

CODERHOUSE
DATA ANALYTICS - Turma 45665
Documentação do Projeto Final

Equipe: Jessica Hemily, João de Mantova, Julia Lima e Kamylla de Lima Alves .

PARTICIPAÇÃO FEMININA NOS JOGOS OLÍMPICOS

Uma análise da participação feminina nos jogos olímpicos ao longo das edições

Data da apresentação: 17/05/2023

ÍNDICE

1 Introdução	3
2 Fonte de dados	3
3 Tabela de controle de versão	3
4 Objetivo	4
Análise da evolução, de 1896 até 2022, da participação de mulheres em esportes olímpicos ocupados majoritariamente por homens.	4
5 Alcance	4
6 Usuário final e nível de aplicação da análise	4
7 Diagrama de relacionamentos entre entidades das tabelas selecionadas	5
8 Tabelas, com definição dos campos/entidade, tipo dos campos e definição de chave primária e/ou chave estrangeira.	5
9 Transformação	6
10 Análise funcional Dashboard	7
10.1 Filtros: Em ambas as telas é possível ver um conjunto de filtros que permite navegar entre as edições, países, sexo do atleta e período da edição.	7
10.2 Página 1: Visão Geral	7
10.2.1 Nome do gráfico: Conjunto de cards	7
10.2.2 Nome do gráfico: Top 10 países com mais medalhas	8
10.2.3 Nome do gráfico: Relação de medalhas por país	8
10.3 Página 2: Visão Geográfica/Temporal	8
10.4 Página 3: Visão Temporal	9
11 Medidas calculadas	9
11.1 Nome da medida: % de sexo feminino	9
11.2 Nome da medida: % sexo masculino	9
11.3 Nome da medida: % todos mistos	10
11.4 Nome da medida: No de edições	10
12 Colunas calculadas	10
12.1 Nome da coluna calculada: Sex2	10
12.2 Nome coluna calculada: sex num	10
12.3 Nome coluna calculada: Ano	10
12.4 Tabela calendário	11
12.5 Medida calculada com variável (1 medida de agregação)	11
12.6 Medida calculada com duas variáveis (1 função de agregação + inteligência de tempo)	12
12. 7 Medida calculada com parâmetro (1 função de agregação)	12
13 Ferramentas tecnológicas implementadas	13
14 Diretrizes futuras	13

1 Introdução

A participação das mulheres nas Olimpíadas tem passado por uma expressiva evolução ao longo das últimas décadas. Durante grande parte da história dos Jogos Olímpicos, as mulheres enfrentaram desafios e restrições que limitavam seu envolvimento e sua representação no evento.

A primeira participação de mulheres nos jogos ocorreu em 1900, edição que foi sediada em Paris, no entanto, ainda era bastante limitada e restrita a esportes considerados adequados para mulheres na época, como tênis, mergulho e esgrima.

Com o passar das décadas as mulheres foram ganhando mais espaço, como em 1928 com a inclusão do atletismo feminino (jogos sediados na cidade de Amsterdã, nos Países Baixos) e em 1976 com a introdução do basquete feminino (jogos sediados no Canadá).

Nos dias atuais a porcentagem de jogos destinados a mulheres nas olimpíadas é de aproximadamente 41%, não é o ideal, porém os números têm se mostrado bem otimistas, pensando em uma projeção futura, levando em consideração o crescimento da taxa de participação feminina.

2 Fonte de dados

Arquivos no formato csv adquiridos do portal Kaggle.

Link de acesso aos dados utilizados:

[Olympic Historical Dataset From Olympedia.org | Kaggle](https://www.kaggle.com/datasets/olympic-historical-dataset)

Arquivos empregados:

1. Olympic_Athlete_Bio.csv
2. Olympic_Athlete_Event_Results.csv
3. Olympic_Games_Medal_Tally.csv
4. Olympic_Results.csv
5. Olympics_Country.csv
6. Olympics_Games.csv

3 Tabela de controle de versão

Entrega	Data
Modelo relacional e bridge table. Transformações de tipos de dados e eliminação de colunas.	18/04/2023

Medidas calculadas, colunas calculadas e tabela calendário.	20/04/2023
KPI's, títulos, subtítulos, logo, filtros e 5 gráficos	25/04/2023
Medidas calculadas com variáveis e parâmetros e ajustes de layout	02/05/2023
Ajustes de layout e gráficos considerando feedbacks apresentados na Apresentação do projeto	03/06/2023

4 Objetivo

Análise da evolução, de 1896 até 2022, da participação de mulheres em esportes olímpicos ocupados majoritariamente por homens.

