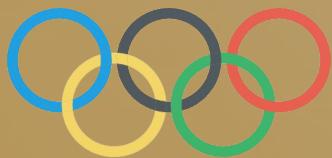


# 126 ANS DE JEUX OLYMPIQUES



Pierre-Alexandre Paget | Février 2026



# PROGRAMME

-  Pourquoi les Jeux Olympiques ?
-  Présentation du Dataset
-  EDA, visualisations pour explorer les pistes
-  OLAP, Modélisation du schéma
-  Dashboard : Analyse multidimensionnelle des performances olympiques
-  Conclusion et ouverture



# POURQUOI LES JEUX OLYMPIQUES ?



**Un thème qui passionne**



**Un sujet qui parle** à une majorité de personnes à travers le monde



**Audience estimée pour les JO de Paris 2024**

≈ 5 milliards de personnes



**Un défi technique**

Manipuler des datasets massifs et variés dans le temps pour en extraire des insights pertinents  
Effectuer du scraping pour extraire des données spécifiques





# EXTRACT, TRANSFORM, LOAD



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES

1

### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES

1

### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES

1

### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



#### Python



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



#### Python

Nettoyage



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



#### Python

Nettoyage  
Création



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



#### Python

Nettoyage  
Création  
Valeurs calculées



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



#### Python

Nettoyage  
Création  
Valeurs calculées  
Normalisation géographique



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



#### Python

Nettoyage  
Création  
Valeurs calculées  
Normalisation géographique  
Distinction des médailles par athlète vs nation



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



#### Python

Nettoyage  
Création  
Valeurs calculées  
Normalisation géographique  
Distinction des médailles par athlète vs nation  
Premières visualisations EDA  
Enregistrement des fichiers .csv



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES



### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

### TRANSFORM



#### Python

Nettoyage  
Création  
Valeurs calculées  
Normalisation géographique  
Distinction des médailles par athlète vs nation  
Premières visualisations EDA  
Enregistrement des fichiers .csv

#### Power Query

Dernières transformations



# NOTRE DATABASE

## COLLECTE, IMPORTATION ET TRAITEMENT DES DONNÉES

1

### EXTRACT



#### Récupération des données

11 tables | 6 via Kaggle - 5 par scraping  
483 740 lignes  
88 colonnes

#### Scraping de données sur Wikipedia

#### Fichier Json

Limites des continents  
Affinage via Mapshaper

2

### TRANSFORM



#### Python

Nettoyage  
Création  
Valeurs calculées  
Normalisation géographique  
Distinction des médailles par athlète vs nation  
Premières visualisations EDA  
Enregistrement des fichiers .csv

#### Power Query

Dernières transformations

3

### LOAD



#### Chargement dans Power BI

Construction du modèle relationnel  
Création du dashboard





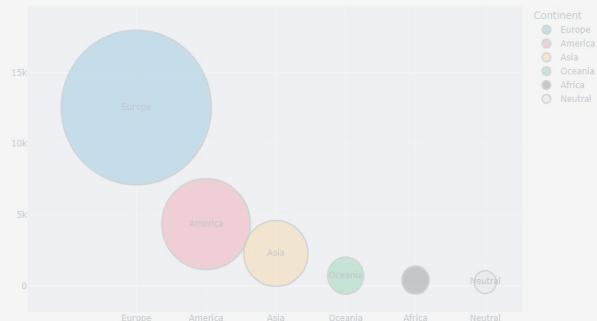
# ANALYSE EXPLORATOIRE



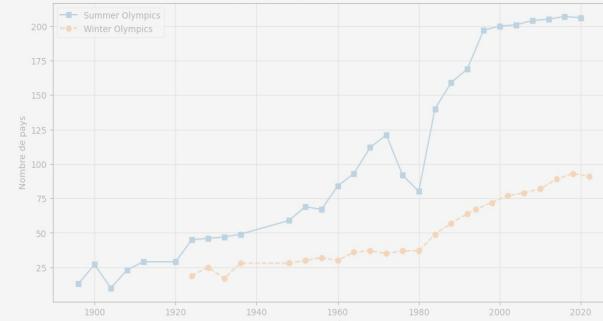
# ANALYSE EXPLORATOIRE

## ORIENTATIONS RETENUES

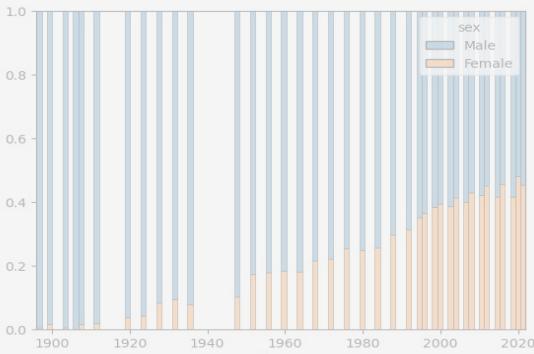
Nombre de médailles total par continent



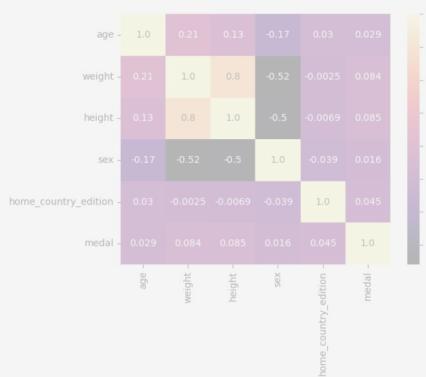
Evolution du nombre de pays participant par édition



Evolution de la mixité par édition



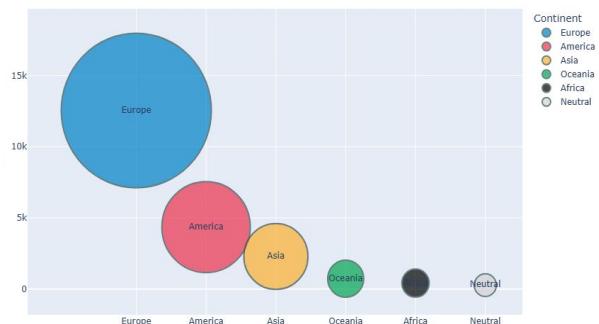
Matrice de corrélation pour l'obtention d'une médaille



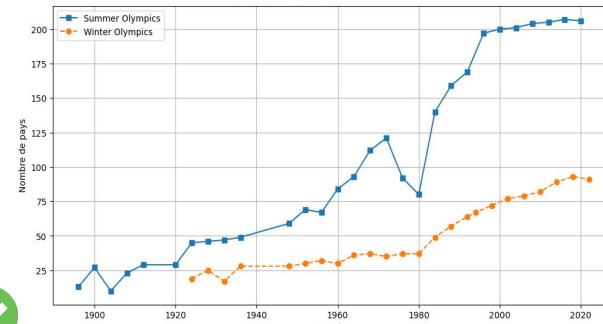
# ANALYSE EXPLORATOIRE

## ORIENTATIONS RETENUES

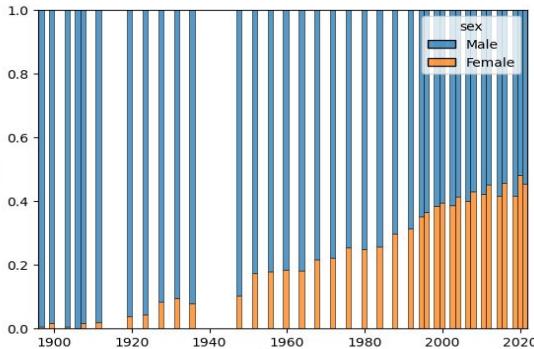
Nombre de médailles total par continent



Evolution du nombre de pays participant par édition



Evolution de la mixité par édition



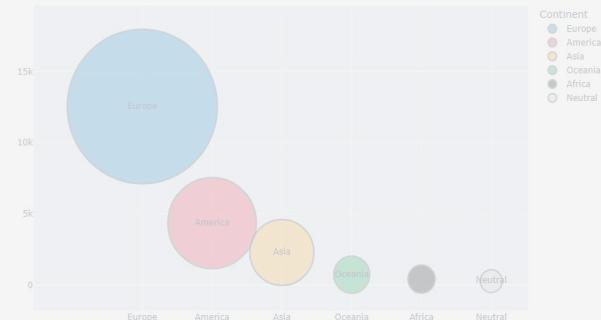
Matrice de corrélation pour l'obtention d'une médaille



