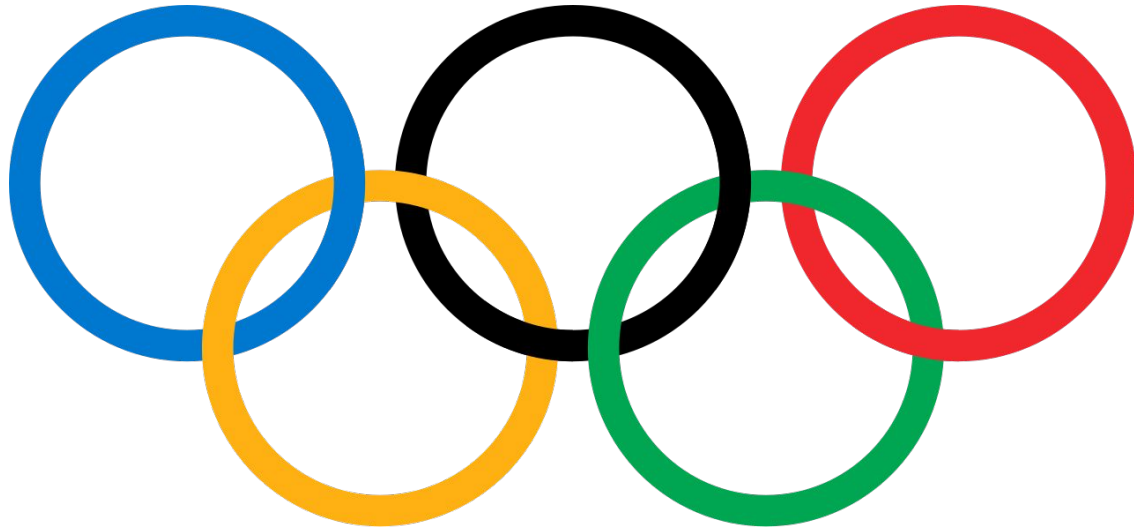


Olympic Analysis

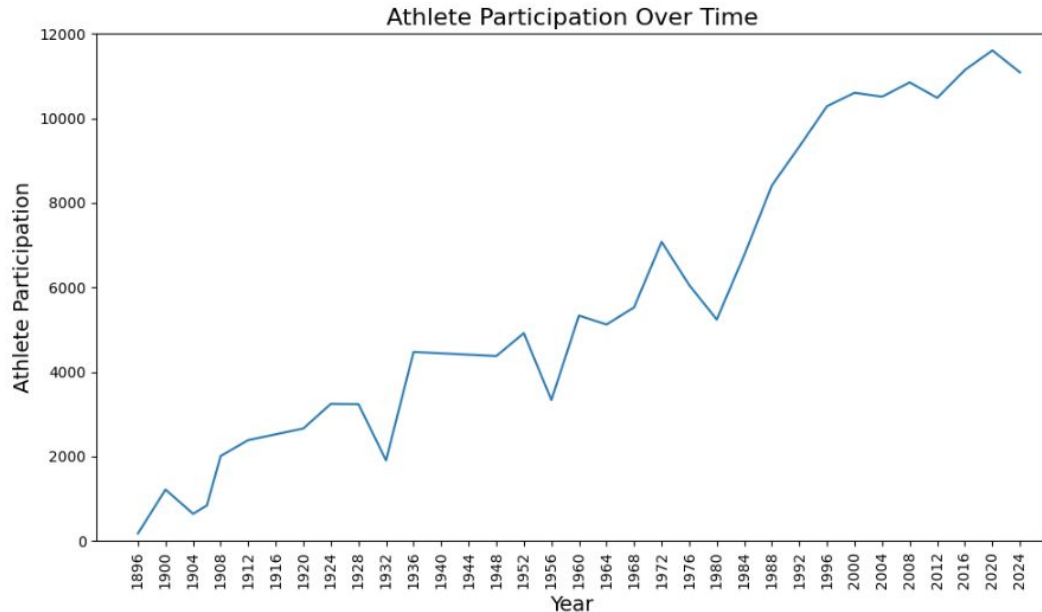


By Maybel Herrera, Sandy Leung, and Kéli Davis

Data Sources

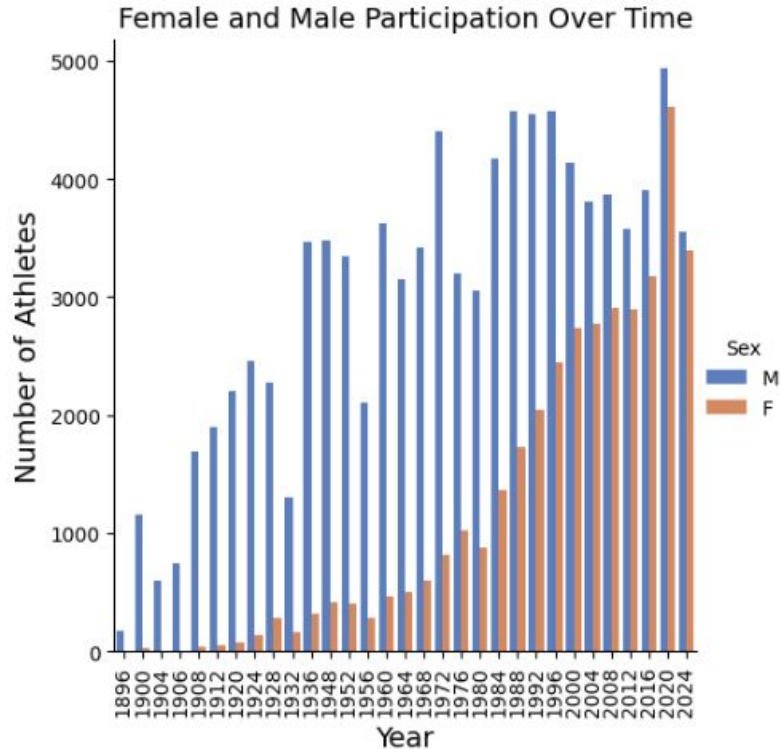
- 5 datasets from Kaggle
- olympics_dataset.csv
 - Columns: Name, Sex, Age, Team, NOC, Games, Year, Season, City, Sport, Event, Medal
 - Contains olympic data from 1896 - 2024
- Medals_2020.xlsx
 - Columns: Rank, NOC, Gold, Bronze, Silver, Total, Rank by Total
 - Accumulation of total medals (gold, bronze, silver) per countries in 2020
- Medals_2024.csv
 - Columns: Rank, NOC, Gold, Bronze, Silver, Total, Rank by Total
