

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA TOÁN - TIN HỌC



BÁO CÁO CUỐI KỲ
NHÓM 03

Project: Paris 2024 Olympic Summer Games

GV: T.S Tô Đức Khánh

Sinh viên thực hiện:

<i>Họ và tên</i>	<i>MSSV</i>	<i>Mã lớp</i>
Trần Ngọc Dễ	21110057	21TTH1
Huỳnh Nguyễn Thế Dân	21110256	21TTH1
Trần Thị Bích Tuyền	21280059	21KDL1
Nguyễn Thị Yến Như	21280082	21KDL1
Huỳnh Công Đức	21280124	21KDL1
Trần Thị Uyên Nhi	21280125	21KDL1

TPHCM, ngày 17 tháng 01 năm 2025

Mục lục

1	Tổng quan project	3
1.1	Nguồn dữ liệu	3
1.2	Phạm vi sử dụng dữ liệu	3
1.3	Các bước chính	4
2	Đề xuất và mục tiêu trực quan mong muốn	4
2.1	Tổng quan sự kiện	4
2.2	Phân tích vận động viên theo môn thể thao	5
2.3	Phân tích các quốc gia tham gia	5
2.4	Phân tích thành tích huy chương	6
2.5	Khám phá lịch trình và địa điểm	7
3	Xử lý dữ liệu	7
3.1	Quá trình biến đổi dữ liệu	7
3.2	Quá trình thêm dữ liệu trong Data View	8
3.3	Các measures sử dụng	9
3.4	Chọn dữ liệu cho mỗi Dashboard	13
4	Xây dựng Dashboard	14
4.1	Dashboard Overview	14
4.1.1	Tổng quan	14
4.1.2	Các thành phần của dashboard	15
4.1.3	Nhận xét và kết luận chung	17
4.2	Dashboard Athletes	18
4.2.1	Tổng quan	18
4.2.2	Các thành phần của dashboard	18
4.2.3	Nhận xét và kết luận chung	21
4.3	Dashboard Country	21
4.3.1	Tổng quan	21
4.3.2	Các thành phần của dashboard	22
4.3.3	Nhận xét và kết luận chung	24

4.4	Dashboard Medal	25
4.4.1	Tổng quan	25
4.4.2	Các thành phần của dashboard	25
4.4.3	Nhận xét và kết luận chung	28
4.5	Dashboard Event	28
4.5.1	Tổng quan	28
4.5.2	Các thành phần của dashboard	29
4.5.3	Nhận xét và kết luận chung	31
5	Tổng kết chung	31

1 Tổng quan project

Project tập trung vào trực quan hóa dữ liệu liên quan đến Thế vận hội Olympic mùa hè Paris 2024. Mục tiêu là tạo ra các Dashboard tổng hợp để minh họa thông tin chi tiết về các quốc gia tham gia, vận động viên, môn thể thao, bảng xếp hạng huy chương và các sự kiện. Kết quả sẽ giúp người xem dễ dàng nắm bắt các xu hướng, thành tích nổi bật và những khía cạnh đáng chú ý của sự kiện.

1.1 Nguồn dữ liệu

Bộ dữ liệu bao gồm 13 tệp CSV, mỗi tệp đóng vai trò cung cấp một khía cạnh khác nhau về Thế vận hội:

1. `athletes.csv`: Thông tin cá nhân của tất cả vận động viên (tên, tuổi, quốc gia, giới tính, môn thể thao).
2. `coaches.csv`: Thông tin cá nhân của tất cả huấn luyện viên (tên, quốc gia, môn thể thao).
3. `events.csv`: Thông tin về tất cả các sự kiện và địa điểm tổ chức.
4. `medals.csv`: Chi tiết về các vận động viên/đội nhận huy chương.
5. `medals_total.csv`: Tổng hợp số lượng huy chương (vàng, bạc, đồng) theo quốc gia.
6. `medallists.csv`: Danh sách chi tiết những người giành huy chương.
7. `nocs.csv`: Thông tin về quốc gia (mã NOC, tên quốc gia).
8. `schedule_preliminary.csv`: Lịch trình sơ bộ các sự kiện.
9. `schedule.csv`: Lịch trình chính thức theo ngày của các sự kiện.
10. `teams.csv`: Thông tin các đội thi đấu.
11. `technical_officials.csv`: Danh sách các quan chức kỹ thuật.
12. `torch_route.csv`: Địa điểm rước đuốc Olympic.
13. `venues.csv`: Danh sách các địa điểm tổ chức sự kiện.

1.2 Phạm vi sử dụng dữ liệu

- `athletes.csv`
- `medals.csv`
- `medals_total.csv`
- `nocs.csv`

- events.csv
- schedule.csv
- venues.csv
- torch_route.csv
- teams.csv

1.3 Các bước chính

1. **Tích hợp các tệp dữ liệu đã cung cấp vào Power BI:** Đảm bảo rằng các tệp dữ liệu được tải lên Power BI và sắp xếp hợp lý để phục vụ các bước phân tích tiếp theo.
2. **Làm sạch và xử lý dữ liệu để đảm bảo tính nhất quán:** Xử lý các giá trị bị thiếu, loại bỏ dữ liệu trùng lặp và chuẩn hóa định dạng để tăng độ chính xác của phân tích.
3. **Thiết kế các biểu đồ phù hợp với mục tiêu đề ra:** Lựa chọn các loại biểu đồ như biểu đồ cột, biểu đồ đường, hoặc biểu đồ tròn để biểu diễn thông tin theo cách trực quan và dễ hiểu.
4. **Xây dựng dashboard trực quan, dễ hiểu và thẩm mỹ:** Sắp xếp các thành phần trên dashboard hợp lý, sử dụng màu sắc và nhãn chú thích để tăng cường trải nghiệm người dùng.

2 Đề xuất và mục tiêu trực quan mong muốn

2.1 Tổng quan sự kiện

Đề xuất trực quan dữ liệu:

1. Card
 - Hiển thị số lượng vận động viên, đội tuyển, quốc gia tham dự, và tổng số huy chương.
 - Lý do: Cung cấp tóm tắt thông tin chính một cách rõ ràng, dễ hiểu.
2. Stacked column chart
 - Thống kê tổng số huy chương (chia thành vàng, bạc, đồng) được trao.
 - Lý do: Giúp người xem hiểu rõ phân bố các loại huy chương.
3. Map

- Minh họa các quốc gia tham dự trên bản đồ thế giới với màu sắc đại diện cho số lượng vận động viên.
- Lý do: Tăng tính trực quan, hiển thị sự tham gia toàn cầu.

Mục tiêu trực quan:

- Số lượng vận động viên, đội tuyển, quốc gia tham dự.
- Tổng số huy chương được trao.
- Hiển thị bản đồ các quốc gia tham dự.

2.2 Phân tích vận động viên theo môn thể thao

Đề xuất trực quan dữ liệu:

1. Clustered column chart

- Phân bố độ tuổi của vận động viên theo giới tính.
- Lý do: Minh họa độ tuổi phổ biến của vận động viên, chia theo nam và nữ.

2. Stacked column chart

- Số lượng vận động viên đạt huy chương, phân theo loại huy chương (vàng, bạc, đồng).
- Lý do: Phân tích chi tiết đóng góp của vận động viên.

3. Filter/Search Box

- Tìm kiếm thông tin của từng vận động viên, bao gồm thành tích và môn thi đấu.
- Lý do: Tăng tính tương tác và cá nhân hóa trải nghiệm người dùng.

Mục tiêu trực quan:

- Thông tin chi tiết về vận động viên (tuổi, giới tính, huy chương).
- Phân loại vận động viên theo môn thi đấu.

2.3 Phân tích các quốc gia tham gia

Đề xuất trực quan dữ liệu:

1. Stacked bar chart

- Số lượng huy chương của các quốc gia, tập trung vào top 10 quốc gia.

- Lý do: Dễ dàng so sánh số lượng huy chương giữa các quốc gia.

2. Map Visualization

- Bản đồ hiển thị số lượng huy chương của các quốc gia với độ đậm nhạt của màu sắc.
- Lý do: Thể hiện mức độ tham gia toàn cầu một cách trực quan.

3. Stacked bar chart

- Số lượng quốc gia tham gia theo từng môn thể thao.
- Lý do: Nhấn mạnh môn thể thao phổ biến tại các quốc gia khác nhau.

Mục tiêu trực quan:

- Phân tích số lượng huy chương theo quốc gia.
- Thống kê số lượng quốc gia theo từng môn thể thao.

2.4 Phân tích thành tích huy chương

Đề xuất trực quan dữ liệu:

1. Pie chart:

- Thống kê tỷ lệ huy chương theo giới tính.
- Lý do: Dễ dàng so sánh thành tích theo giới tính.

2. Tree Map

- Phân tích số lượng huy chương theo từng môn thể thao.
- Lý do: Nhấn mạnh vai trò và thành tích của các môn thể thao quan trọng.

3. Area chart

- Phân tích số huy chương của từng quốc gia, chia thành vàng, bạc, đồng.

Mục tiêu trực quan:

- Thống kê huy chương (vàng, bạc, đồng) theo quốc gia, môn thể thao, và vận động viên.
- So sánh hiệu suất giữa các quốc gia.

2.5 Khám phá lịch trình và địa điểm

Đề xuất trực quan dữ liệu:

1. Gantt Chart

- Lịch thi đấu theo từng ngày với các môn thi đấu được phân loại rõ ràng.
- Lý do: Giúp người xem theo dõi lịch trình dễ dàng.

2. Gantt Chart

- Lộ trình rước đuốc minh họa trên dòng thời gian với các địa điểm chính.
- Lý do: Nhấn mạnh hành trình và biểu tượng của sự kiện Olympic.

Mục tiêu trực quan:

- Hiển thị lịch thi đấu chi tiết.
- Minh họa địa điểm tổ chức trên bản đồ.
- Lộ trình rước đuốc theo thời gian.

3 Xử lý dữ liệu

3.1 Quá trình biến đổi dữ liệu

Trong quá trình biến đổi dữ liệu, nhóm có sử dụng thêm 2 files csv để tăng màu sắc và hoàn thiện cho Dashboard, bao gồm một file **Icons** của các môn thi đấu và một file **Flags** chứa cờ của các quốc gia.

1. Athletes

Trong table **Athletes**, tiến hành **Merge Queries** đến table **Flags** thông qua giá trị **Join** vào là “Country code”.

2. Medallists

Trong table **Medallists**, tiến hành **Merge Queries** đến table **Icons** thông qua giá trị **Join** vào là “discipline”. Sau đó, ta **Merge Queries** đến table **Flags** thông qua giá trị **Join** vào là “Country code”.

3. Medals

Trong table **Medals**, ta cũng tiến hành tương tự như table **Medallists** là **Merge Queries** đến table **Icons** thông qua giá trị **Join** vào là “discipline”. Sau đó, ta **Merge Queries** đến table **Flags** thông qua giá trị **Join** vào là “Country code”.