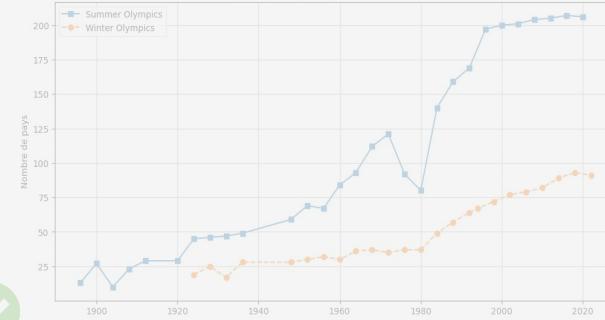
# ANALYSE EXPLORATOIRE

## ORIENTATIONS RETENUES

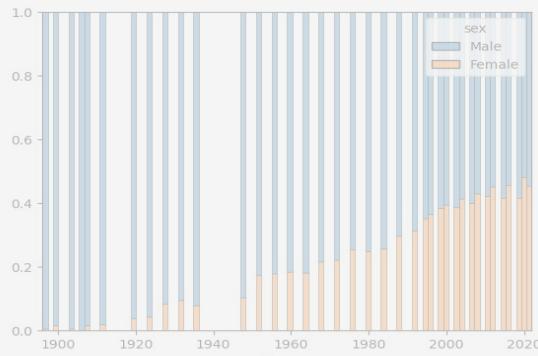
Nombre de médailles total par continent



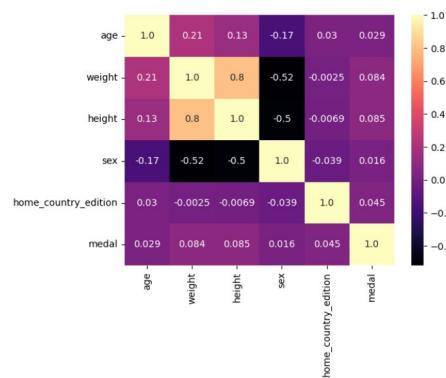
Evolution du nombre de pays participant par édition