 - Accumulation of total medals (gold, bronze, silver) per countries in 2024
- Athletes_2024.csv
 - Columns: name, gender, function, country, disciplines, events, birth_date,
 - 2024 Athlete information
- Tokyo_2020_tweets: Close to 8,000 tweets

How has athlete participation changed over time?

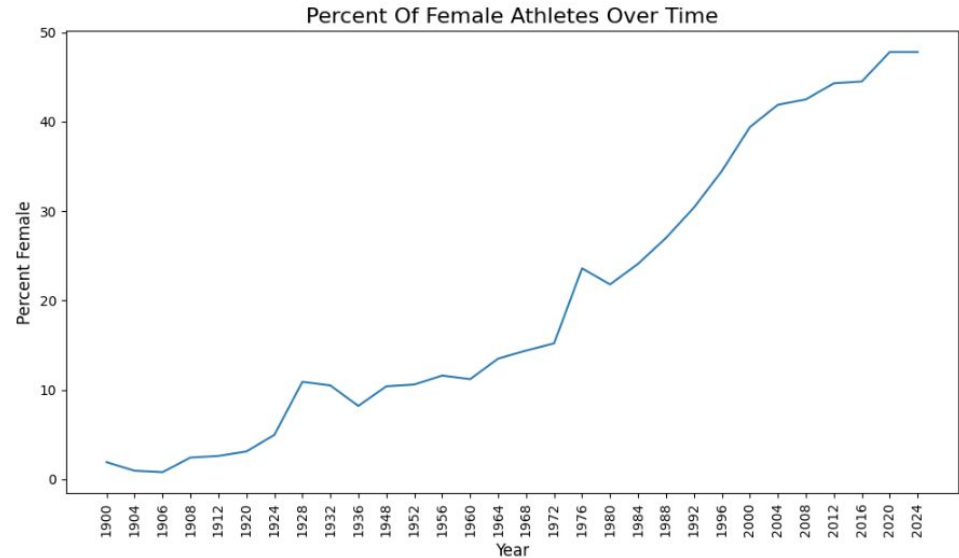


- First modern olympics in 1896
- Between 1896 and 2024: 6200% growth in participation
- 1932 - 68% drop due to Great Depression
- 1956 - 48% drop due to boycotts of Soviet Union invasion of Hungary
- 1976 - 17% drop due to boycotts of New Zealand
- 1980 - another 16% drop due to boycotts of Soviet Union invasion of Afghanistan

