

# Pré-rapport de projet en base de données : Recueil de performances aux Jeux Olympiques

Marc Chehire - Rodolphe Rouyrre

7 avril 2020

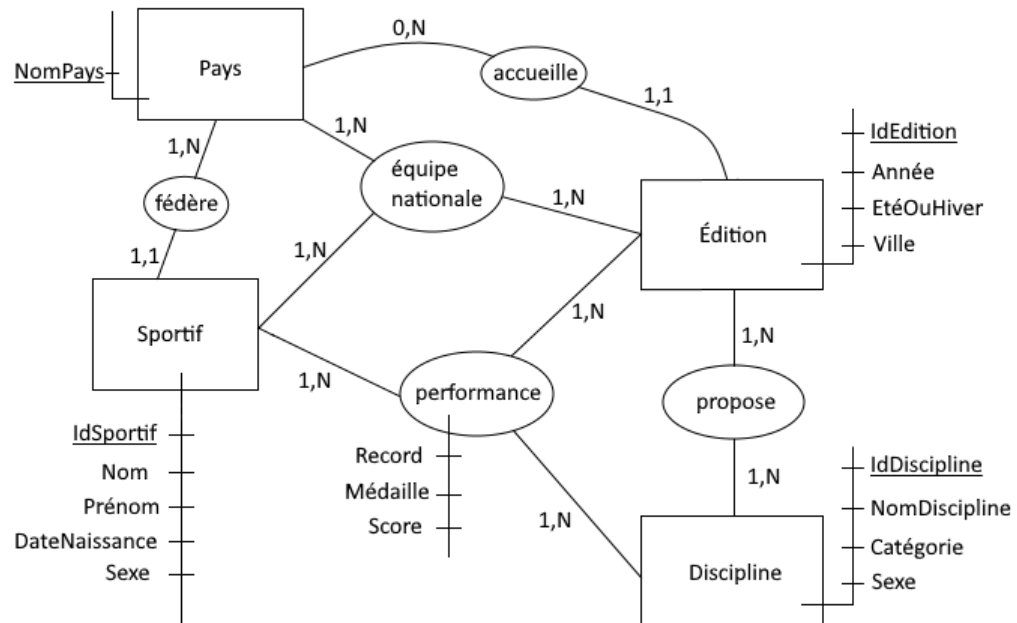
## Introduction

Le but de notre projet est de constituer une base de données des scores enregistrés par les athlètes aux Jeux Olympiques. Nous ne voulons pas nous limiter par rapport à la période étudiée et nous voulons considérer les Jeux d'été et d'hiver.

Nous aimerions proposer un projet permettant de tracer les performances d'un athlète au fil des années, d'observer les résultats par édition, par discipline ou par équipe nationale. Nous pourrions ainsi dénombrer le nombre de médailles récoltées par un pays pour une instance des Jeux. Par la suite, nous parlerons de performance sur une épreuve, comme l'association d'un score et d'une appréciation qualitative du résultat ("est-ce un record ?", "délivre-t-il un médaille?").

Nous adoptons le point de vue d'un spectateur souhaitant s'informer sur l'historique des performances réalisées durant les Jeux Olympiques.

## Modèle Entité/Association



### Définition des entités

Nous organisons notre modélisation autour de **pays**, puisqu'ils servent à la fois de lieu d'accueil des Jeux et présentent chacun une équipe nationale. Cette entité contient tous les pays ayant participé à au moins une édition des Jeux Olympiques, d'hiver ou d'été.

Les acteurs essentiels de ces Jeux sont les **sportifs**, regroupés en équipes nationales. Nous ne considérerons pas le cas des sportifs ayant participé dans une équipe neutre, sans pouvoir représenter de pays. Nous supposons également qu'un sportif n'a pu représenter qu'un seul pays au fil de ses participations. Nous avons intégré un IdSportif abstrait car il serait trop complexe de considérer à la fois le Nom, Prénom, et la date (et/ou lieu) de naissance comme clés pour distinguer les sportifs.

La notion d'**édition** permettra une recherche ciblée par événement. L'année n'est pas suffisante pour faire une recherche de façon univoque, puisque les Jeux Olympiques d'hiver ont été disputés la même année que ceux d'été jusqu'en 1992. Nous pouvons prendre un identifiant comme clé primaire, cela facilitera notre modèle relationnel.

Chaque édition des Jeux propose des sports, subdivisés en épreuves. C'est ce que nous représentons au travers de la classe **discipline** : la dénomination d'un sport (ex : athlétisme), la catégorie (ex : 400m haies), ainsi que le sexe pour distinguer les podiums homme/femme/mixte. Nous utiliserons la clé IdDiscipline pour ne garder qu'une variable.

Le type d'entité discipline comprend toutes les disciplines ayant été disputées lors d'au moins une édition.

### Explication des associations et du choix des cardinalités :

Un pays **federe** 1,N sportifs : la classe Pays comporte uniquement les pays ayant déjà enregistré une participation aux Jeux et donc présenté une équipe nationale. Un sportif sera vu comme un représentant d'un seul et unique pays, ne pouvant changer aux cours des participations. Cette association permet donc de trouver tous les sportifs ayant représenté un même Pays.