Evolution de la mixité par édition



Matrice de corrélation pour l'obtention d'une médaille



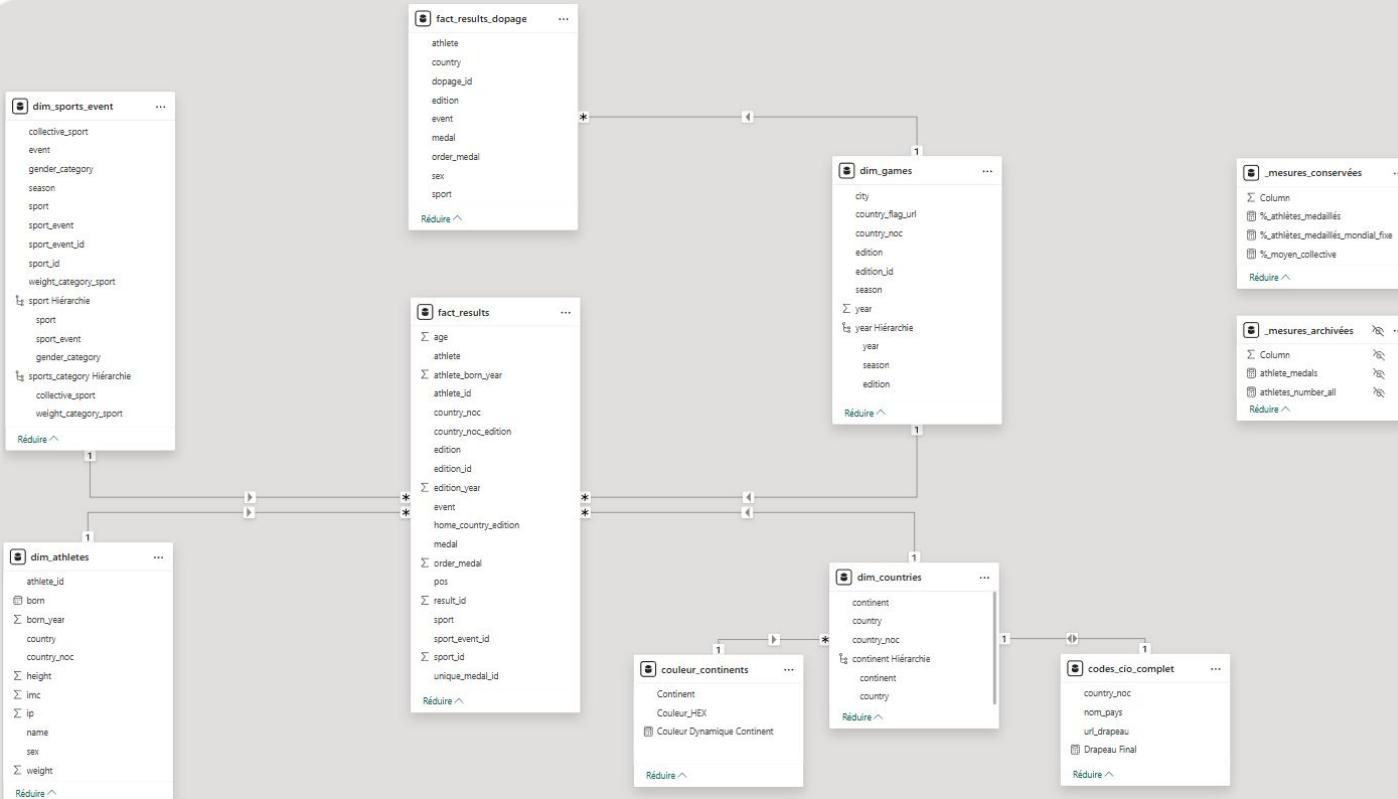


# MODÈLE RELATIONNEL



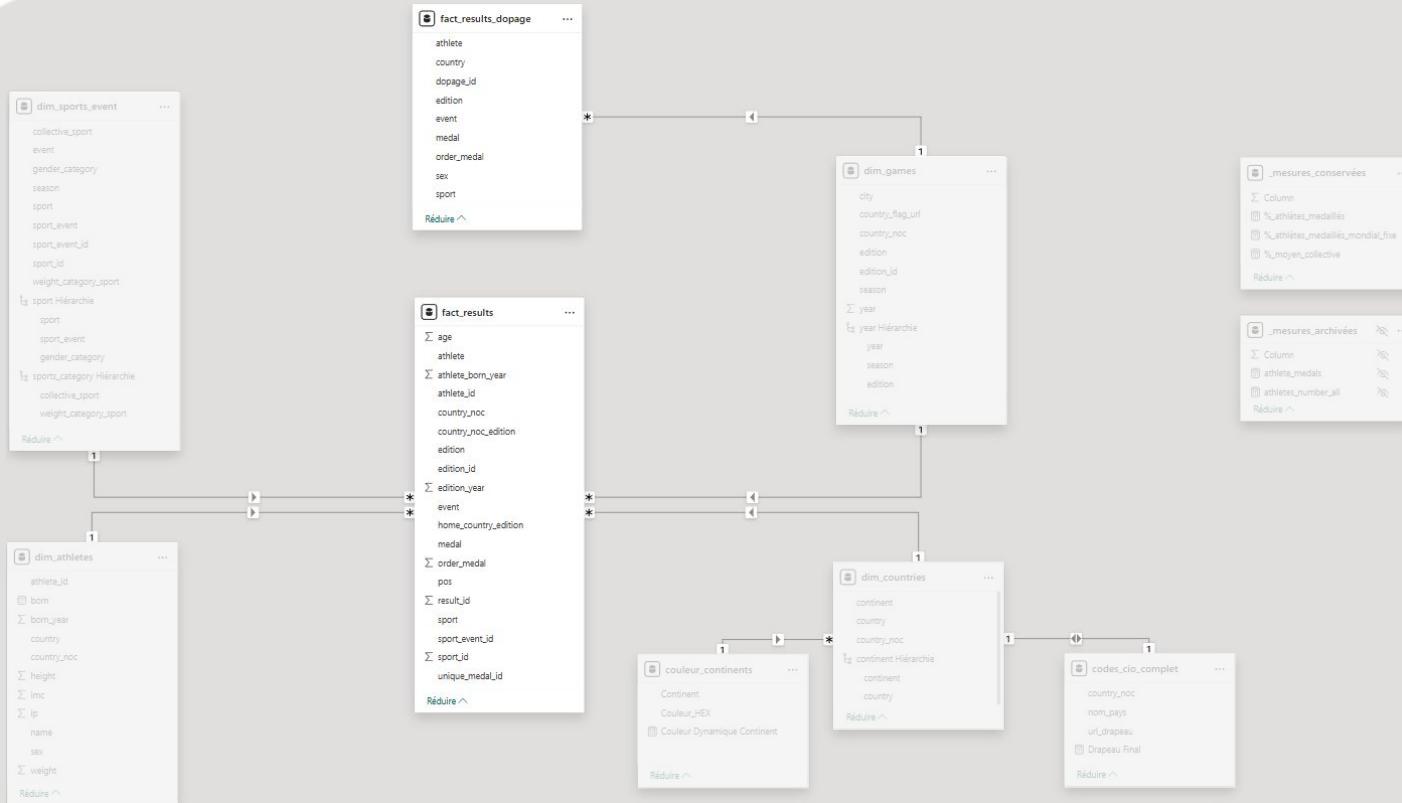
# MODÉLISATION

## SCHÉMA CONSTELLATION OLAP



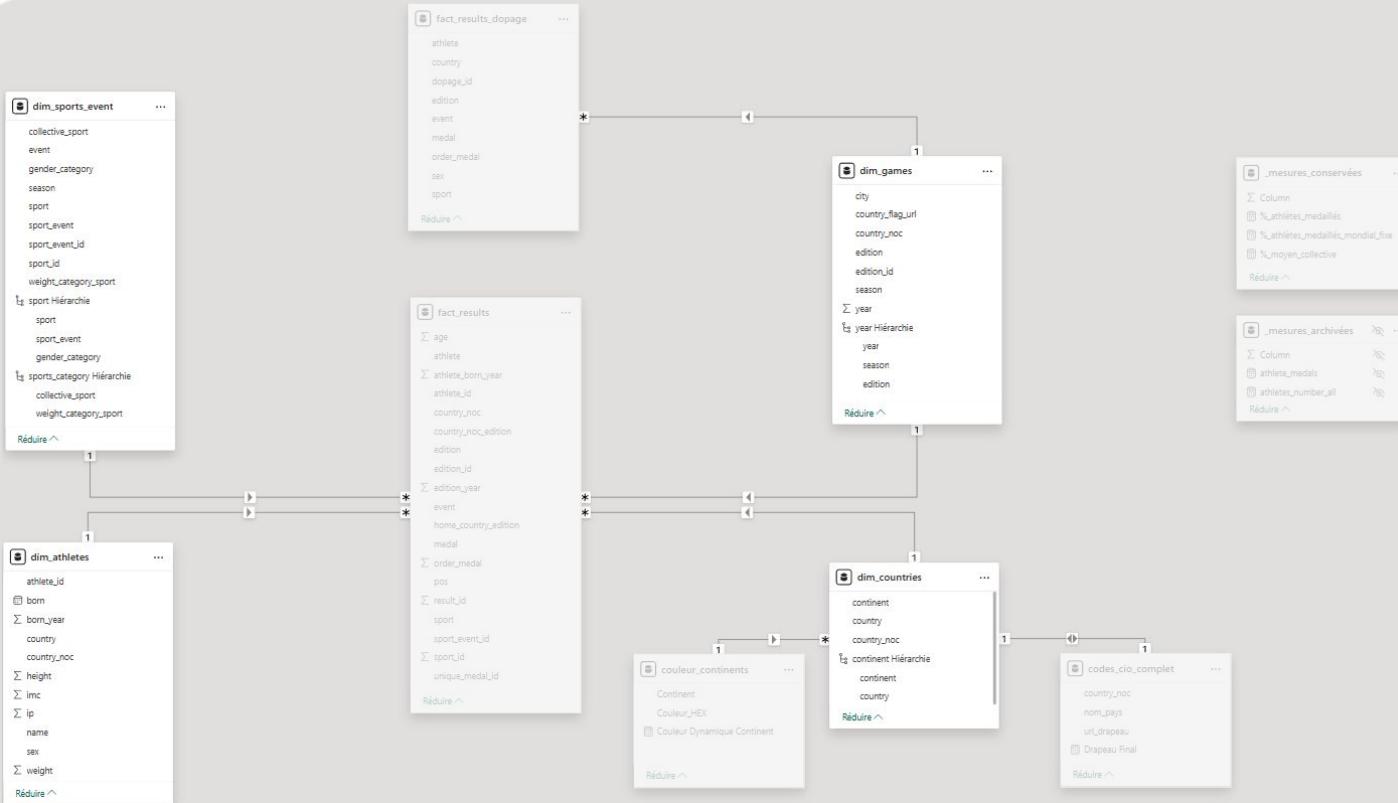
# MODÉLISATION

## SCHÉMA CONSTELLATION OLAP



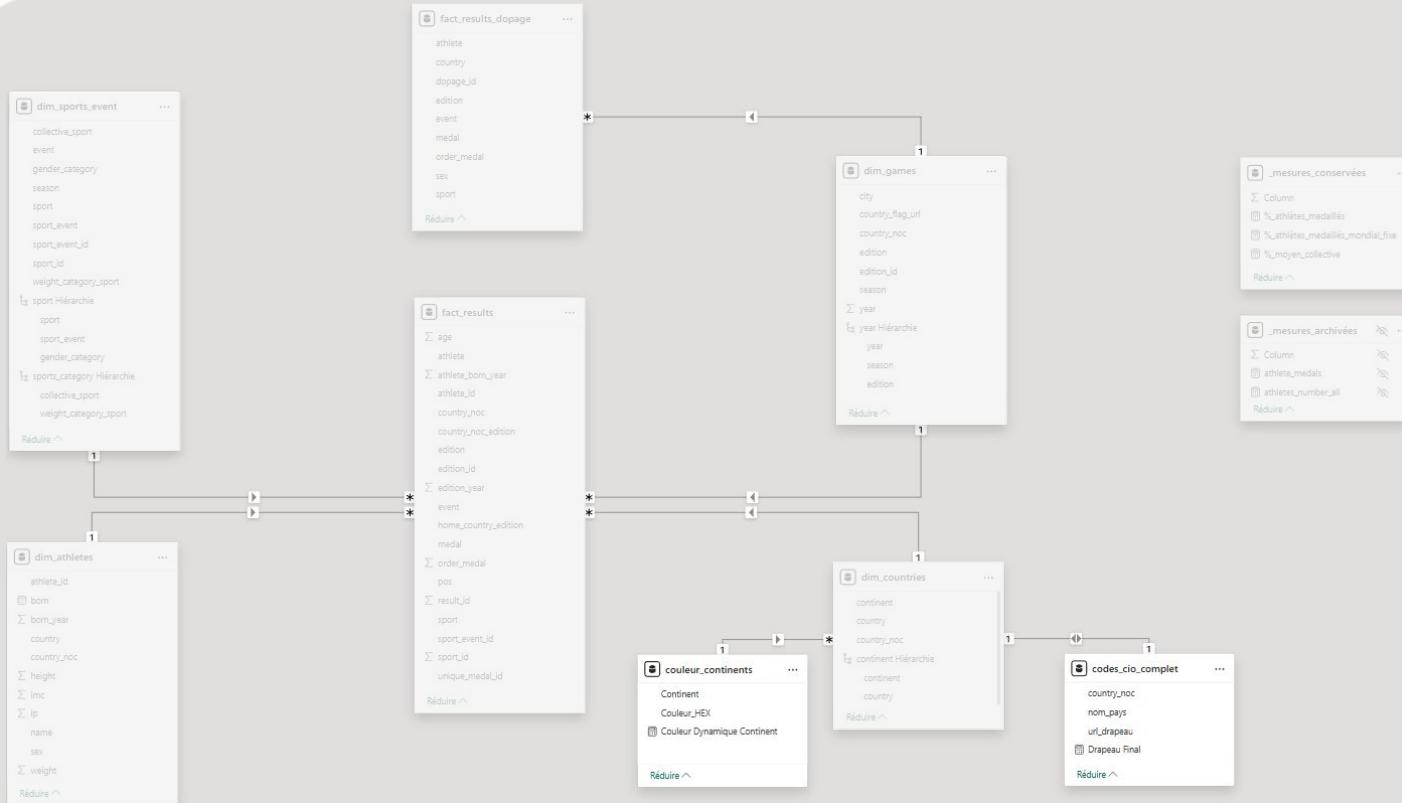
# MODÉLISATION

## SCHÉMA CONSTELLATION OLAP



# MODÉLISATION

## SCHÉMA CONSTELLATION OLAP





# DASHBOARD



# DASHBOARD

## ANALYSE MULTIDIMENSIONNELLE DES JEUX OLYMPIQUES

-  **Synthèse et chiffres clés** de 126 ans d'Olympisme
-  **Evolution des athlètes** à travers les différentes éditions
-  **Multiplication des disciplines** pour une spécialisation sportive
-  **Un palmarès très inégalitaire** entre le nord et le sud
-  **Dopage & tricherie** : Prise de conscience de l'impact géopolitique sur le sport ?



# VUE D'ENSEMBLE

## SYNTHÈSE ET CHIFFRES CLÉS DE 126 ANS D'OLYMPISME



25

Pays organisateurs



112

Disciplines



231

Délégations



20 427

Médailles



31 585

Athlètes médaillés



155 867

Athlètes

### RECORD DE TITRES

TOP 5

Athlètes	Titres
Michael Phelps	23
Carl Lewis	9
Jenny Thompson	9
Larisa Latynina	9
Mark Spitz	9

### RECORD DE MÉDAILLES

TOP 5

Athlète	Médailles
Michael Phelps	28
Larisa Latynina	18
Marit Bjørgen	15
Nikolay Andrianov	15
Boris Shakhlin	13