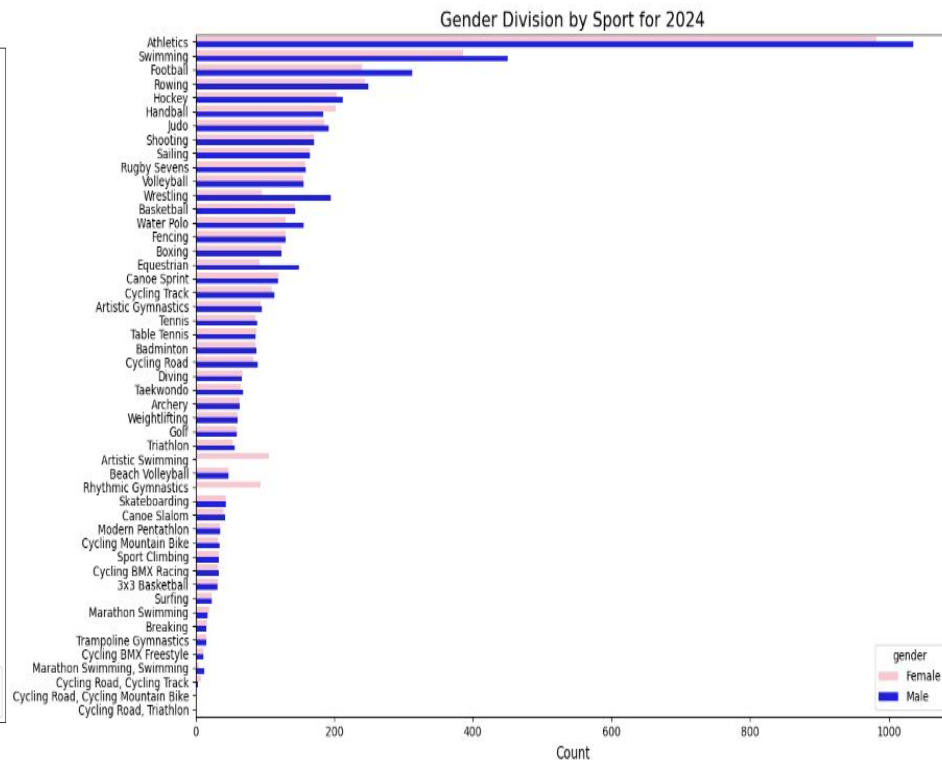
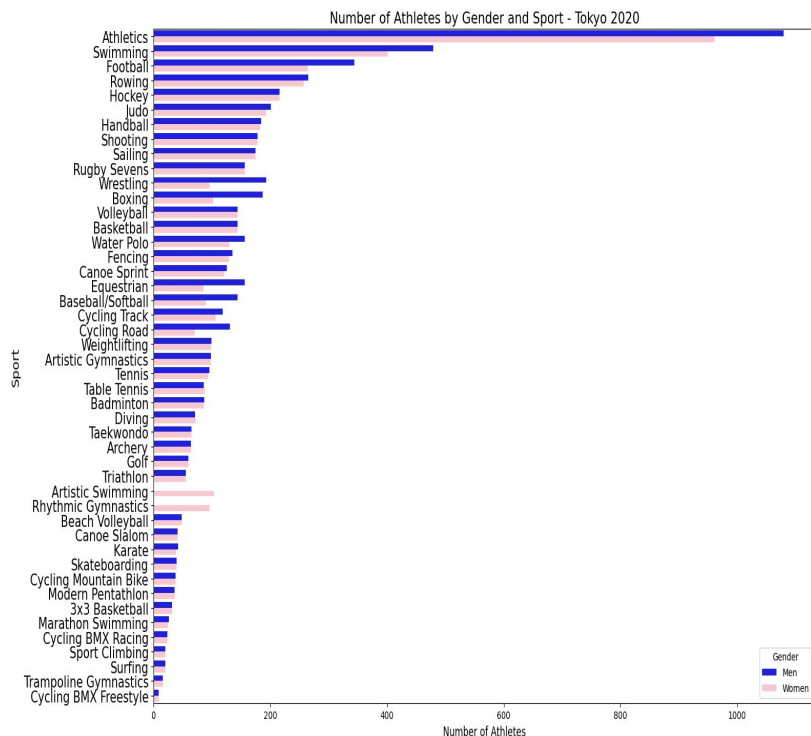
How has athlete gender distribution changed over time?



