

Case study

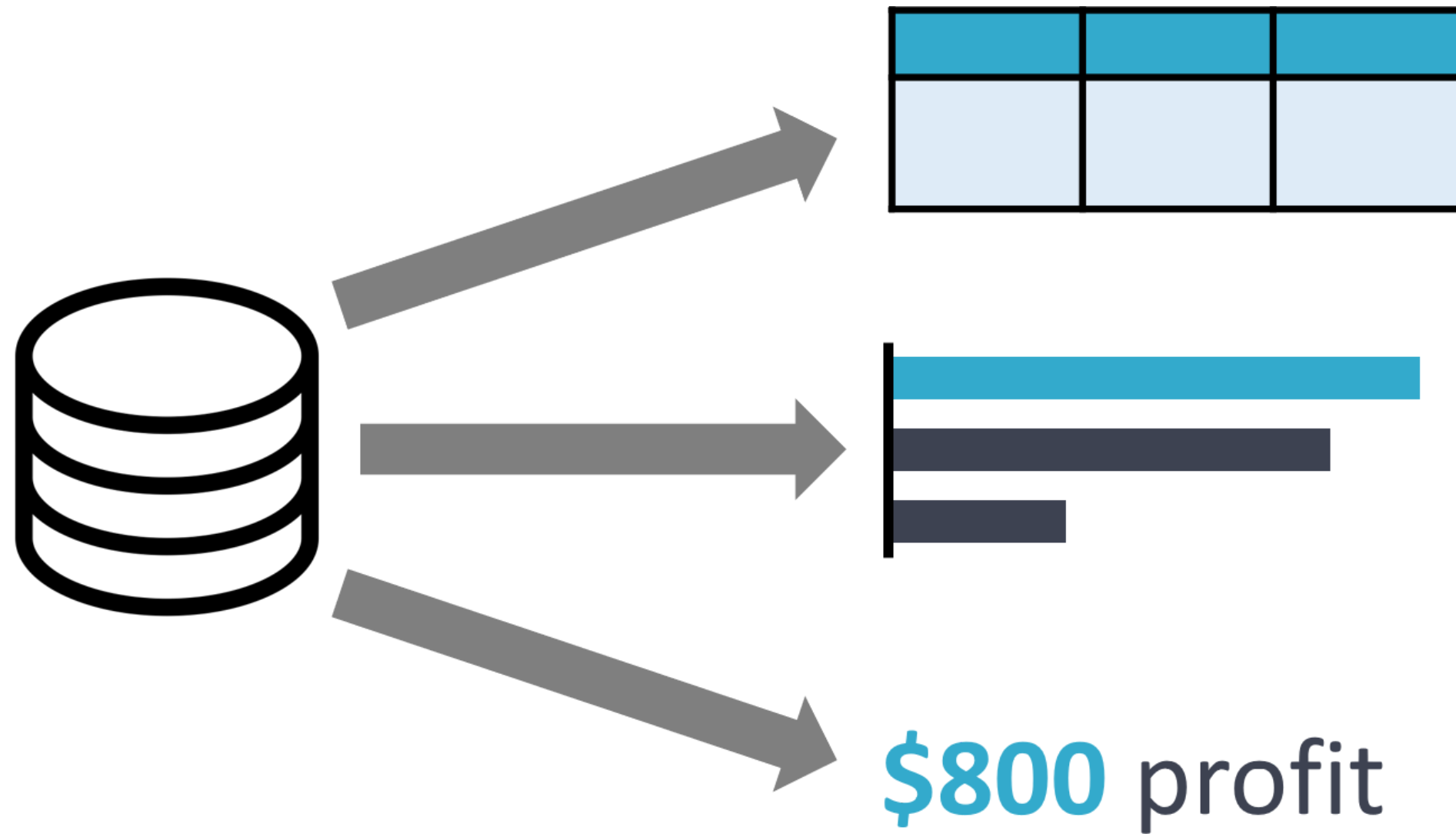
REPORTING IN SQL



Tyler Pernes

Learning & Development Consultant

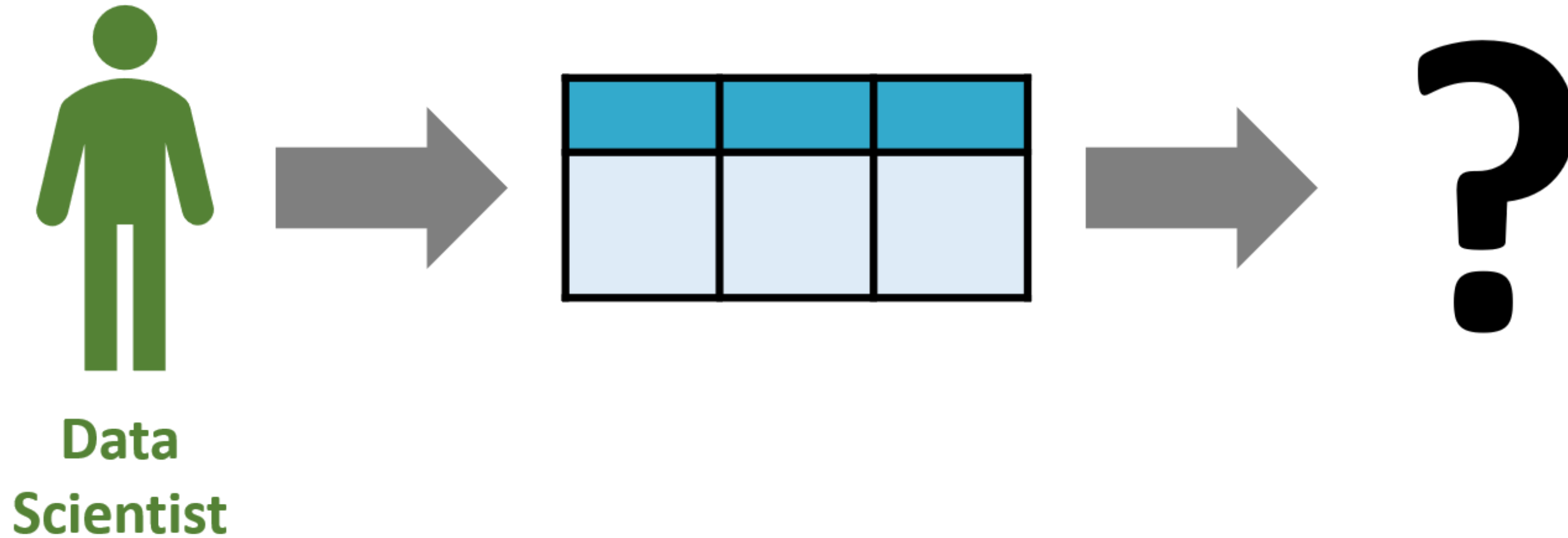
What is reporting?