4. Disciplines

Việc sử dụng thêm cột này lý do là vì ở table **Athletes**, có một vài vận động viên tham dự cả nhiều hơn 1 nội dung, làm ta bị khó khăn trong việc tổng hợp số lượng vận động viên trong một nội dung bất kỳ. Đây là một table mới, được Duplicate từ table **Athletes**, giữ lại các cột là “code”, “disciplines”, “country_code”, “country”. Sau đó, ta tiến hành sửa lại nội dung cột “discipline” để loại bỏ những ký tự như “[”, “]”, “’”. Tiếp theo đó, ta sẽ Split cột “discipline” theo Delimiter là “’”. Cuối cùng là Merge Queries đến table **Icons** thông qua giá trị Join vào là “discipline”.

3.2 Quá trình thêm dữ liệu trong Data View

1. Athletes

Trong table **Athletes**, ta tiến hành thêm cột mới là **Age** vào thông qua câu lệnh:

```
Age = DATEDIFF(Athletes[birth_date], TODAY(), YEAR)
```

Sau đó, ta tiến hành tạo cột mới **Age Category** dựa theo cột **Age** thông qua câu lệnh:

```
Age Category = SWITCH(
    TRUE(),
    'Athletes'[Age] >= 15 && 'Athletes'[Age] <= 20, "15-20",
    'Athletes'[Age] > 20 && 'Athletes'[Age] <= 25, "21-25",
    'Athletes'[Age] > 25 && 'Athletes'[Age] <= 30, "26-30",
    'Athletes'[Age] > 30 && 'Athletes'[Age] <= 35, "31-35",
    'Athletes'[Age] > 35 && 'Athletes'[Age] <= 40, "36-40",
    'Athletes'[Age] > 40 && 'Athletes'[Age] <= 45, "41-45",
    'Athletes'[Age] > 45 && 'Athletes'[Age] <= 50, "46-50",
    'Athletes'[Age] > 50, "51+",
    "Under 15"
)
```

2. Medallists

Tại table **Medallists**, ta thêm một cột mới là **Full Event** thông qua lệnh:

```
Full Event = Medallists[discipline] & " - " & Medallists[event]
```

3. Medals

Tại table **Medals**, ta thêm một cột mới là **Full Gender** dựa trên cột **gender** thông qua lệnh:

```
Full Gender = SWITCH(
    TRUE(),
    Medals[gender] == "M", "Men",
    Medals[gender] == "W", "Women",
    Medals[gender] == "X", "Mixed",
    "None Gender"
)
```

)

4. Schedules

Tại table Schedules, ta thêm một cột mới là Full Event thông qua lệnh:

```
Full Event = Schedules[discipline] & " - " & Schedules[phase]
```

Cột mới này nhằm mục đích thể hiện rõ giữa môn thi đấu và nội dung thi đấu của từng môn đó.

Tiếp đến, ta tiến hành tạo thêm cột mới là Day Playing thông qua lệnh:

```
Day Playing = DATE(YEAR(Schedules[start_date]), MONTH(Schedules[start_date]), DAY(Schedules[start_date]))
```

Cột mới này nhằm mục đích định dạng lại cho dễ theo dõi so với cột cũ là start_date.

3.3 Các measures sử dụng

- Country_Participate

```
Country_Participate =  
COUNTROWS(SUMMARIZE(Disciplines, Disciplines[country_code]))
```

- Country_Playing

```
Country_Playing = DISTINCTCOUNT(Athletes[country])
```

- Country_Win

```
Country_Win = DISTINCTCOUNT(Medals[country])
```

- Highlight

```
Highlight =  
"Total athletes: " & [Total_Athletes] &  
", with " & [Total_Male_Athletes] & " male and " &  
[Total_Female_Athletes] & " female participants. The males won "  
&  
[Total_Male_Athletes_With_Gold_Medal] & " gold medals, and the  
females won " &  
[Total_Female_Athletes_With_Gold_Medal] & " gold medals in the  
tournament.  
Country " & [Top_Performance_Country] & " is the highest-  
performing country in the tournament."
```

- Highlight_Bronze_Medallists

```
Highlight_Bronze_Medallists =  
"Total " & [Total_Bronze_Medallists] & " Medals where Male got "  
    &  
[Total_Male_Athletes_With_Bronze_Medal] & " Medals and Female got "  
    &  
[Total_Female_Athletes_With_Bronze_Medal] & " Medals."
```

- **Highlight_Country_Summary**

```
Highlight_Country_Summary =  
"A total of " & [Country_Playing] &  
" countries have participated in the 2024 Olympics Summer Games.  
    Among them, the " &  
[Top_Performance_Country] & " has the highest performance with "  
    &  
[Top_Performance_Country_Medals] & " medals."
```

- **Highlight_Gold_Medallists**

```
Highlight_Gold_Medallists =  
"Total " & [Total_Gold_Medallists] & " Medals where Male got " &  
[Total_Male_Athletes_With_Gold_Medal] & " Medals and Female got "  
    &  
[Total_Female_Athletes_With_Gold_Medal] & " Medals."
```

- **Highlight_Silver_Medallists**

```
Highlight_Silver_Medallists =  
"Total " & [Total_Silver_Medallists] & " Medals where Male got "  
    &  
[Total_Male_Athletes_With_Silver_Medal] & " Medals and Female got "  
    &  
[Total_Female_Athletes_With_Silver_Medal] & " Medals."
```

- **Playing_Teams**

```
Playing_Teams = DISTINCTCOUNT(Teams[code])
```

- **Playing_Teams_Win**

```
Playing_Teams_Win =  
CALCULATE(  
    DISTINCTCOUNT(Medallists[code_team])  
)
```

- **Top_Performance_Country**

```
Top_Performance_Country =  
VAR MaxCountMedals = MAX(Medals_total[Total])  
RETURN  
CALCULATE(VALUE(Medals_total[country_code]), Medals_total[Total]  
= MaxCountMedals)
```

- **Top_Performance_Country_Medals**

```
Top_Performance_Country_Medals =  
VAR TopCountry = [Top_Performance_Country]  
RETURN  
CALCULATE(  
[Total_Medallists],  
medallists[country_code] = TopCountry  
)
```

- **Total_Athletes**

```
Total_Athletes = DISTINCTCOUNT(Athletes[code])
```

- **Total_Athletes_Win**

```
Total_Athletes_Win = DISTINCTCOUNT(Medals[code])
```

- **Total_Bronze_Medallists**

```
Total_Bronze_Medallists = CALCULATE([Total_Medallists],  
Medallists[medal_type]="Bronze Medal")
```

- **Total_Bronze_Medals**

```
Total_Bronze_Medals = CALCULATE([Total_Medals], Medals[medal_type]  
="Bronze Medal")
```

- **Total_Events**

```
Total_Events = DISTINCTCOUNT(Events[event])
```

- **Total_Female_Athletes**

```
Total_Female_Athletes = var obj = CALCULATE([Total_Athletes],  
Athletes[gender]="Female")  
RETURN  
IF(ISBLANK(obj), 0, obj)
```

- **Total_Female_Athletes_With_Bronze_Medal**

```
Total_Female_Athletes_With_Bronze_Medal = CALCULATE([
    Total_Bronze_Medallists], Medallists[Gender]="Female")
```

- **Total_Female_Athletes_With_Gold_Medal**

```
Total_Female_Athletes_With_Gold_Medal = CALCULATE([
    Total_Gold_Medallists], Medallists[Gender]="Female")
```

- **Total_Female_Athletes_With_Silver_Medal**

```
Total_Female_Athletes_With_Silver_Medal = CALCULATE([
    Total_Silver_Medallists], Medallists[Gender]="Female")
```

- **Total_Gold_Medallists**

```
Total_Gold_Medallists = CALCULATE([Total_Medallists], Medallists[
    medal_type]="Gold Medal")
```

- **Total_Gold_Medals**

```
Total_Gold_Medals = CALCULATE([Total_Medals], Medals[medal_type
    ]="Gold Medal")
```

- **Total_Male_Athletes**

```
Total_Male_Athletes = var obj = CALCULATE([Total_Athletes],
    Athletes[gender]="Male")
RETURN
IF(ISBLANK(obj), 0, obj)
```

- **Total_Male_Athletes_With_Bronze_Medal**

```
Total_Male_Athletes_With_Bronze_Medal = CALCULATE([
    Total_Bronze_Medallists], Medallists[Gender]="Male")
```

- **Total_Male_Athletes_With_Gold_Medal**

```
Total_Male_Athletes_With_Gold_Medal = CALCULATE([
    Total_Gold_Medallists], Medallists[Gender]="Male")
```

- **Total_Male_Athletes_With_Silver_Medal**

```
Total_Male_Athletes_With_Silver_Medal = CALCULATE([
    Total_Silver_Medallists], Medallists[Gender]="Male")
```

- **Total_Medallists**

```
Total_Medallists = var obj = COUNT(Medallists[code_athlete])  
RETURN  
IF(ISBLANK(obj), 0, obj)
```

- **Total_Medals**

```
Total_Medals = var obj = COUNT(Medals[medal_code])  
RETURN  
IF(ISBLANK(obj), 0, obj)
```

- **Total_Silver_Medallists**

```
Total_Silver_Medallists = CALCULATE([Total_Medallists],  
Medallists[medal_type]="Silver Medal")
```

- **Total_Silver_Medals**

```
Total_Silver_Medals = CALCULATE([Total_Medals], Medals[medal_type]  
="Silver Medal")
```

3.4 Chọn dữ liệu cho mỗi Dashboard

Trước tiên ta cần nắm rõ về số lượng huy chương giữa **Medals** và **Medallists**. Một nội dung sẽ có thể là nội dung cá nhân, hoặc nội dung đồng đội, tổng hợp, phối hợp. Khi thống kê về huy chương, sẽ có sự khác biệt giữa số lượng huy chương được phát cho các vận động viên. Từ đó ta có thể hiểu được rằng, tại bảng **Medals**, mỗi nội dung sẽ có 1 huy chương vàng, 1 huy chương bạc và 1 huy chương đồng. Ở bảng **Medallists** thì khác, mỗi nội dung nhóm sẽ có nhiều hơn 1 huy chương vàng, và số lượng sẽ được tính theo số lượng thành viên trong nhóm. Từ đây chúng ta đã nắm được thống kê giữa **Medals** và **Medallists**.

1. Overview

Vì đây là Dashboard Overview, nơi thể hiện toàn diện các nước đạt giải và số lượng huy chương tóm tắt cho một nội dung, nên hầu như dữ liệu trong Dashboard này được lấy từ bảng **Medals** và các Measures được sử dụng cũng lấy dữ liệu từ **Medals**.

2. Athletes

Vì đây là Dashboard về vận động viên, nên các dữ liệu được lấy chủ yếu từ bảng **Athletes**, và số huy chương sẽ được lấy từ **Medallists** để thống kê được số lượng huy chương cho từng vận động viên. Chỉ riêng biểu đồ *Athletes by Discipline*, tại đây có những vận động viên tham gia hơn một nội dung, nên ta phải lấy dữ liệu từ bảng **Disciplines** để thống kê đủ số lượng vận động viên cho từng nội dung cụ thể.