- 1896 - women not allowed to compete
- 1900 - women less than 2% of competitors
- 2024 - women made up over 47%

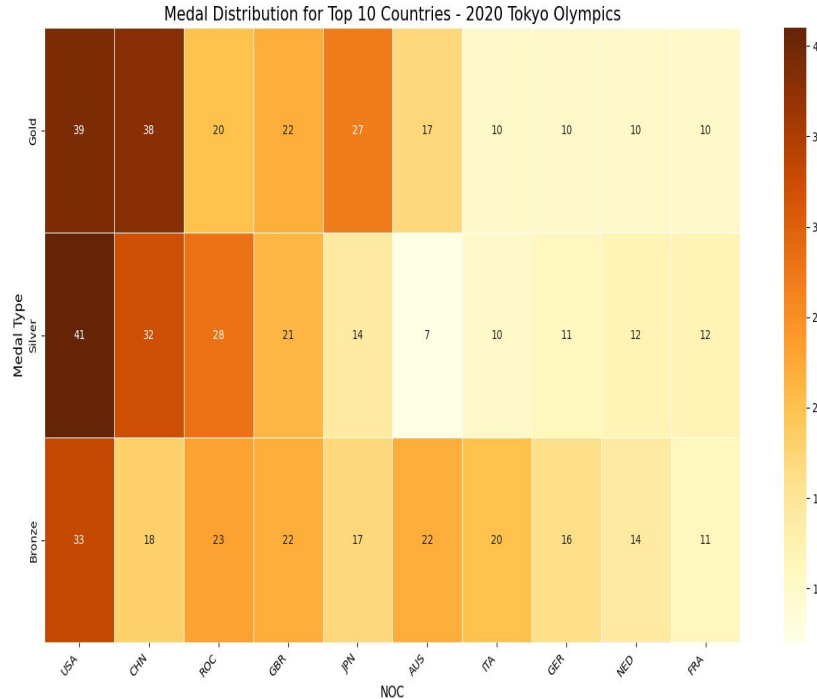


What was the gender distribution in each sport for 2020 ? Are there sports more male-dominated or female-dominated?



2020 Tokyo Olympics Top 10 Countries

- USA stands out as the top performing country with 113 medals in total
- China follows closely behind securing more gold and silver medals than bronze.
- ROC (Russian Olympic Committee) came 5th with only 341 athletes competing.
- Japan the (host) secured 5th place with 58 medals in total but had 613 athletes competing.



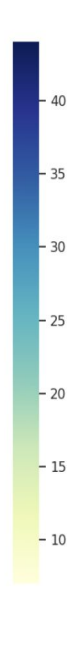
NOC	Athlete Count
USA	633
JPN	613
AUS	489
CHN	417
GER	415
FRA	396
GBR	392
CAN	386
ITA	377
ROC	341

Comparison with recent 2024 Paris Olympics

- The United States and China have maintained dominance, with slight increases in their overall medal counts from the 2020 to 2024 games.
- France went from number 10th place in 2020 to 4th place, they Increased total medals by 31 (from 33 in 2020 to 64 in 2024)
- Even Though Japan did not have the advantage of hosting they still made it to the top 10 with 6th place.