63

Éditions



35

Été



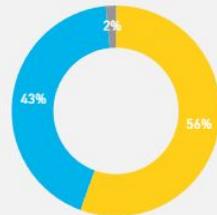
27

Hiver



1

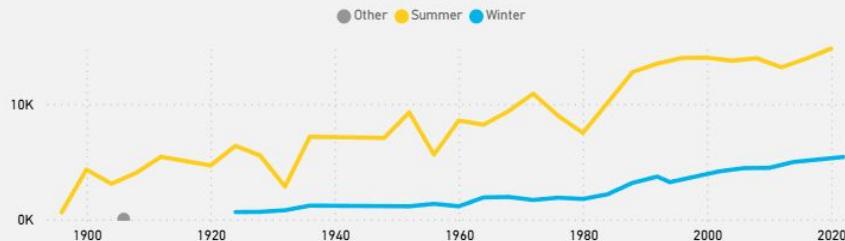
Intercalaire



● Summer ● Winter ● Other

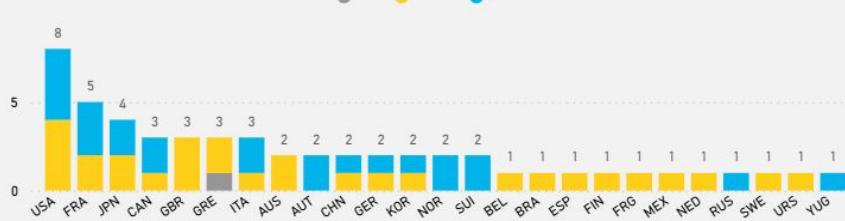
### CROISSANCE DE LA PARTICIPATION OLYMPIQUE

● Other ● Summer ● Winter



### NOMBRE D'ORGANISATIONS PAR PAYS

● Other ● Summer ● Winter



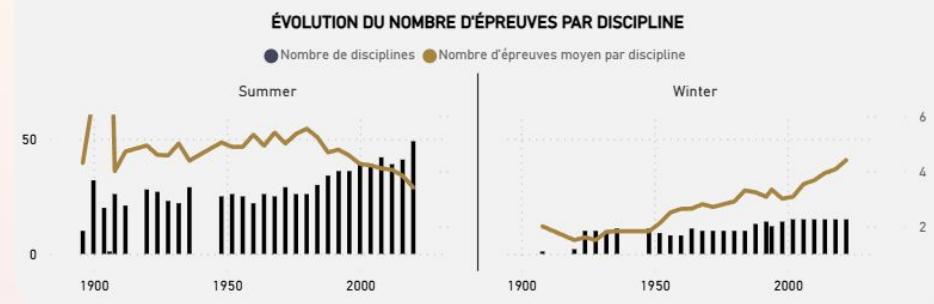
# ANALYSE DES ATHLÈTES

## DU SPORT AMATEUR AU SPORT DE HAUT NIVEAU



# ANALYSE DES DISCIPLINES

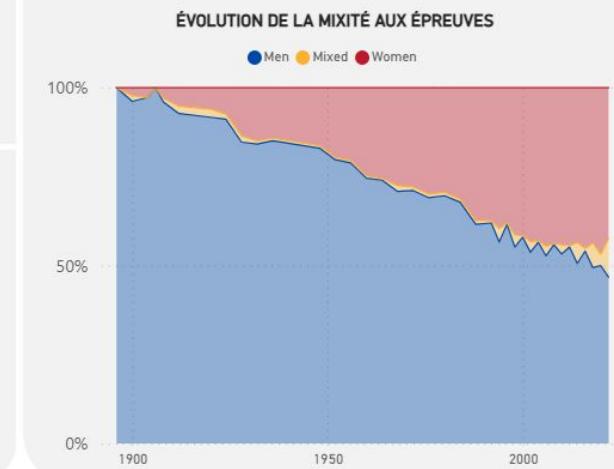
## ENTRE SPÉCIALISATION SPORTIVE ET ÉQUITÉ DE GENRE



### VARIÉTÉ D'ÉPREUVES SELON LA DISCIPLINE

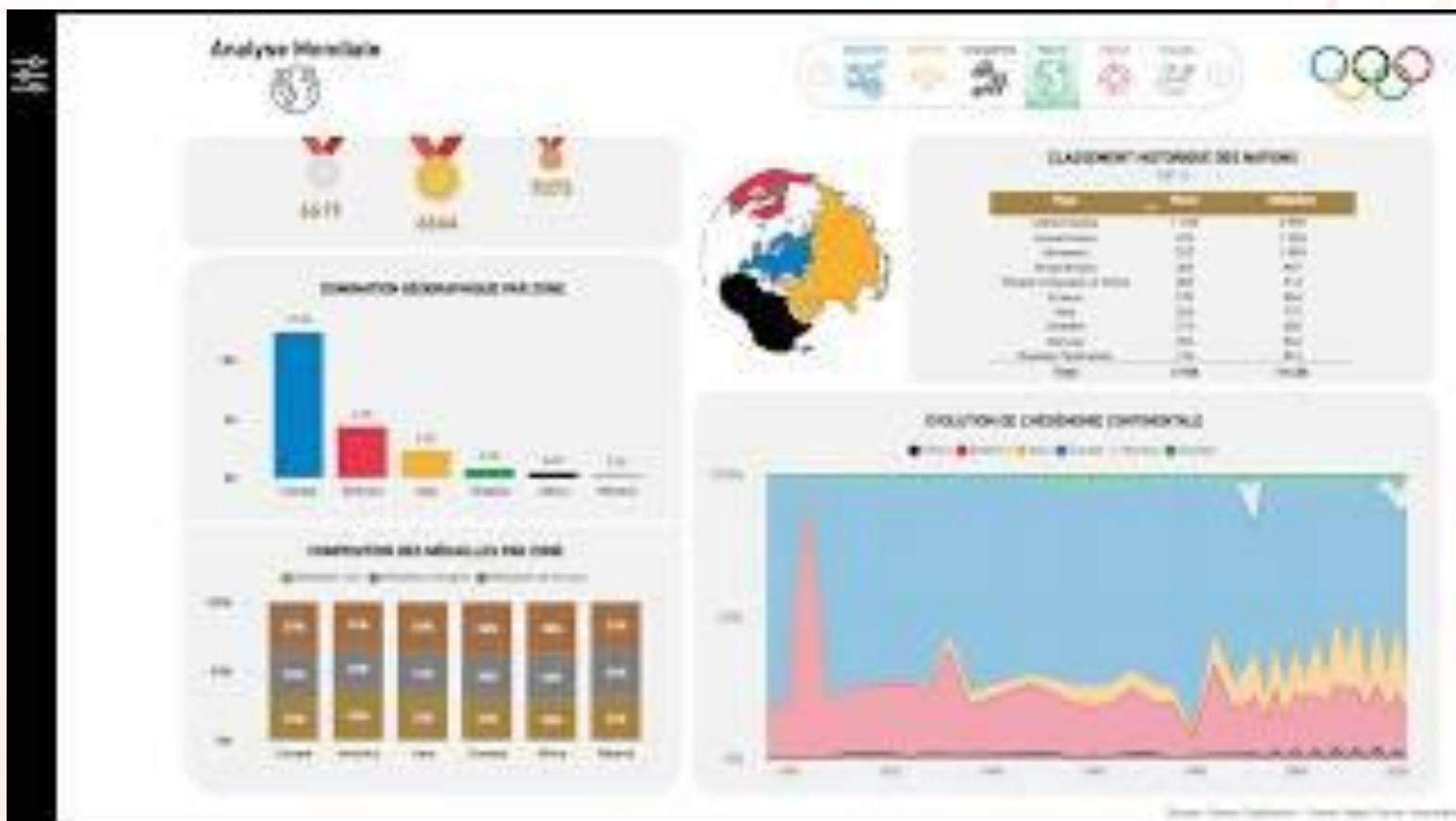
TOP 10

Discipline	Nombre d'épreuves
Athletics	114
Shooting	63
Cycling Track	53
Swimming	42
Artistic Gymnastics	34
Archery	31
Sailing	28
Art Competitions	27
Ballooning	26
Rowing	24



# ANALYSE À TRAVERS LE MONDE

## UNE UNIVERSALITÉ ILLUSOIRE ?



# DOPAGE ET TRICHERIE

## PRISE DE CONSCIENCE DE L'IMPACT GÉOPOLITIQUE SUR LE SPORT



166

Médailles retirées pour dopage



56



52



58

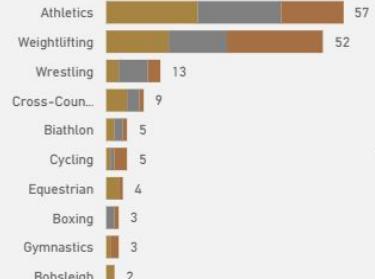
### RECORDS DE RETRAITS PAR ATHLÈTE

TOP 5 DES ATHLÈTES

Athlètes	Médailles retirées	Pays
Marion Jones	5	United States
Johann Mühlegg	3	Spain
Larisa Lazutina	3	Russian Federation
Artur Taymazov	2	Uzbekistan
Elvan Abeylegesse	2	Türkiye
Ilya Ilyin	2	Kazakhstan
Jack Egan	2	United States

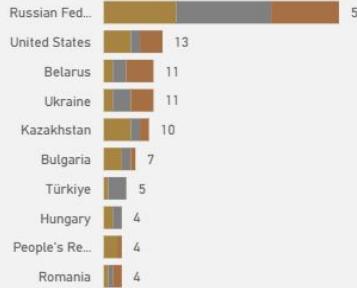
### MÉDAILLES RETIRÉES POUR DOPAGE

PAR DISCIPLINE



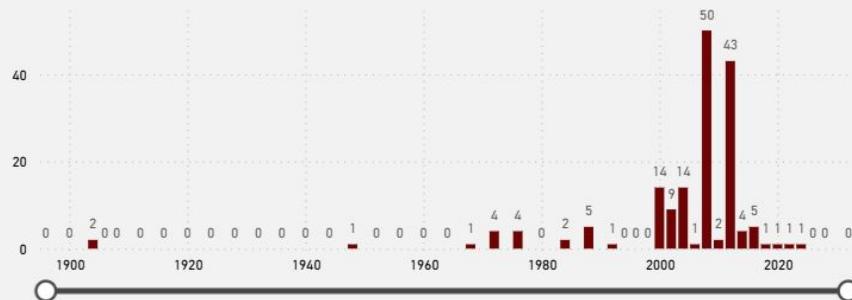
### MÉDAILLES RETIRÉES POUR DOPAGE

PAR PAYS



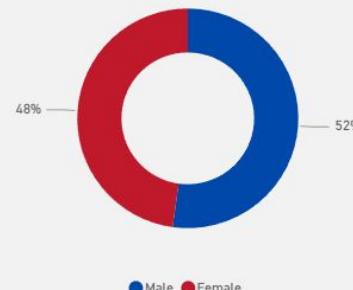
### HISTORIQUE DE LA LUTTE CONTRE LE DOPAGE

NOMBRE DE MÉDAILLES RETIRÉES DANS LE TEMPS



### PROPORTION DES RETRAITS

FEMMES VS HOMMES



# QUELQUES CHIFFRES POUR CONCLURE



**0** Le nombre de JO organisés par tout le continent africain



**7** Le nombre de JO organisés par la France (incluant les futurs JO d'Hiver 2030)



**60** Le nombre de pays n'ayant jamais obtenu de médailles (tous dans l'hémisphère sud)



**5 000 000 000** L'audience pour les JO de Paris 2024



**12 000 000** Le nombre de tickets vendus pour les JO 2024



**+150%** Le dépassement moyen du budget initialement prévu pour chaque organisation de JO depuis 1960



**5 200 000 000\$** Le coût global moyen par organisation de JO depuis 1960



**50 000 000 000\$** Le coût global record pour l'organisation des JO de Sotchi en 2014



# QUELQUES CHIFFRES POUR CONCLURE



**0** Le nombre de JO organisés par tout le continent africain



**7** Le nombre de JO organisés par la France (incluant les futurs JO d'Hiver 2030)



**60** Le nombre de pays n'ayant jamais obtenu de médailles (tous dans l'hémisphère sud)



**5 000 000 000** L'audience pour les JO de Paris 2024



**12 000 000** Le nombre de tickets vendus pour les JO 2024



**+150%** Le dépassement moyen du budget initialement prévu pour chaque organisation de JO depuis 1960



**5 200 000 000\$** Le coût global moyen par organisation de JO depuis 1960



**50 000 000 000\$** Le coût global record pour l'organisation des JO de Sotchi en 2014



# QUELQUES CHIFFRES POUR CONCLURE



**0** Le nombre de JO organisés par tout le continent africain



**7** Le nombre de JO organisés par la France (incluant les futurs JO d'Hiver 2030)



**60** Le nombre de pays n'ayant jamais obtenu de médailles (tous dans l'hémisphère sud)



**5 000 000 000** L'audience pour les JO de Paris 2024



**12 000 000** Le nombre de tickets vendus pour les JO 2024



**+150%** Le dépassement moyen du budget initialement prévu pour chaque organisation de JO depuis 1960



**5 200 000 000\$** Le coût global moyen par organisation de JO depuis 1960



**50 000 000 000\$** Le coût global record pour l'organisation des JO de Sotchi en 2014





# CONCLUSION



# PISTES DE RÉFLEXION À PARTIR DU DASHBOARD

## CONCLUSION



**Universalité et les enjeux financiers comme défi**

### Évolutions techniques du dashboard

- Ajout des nouvelles éditions et des résultats
- Ajout & analyse des JO Paralympiques