3. Country

Tại Dashboard này, ta cần tổng hợp lại các nước tham dự, nên sẽ sử dụng dữ liệu từ bảng *Athletes*. Còn về biểu đồ *Map* và *Athletes by Country* có chứa huy chương, ta sẽ kết hợp thêm dữ liệu từ *Medallists* để hiển thị rõ số lượng huy chương mà nước đó thực sự nhận được.

4. Medal

Tại đây, ta sẽ chỉ sử dụng dữ liệu từ *Medals* để thể hiện tổng quan số huy chương cho từng bộ môn thi đấu, cho từng nội dung về giới tính, và về màu huy chương cho từng nước tham dự.

5. Event

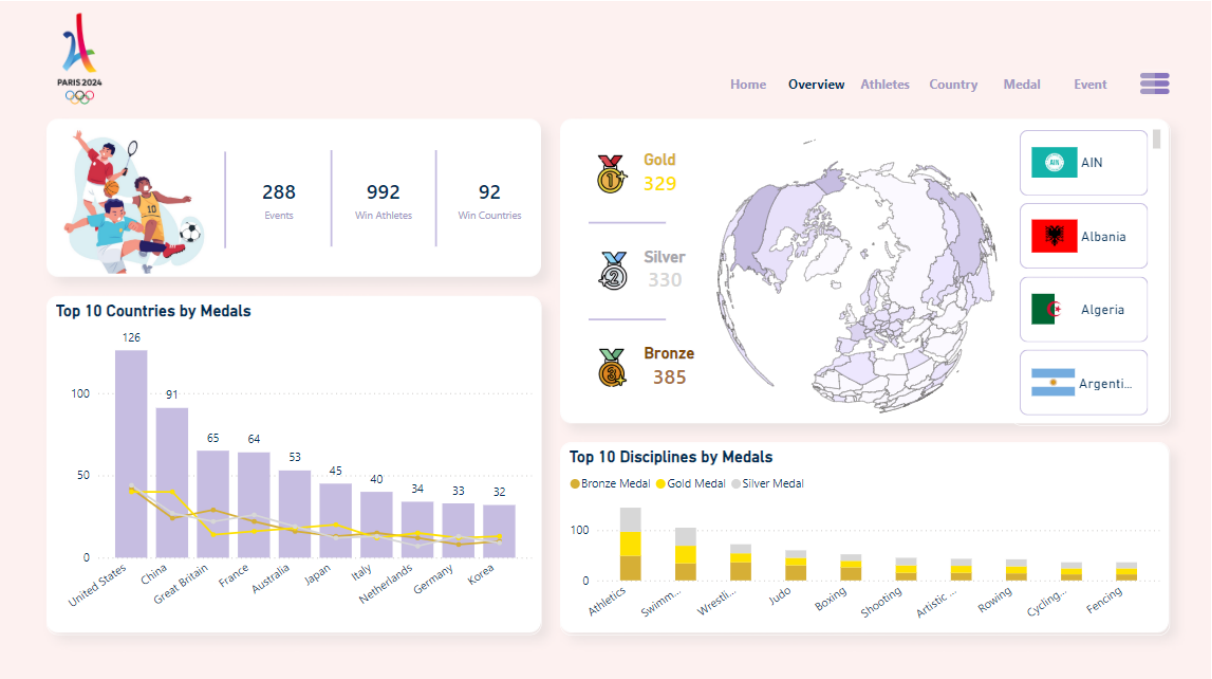
Cuối cùng, tại Dashboard Event, ta sẽ sử dụng 2 bảng biểu khác nhau từ 2 bảng *Schedules* và *Torch_route* để thể hiện các timelines tham dự các nội dung và lịch trình rước đuốc tại các địa điểm tham gia.

4 Xây dựng Dashboard

4.1 Dashboard Overview

4.1.1 Tổng quan

Dashboard giúp hiển thị thông tin tổng quan về các khía cạnh chính của thể vận hội, bao gồm số lượng sự kiện, vận động viên chiến thắng, quốc gia chiến thắng, cũng như các thống kê liên quan đến huy chương và môn thi đấu.



Hình 1: Dashboard Overview

4.1.2 Các thành phần của dashboard

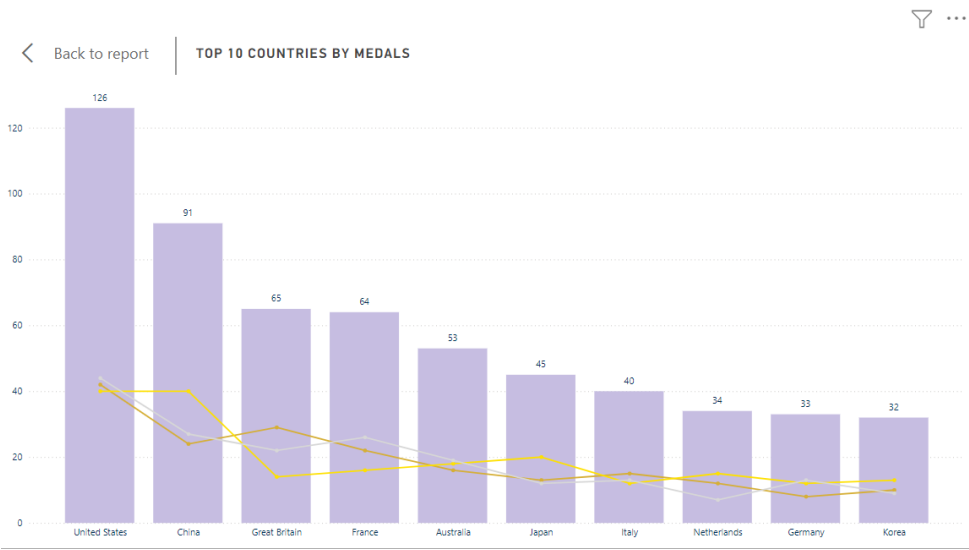
1. Phần tổng quan



Hình 2: Tổng quan Thế vận hội

- 288 sự kiện (Events): Tổng số môn thi đấu tại Thế vận hội Paris 2024.
- 992 vận động viên chiến thắng (Win Athletes): Tổng số vận động viên đã giành được huy chương.
- 92 quốc gia chiến thắng (Win Countries): Tổng số quốc gia có vận động viên giành huy chương.
- Nhận xét:
 - Các chỉ số này cung cấp cái nhìn nhanh về quy mô và mức độ cạnh tranh của sự kiện. Số lượng lớn sự kiện và vận động viên chiến thắng cho thấy mức độ phổ biến và toàn cầu hóa của Thế vận hội.

2. Line and Stacked column chart - Top 10 Countries by Medals

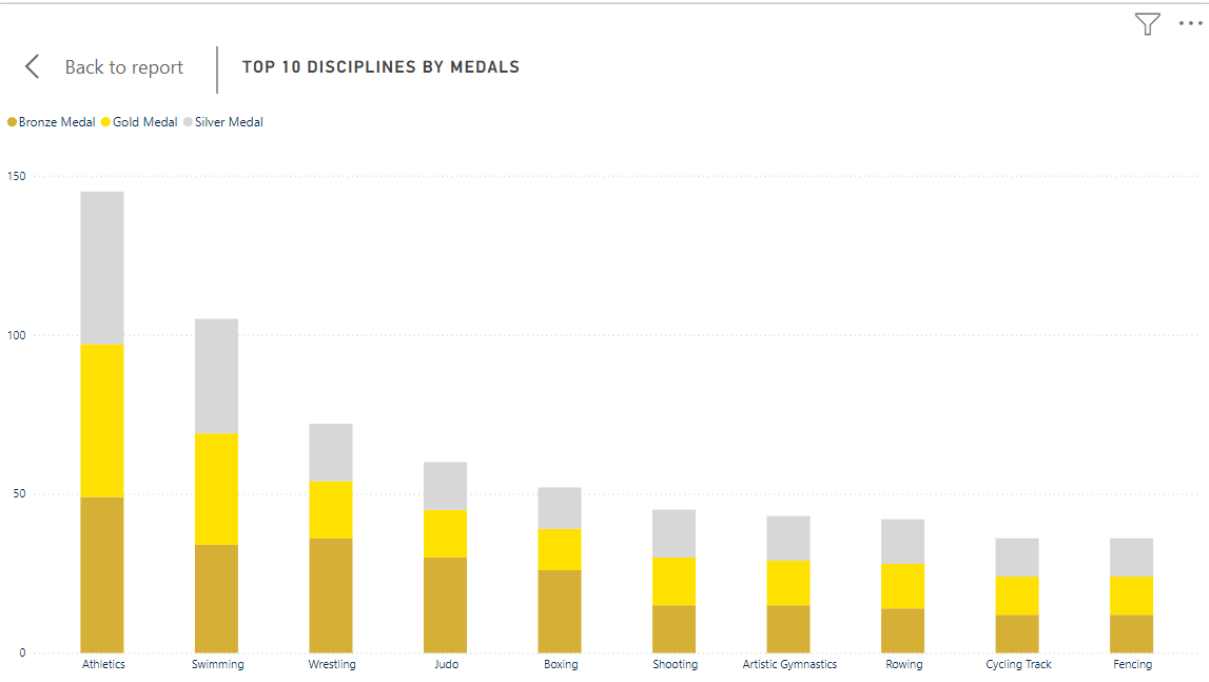


Hình 3: Top 10 countries by medals

- Biểu đồ này hiển thị số lượng huy chương của 10 quốc gia hàng đầu

- Trục X là các quốc gia, trục Y biểu thị số lượng huy chương.
- Mỗi cột được chia theo loại huy chương: Vàng, Bạc, Đồng.
- Nhận xét:
 - Biểu đồ giúp xác định các quốc gia có thành tích cao nhất. Hoa Kỳ vượt trội với 126 huy chương, phản ánh sự đầu tư mạnh mẽ vào thể thao.
 - Trung Quốc, Anh, và Pháp cũng đạt thành tích cao, cho thấy sự tập trung vào thể thao ở các quốc gia này.
 - Sự chênh lệch giữa quốc gia dẫn đầu và các quốc gia khác phản ánh sự khác biệt về đầu tư thể thao và cơ sở vật chất.

