

v. 1.0

Athletic Competitions

DOCUMENT ANALYSIS
ROCCO CALIANDRO

Sommario

CONCEPTUAL DESIGN	2
Requirements gathering.....	2
Requirements analysis.....	2
Choose the right level of abstraction	2
Linearize phrases and divide those articulated	2
Making explicit reference between terms	2
Create a glossary and identify synonyms	2
Reorganizing for keyword phrases	3
Project strategy: Hybrid design	3
Constraints.....	5
Operations on data.....	6
LOGICAL DESIGN	6
Logical Model.....	7
Data type	7
Volume Table.....	8
Operations Table	8
PHYSICAL DESIGN	8
Database technology	8
Optimization of operations	8
Result	9
BUSINESS INTELLIGENCE	10
Data model	11
Dashboard	11

CONCEPTUAL DESIGN

Requirements gathering

Each athlete is assigned a number before the start of the competition. This number is never changed. We must be able to provide daily a list of the competitions that will take place during the day indicating the time, the specialty (one hundred meters, high jump, etc.) and, where relevant, the level of competition (semi-final, final , etc...). Participants in a race can be individual athletes or teams (composed of athletes). The system must keep the information relating to the winners of the various competitions with the corresponding results. We need to know who is registered for the various competitions, and for each competition in which order and how many participants participate. For each specialty, we must have access to world records. For each participating country, it is necessary to know the name, the flag, a brief historical review, and as much information as possible relating to the athletes of that country participating in the competitions (personal data, disciplines, records, ...). The results of the competitions must be stored in the database so that it is possible to consult the classification of the participants in a specific competition after this has taken place. In particular, it must be possible to know who won the gold, silver, and bronze medals in each specialty.

Requirements analysis

Choose the right level of abstraction

- Use of *“records”* rather than *“world records”*
- Use of *“Country”* rather than *“participating Country”*
- Use of *“historical review”* rather than *“brief historical review”*

Linearize phrases and divide those articulated

- Deletion of the sentence *“Each athlete is assigned a number before the start of the competition. This number is never changed”* because we use an incremental ID to identify the athlete in the Database.
- Use of *“a daily list of the competition”* rather than *“daily a list of the competitions that will take place during the day indicating the time, the specialty (one hundred meters, high jump, etc.) and, where relevant, the level of competition (semi-final, final , etc...)”*.
- Use of *“We need to know who and how many participants are registered for the various competitions”* rather than *“We need to know who is registered for the various competitions, and for each competition in which order and how many participants participate”*.

Making explicit reference between terms

- The term *“Country”* to the line 8 refers to the Country of the competition. The term *“Country”* to the line 9 refers to the Nationality of the athletes

Create a glossary and identify synonyms

TERM	TYPE	DESCRIPTION	SYNONYMS	CONNECTIONS
Competition	Entity	An event composed by one or more matches in which each participant tries to win a prize by being the fastest and break the record.		Race, Ranking, Location
Ranking	Relation	The position or level that participant has in a list that compares.	Classification, leader board	Competition, Result
Race	Relation	An organized match of the concerned competition in a specific date.	Match, round	Level, Competition, Participant
Level	Entity	Type of the race, can be single race, final, semi-final, quarter finals etc...		Race

Location	Relation	Association between Competition and Country of the same		Competition, Country
Country	Entity	An area of land that has its own government and official borders. For example, England, Ireland, Italy etc...	Place, Nation	Location, Nationality
Nationality	Relation	Association between Participant and Country where he/she was born.		Country, Participant
Participant	Entity	Athlete or team that takes part and run for a competition.	Player	Score, Team, Nationality, Race
Athlete	Entity	A person who is very good at sports or physical exercise. He/she is ready for the competition		Team
Team	Relation	Number of athletes who race together as a group in the competition		Athlete, Participant
Score	Relation	Association between Result and Participant.		Result, Participant
Result	Entity	It is the score and the name of the winner for the competition. A result can be defined <i>record</i> if the score is the best one for the specialty		Score, Ranking
Specialty	Entity	The type of the specific competition. For example, one hundred meters, high jump etc...		
CompType	Relation	Association between Competition and Specialty.		

Reorganizing for keyword phrases

Phrases related to competitions

For the competition we represent the Country, the date and time, the specialty (one hundred meters, high jump, etc.), the level of competition (semi-final, final, etc...) and the world records. The system must keep the information relating to the winners of the various competitions with the corresponding results. We need to know who and how many participants are registered for the various competitions. We must be able to provide a daily list of the competitions.