# PISTES DE RÉFLEXION À PARTIR DU DASHBOARD

## CONCLUSION



**Universalité et les enjeux financiers comme défi**

### Évolutions techniques du dashboard

- Ajout des nouvelles éditions et des résultats
- Ajout & analyse des JO Paralympiques

## OUVERTURE



**La richesse et la densité de population**  
dictent-elles réellement le palmarès ?



# PISTES DE RÉFLEXION À PARTIR DU DASHBOARD

## CONCLUSION



**Universalité et les enjeux financiers comme défi**

### Évolutions techniques du dashboard

- Ajout des nouvelles éditions et des résultats
- Ajout & analyse des JO Paralympiques

## OUVERTURE



**La richesse et la densité de population**  
dictent-elles réellement le palmarès ?

**Quelle est la viabilité du modèle économique**  
face à la montée des plateformes de streaming et au  
désintérêt relatif des jeunes générations pour la télévision  
linéaire ?



MERCI DE VOTRE ATTENTION



À VOS QUESTIONS





# ANNEXES

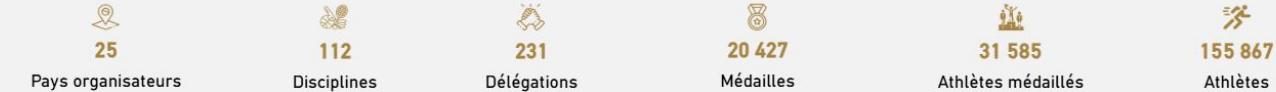


# VUE D'ENSEMBLE

## SYNTHÈSE ET CHIFFRES CLÉS DE 126 ANS D'OLYMPISME



### Vue d'ensemble



#### RECORD DE TITRES

TOP 5

Athlètes	Titres
Michael Phelps	23
Carl Lewis	9
Jenny Thompson	9
Larisa Latynina	9
Mark Spitz	9

#### RECORD DE MÉDAILLES

TOP 5

Athlètes	Médailles
Michael Phelps	28
Larisa Latynina	18
Marit Bjørgen	15
Nikolay Andrianov	15
Boris Shakhlin	13