3. Stacked column chart - Top 10 Disciplines by Medals

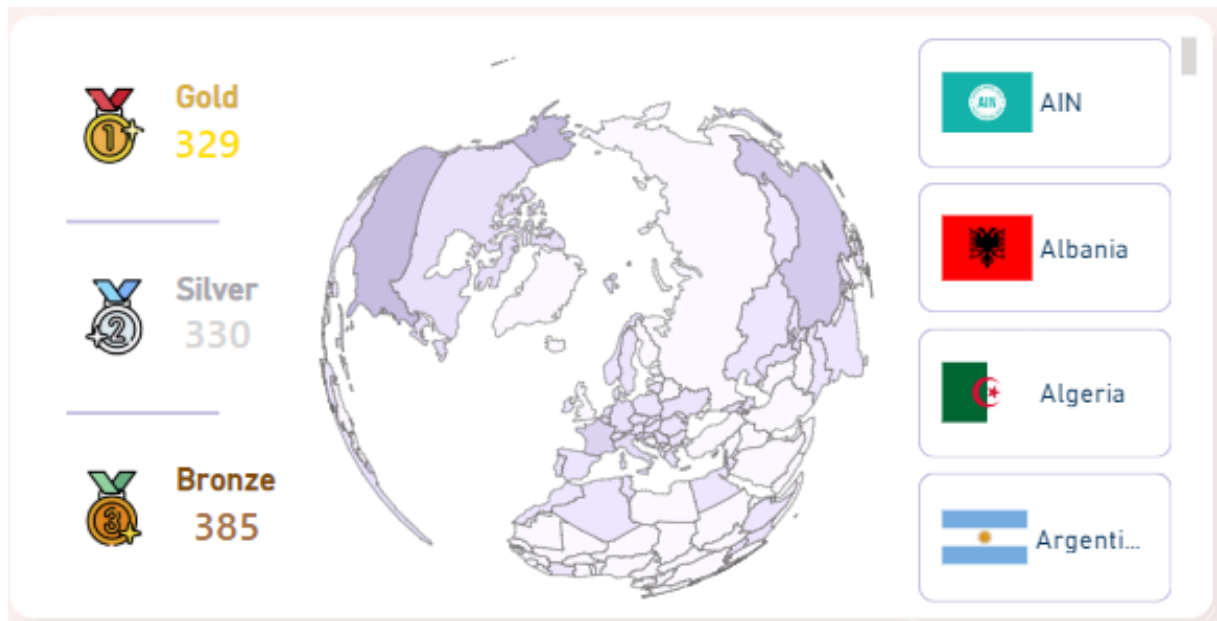


Hình 4: Top 10 disciplines by medals

- Hiển thị số lượng huy chương của 10 môn thể thao hàng đầu.
- Trục X là các môn thể thao, trục Y là số lượng huy chương, phân loại thành Vàng, Bạc, Đồng.
- Nhận xét:
 - Điền kinh và bơi lội là hai môn có số lượng huy chương cao nhất, cho thấy tầm quan trọng và tính cạnh tranh lớn của chúng tại Thế vận hội.
 - Các môn thể thao đồng đội không xuất hiện trong top 10, do số lượng huy chương hạn chế cho mỗi đội.

4. Map

- Hiển thị sự phân bố các quốc gia tham gia Olympic.
- Bản đồ tương tác cho phép người dùng chọn quốc gia cụ thể để xem thông tin chi tiết về huy chương.



Hình 5: World Map

- Tổng số lượng huy chương từng loại được trao gồm 329 huy chương vàng, 330 huy chương bạc, và 385 huy chương đồng.
- Nhận xét:
 - Olympic tạo cơ hội cho tất cả các quốc gia tham gia, không phân biệt quy mô hay vị trí địa lý.
 - Huy chương đồng chiếm tỷ lệ lớn hơn, thể hiện sự cạnh tranh mạnh mẽ ở các nội dung không có sự áp đảo rõ ràng.

4.1.3 Nhận xét và kết luận chung

1. Quy mô và cạnh tranh

- Với 288 sự kiện và 92 quốc gia chiến thắng, Thế vận hội Paris 2024 là một sự kiện thể thao toàn cầu với sự tham gia đông đảo.
- Phân bố huy chương vàng, bạc, đồng khá đồng đều, thể hiện mức độ cạnh tranh khốc liệt.

2. Thành tích quốc gia

- Mỹ và Trung Quốc tiếp tục duy trì vị thế cường quốc thể thao với thành tích vượt trội.
- Các nước châu Âu (Anh, Pháp, Đức) và châu Á (Nhật, Hàn Quốc) cũng có thành tích đáng ghi nhận.

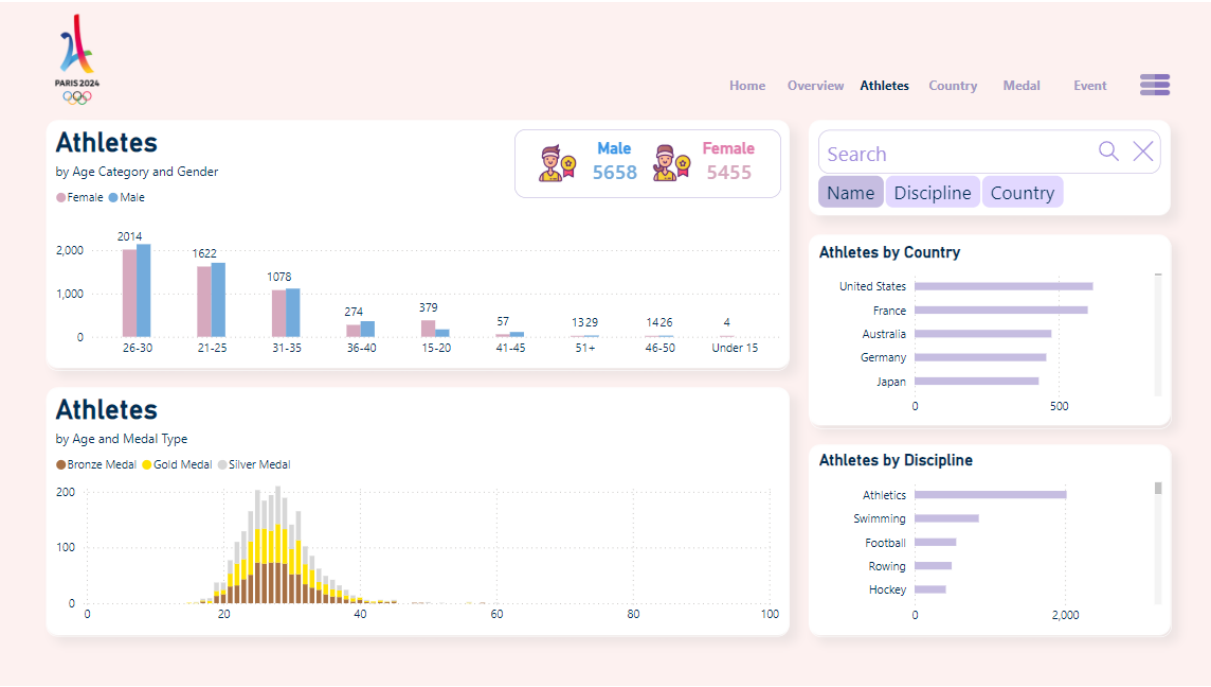
3. Hiệu suất môn thể thao

- Điền kinh và bơi lội là hai môn thi đấu cốt lõi của Thế vận hội, thu hút nhiều vận động viên và đóng góp lớn vào tổng số huy chương.

4.2 Dashboard Athletes

4.2.1 Tổng quan

Dashboard này tập trung vào phân tích thông tin chi tiết liên quan đến vận động viên tham gia Thế vận hội Mùa hè Paris 2024



Hình 6: Dashboard Athletes

4.2.2 Các thành phần của dashboard

1. Tổng quan số lượng vận động viên

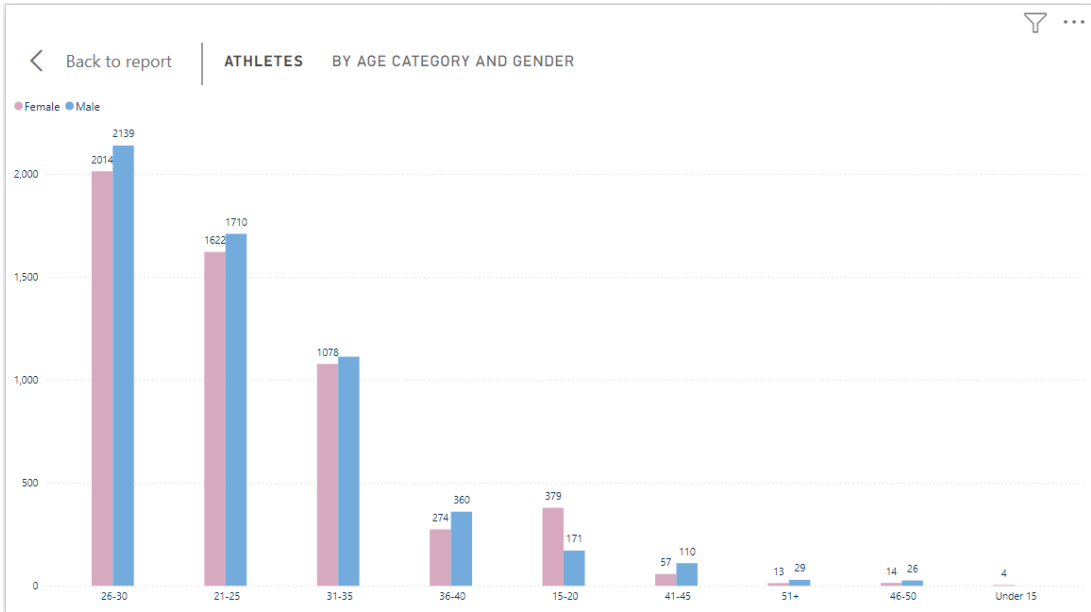


Hình 7: Tổng quan số lượng vận động viên

- Số lượng vận động viên nam: 5,658 người.
- Số lượng vận động viên nữ: 5,455 người.
- Nhận xét:

- Tỷ lệ vận động viên nam và nữ gần như cân bằng, cho thấy tính bình đẳng giới tại Thế vận hội.
- Quy mô lớn với hơn 11,000 vận động viên tham gia, thể hiện sự góp mặt đa dạng trên toàn cầu.

2. Clustered column chart - Athletes by Age and Gender

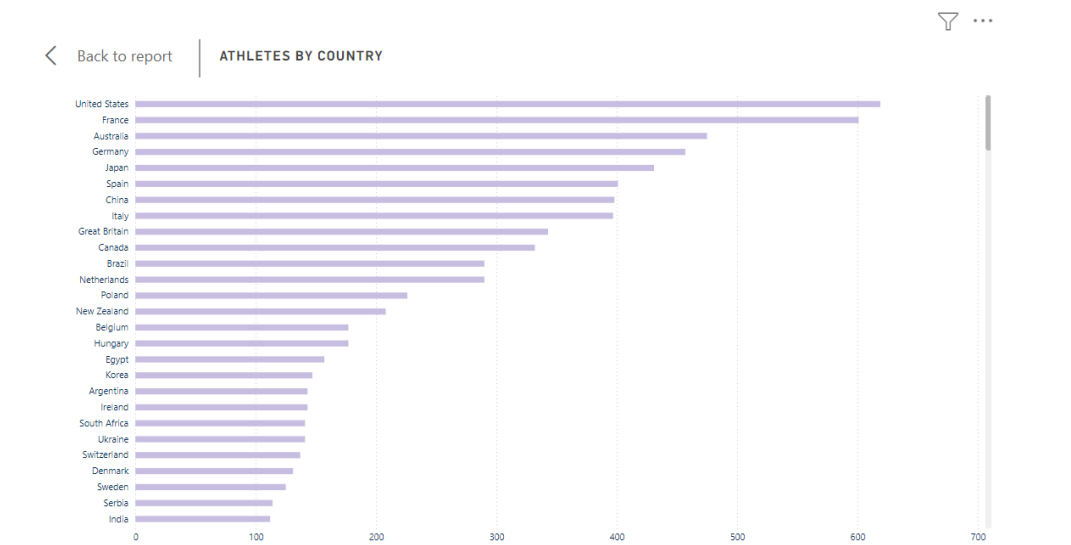


Hình 8: Athletes by Age and Gender

- Hiển thị số lượng vận động viên theo từng nhóm tuổi (dưới 15, 15-20, 21-25, v.v.) và giới tính.
- Trục X là nhóm tuổi, trục Y là số lượng vận động viên.
- Nhận xét:
 - Nhóm tuổi 26-30 có số lượng vận động viên cao nhất (nam: 2,139, nữ: 2,014). Tiếp theo là nhóm tuổi 21-25 và 31-35. Cho thấy độ tuổi 26-30 là giai đoạn vận động viên đạt phong độ tốt nhất.
 - Số lượng vận động viên giảm mạnh sau 35 tuổi. Số lượng vận động viên trẻ dưới 20 tuổi cũng tương đối thấp, có thể do yêu cầu kinh nghiệm và sức bền.