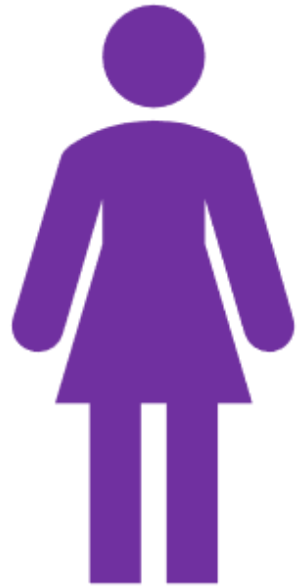
What is reporting?



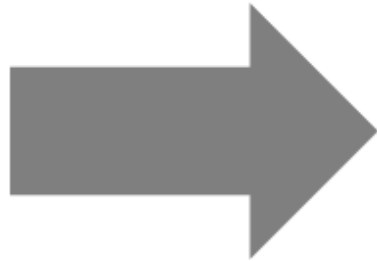
What is reporting?

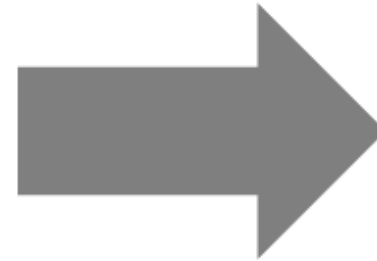


What is reporting?

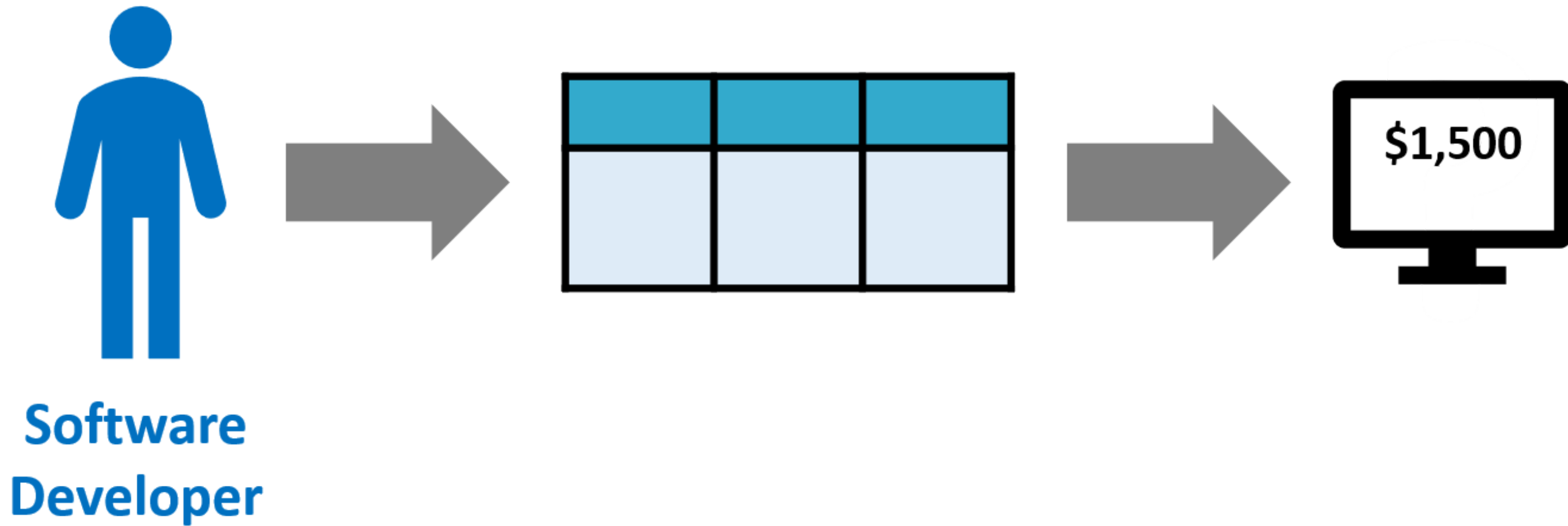


**Visualization
Specialist**





What is reporting?