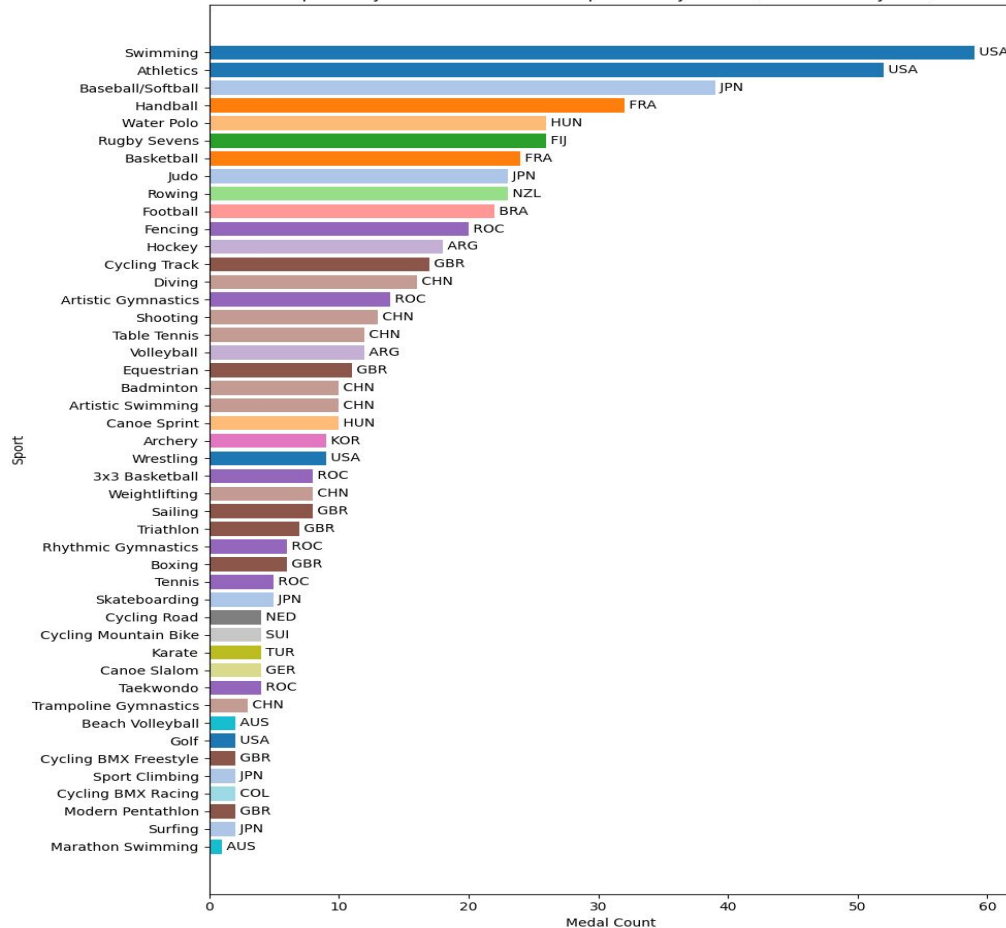
Medal Distribution for Top 10 Countries - 2024 Summer Olympics

Country Code	USA	CHN	GBR	FRA	AUS	JPN	ITA	NED	GER	KOR
	40	40	14	16	18	20	12	15	12	13
	44	27	22	26	19	12	13	7	13	9
	42	24	29	22	16	13	15	12	8	10
Medal Type										
Gold Medal										
Silver Medal										
Bronze Medal										



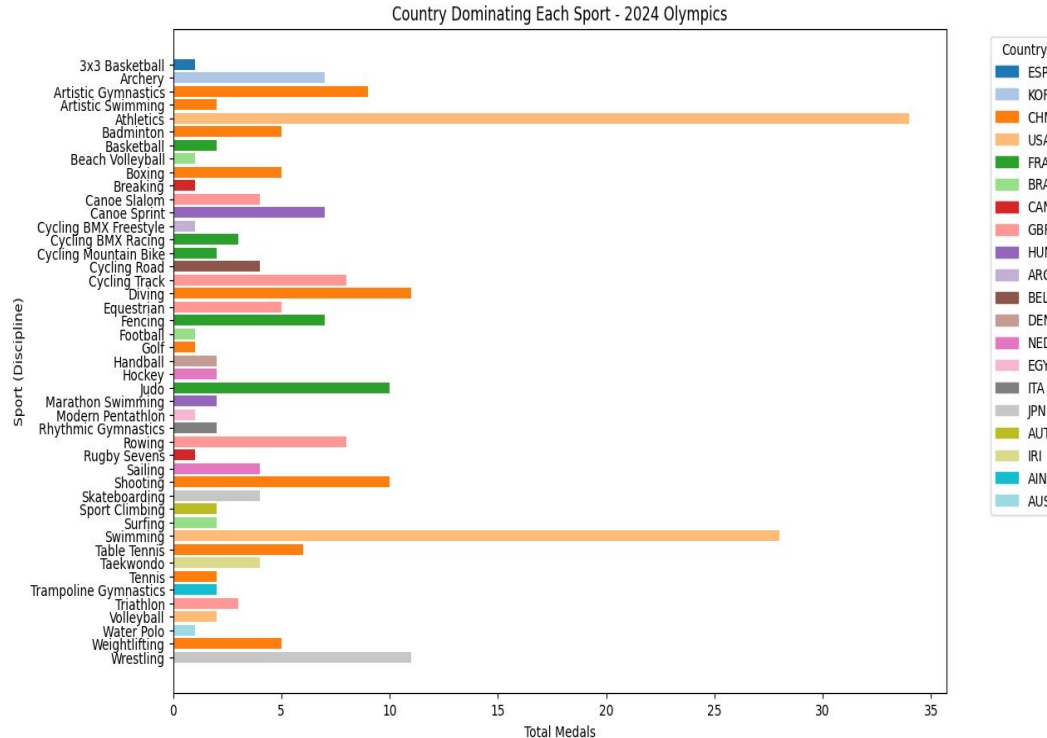
country_code	athlete_count
USA	619
FRA	600
AUS	475
GER	457
JPN	431
ESP	401
CHN	398
ITA	397
GBR	342
CAN	332