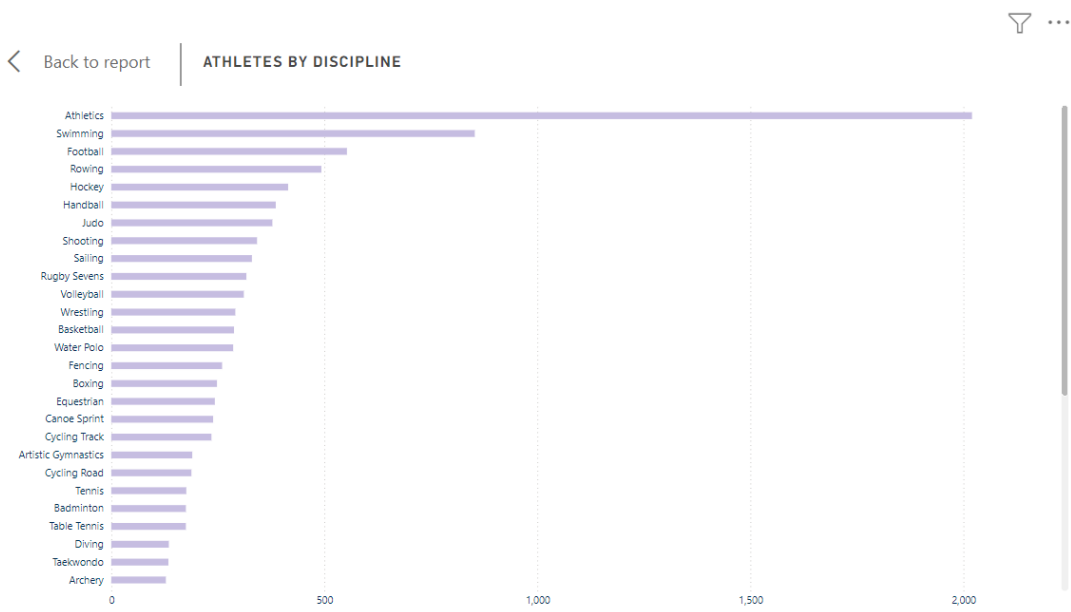
3. Stacked bar chart - Athletes by Country

- Hiển thị số lượng vận động viên từ từng quốc gia.
- Trục Y là tên quốc gia, trục X là số lượng vận động viên.
- Nhận xét:
 - Hoa Kỳ và Pháp có số lượng vận động viên áp đảo, thể hiện sự đầu tư mạnh mẽ vào thể thao.
 - Các quốc gia phát triển chiếm ưu thế trong việc gửi vận động viên đến Thế vận hội.



Hình 9: Athletes by Country

4. Stacked bar chart - Athletes by Discipline



Hình 10: Athletes by Discipline

- Thể hiện số lượng vận động viên tham gia từng môn thể thao.
- Trục Y là môn thể thao, trục X là số lượng vận động viên.
- Nhận xét:
 - Các quốc gia phát triển chiếm ưu thế trong việc gửi vận động viên đến Thế vận hội.
 - Một số môn có tính đặc thù cao (ví dụ: bắn cung, taekwondo) có số lượng vận động viên ít hơn.

4.2.3 Nhận xét và kết luận chung

1. Phân bố độ tuổi và giới tính

- Vận động viên nam và nữ có sự tham gia đồng đều.
- Độ tuổi từ 20-30 chiếm ưu thế, là nhóm tuổi vàng trong thi đấu thể thao.

2. Số lượng theo quốc gia và môn thi đấu

- Các quốc gia phát triển tiếp tục dẫn đầu trong số lượng vận động viên tham gia.
- Điền kinh và bơi lội là những môn thi đấu cốt lõi với số lượng vận động viên đông đảo.

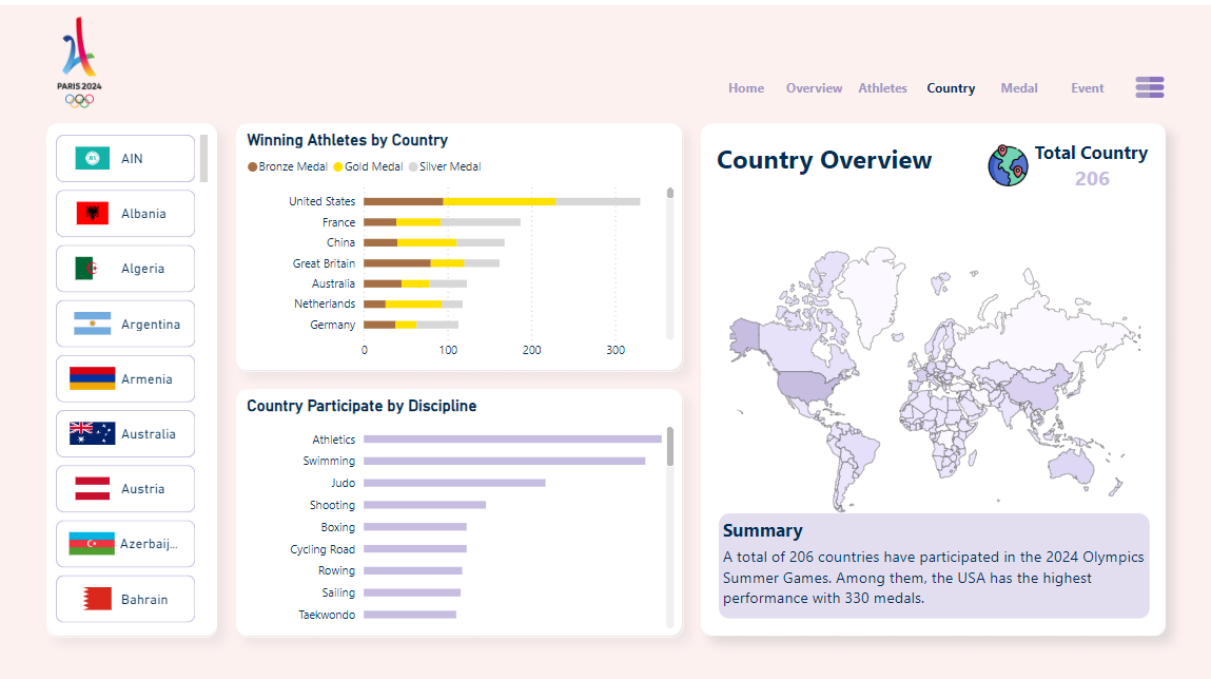
3. Hiệu quả huy chương theo độ tuổi

- Độ tuổi 20-30 có hiệu suất huy chương cao, là nhóm vận động viên cần được đầu tư tối ưu.

4.3 Dashboard Country

4.3.1 Tổng quan

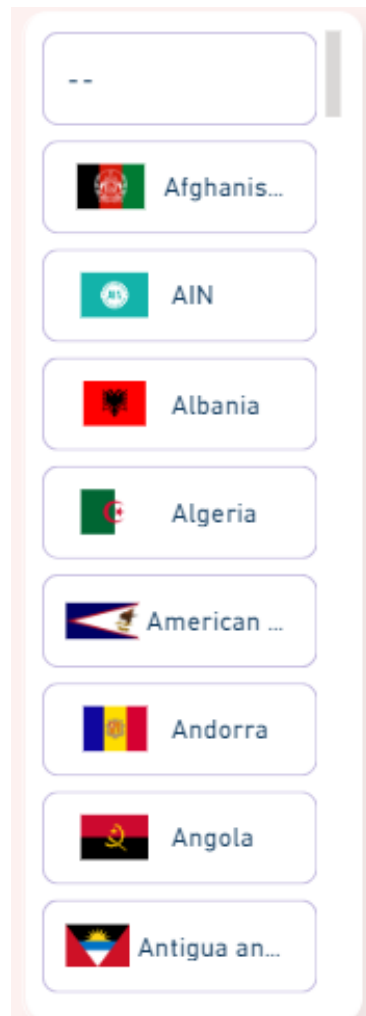
Dashboard này cung cấp cái nhìn toàn diện về sự tham gia và thành tích của các quốc gia tại Thế vận hội Mùa hè Paris 2024.



Hình 11: Dashboard Country

4.3.2 Các thành phần của dashboard

1. Tile slicer - Danh sách quốc gia



Hình 12: Danh sách quốc gia

- Danh sách các quốc gia tham gia với biểu tượng lá cờ để dễ nhận diện.
- Người dùng có thể tìm kiếm hoặc chọn một quốc gia cụ thể để xem chi tiết về thành tích hoặc dữ liệu liên quan.
- Tăng tính tương tác và khả năng tìm kiếm thông tin nhanh chóng cho người dùng.

2. Tổng quan số quốc gia

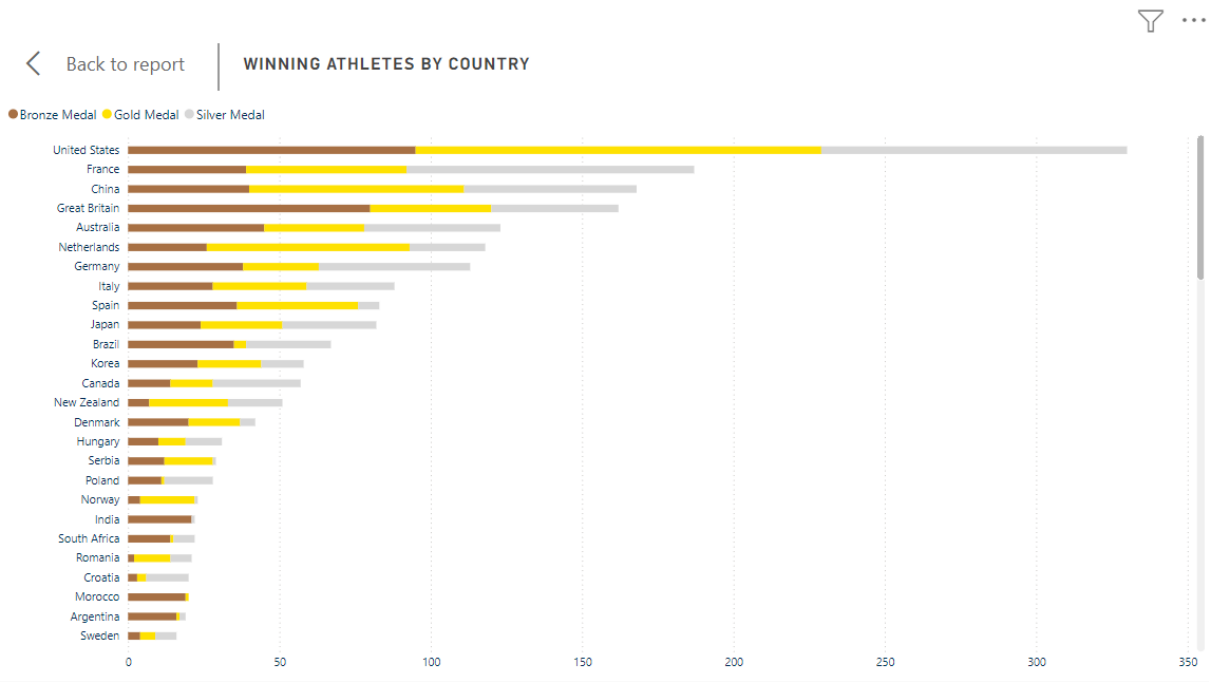
- Tổng số quốc gia tham gia: 206 quốc gia.
- Hoa Kỳ là quốc gia có thành tích tốt nhất với 330 huy chương.
- Nhận xét:
 - Quy mô tham gia rộng khắp toàn cầu, thể hiện tính chất quốc tế và sức hấp dẫn của sự kiện.