Phrases related to results

For the results of each competition we represent the rank, the participant and a description (gold, silver, bronze etc...).

Phrases related to participants

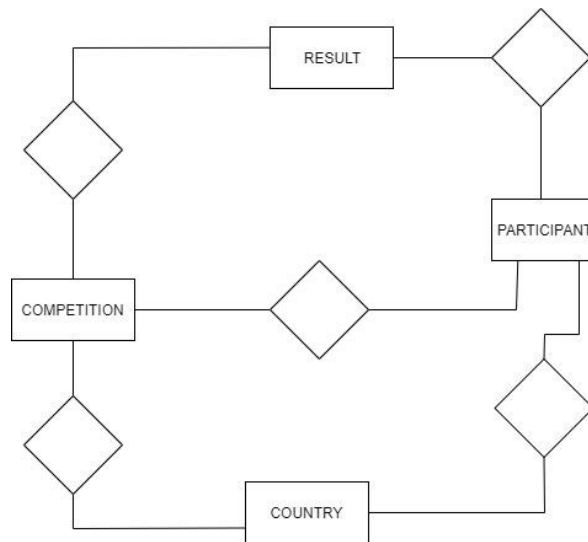
A participant in a competition can be individual athlete or team of athletes. For the athletes we represent the personal data and the Country.

Phrases related to Countries

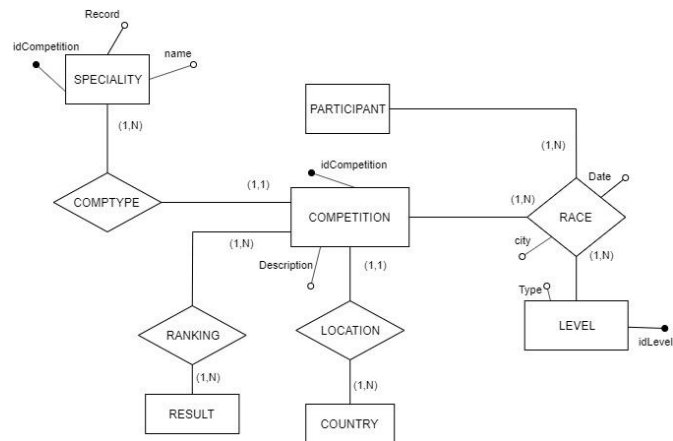
For each Country we represent the name, the flag and a historical review.

Project strategy: Hybrid design

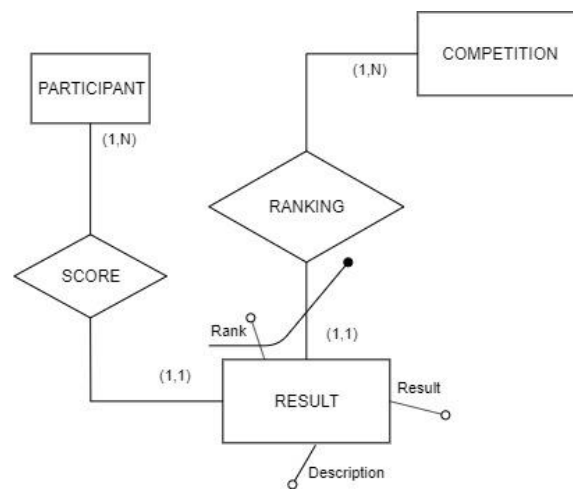
Skeleton Scheme:



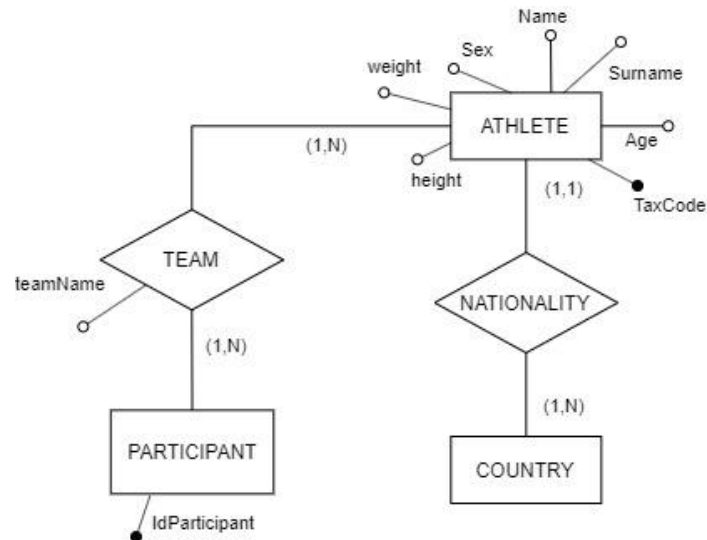
Competitions



Results



Participants

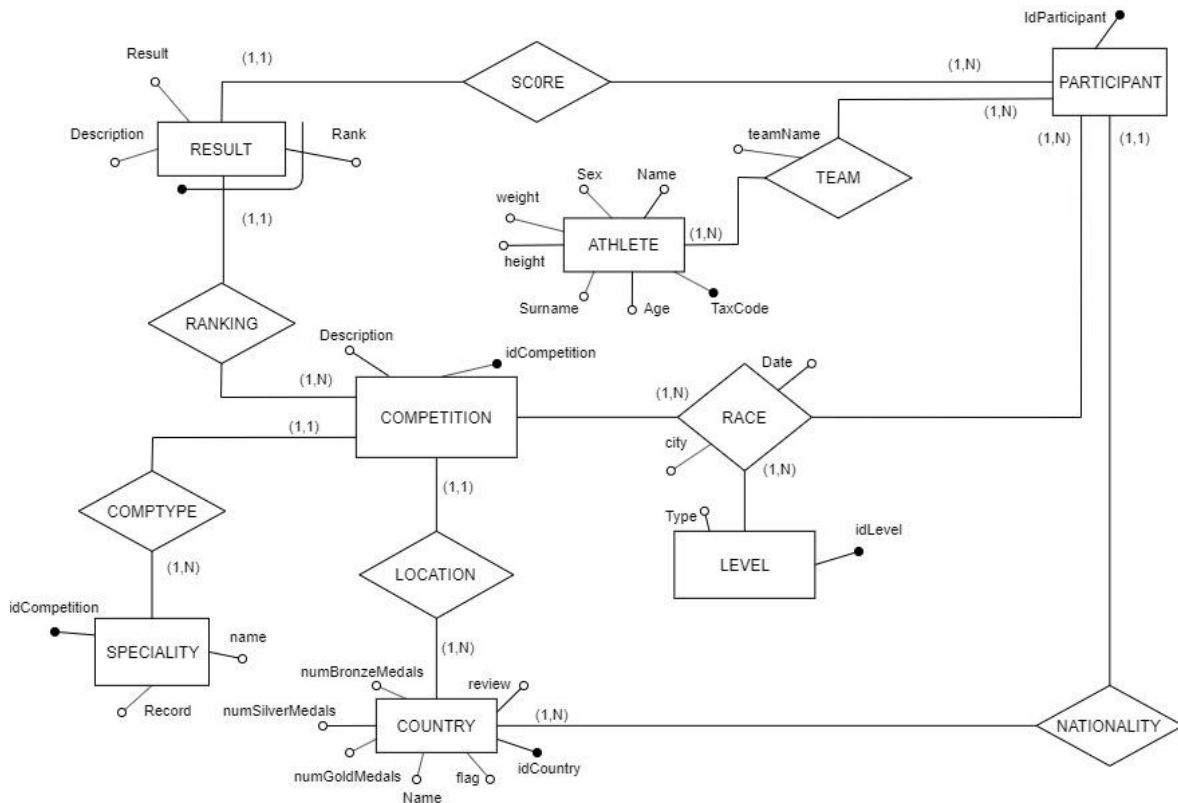


Countries

In order to keep track about the number of medals won for a Country, we add three attributes on the Entity Country as follows:



Final ER



Constraints

Constraint	Description
------------	-------------

Constr_01	The entity <i>Level</i> must contain at least the following instances: <ul style="list-style-type: none"> - final - semi-final - quarterfinals - single race
Constr_02	The attribute <i>city</i> of the Relation <i>Race</i> must be a city of the Competition Country.
Constr_03	The dates of the Races must be ordered chronologically. For example, final most recent than semi-finals; Semi-finals most recent than quarter of finals etc...
Constr_04	The <i>Result</i> entity must contain an instance for the total number of participants for the competition. Let suppose a competition with 2 players in final, 4 players in semi-finals. This implies that <i>Result</i> must contain 6 rows.
Constr_05	The attribute <i>Rank</i> must be a counter started from 1 for the competition. Smaller is the rank, highest is the <i>Result</i> .
Constr_06	When the attribute <i>rank</i> of the <i>Result</i> entity equals 1 then the description of the same entity is " <i>Gold</i> ", when rank equals 2 then the description is " <i>Silver</i> ", when rank equals 3 then the description is " <i>Bronze</i> ".
Constr_07	The attribute <i>Record</i> of the specialty must be of the same type of the attribute <i>Result</i> of the Entity <i>Result</i> .
Constr_08	If the <i>Result</i> of the winner is higher than the record of concern competition, the <i>record</i> attribute must be updated with the new score.
Constr_09	If the participant is composed by at least 2 athletes, then the attribute <i>teamName</i> must be enhanced. Otherwise, the attribute is NULL.
Constr_10	The attribute <i>flag</i> of the entity <i>Country</i> is an image, so it is a binary attribute.
Constr_11	The attribute <i>TaxCode</i> of the entity <i>Athlete</i> contains 16 Chars
Constr_12	The value of <i>sex</i> of the entity <i>Athlete</i> must be ' <i>M</i> ' or ' <i>F</i> '.

Operations on data

- OP1:** Creation of a new competition
- OP2:** Print the list of participants for a given competition
- OP3:** Modification of the world record given a ranking (result entity)
- OP4:** Print the ranking (result entity) at the end of the competition
- OP5:** Enter the winners of a competition (gold, silver and bronze)
- OP6:** Print the countries with most medals for a specific specialty
- OP7:** Print the athlete with the highest record for a specific specialty

LOGICAL DESIGN

We use a Relational Database as a logical mode.

The attribute *teamName* in the relation *Team* is redundant, we choose to avoid the redundancy moving in this phase the attribute in the entity *Participant*. In fact, with the *Constr_09* we know if the participant is an athlete or a team, just checking if the attribute is NULL.

The attributes *Date* and *City* of the relation *Race* are redundant too. We can decide if create concrete instances of levels with the specific date and location. So, in this case we choose to leave this redundancy for simplicity. Let suppose a final in Rome at 24/02/2026, then we have two rows: one row contains the first finalist while the other one contains the second athlete. Both rows have the same city and the same date.

Logical Model

Country (idCountry, flag, name, review, numGoldMedals, numSilverMedals, numBronzeMedals)

Competition (idCompetition, description, **Specialty**, **Country**)

Specialty (idSpecialty, name, record)

Level (idLevel, type)

Participant (idParticipant, **Nation**, teamName)

Result (**Competition**, **Rank**, **Participant**, result, description)

Athlete (TaxCode, name, surname, age, sex, weight, height)

Team (**Participant**, **Athlete**)

Race (**Competition**, **Level**, **Participant**, date, city)

Data type

TABLE	FIELD	DATATYPE
Country	idCountry	NUMBER
	flag	BLOB
	name	VARHCHAR (35)
	review	VARCHAR (300)
	numGoldMedals	NUMBER
	numSilverMedals	NUMBER
	numBronzeMedals	NUMBER
Competition	idCompetition	NUMBER
	Description	VARCHAR (30)
	speciality	NUMBER
	Country	NUMBER
Specialty	idSpecialty	NUMBER
	name	VARCHAR (20)
	record	NUMBER (10,4)
Level	idLevel	NUMBER
	type	VARCHAR (20)
Participant	idParticipant	NUMBER
	Nation	NUMBER
	teamName	VARCHAR (20)
Result	Competition	NUMBER
	Rank	NUMBER
	Participant	NUMBER
	result	NUMBER (10,4)
Athlete	TaxCode	VARCHAR (16)
	name	VARCHAR (25)
	surname	VARCHAR (25)
	age	NUMBER
	sex	CHAR
	weight	NUMBER
	height	NUMBER
Team	Participant	NUMBER
	Athlete	VARCHAR (16)
Race	Competition	NUMBER

	Level	NUMBER
	Participant	NUMBER
	date	DATETIME
	city	VARCHAR (30)

Volume Table

Concept	TYPE	Volume
Competition	Entity	62
Ranking	Relation	496
Race	Relation	868
Level	Entity	4
Location	Relation	62
Country	Entity	195
Nationality	Relation	1100
Participant	Entity	1100
Athlete	Entity	5400
Team	Relation	3300
Score	Relation	496
Result	Entity	496
CompType	Relation	62
Specialty	Entity	23