SQL and reporting



Course goals

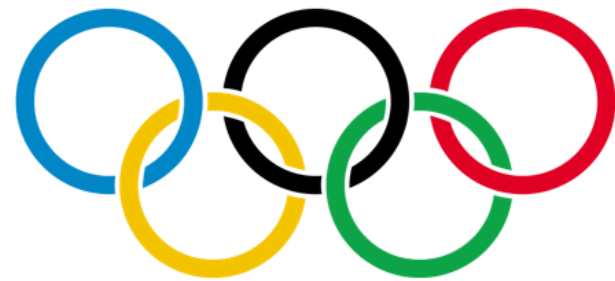
- Create **real-world reports**
- Use **real-world data**
- Overcome **real-world obstacles**

Case study



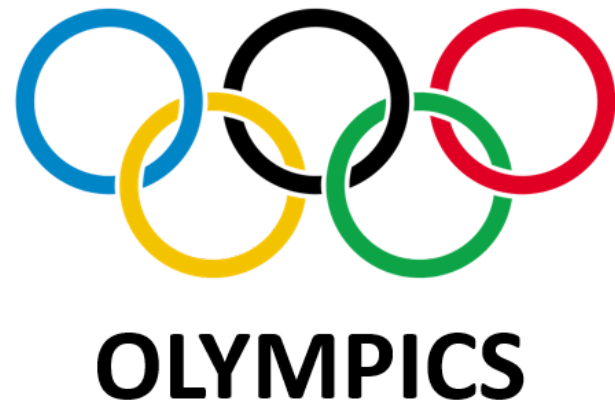
ONLY VIPs

Case study



OLYMPICS

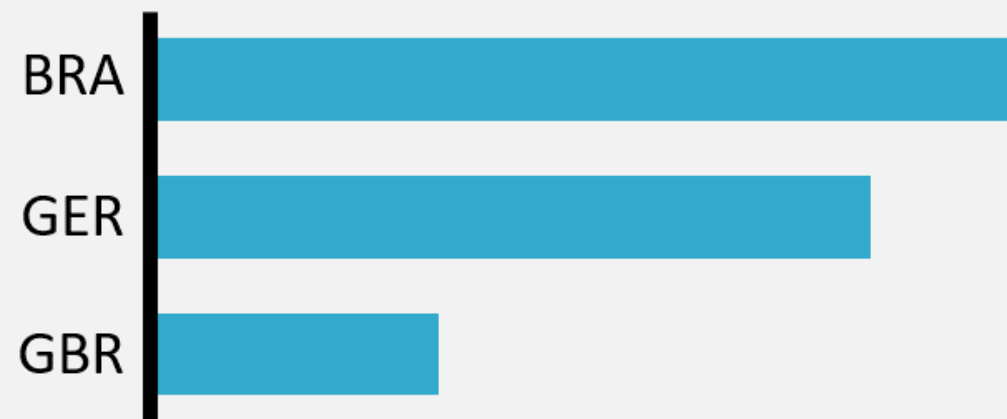
Case study



What is a dashboard?

2018 International Soccer Tournament

Most wins by country



Game Log

Home	Away	Score
GER	MEX	3 - 2
GBR	BRA	1 - 2
USA	JAP	3 - 4

What is a base report?

BASE REPORT

Country	Wins
BRA	12
GER	11
GBR	5



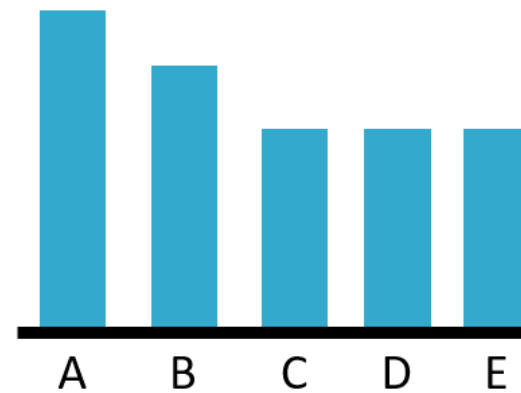
VISUALIZATION



```
SELECT
  country,
  SUM(wins) AS wins
FROM game_log
GROUP BY country
ORDER BY wins DESC
LIMIT 3;
```

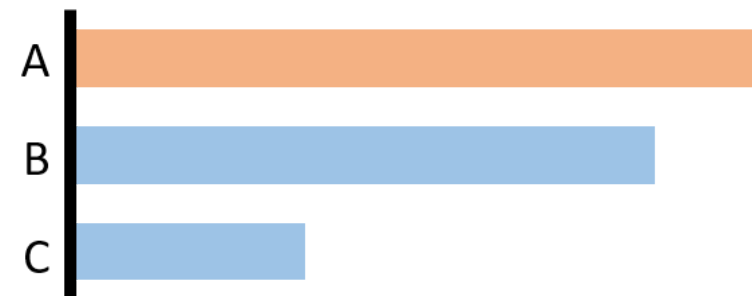
Your dashboard

Most Decorated Athletes



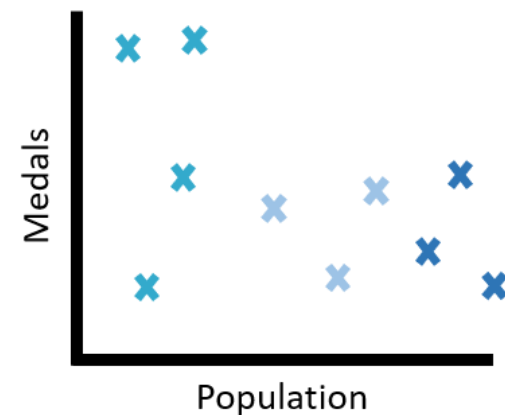
Top Athletes in Nobel-Prized Countries

By Gender



Medal vs Population Rate

By country



Average Tallest Athlete and GDP

By region

region	avg_tallest	perc_world_gdp
A	172	45%
B	168	12%
C	155	18%

Querying time!