Hình 13: Tổng quan số quốc gia

– Hoa Kỳ tiếp tục khẳng định vị thế dẫn đầu trong lĩnh vực thể thao.

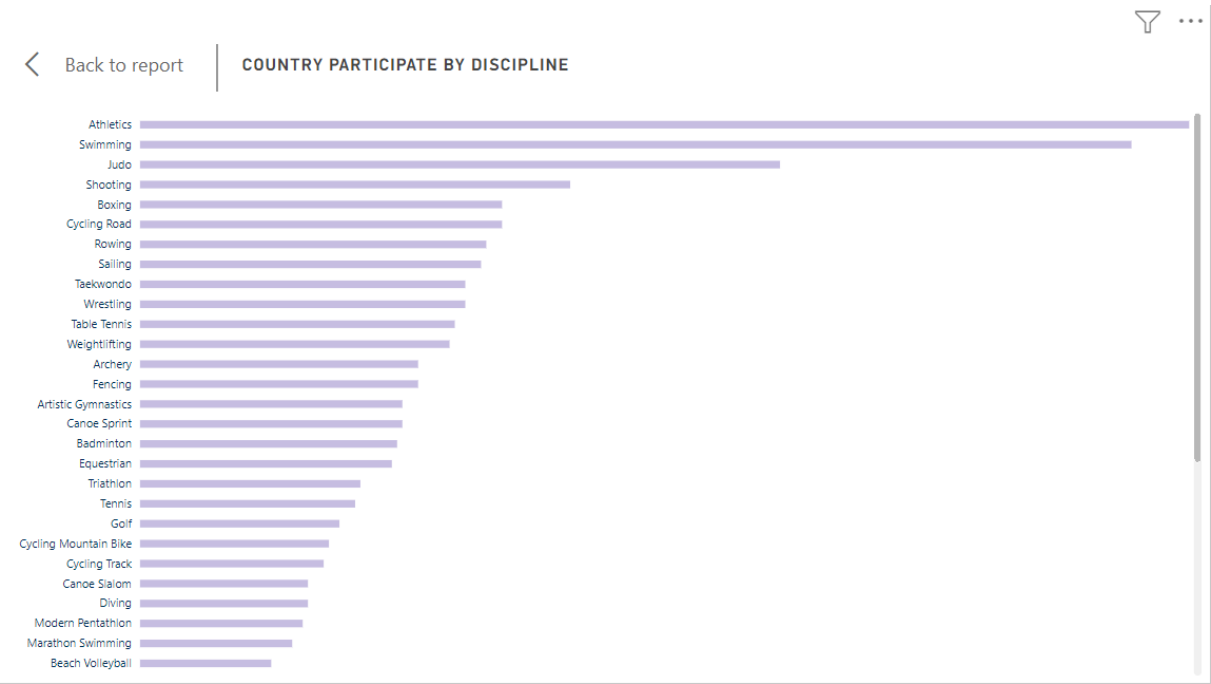
3. Stacked bar chart - Winning Athletes by Country



Hình 14: Winning Athletes by Country

- So sánh số lượng huy chương (Vàng, Bạc, Đồng) giữa các quốc gia.
- Trục Y là các quốc gia, trục X là số lượng huy chương.
- Nhận xét:
 - Hoa Kỳ dẫn đầu với khoảng cách lớn, tiếp theo là các quốc gia có truyền thống thể thao mạnh như Pháp và Trung Quốc.
 - Phân tích này giúp xác định các quốc gia có thành tích vượt trội và chiến lược phát triển thể thao hiệu quả.

4. Stacked bar chart - Country Participate by Discipline



Hình 15: Country Participate by Discipline

- Hiện thị số quốc gia tham gia từng môn thể thao.
- Trục Y là môn thể thao, trục X là số quốc gia.
- Nhận xét:
 - Điền kinh và bơi lội tiếp tục là các môn thu hút nhiều quốc gia tham gia nhất, nhờ tính phổ biến và quy mô lớn.
 - Các môn thi đấu đặc thù như bắn cung, xe đạp địa hình có số quốc gia tham gia ít hơn.

4.3.3 Nhận xét và kết luận chung

1. Quy mô toàn cầu

- Với 206 quốc gia tham gia, Thế vận hội Paris 2024 khẳng định là sự kiện thể thao quốc tế lớn nhất.
- Bản đồ và danh sách quốc gia cung cấp cái nhìn trực quan về sự tham gia rộng khắp.

2. Thành tích vượt trội

- Hoa Kỳ là quốc gia dẫn đầu với 330 huy chương, thể hiện sự đầu tư và phát triển mạnh mẽ trong thể thao.
- Các quốc gia châu Âu như Pháp, Anh và châu Á như Trung Quốc cũng đạt được những thành tích đáng kể.

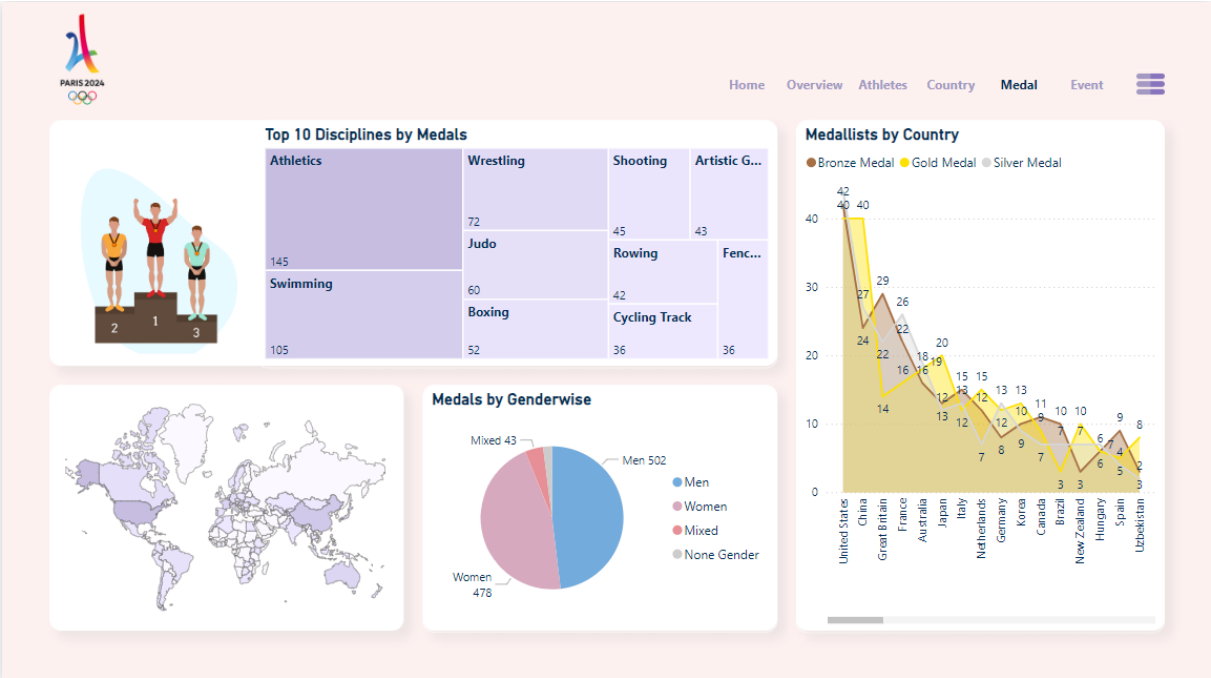
3. Phân bố theo môn thi đấu

- Điền kinh và bơi lội là hai môn có sức hút mạnh mẽ nhất đối với các quốc gia tham gia.
- Các môn thi đấu đòi hỏi cơ sở vật chất đặc thù (ví dụ: đua xe đạp, bắn cung) có ít quốc gia tham gia hơn.

4.4 Dashboard Medal

4.4.1 Tổng quan

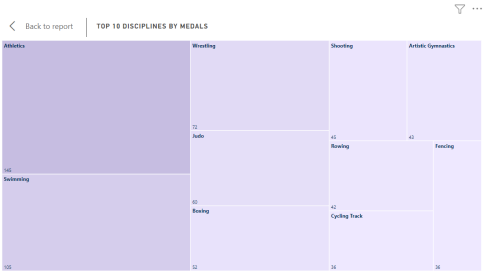
Dashboard này tập trung vào phân tích số lượng huy chương theo các khía cạnh như quốc gia, giới tính, và môn thể thao, giúp cung cấp cái nhìn toàn diện về kết quả của Thế vận hội Mùa hè Paris 2024.