63

Éditions



35

Été



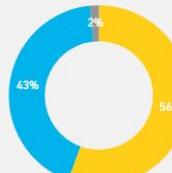
27

Hiver



1

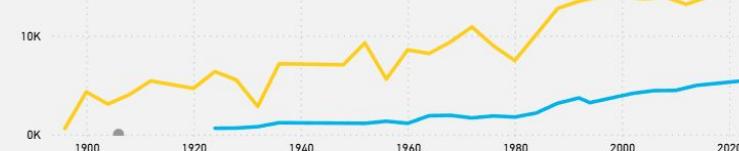
Intercalaire



● Summer ● Winter ● Other

#### CROISSANCE DE LA PARTICIPATION OLYMPIQUE

● Other ● Summer ● Winter



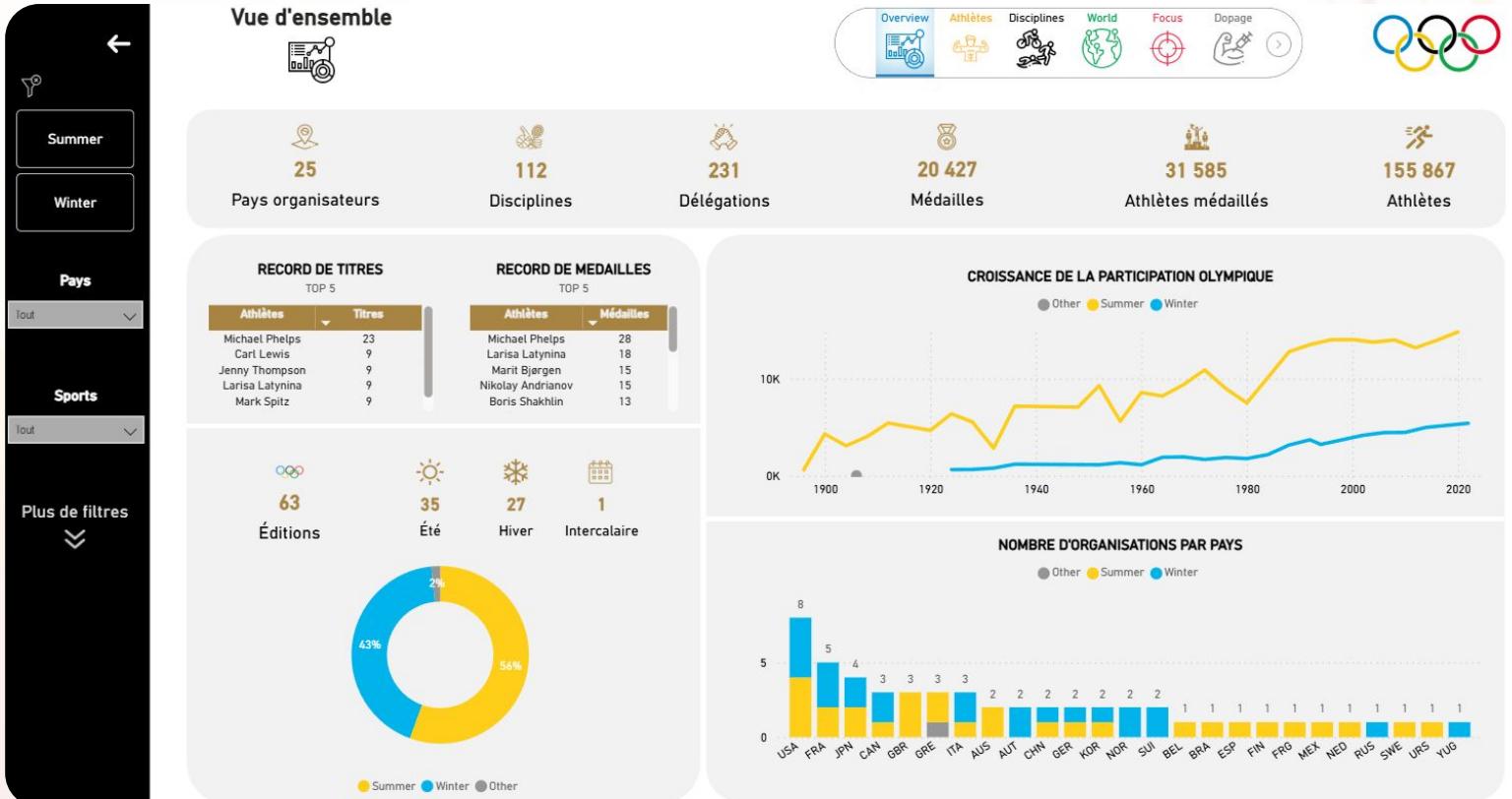
#### NOMBRE D'ORGANISATIONS PAR PAYS

● Other ● Summer ● Winter



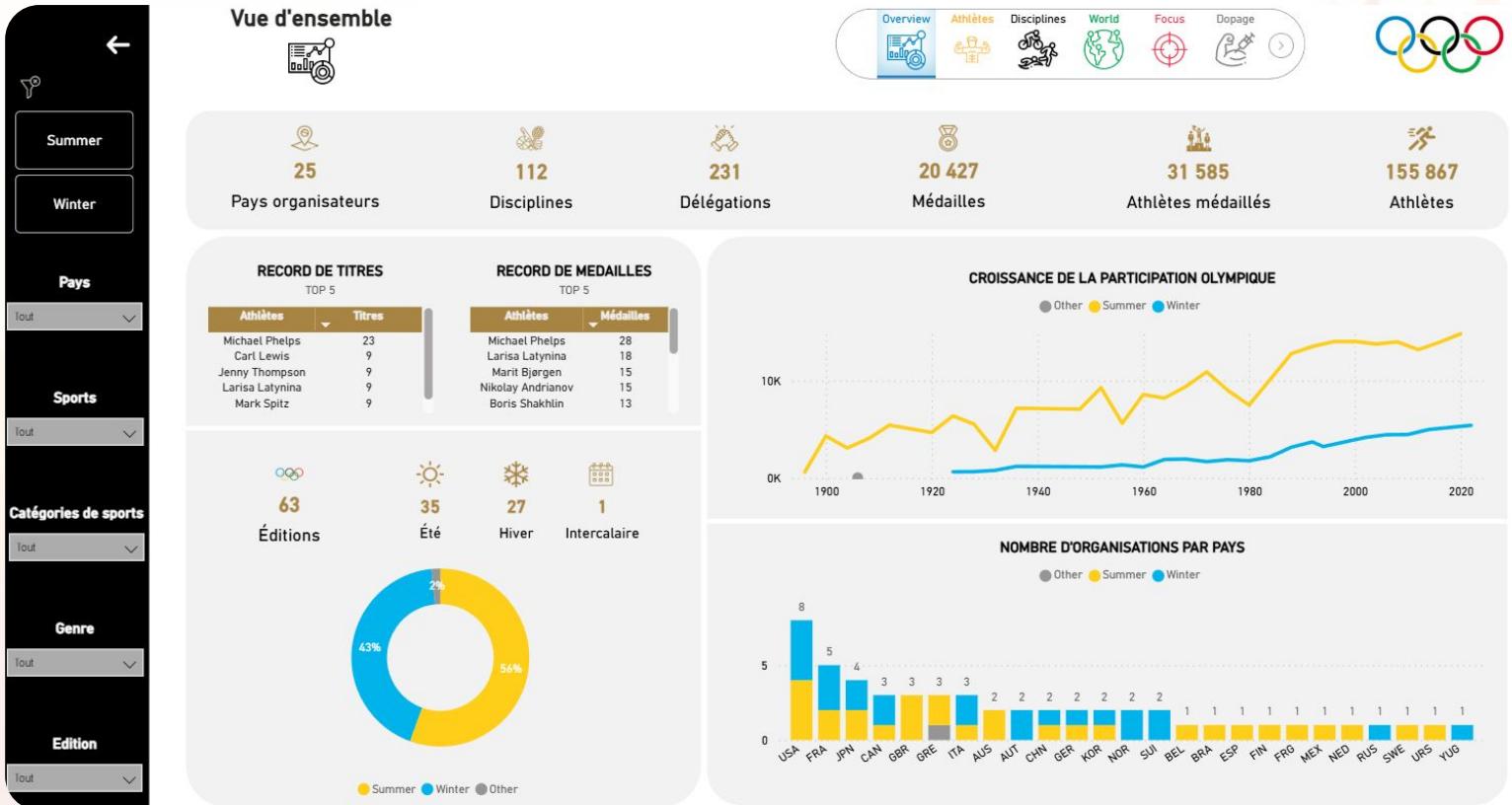
# VUE D'ENSEMBLE

## SYNTHÈSE ET CHIFFRES CLÉS DE 126 ANS D'OLYMPISME



# VUE D'ENSEMBLE

## SYNTHÈSE ET CHIFFRES CLÉS DE 126 ANS D'OLYMPISME

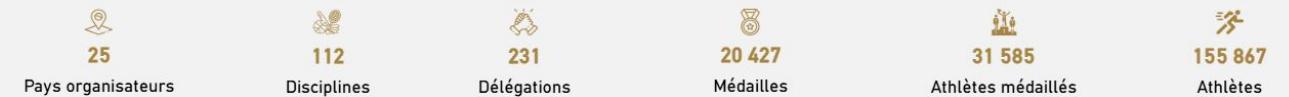


# VUE D'ENSEMBLE

## SYNTHÈSE ET CHIFFRES CLÉS DE 126 ANS D'OLYMPISME



### Vue d'ensemble



#### RECORD DE TITRES

TOP 5

Athlètes	Titres
Michael Phelps	23
Carl Lewis	9
Jenny Thompson	9
Larisa Latynina	9
Mark Spitz	9

#### RECORD DE MÉDAILLES

TOP 5

Athlètes	Médailles
Michael Phelps	28
Larisa Latynina	18
Marit Bjørgen	15
Nikolay Andrianov	15
Boris Shakhlin	13

63  
Éditions



35



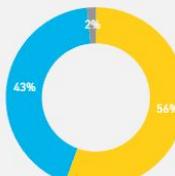
27



1

Été

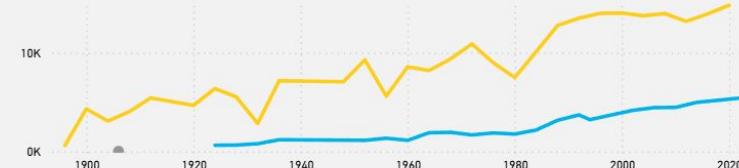
Hiver Intercalaire



● Summer ● Winter ● Other

#### CROISSANCE DE LA PARTICIPATION OLYMPIQUE

● Other ● Summer ● Winter



#### NOMBRE D'ORGANISATIONS PAR PAYS

● Other ● Summer ● Winter



# ANALYSE DES ATHLÈTES

## DU SPORT AMATEUR AU SPORT DE HAUT NIVEAU



### Analyse des athlètes



#### Taille moyenne



226

#### Poids moyen



25

#### Âge moyen

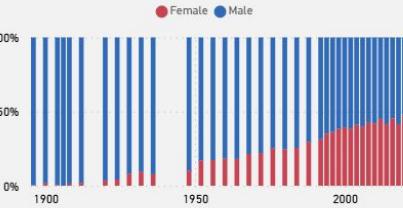


11

#### Ratio par genre



#### PROGRESSION DE LA PARITÉ DANS LE TEMPS



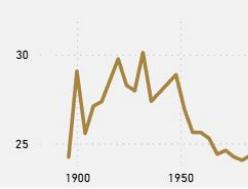
#### ÉVOLUTION DE LA TAILLE



#### ÉVOLUTION DU POIDS



#### ÉVOLUTION DE L'ÂGE



#### LONGEVITÉ : RECORD DE PARTICIPATIONS

TOP 10 DES ATHLETES

Athlètes	Participations
Ian Millar	10
Afanasij Kuzmins	9
Andrew Nicholson	9
Antonio García	9
Hubert Raudaschl	9
José Ramos	9
Luis Martínez	9
Nino Salukvadze	9
Sven Johansson	9
Vladimir Andreyev	9
Vladimir Morozov	9



# ANALYSE DES DISCIPLINES

## ENTRE SPÉCIALISATION SPORTIVE ET ÉQUITÉ DE GENRE

### Analyse des disciplines



#### Individuelle VS Collective



112



739

#### Moyenne d'épreuves par discipline

7

#### COMPARAISON DU VOLUME D'ÉPREUVES : ÉTÉ VS HIVER



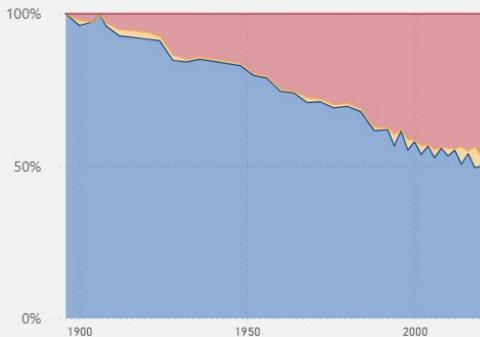
#### VARIÉTÉ D'ÉPREUVES SELON LA DISCIPLINE

TOP 10

Discipline	Nombre d'épreuves
Athletics	114
Shooting	63
Cycling Track	53
Swimming	42
Artistic Gymnastics	34
Archery	31
Sailing	28
Art Competitions	27
Ballooning	26
Rowing	24