REPORTING IN SQL

The Olympics dataset

REPORTING IN SQL



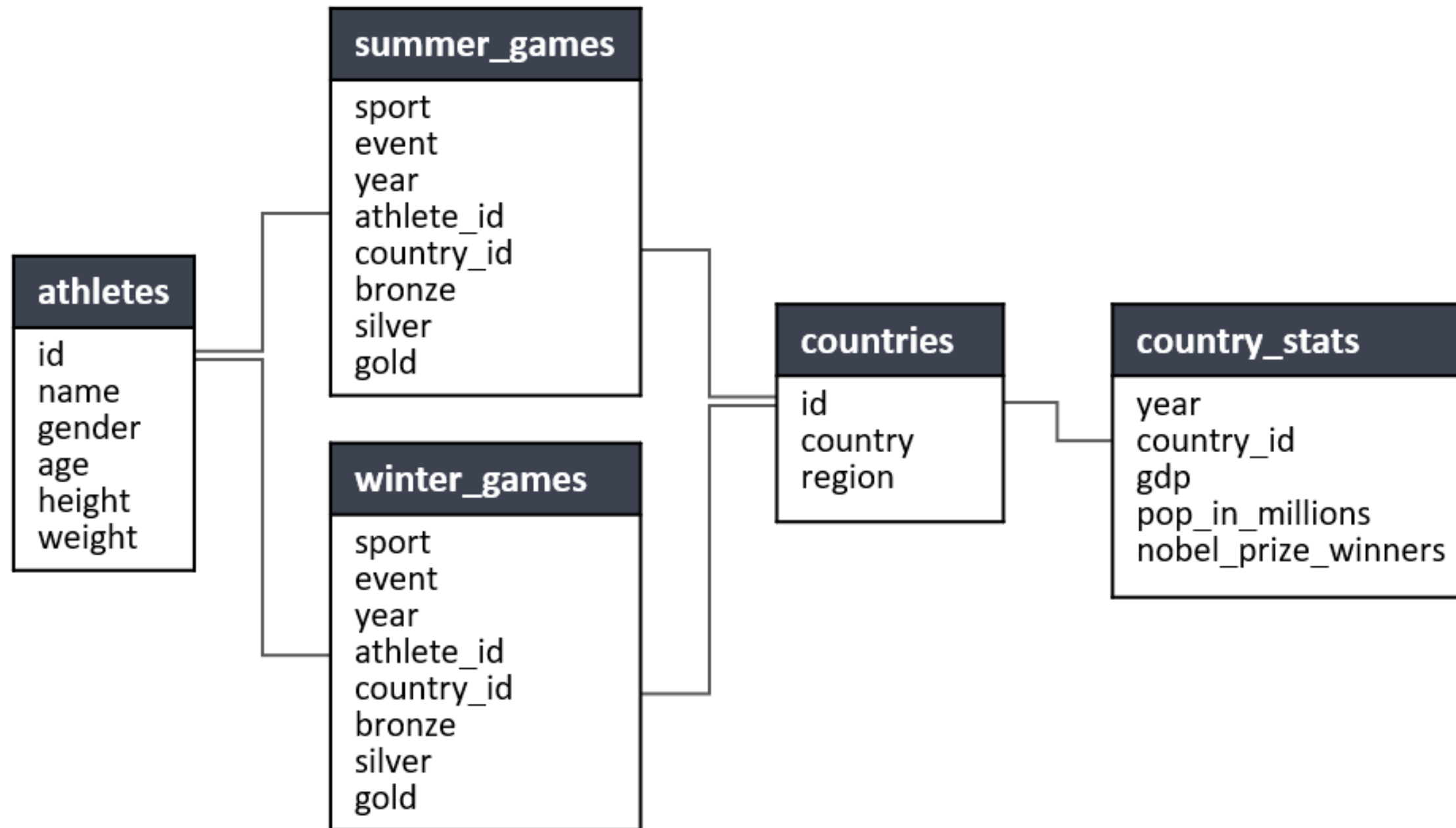
Tyler Pernes

Learning & Development Consultant

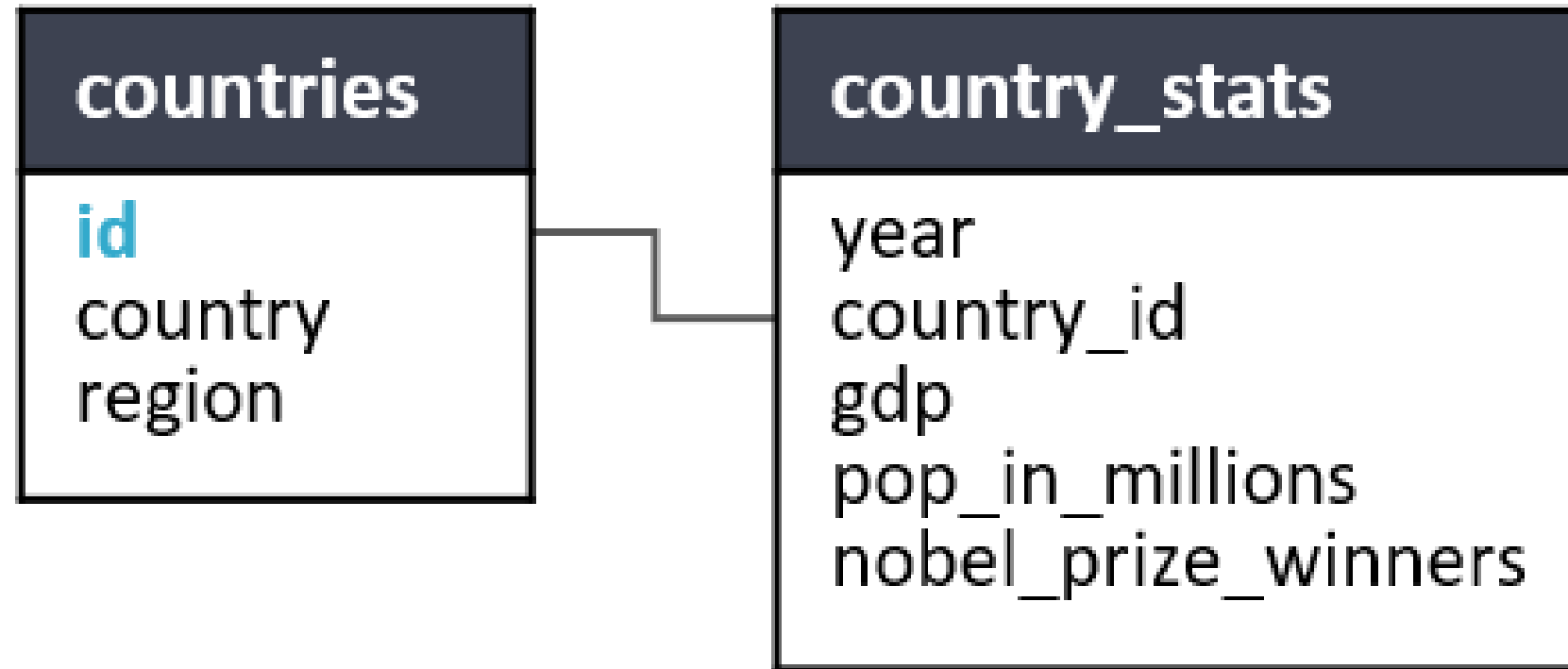
Entity-Relationship (E:R) diagram

- Visual representation of **database structure**
- Lays out **tables** and **fields**
- Identifies **relationships**

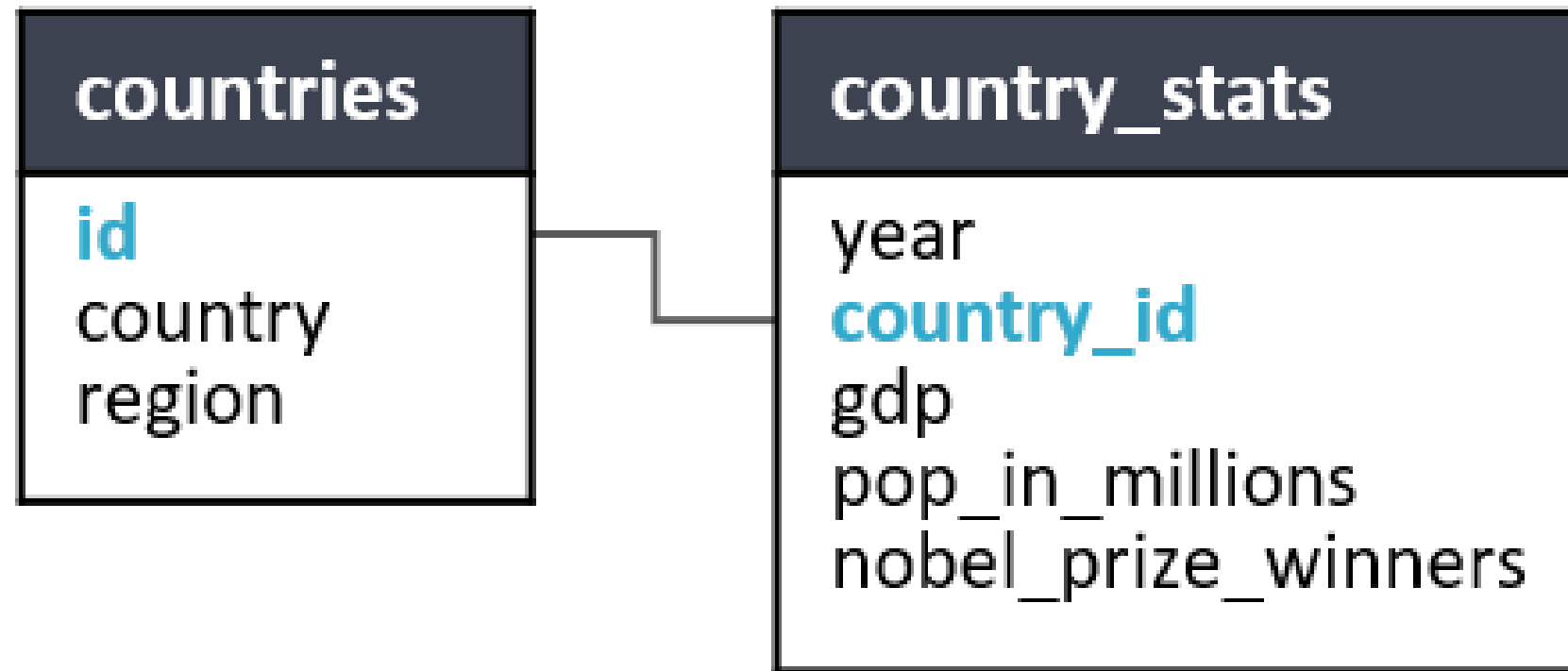
Olympics dataset E:R diagram



Olympics dataset E:R diagram



Olympics dataset E:R diagram



```
SELECT *  
FROM countries AS c1  
JOIN country_stats AS c2  
ON c1.id = c2.country_id;
```

Tables and fields

athletes
id
name
gender
age
height
weight

Tables and fields

summer_games	winter_games
sport	sport
event	event
year	year
athlete_id	athlete_id
country_id	country_id
bronze	bronze
silver	silver
gold	gold