Hình 16: Dashboard Medal

4.4.2 Các thành phần của dashboard

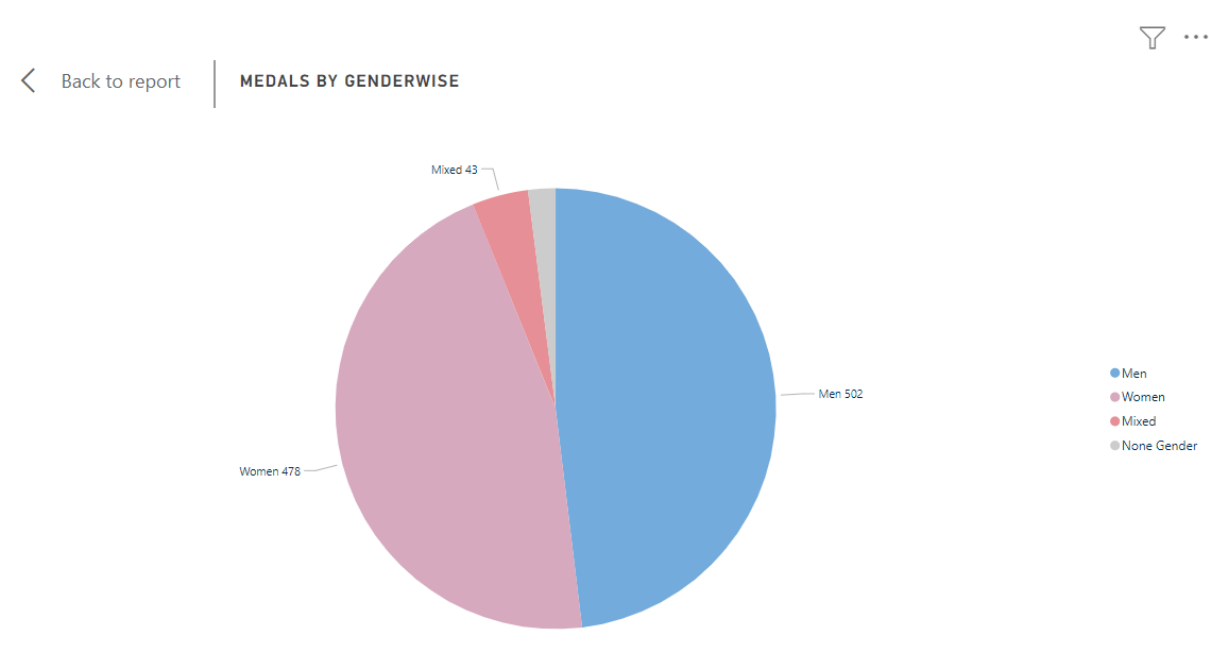
1. TreeMap - Top 10 Disciplines by Medals



Hình 17: Top 10 Disciplines by Medals

- Biểu đồ dạng hình vuông phân chia theo số lượng huy chương của 10 môn thể thao hàng đầu.
- Mỗi ô đại diện cho một môn thể thao, kích thước thể hiện số huy chương.
- Nhận xét:
 - Điền kinh và bơi lội tiếp tục khẳng định vị thế là hai môn thể thao chủ lực tại Thế vận hội.
 - Các môn thể thao đặc thù như Đấu kiếm và Đua xe đạp vẫn giữ được vị trí quan trọng.

2. Pie chart - Medals by Genderwise

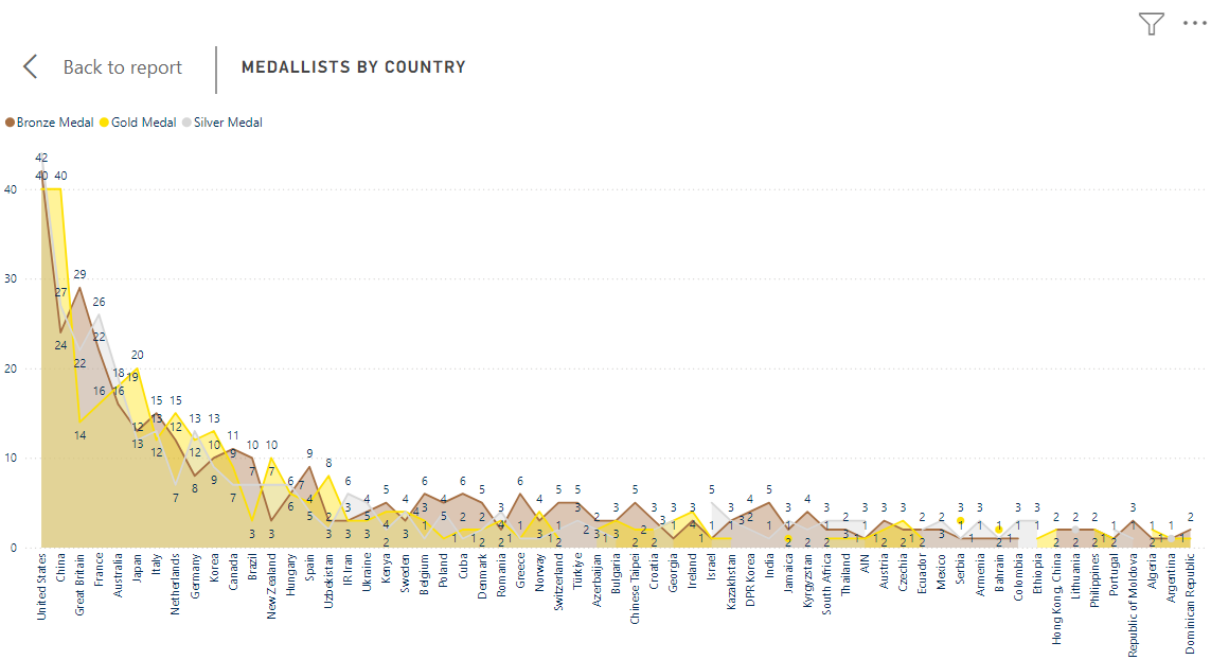


Hình 18: Medals by Genderwise

- Hiển thị tỉ lệ huy chương theo giới tính (nam, nữ, hỗn hợp).
- Phần trăm từng loại được hiển thị trong từng phần tròn.
- Nhận xét:
 - Số lượng huy chương giữa vận động viên nam và nữ gần như cân bằng, thể hiện sự bình đẳng trong cơ hội thi đấu.
 - Các nội dung thi đấu hỗn hợp (Mixed) ngày càng phổ biến, thúc đẩy sự đa dạng trong các môn thể thao.

3. Area chart - Medallists by Country

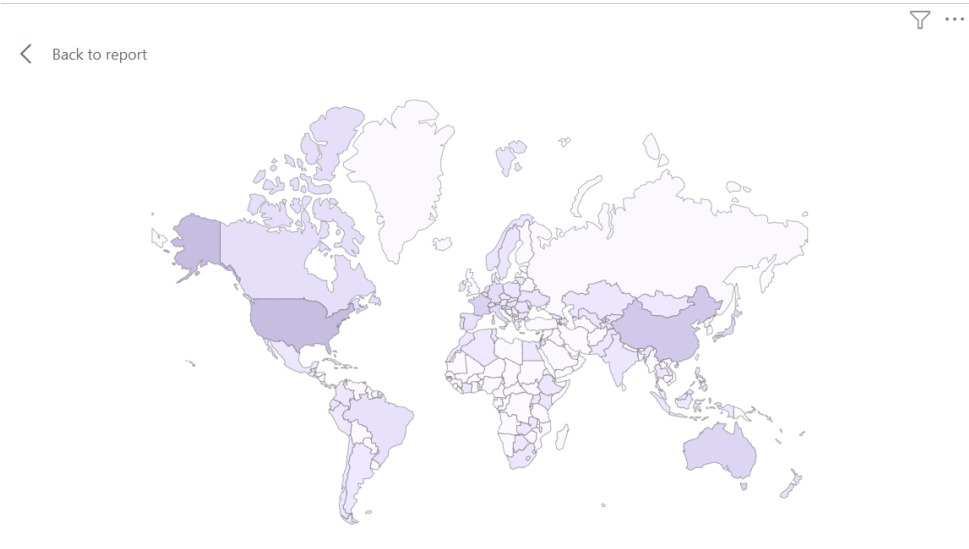
- Hiển thị số lượng huy chương của các quốc gia, phân chia theo loại huy chương.
- Trục X là các quốc gia, trục Y là số lượng huy chương.
- Nhận xét:
 - Hoa Kỳ và Trung Quốc tiếp tục thống trị bảng xếp hạng huy chương.



Hình 19: Medallists by Country

- Sự phân bố huy chương rộng khắp cho thấy mức độ cạnh tranh toàn cầu tại Thế vận hội.

4. Map



Hình 20: Map Visualization

- Hiển thị số lượng huy chương của các quốc gia trên toàn thế giới, với màu sắc và cường độ phản ánh hiệu suất.
- Nhận xét:
 - Các quốc gia Bắc Mỹ, châu Âu và châu Á chiếm ưu thế trong việc giành huy chương.

4.4.3 Nhận xét và kết luận chung

1. Thành tích vượt trội

- Điền kinh và bơi lội là hai môn thi đấu cốt lõi, góp phần lớn vào tổng số huy chương.

2. Sự cân bằng giới tính

- Huy chương được phân bổ khá đồng đều giữa nam và nữ, phản ánh sự bình đẳng giới trong thể thao.
- Nội dung thi đấu hỗn hợp là một điểm nhấn mới, thúc đẩy tinh thần hợp tác và đa dạng.

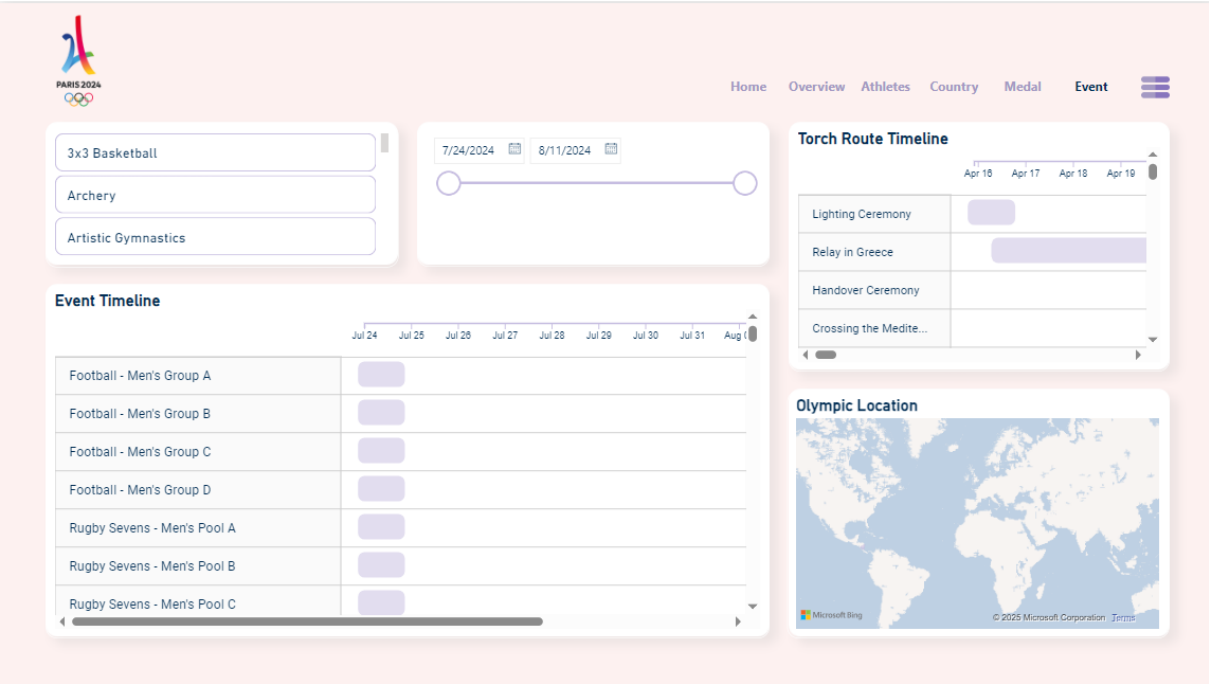
3. Đóng góp của các môn thể thao

- Các môn như Điền kinh, Bơi lội và Đấu vật đóng vai trò quan trọng trong việc định hình kết quả tổng thể của Thế vận hội.