5 Alcance

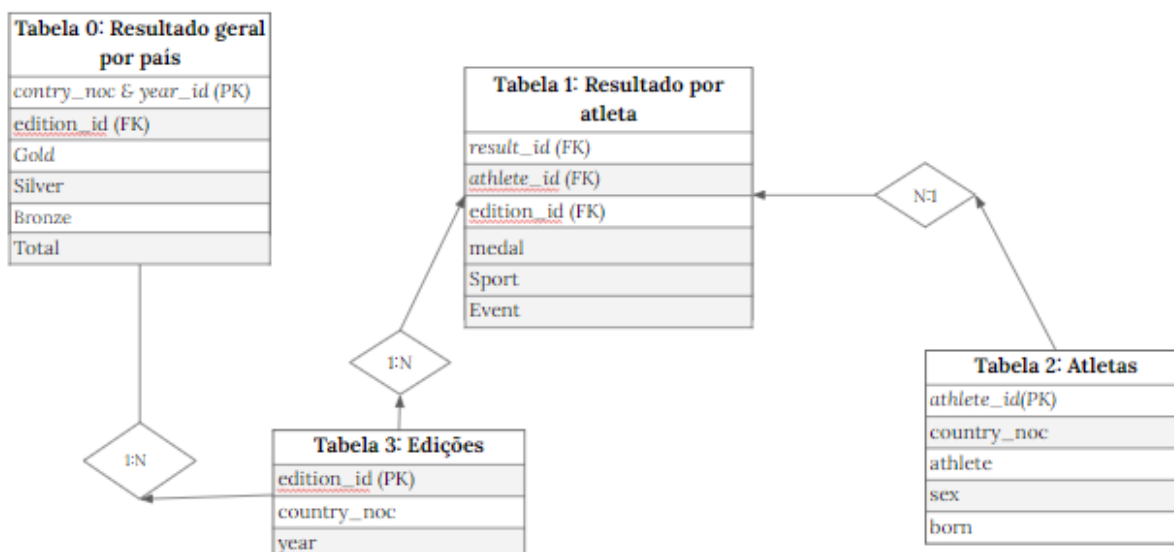
Turma de Data Analytics da Coderhouse e público livre do GitHub.

6 Usuário final e nível de aplicação da análise

Pessoas interessadas em esportes, análise de dados e equidade de gênero.
Nosso nível de aplicação é estratégico.

7 Diagrama de relacionamentos entre entidades das tabelas selecionadas

Diagrama E-R: Tema 1 (esportes)



8 Tabelas, com definição dos campos/entidade, tipo dos campos e definição de chave primária e/ou chave estrangeira.

Tabela 0: Resultado geral por país		
Tipo de chave	Campo	Tipo de campo
PK	contry_noc & year_id (PK)	Varchar (10)
FK	edition_id (FK)	int (2)
-	Gold	int (2)
-	Silver	int (2)
-	Bronze	int (2)
-	Total	int (3)

Tabela 1: Resultado por atleta		
Tipo de chave	Campo	Tipo de campo
FK	result_id (FK)	int (8)
FK	athlete_id (FK)	int (8)
FK	edition_id (FK)	int (2)
-	medal	int (3)
-	Sport	Varchar (50)

-	Event	Varchar (100)
---	-------	---------------

Tabela 2: Atletas		
<i>Tipo de chave</i>	<i>Campo</i>	<i>Tipo de campo</i>
PK	athlete_id(PK)	int (8)
-	country_noc	texto (3)
-	athlete	texto (50)
-	sex	Varchar (6)
-	born	date

Tabela 3: Edições		
<i>Tipo de chave</i>	<i>Campo</i>	<i>Tipo de campo</i>
PK	edition_id (PK)	int (2)
-	contry_noc	texto (3)
-	year	year (4)

9 Transformação

Transformações feitas no Power BI foram:

Em todas as tabelas: Foi alterada a primeira linha como cabeçalho.

Tabela 0: Resultado Geral

- *Edition_id* de número para texto

Tabela 1: Resultado por atleta

- *Athlete_id*, *Result_id* e *Edition_id* de número para texto

Tabela 2: Atleta

- *Athlete_id* de número para texto

Tabela 3: Esporte

- *Result_id* e *Athlete_id* de número para texto

Tabela 4: Edições

- *Edition_id* de números para texto

10 Análise funcional Dashboard

Informação à ser analisada por página/gráfico.

Considerações importantes: a tabela Tabela2_Atleta possui muitos registros faltantes referente aos atletas que participaram dos jogos. Sendo assim, ao buscar o sexo do participante (por intermédio do identificador athlete_id), para identificar se o resultado apresentado na Tabela1_Resultadoporatleta foi obtido por uma mulher ou homem não se tem um retorno favorável para a análise, devido ao grande número de campos vazios retornados.

Por este motivo, o número de homens e mulheres foram determinados pela definição do evento do esporte realizado, em que traz em sua descrição se a modalidade foi disputada por um competidor do sexo masculino ou feminino, sendo assim, o número de atletas homens e mulheres apresentados no projeto podem ser menor do que o número real, uma vez que os atletas dos jogos mistos foram desconsiderados nessa contagem.