Un pays **accueille** 0,N édition(s) : des pays comme les Etats-Unis ou la France ont déjà accueilli plusieurs éditions, et la grande majorité des pays invités n'ont pas eu cette occasion. Il n'y a qu'une nation d'accueil par édition donnée (1,1).

Chaque édition **propose** 1,N disciplines. Les disciplines présentées varient selon l'édition, chacune peut apparaître une ou plusieurs fois dans le registre (1,N). Le croquet fut par exemple présenté une seule fois aux JO de Paris, en 1900.

**Performance** est une association ternaire : une performance représente le résultat d'un sportif (lorsqu'il est distinguable de ceux de ses partenaires), pour une participation à une discipline, lors d'une édition donnée. Nous souhaitons quantifier le score, que ce soit un temps réalisé, une distance ou un poids soulevé, ainsi que de le quantifier. Était-ce un nouveau record au moment de sa réalisation ? A-t-il donné lieu à une médaille ? Si oui, laquelle ? L'attribut médaille aura donc une valeur neutre.

- chaque sportif apparaîtra au moins une fois dans l'historique des performances : 1,N.
- chaque édition apparaîtra plusieurs fois car plusieurs disciplines y sont disputées (et chacune donnant lieu à plusieurs performances individuelles) : 1,N.
- chaque discipline apparaîtra plusieurs fois car chacune donne lieu à plusieurs performances : 1,N.

Nous avons choisi d'ajouter une deuxième association ternaire : **équipe nationale**. Une équipe nationale est éphémère, c'est l'association d'un sportif à un pays, pour le temps d'une édition donnée.

- chaque pays apparaîtra plusieurs fois (plusieurs sportifs pour au moins une édition) : 1,N.
- chaque sportif apparaîtra au moins une fois : 1,N.
- chaque édition apparaîtra plusieurs fois puisque plusieurs équipes nationales y participent : 1,N.

## Justification des associations ternaires

**Equipe nationale** est nécessairement une association ternaire car un sportif d'un pays, (qui n'est associé qu'à un seul pays), peut participer à plusieurs éditions. De même un pays va présenter une équipe nationale pour chaque édition, et qui sera normalement différente.

**Performance** est nécessairement une association ternaire car un sportif qui ne concoure que dans une discipline, peut se représenter dans plusieurs éditions. Un sportif qui participe à une seule édition, peut concourir dans plusieurs disciplines.

## Modèle relationnel

### Nouvelles entités

Pays (NomPays)

Sportif (IdSportif, Nom, Prénom, DateNaissance, Sexe)

Édition(IdEdition, Année, EtéOuHiver, Ville)

Discipline(IdDiscipline, NomDiscipline, Catégorie, Sexe)

### Relations binaires 1.N

Les associations binaires 1.N ne définissent pas de tables relationnelles intrinsèquement. Elles enrichissent les tables qui présentent une cardinalité 1.1

Nous obtenons les tables actualisées :

Sportif (IdSportif, Nom, Prénom, DateNaissance, Sexe, NomPays). NomPays est ajouté car une cardinalité 1,1 relie Sportif à Pays. C'est une clé étrangère.

Édition(IdEdition, Année, EtéOuHiver, Ville, NomPays). Une cardinalité 1,1 est présente ici également.

### Relations binaires N.N

Propose(IdEdition, IdDiscipline)

### Relations ternaires

Performance(IdPerformance, IdSportif, IdDiscipline, IdEdition, Record, Médaille, Score)

Équipe nationale(IdEquipeNationale, IdSportif, IdEdition, NomPays)

Nous avons ajouté un identifiant propre à chaque relation ternaire, en considérant qu'avoir 2 clés primaires est le nombre maximal admissible.

### Justification des clés

Nous avons choisi d'ajouter un attribut IdSportif pour l'entité Sportif et un attribut IdDiscipline pour l'entité Discipline car sinon nous supposons qu'il faudrait prendre en compte trois attributs pour faire une clé. En effet, on admet qu'il est possible que deux sportifs aient le même nom et le même prénom et il est évident que le sexe est nécessaire pour distinguer certaines épreuves. L'ajout de ces attributs permet donc de simplifier grandement le modèle. Quant à l'attribut IdEdition, il n'est pas nécessaire puisque les attributs Année et EtéOuHiver suffisent à déterminer de façon univoque un tuple d'une Edition. Nous aurions pu choisir 2 clés primaires mais un identifiant simplifie notre modèle relationnel.

### Dénormalisation

L'association équipe nationale, apporte une connexion directe entre Sportif et Edition. De plus Pays possède un seul attribut (et des cardinalités 1,1), nous pouvons nous demander s'il faut retirer cette entité. Nous avons choisi de la garder pour permettre 2 choix d'approche de recherche :

- sélectionner un pays pour mettre en valeur tous les sportifs qu'il a fédéré, ou toutes les éditions accueillies ; (par associations binaires).
- croiser le choix d'un pays et d'une édition pour récolter tous les IdSportif d'une équipe nationale ; (association ternaire).

Un argument supplémentaire consiste à voir que retirer cette entité donnerait naissance à 2 attributs au lieu d'un seul : pays d'accueil (pour une édition), et pays représenté (pour un sportif).