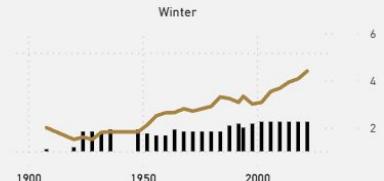
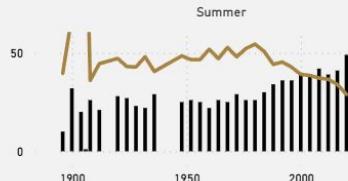
#### ÉVOLUTION DE LA MIXITÉ AUX ÉPREUVES

Men   Mixed   Women



#### ÉVOLUTION DU NOMBRE D'ÉPREUVES PAR DISCIPLINE

● Nombre de disciplines ● Nombre d'épreuves moyen par discipline



# ANALYSE À TRAVERS LE MONDE

## UNE UNIVERSALITÉ ILLUSOIRE ?



### Analyse Mondiale

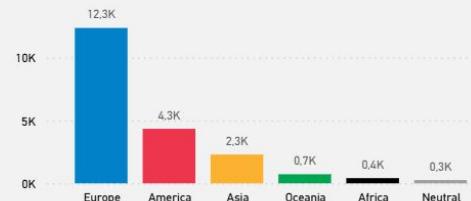


6619

6664

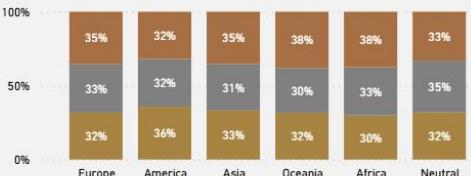
7075

#### DOMINATION GÉOGRAPHIQUE PAR ZONE



#### COMPOSITION DES MÉDAILLES PAR ZONE

● Médailles d'or ● Médailles d'argent ● Médailles de bronze



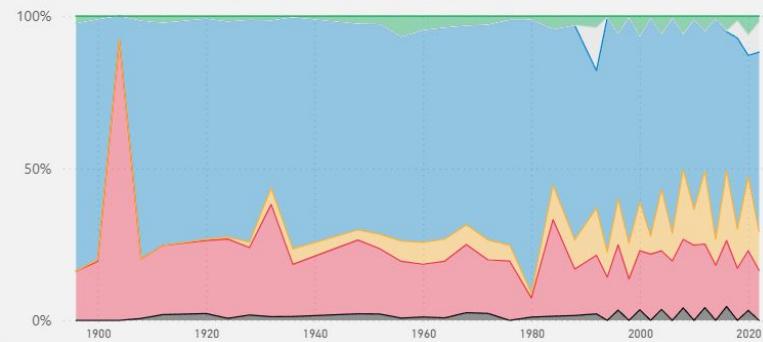
#### CLASSEMENT HISTORIQUE DES NATIONS

TOP 10

Pays	Titres	Médailles
United States	1 185	2 990
Soviet Union	473	1 204
Germany	352	1 083
Great Britain	306	967
People's Republic of China	285	713
France	270	904
Italy	264	773
Sweden	215	684
Norway	204	564
Russian Federation	194	551
Total	3 748	10 433

#### ÉVOLUTION DE L'HÉGEMONIE CONTINENTALE

● Africa ● America ● Asia ● Europe ● Neutral ● Oceania



# ANALYSE À TRAVERS LE MONDE

## UNE UNIVERSALITÉ ILLUSOIRE ?

Analyse spécifique :  
France



Athlètes médaillés en discipline individuelle



Athlètes médaillés en discipline collective

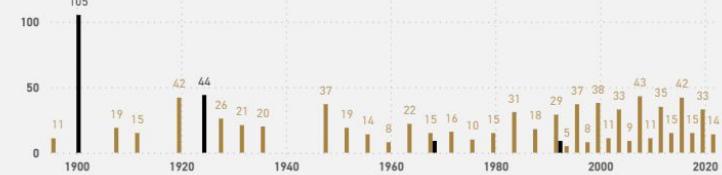


France



### HISTORIQUE DES MÉDAILLES : L'EFFET "À DOMICILE"

Édition à domicile ● False ● True



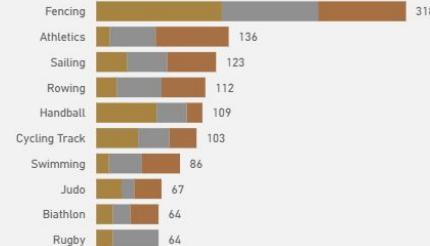
### PROPORTION D'ATHLÈTES MÉDAILLÉS : FRANCE VS MONDE

Édition à domicile ● False ● True ● Moyenne mondiale



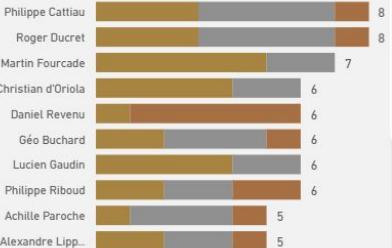
### DISCIPLINES HISTORIQUEMENT DOMINANTES

TOP 10 - PAR DISCIPLINE



### LES LÉGENDES DU SPORT NATIONAL

TOP 10 - PAR ATHLETE



# DOPAGE ET TRICHERIE

## PRISE DE CONSCIENCE DE L'IMPACT GÉOPOLITIQUE SUR LE SPORT

### Analyse du dopage



166  
Médailles retirées pour dopage



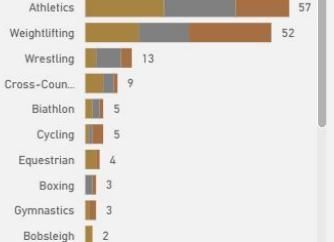
#### RECORDS DE RETRAITS PAR ATHLÈTE

TOP 5 DES ATHLETES

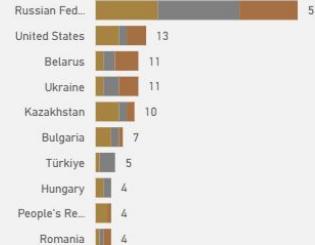
Athlètes	Médailles retirées	Pays
Marion Jones	5	United States
Johann Mühlegg	3	Spain
Larisa Lazutina	3	Russian Federation
Artur Taymazov	2	Uzbekistan
Evan Abeylegesse	2	Türkiye
Ilya Ilyin	2	Kazakhstan
Jack Egan	2	United States



#### MÉDAILLÉES RETIRÉES POUR DOPAGE PAR DISCIPLINE

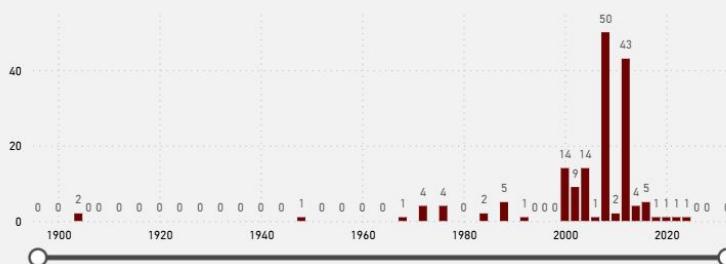


#### MÉDAILLÉES RETIRÉES POUR DOPAGE PAR PAYS



#### HISTORIQUE DE LA LUTTE CONTRE LE DOPAGE

NOMBRE DE MÉDAILLÉES RETIRÉES DANS LE TEMPS



#### PROPORTION DES RETRAITS

FEMMES VS HOMMES