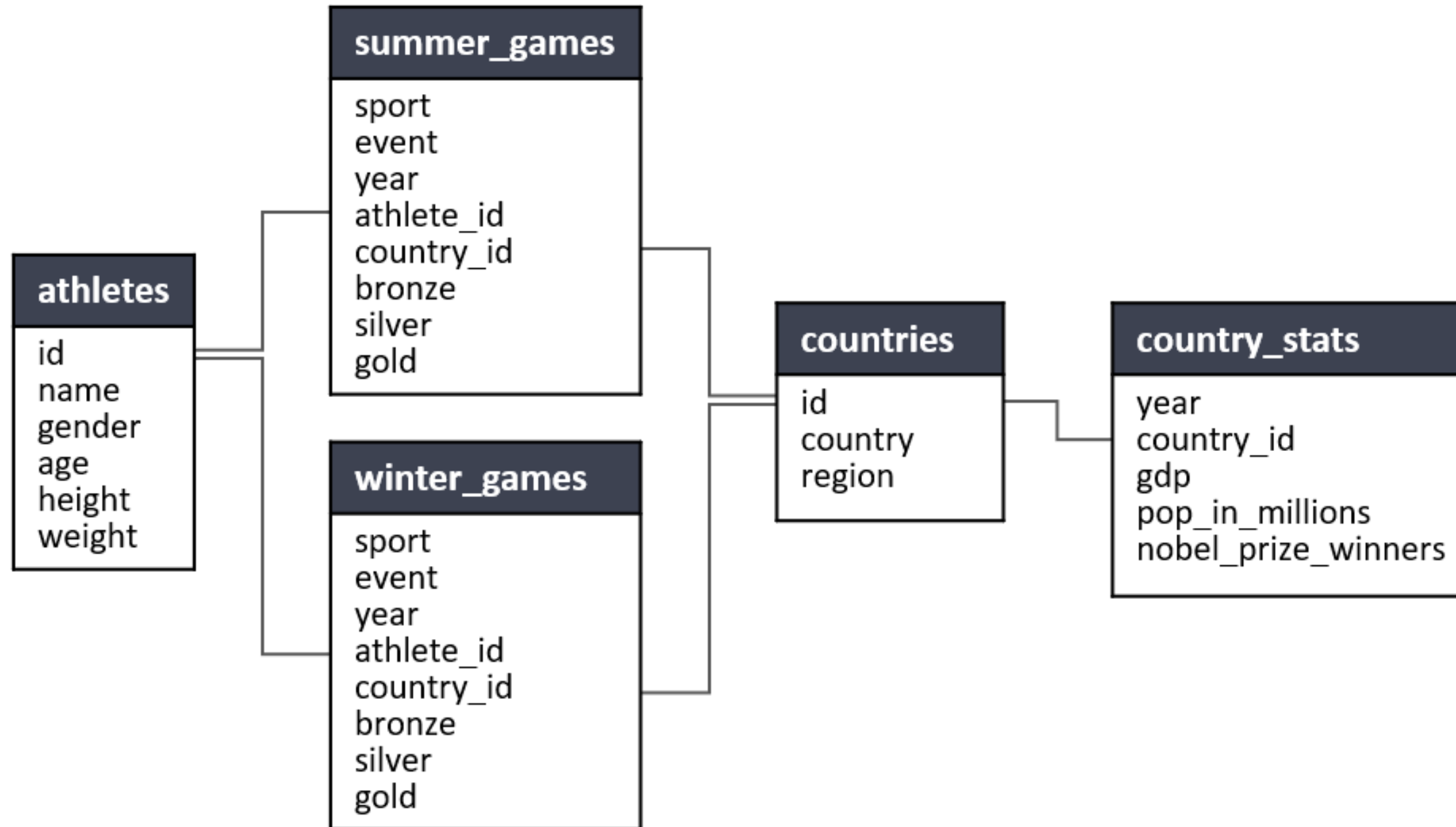
Tables and fields

countries
id
country
region

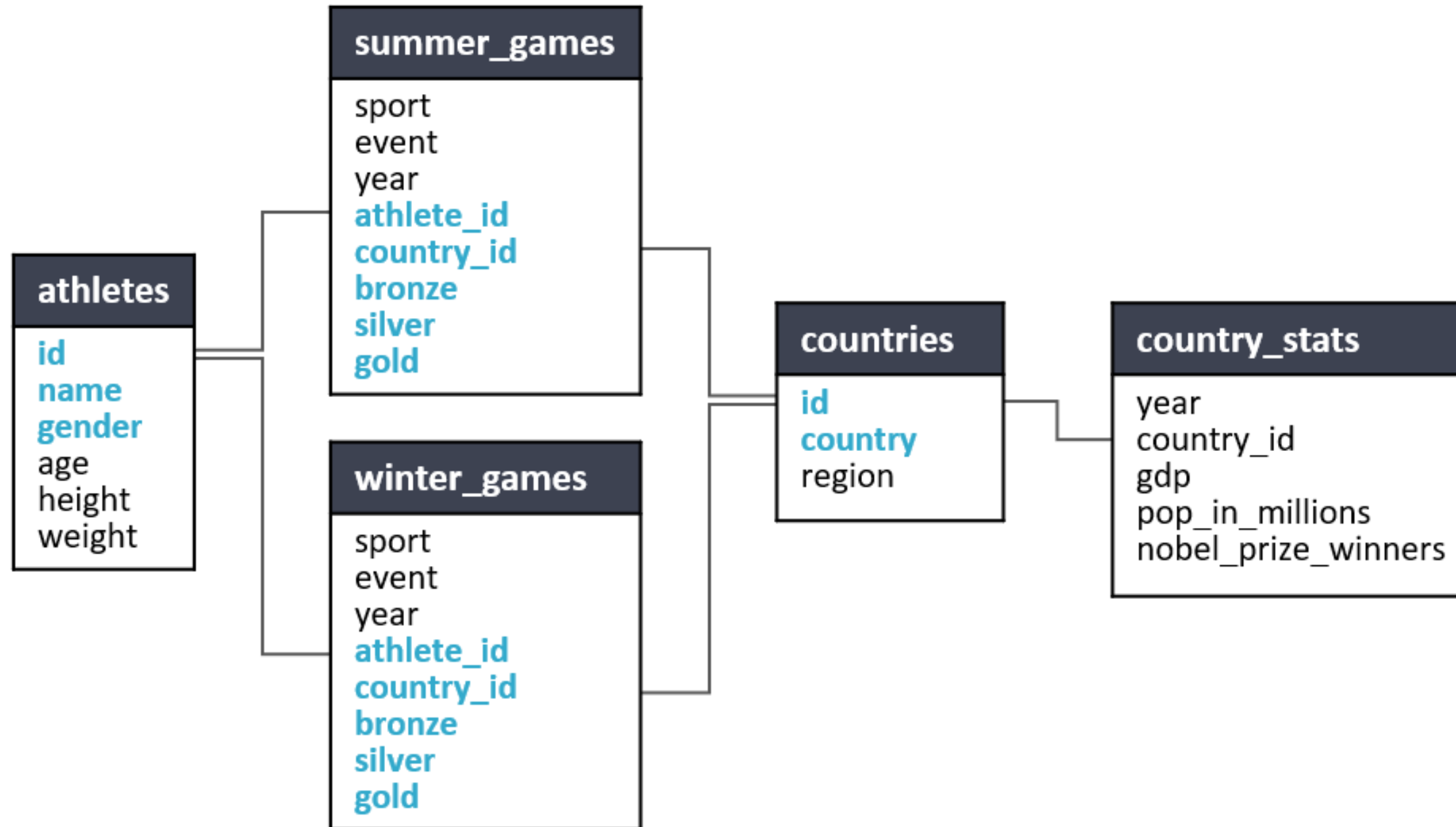
Tables and fields

country_stats
year
country_id
gdp
pop_in_millions
nobel_prize_winners

Reference for queries



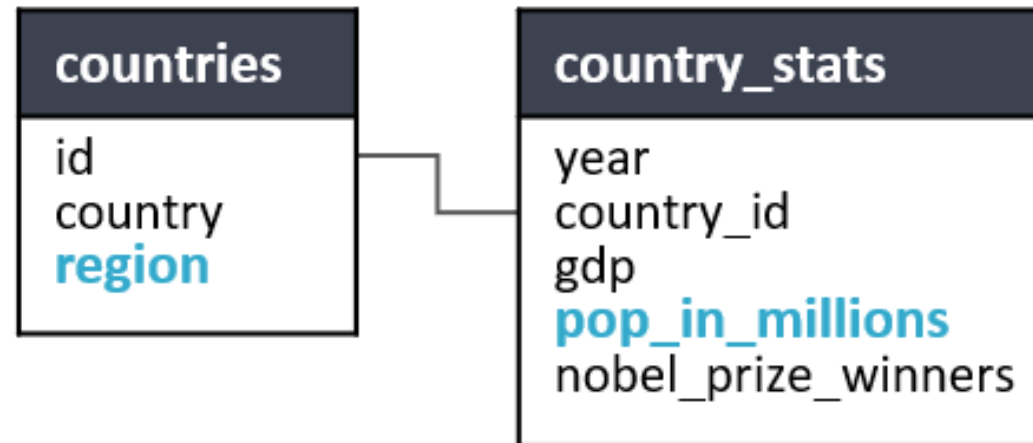
Reference for queries



Reference for queries

Population by Region

region	pop_in_millions
A	172
B	168
C	155



```
SELECT
    region,
    SUM(pop_in_millions) AS pop_in_millions
FROM countries AS c1
JOIN country_stats AS c2
ON c1.id = c2.country_id
GROUP BY region;
```

Practice time!

REPORTING IN SQL

Exploring our data

REPORTING IN SQL




Tyler Pernes


Learning & Development Consultant

Exploring with the console

query.sql

1



 Run Code Submit Answer

query result

countries

summer_games

athletes

country_stats

▼

id	country	region
1	AFG - Afghanistan	ASIA (EX. NEAR EAST)
2	ALB - Albania	EASTERN EUROPE
3	ALG - Algeria	NORTHERN AFRICA

Exploring with the console

- May show an **inaccurate picture**

```
preview of table: summer_games
```

```
+-----+-----+-----+
| sport      | event                                     | bronze |
+-----+-----+-----+
| Gymnastics | Gymnastics Men's Individual All-Around | null   |
