	No se ha conseguido medalla	Cross Country Skiing Men's 10/15 kilometres Pursuit
<input type="checkbox"/> Art Competitions		24	USA	USA	John Aalberg	M	33.00	183.00	72.00	No se ha conseguido medalla	Cross Country Skiing Men's 4 x 10 kilometres Relay
<input type="checkbox"/> Athletics		25	FIN	Finland	Antti Sami Aalto	M	26.00	186.00	96.00	No se ha conseguido medalla	Ice Hockey Men's Ice Hockey
Disciplina		26	FIN	Finland	Jorma Ilmarinen Aalto	M	22.00	182.00	785.00	No se ha conseguido medalla	Cross Country Skiing Men's 30 kilometres
<input type="checkbox"/> Seleccionar todo		27	FIN	Finland	Jari Tapio Aalto	M	31.00	172.00	70.00	No se ha conseguido medalla	Badminton Men's Singles
<input type="checkbox"/> (En blanco)		28	FIN	Finland	Minna Maarit Aalto	F	30.00	159.00	555.00	No se ha conseguido medalla	Sailing Women's Windsurfer
<input type="checkbox"/> 1,500 metres		29	FIN	Finland	Minna Maarit Aalto	F	34.00	159.00	555.00	No se ha conseguido medalla	Sailing Women's Windsurfer
<input type="checkbox"/> 10,000 metres		30	FIN	Finland	Piyo Hannele Aalto (Mattila)	F	32.00	171.00	65.00	No se ha conseguido medalla	Biatlon Women's 7.5 kilometres Sprint
<input type="checkbox"/> 100 metres		31	FIN	Finland	Juhani Tapio Aaltonen	M	28.00	184.00	85.00	Medalla de Bronce	Ice Hockey Men's Ice Hockey
<input type="checkbox"/> 100 metres hurdles		32	FIN	Finland	Paavo Johannes Aaltonen	M	28.00	175.00	64.00	Medalla de Bronce	Gymnastics Men's Individual All-Around
		33	FIN	Finland	Paavo Johannes Aaltonen	M	28.00	175.00	64.00	Medalla de Oro	Gymnastics Men's Team All-Around
		34	FIN	Finland	Paavo Johannes Aaltonen	M	28.00	175.00	64.00	No se ha conseguido medalla	Gymnastics Men's Floor Exercise
		35	FIN	Finland	Paavo Johannes Aaltonen	M	28.00	175.00	64.00	Medalla de Oro	Gymnastics Men's Horse Vault
		36	FIN	Finland	Paavo Johannes Aaltonen	M	28.00	175.00	64.00	No se ha conseguido medalla	Gymnastics Men's Parallel Bars
		37	FIN	Finland	Paavo Johannes Aaltonen	M	28.00	175.00	64.00	No se ha conseguido medalla	Gymnastics Men's Horizontal Bar
		38	FIN	Finland	Paavo Johannes Aaltonen	M	28.00	175.00	64.00	No se ha conseguido medalla	Gymnastics Men's Rings