Operations Table

OPERATION	DESCRIPTION	TYPE	FREQUENCY
OP1	Creation of a new competition	I	7 / day
OP2	Print the list of participants for a given competition	I	30 / day
OP3	Modification of the world record	I	1 / 2 weeks
OP4	Print all the rankings at the end of the competition	I	1 / week
OP5	Enter the winners of a competition (gold, silver and bronze)	I	5 / day
OP6	Show the countries with most medals for a specific specialty	B	6 / month
OP7	Show the athlete with the highest record for a specific specialty	B	2 / month

PHYSICAL DESIGN

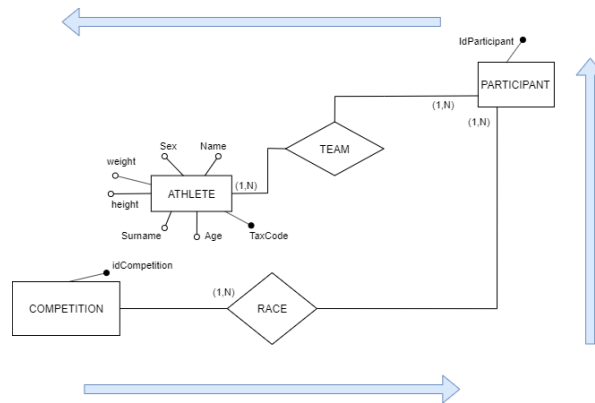
Database technology

The chosen technology for the DBMS is: Oracle Database 11g. In Oracle, an index is a schema object that contains an entry for each value that appears in the indexed column(s) of the table or cluster and provides direct, fast access to rows. Oracle Database supports several types of index:

- Heap files, hash structure, cluster (also with multiple tables) also ordered (with dense B-tree) for the primary structure
- B-tree, bitmap, hash functions as secondary indices

Optimization of operations

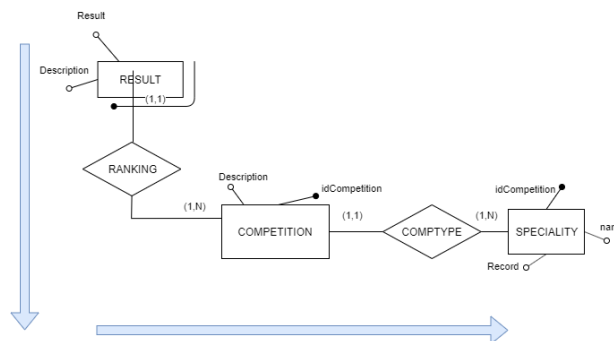
- **OP1:** Insert, not to be optimized
- **OP2:**



Involved tables: RACE, PARTICIPANT, TEAM, ATHLETE

The candidate attributes for optimize this operation are: RACE.competition, RACE.participant, PARTICIPANT.idParticipant, TEAM.participant, TEAM.athlete, ATHLETE.TaxCode.

- **OP3:**



Involved tables: RESULT, COMPETITION, SPECIALTY

The candidate attributes for optimize this operation are: RESULT.competition, RESULT.rank, COMPETITION.idCompetition, COMPETITION.specialty, SPECIALTY.idSpecialty

- **OP4:** it is enough to print only the result table given the ID of the competition. So, the candidate attribute for optimize this operation is RESULT.Competition and RESULT.rank (for the ORDER BY clause)
- **OP5:** Insert, not to be optimized
- **OP6** and **OP7** not to be optimized because they are complex OLAP queries performed into a Datawarehouse.

Result

The following picture shows the response time in Oracle for the operations 2, 3 and 4 performed 23 times, both with and without B-tree index definition:

```

select
  'Operation number: ' || operationNum as operation,
  case when indexedOperation = 1 then 'true' else 'false' end as use_of_index,
  count(*) as num_of_launches,
  Sum(extract( second from endTimeStamp - startTimeStamp )) as seconds
from op_exec_time
group by
  operationNum,
  indexedOperation,
  'Operation number: ' || operationNum,
  case when indexedOperation = 1 then 'true' else 'false' end
order by operationNum, indexedOperation;

```

Script Output x Query Result x Query Result 1 x Query Result 2 x Query Result 3 x Query Result 4 x

SQL | All Rows Fetched: 6 in 0.017 seconds

OPERATION	USE_OF_INDEX	NUM_OF_LAUNCHES	SECONDS
1 Operation number: 2	false	23	0.116
2 Operation number: 2	true	23	0.111
3 Operation number: 3	false	23	0.033
4 Operation number: 3	true	23	0.02
5 Operation number: 4	false	23	0.02
6 Operation number: 4	true	23	0.012

It is easy to see that for each operation the execution of the query is faster when the attribute `use_of_index` is true.