Top NOC by Medal Count for Each Sport in Tokyo 2020 (Color-coded by NOC)



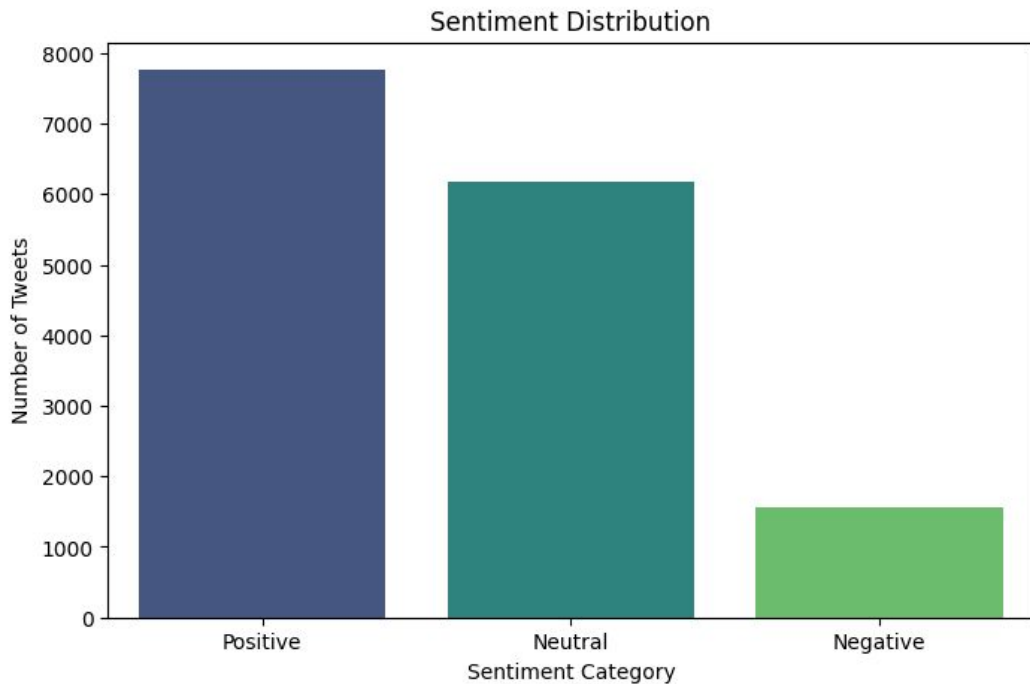
- **18 Unique NOC**
- **USA and Japan** emerged as key leaders in dominating the top 3 sports .
Swimming , Athletics and Baseball/Softball.
- **USA** dominating the two 2 sports
- **China** dominating more sports (7) but each had less medal counts.
- Several NOCs dominated traditional team sports: **Fiji** in **Rugby Sevens**, **Brazil (BRA)** in **Football**, and **New Zealand (NZL)** in **Rowing**.
- **Great Britain** dominated 7 sports: cycling, equestrian, Sailing, triathlon, boxing , cycling BMX freestyle and modern pentathlon. Cycling track winning the most medals.
- **ROC** dominated most gymnastic sports securing 3rd place : Artistic gymnastic, rhythmic gymnastic and trampoline gymnastic.

Which NOC dominated which sport in 2024 Olympics and what are some similarity and differences.



- 20 Unique NOC
- More diversity in countries nominating different sports.
- Spain made the list for dominating 3 x 3 Basketball.
- France, as the host country, made notable strides in sports like Judo, Fencing Cycling, surfing, and Triathlon, which might reflect a home advantage
- USA maintained their dominance with Athletics and Swimming.