4.5 Dashboard Event

4.5.1 Tổng quan

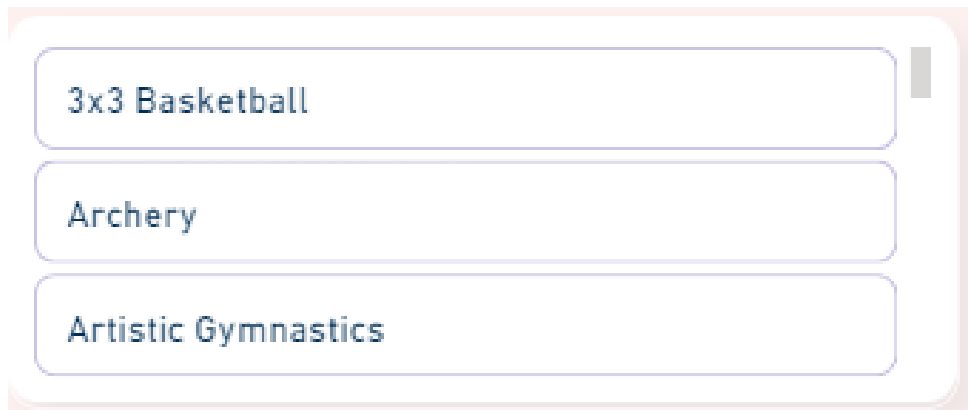
Dashboard này tập trung vào lịch trình và thông tin chi tiết về các sự kiện tại Thế vận hội Mùa hè Paris 2024, bao gồm các hoạt động chính, thời gian và địa điểm.



Hình 21: Dashboard Event

4.5.2 Các thành phần của dashboard

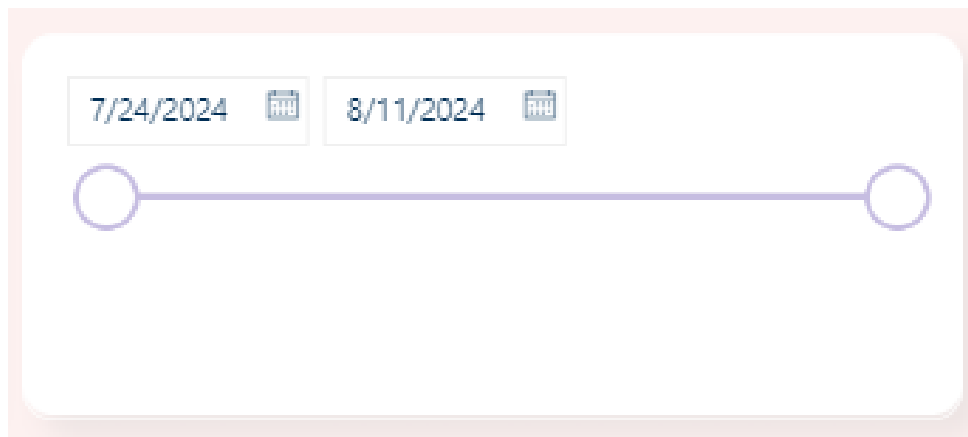
1. Tile slicer by Discipline



Hình 22: *Tile slicer by Discipline*

- Người dùng có thể chọn các môn thi đấu cụ thể (ví dụ: 3x3 Basketball, Archery, Artistic Gymnastics) để hiển thị dữ liệu liên quan.

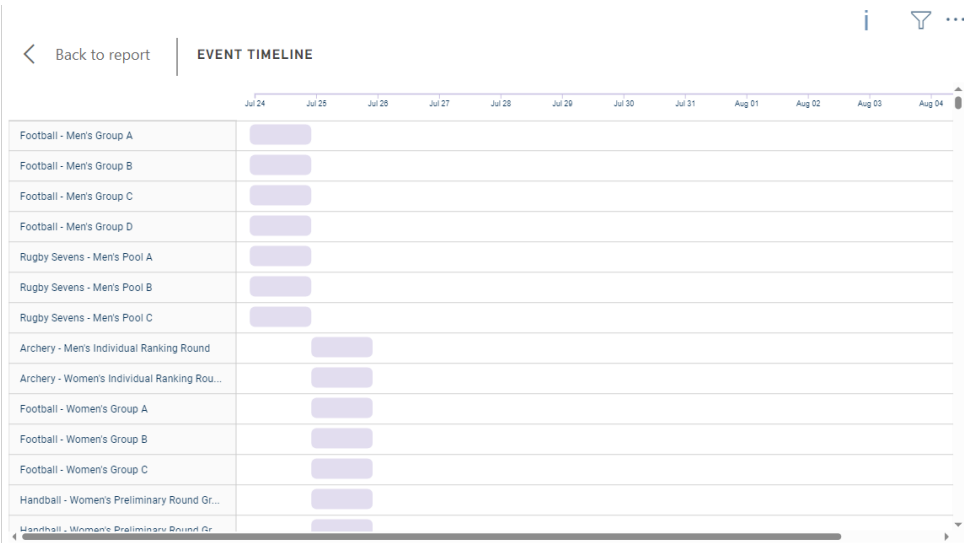
2. Date Slicer



Hình 23: *Date Slicer*

- Cho phép người dùng chọn khoảng thời gian cụ thể để theo dõi các sự kiện diễn ra trong giai đoạn đó.
- Khoảng thời gian mặc định: từ 24/7/2024 đến 11/8/2024.

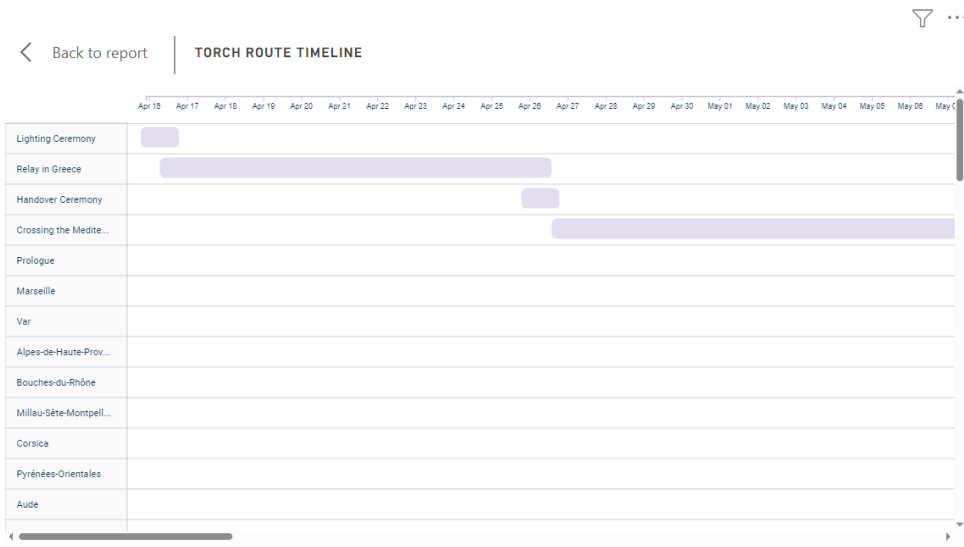
3. Gantt chart - Event Timeline



Hình 24: Event Timeline

- Biểu đồ Gantt hiển thị các sự kiện Olympic theo thời gian.
- Trục X là các ngày diễn ra sự kiện, trục Y là tên các sự kiện (ví dụ: Bóng đá, Rugby, Archery, Handball).
- Mỗi hàng đại diện cho một sự kiện, với các thanh màu tím thể hiện khoảng thời gian diễn ra.
- Nhận xét:
 - Cung cấp một cái nhìn tổng quan về toàn bộ lịch trình Thế vận hội.
 - Giúp theo dõi lịch trình và thời gian cụ thể của từng sự kiện.

4. Gantt chart - Torch Route Timeline



Hình 25: Torch Route Timeline

- Biểu đồ Gantt này hiển thị lộ trình rước đuốc từ 16/4/2024 đến 11/8/2024.
- Các thanh màu tím đại diện cho các giai đoạn của lộ trình ví dụ như Lighting Ceremony, Relay in Greece, Handover Ceremony,...

- Nhận xét:
 - Cho thấy quá trình tổ chức và sự chuẩn bị kỹ lưỡng trước khi khai mạc Thế vận hội.
 - Lộ trình rước đuốc tạo không khí lễ hội và kết nối tinh thần Olympic giữa các quốc gia.

4.5.3 Nhận xét và kết luận chung

- Cung cấp một bức tranh đầy đủ về lịch trình và các sự kiện của Thế vận hội, từ thi đấu đến hành trình rước đuốc.
- Việc tổ chức sự kiện trong khoảng thời gian dài đảm bảo rằng tất cả các môn thể thao đều có đủ thời gian thi đấu.
- Dashboard phù hợp với các nhà quản lý, khán giả hoặc truyền thông để theo dõi sự kiện và lập kế hoạch tham gia hoặc đưa tin.

5 Tổng kết chung

Dựa trên toàn bộ các phân tích và trực quan hóa dữ liệu từ dự án, báo cáo này đã mang đến một cái nhìn toàn diện về quy mô, thành tựu và ý nghĩa của kỳ Thế vận hội lần này. Với sự tham gia của 206 quốc gia, hơn 10.000 vận động viên, và gần 300 sự kiện thi đấu, Thế vận hội mùa hè Paris 2024 không chỉ là một sự kiện thể thao quốc tế mà còn là biểu tượng của sự đoàn kết và giao lưu văn hóa giữa các quốc gia. Các phân tích dữ liệu đã cho thấy:

- **Quy mô toàn cầu:** Với sự tham gia của 206 quốc gia, Thế vận hội mùa hè Paris 2024 không chỉ là sự kiện thể thao mà còn là biểu tượng của sự đoàn kết và hòa bình quốc tế. Sự đa dạng về địa lý và văn hóa đã tạo nên một sân chơi công bằng và ý nghĩa, nơi các quốc gia và vận động viên giao lưu, học hỏi và thúc đẩy tinh thần thi đấu.
- **Thành tích nổi bật của các quốc gia và vận động viên:** Hoa Kỳ đã khẳng định vị thế dẫn đầu với số lượng huy chương cao nhất, theo sát là Trung Quốc, Anh, Pháp và Australia. Điều này phản ánh sự đầu tư dài hạn và chiến lược phát triển thể thao tại các quốc gia này. Các vận động viên, đặc biệt trong độ tuổi 26-30, đã thể hiện phong độ đỉnh cao, mang về thành tích đáng tự hào, đồng thời cho thấy sự cân bằng giữa vận động viên nam và nữ, thúc đẩy bình đẳng giới trong thể thao.
- **Sự đa dạng và phát triển của các môn thể thao:** Các môn thi đấu như Điền kinh, Bơi lội, và Đấu vật tiếp tục là tâm điểm thu hút, đồng thời khẳng định vai trò cốt lõi trong hệ thống thi đấu Olympic. Ngoài ra, những môn thể thao mới hoặc ít phổ biến hơn cũng được chú ý, cho thấy nỗ lực mở rộng phạm vi và thu hút sự tham gia của các vận động viên từ khắp nơi trên thế giới.