BUSINESS INTELLIGENCE

We remember that Operation 6 (Show the countries with most medals for a specific specialty) and Operation 7 (Show the athlete with the highest record for a specific specialty) are Business OLAP Request. For this scope it is available a data warehouse (DWH) with the purpose of load and analyse data from the athletics DB. The DWH is scheduled and perform 4 steps:

1. Import phase
2. Staging Area
3. Data Mart
4. Export

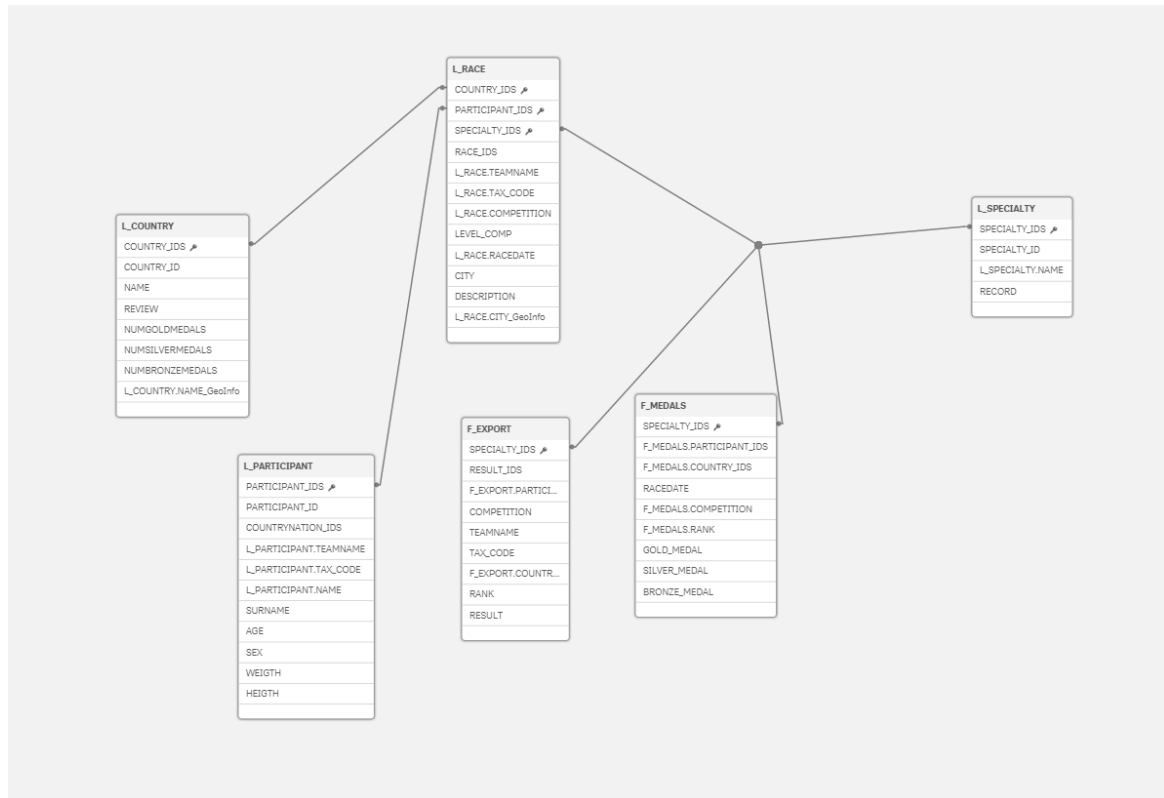
The step 1 is the import phase. In this phase SQL can be used to extract the data from database or other source. Extracting data often involves the transfer of large amounts of data from source operational systems. Such operations should be performed during a period of relatively low system load or overnight.

The next step regards the creation of Staging Area (SA), an intermediate storage area used for data processing during ETL process. The data in SA splits between data sources and data marts (DM). In Staging Area, we implement the Business Logic for the creation of Star Schema.

After that, we Load the elaborated data into DM that are small, faster and more flexible compared to a single data warehouse. In the last phase the data is ready for the visualization.

Qlik-Sense is a business analytics service by Microsoft. It aims to provide interactive visualizations and BI capabilities with an interface simple enough for end users to create their own reports and dashboards.

Data model



Dashboard

Following an example of report that shows 3 KPI for each country.