10.1 Filtros: Em ambas as telas é possível ver um conjunto de filtros que permite navegar entre as edições, países, sexo do atleta e período da edição.

- Página Visão Geral:
 - Filtros por País;
 - Período.
- Página Geográfica/Temporal:
 - Filtro por período;
 - Uso de botões indicadores (HOMENS e Mulheres) para selecionar os elementos que serão mostrados na visualização.
- Página Temporal:
 - Filtro por período;
 - Botões “% de homens” e “% de mulheres”.

10.2 Página 1: Visão Geral

10.2.1 Nome do gráfico: Conjunto de cards

- **Gráfico & Informação à ser analisada:** Fileira de KPIs, da esquerda para a direita, contendo cards com as porcentagem de participantes por sexo, quantidade de medalhas ganhas por tipo e número de edições.

- % sexo feminino: porcentagem de mulheres que participaram dos jogos exclusivos para mulheres
`percent_F = [esporte_exclusivo_sexo_F]/COUNTROWS(Tabela1_Resultadoporatleta)`
- % jogos mistos:
`percent_openmix = [esporte_sexo_OpenMix]/COUNTROWS(Tabela1_Resultadoporatleta)`
- % sexo masculinos: porcentagem de homens que participaram dos jogos exclusivos para homens.
`percent_M = [esporte_exclusivo_sexo_M]/COUNTROWS(Tabela1_Resultadoporatleta)`
- Ouro ganhos: Soma dos valores da coluna Gold
- Pratas ganhas: Soma dos valores da coluna Silver
- Bronzes ganhos: Soma dos valores da coluna Bronze
- N° Edições:
`Qtd Edições = CALCULATE(DISTINCTCOUNT(Tabela0_Resultadogeral[edition_id]))`

10.2.2 Nome do gráfico: Top 10 países com mais medalhas

- **Gráfico & Informação à ser analisada:** Matriz com apresentação do quantitativo de medalhas (ouro, prata e bronze) por país

10.2.3 Nome do gráfico: Relação de medalhas por país

- **Gráfico & Informação à ser analisada:** Gráfico de barras com quantidade total e tipos de medalhas ganhas por país

10.3 Página 2: Visão Geográfica/Temporal

10.3.1 Nome do gráfico: Países com a maior quantidade de atletas do sexo feminino

- **Gráfico & Informação à ser analisada:** Mapa com bolha indicado os países com a maior quantidade de atletas do sexo feminino

10.3.2 Nome do gráfico: % de participação de MULHERES/HOMENS por ano.

- **Gráfico & Informação à ser analisada:** Mapa de árvore com a % de atletas por sexo em cada ano de participação. É possível ajustar o sexo da visualização de acordo pelo parametro aplicado e indicador acionado por intermédio de botões.

10.4 Página 3: Visão Temporal

10.4.1 Nome do gráfico: “Porcentagem - sexo masculino” e “Porcentagem - sexo feminino”. Nome do gráfico varia conforme o parâmetro

- **Gráfico & Informação a ser analisada:** Mapa de linhas com parâmetro aplicado que permite navegar entre as porcentagens de atletas masculinos e femininos.

10.4.2 Nome do gráfico: Porcentagem de mulheres

- **Gráfico & Informação a ser analisada:** Gráfico de rosca que informa a participação acumulada de mulheres, em %. Valor pode ser ajustado de acordo com o filtro de período aplicado.
 - % sexo feminino
`percent_F =
[esporte_exclusivo_sexo_F]/COUNTROWS(Tabela1_Resultadoporatleta)`

11 Medidas calculadas

11.1 Nome da medida: % de sexo feminino

- **Campos que compõem a medida:**
Tabela1_Resultadoporatleta
- **Finalidade da medida/análise:** Informar o valor percentual de pessoas do sexo masculinos que participaram dos jogos

11.2 Nome da medida: % sexo masculino

- **Campos que compõem a medida:**
Tabela1_Resultadoporatleta
- **Finalidade da medida/análise:** Informar o valor percentual de pessoas do sexo masculino que participaram dos jogos

11.3 Nome da medida: % todos mistos

- **Campos que compõem a medida:**
Tabela1_Resultadoporatleta
- **Finalidade da medida/análise:** Informar o valor percentual de pessoas que participaram de times mistos

11.4 Nome da medida: No de edições

- **Campos que compõem a medida:**
Tabela0_Resultadogeral[edition_id]
- **Finalidade da medida/análise:** Informar a quantidade total de edições realizadas até 2022