| Gymnastics | Gymnastics Men's Floor Exercise         | null   |
| Gymnastics | Gymnastics Men's parallel Bars          | null   |
+-----+-----+-----+
```

Exploring with the console

- No insight into **distributions**

```
preview of table: clients
```

```
+-----+
| country_of_client |
+-----+
| United States     |
| United States     |
| Mexico            |
| United States     |
| Canada            |
| Canada            |
+-----+
```


Exploring with queries

```
SELECT DISTINCT region
FROM countries;
```

```
+-----+
| region          |
+-----+
| WESTERN EUROPE  |
| null            |
| C.W. IF IND. STATES |
| OCEANIA         |
| NEAR EAST       |
| SUB-SAHARAN AFRICA |
+-----+
```

Exploring with queries

```
SELECT region
FROM countries
GROUP BY region;
```

```
+-----+
| region                |
+-----+
| WESTERN EUROPE        |
| null                  |
| C.W. IF IND. STATES   |
| OCEANIA               |
| NEAR EAST             |
| SUB-SAHARAN AFRICA    |
+-----+
```

Field-level aggregations

```
SELECT region, COUNT(*) AS row_num
FROM countries
GROUP BY region
ORDER BY row_num DESC;
```

region	row_num
SUB-SAHARAN AFRICA	49
LATIN AMER. & CARIB	38
ASIA (EX. NEAR EAST)	26
WESTERN EUROPE	23
OCEANIA	15
EASTERN EUROPE	15