12 Colunas calculadas

12.1 Nome da coluna calculada: Sex2

- **Tabela:** Tabela2_Atleta
- **Finalidade da medida/análise:** Ajuste da nomenclatura do sexo dos participantes com a finalidade de padronização entre tabelas
- **Fórmula DAX:**

```
sex2 =  
IF(Tabela2_Atleta[sex]="Male", "Men", IF(Tabela2  
_Atleta[sex]="Female", "Women"))
```

12.2 Nome coluna calculada: sex num

- **Tabela:** Tabela1_Resultadoporatleta
- **Finalidade da medida/análise:** Atribui característica numérica ao tipo de sexo, facilitando o processo de contagem
- **Fórmula DAX:**

```
sex num = if(CONTAINSSTRING(  
Tabela1_Resultadoporatleta[sex], "Women"), "1",  
if(CONTAINSSTRING(Tabela1_Resultadoporatleta[sex], "  
Men"), "2", if(CONTAINSSTRING(Tabela1_Resultadoporatl  
eta[sex], "Open/Mix"), "3")))
```

12.3 Nome coluna calculada: Ano

- **Tabela:** CALENDARIO

- **Finalidade da medida/análise:** Extrair o ano da coluna de Date da tabela CALENDARIO.
- **Fórmula DAX:**
Ano = `YEAR('CALENDARIO'[Date])`

12.4 Tabela calendário

- **Nome da tabela:** CALENDARIO
 - **Finalidade:** Trazer coluna com datas de 01/01/1896 até data atual, para posterior extração dos anos.
 - **Fórmula DAX:**
`CALENDARIO = CALENDAR(DATE(1896,1,1),TODAY())`

12.5 Medida calculada com variável (1 medida de agregação)

12.5.1 Nome da medida: porc_jogos_exclusivo_F

- **Campos que compõem a medida:** porc_jogos_exclusivo_F
- **Finalidade da medida/análise:** Porcentagem de jogos exclusivos por pessoas do sexo feminino
- **Fórmula DAX:**
`porc_jogos_exclusivo_F =`

```
var quant_jogos_F =
CALCULATE(COUNTROWS(Tabela1_Resultadoporatleta
),Tabela1_Resultadoporatleta[sex]="Women")

return
quant_jogos_F/COUNTROWS(Tabela1_Resultadoporatleta)
```

12.5.2 Nome da medida: porc_jogos_exclusivo_M

- **Campos que compõem a medida:** porc_jogos_exclusivo_M
- **Finalidade da medida/análise:** Porcentagem de jogos exclusivos por masculino
- **Fórmula DAX:**
`porc_jogos_exclusivo_M =`
`var quant_esporte_M = CALCULATE(`
`COUNTROWS(Tabela1_Resultadoporatleta),Tabela1_`
`Resultadoporatleta[sex]="Men"`
`)`

```

var total_atletas =
COUNTROWS (Tabela1_Resultadoporatleta)
return
quant_esporte_M/total_atletas

```

12.6 Medida calculada com duas variáveis (1 função de agregação + inteligência de tempo)

12.6.1 Nome da medida: porcentagem exclusivo de homens e mulheres através do tempo

- **Campos que compõem a medida:** Atleta por sexo e sexo
- **Finalidade da medida/análise:** O mapa de linhas apresenta o aumento ou declínio referente a quantidade de participantes de cada sexo ao passar dos anos
- **Fórmula DAX:**

```

Sexo = {
    ("% exclusiva de sexo masculino",
NAMEOF ('MEDIDAS' [porcent_M]), 0),
    ("% exclusiva de sexo feminino",
NAMEOF ('MEDIDAS' [porcent_F]), 1)
}

```

12.7 Medida calculada com parâmetro (1 função de agregação)

12.7.1 Nome da medida: Treemap

- **Campos que compõem a medida:** porcent_M e Porcent_F
- **Finalidade da medida/análise:** Apresentar/Filtrar de forma comparativa os anos que tiveram maior porcentagem de participantes de cada grupo (sexo feminino e sexo masculino)
- **Fórmula DAX:**

```

Sexo = {
    ("% exclusiva de sexo masculino",
NAMEOF ('MEDIDAS' [porcent_M]), 0),
    ("% exclusiva de sexo feminino",
NAMEOF ('MEDIDAS' [porcent_F]), 1)
}

```

13 Ferramentas tecnológicas implementadas

As ferramentas utilizadas foram SQL Server - preparação e análise de dados - e Microsoft Power BI - transformação, análise, visualização dos dados e elaboração do dashboard.

14 Diretrizes futuras

Embora ainda não tenhamos alcançado a equidade de gênero no esporte, a participação de mulheres nos Jogos Olímpicos têm apresentado uma evolução significativa ao longo das últimas décadas. Para garantir o crescimento contínuo da participação feminina, e buscar a equidade, é necessário fornecer incentivos e investimentos adequados em programas de desenvolvimento esportivo para mulheres.