Field-level aggregations

```
SELECT revenue_source, SUM(revenue) AS revenue
FROM orders
GROUP BY revenue_source
ORDER BY revenue DESC;
```

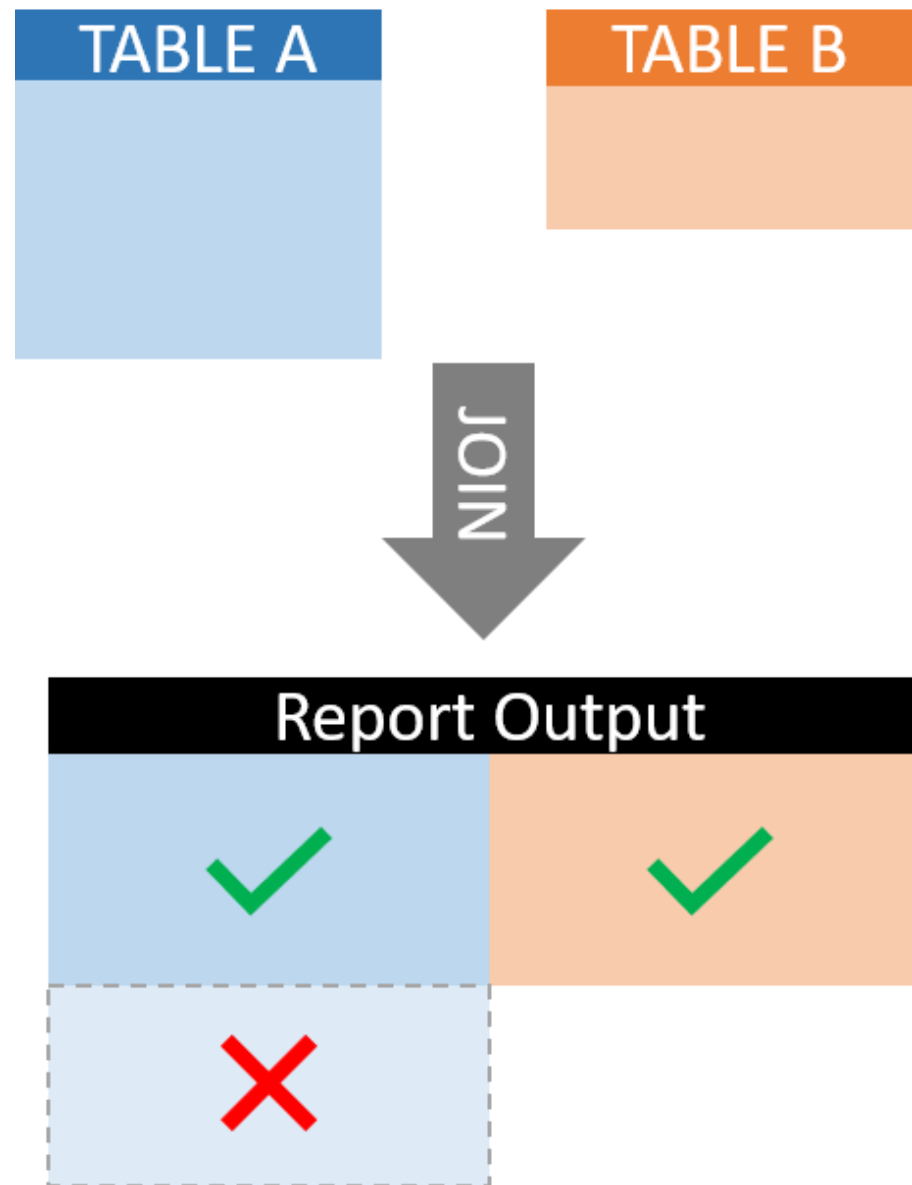
```
+-----+-----+
| revenue_source | revenue |
+-----+-----+
| Olympics      | 122000  |
| NFL           | 80500   |
| MLB           | 300     |
| NBA           | 220     |
| NCAAF         | 120     |
| NCAAB         | 90      |
+-----+-----+
```

Table-level aggregations

```
SELECT COUNT(*)  
FROM country_stats;
```

```
+-----+  
| count |  
+-----+  
| 3451  |  
+-----+
```

Query validation



Query validation

QUERY:

```
SELECT SUM(rev) AS revenue
FROM
  (SELECT country, SUM(rev) AS rev
   FROM orders AS o
   JOIN countries AS c
   ON o.country_id = c.id
   GROUP BY country);
```

```
+-----+
| revenue |
+-----+
| 50.00   |
+-----+
```

ORIGINAL TABLE:

```
SELECT SUM(rev) AS revenue
FROM orders;
```

```
+-----+
| revenue |
+-----+
| 500.00  |
+-----+
```

- Our Query (Left) = **\$50**
- Original Table (Right) = **\$500**
- Lost **90%** of revenue from **JOIN**

Query validation

QUERY:

```
SELECT SUM(rev) AS revenue
FROM
  (SELECT country, SUM(rev) AS rev
   FROM orders AS o
   JOIN country_stats AS cs
   ON o.country_id = cs.country_id
   GROUP BY country);
```

```
+-----+
| revenue |
+-----+
| 5,000.00 |
+-----+
```

ORIGINAL TABLE:

```
SELECT SUM(rev) AS revenue
FROM orders;
```

```
+-----+
| revenue |
+-----+
| 500.00  |
+-----+
```

- Our Query (Left) = **\$5,000**
- Original Table (Right) = **\$500**
- **10x duplication** from JOIN

Let's explore!

REPORTING IN SQL