Sentimental Analysis and Word Cloud





2020 vs. 2024 Top Athletes



Medal Bronze Gold Silver

Name

DRESSEL Caeleb	0	5	0
McKEON Emma	3	4	0
McKEOWN Kaylee	1	3	0
AN San	0	3	0
CARRINGTON Lisa	0	3	0
THOMPSON-HERAH Elaine	0	3	0
TITMUS Ariarne	1	2	1
CAMPBELL Cate	1	2	0
HASSAN Sifan	1	2	0
LAVREYSEN Harrie	1	2	0

Medal

Name

Bronze

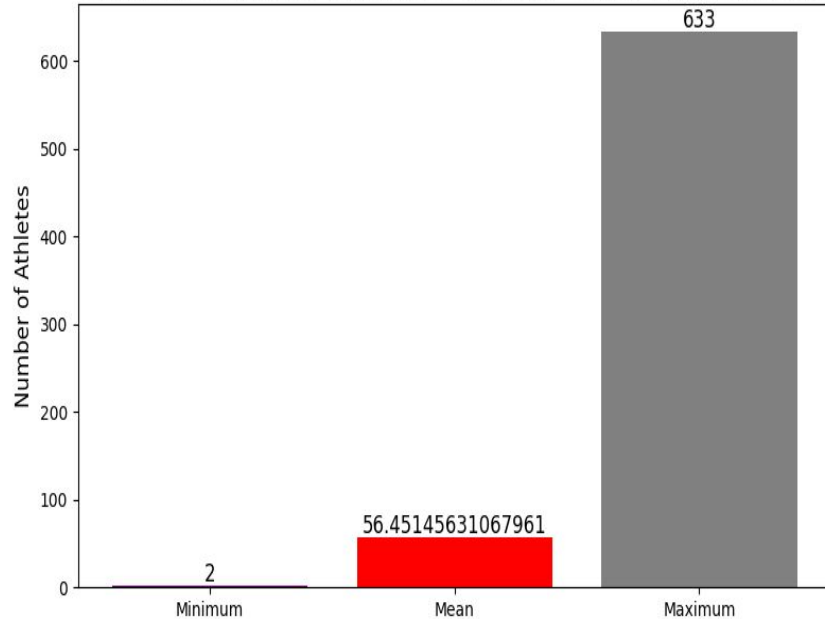
Gold

Silver

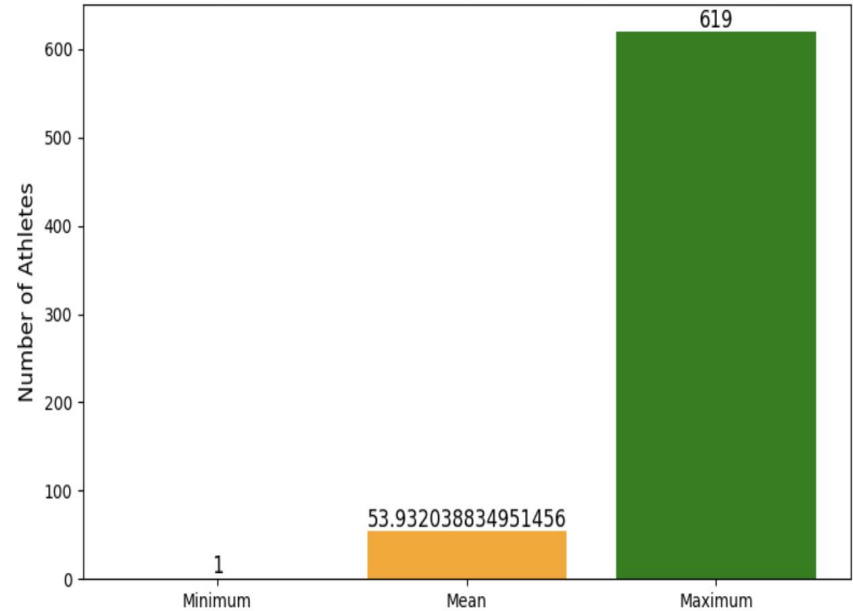
Leon Marchand	1	4	0
Mollie O'callaghan	1	3	1
Shinnosuke Oka	1	3	0
Torri Huske	0	3	2
McINTOSH Summer	0	3	1
...
van Jacob	0	0	1
van Jan	0	0	1
van Jo-ane	0	0	1
van Myrte	0	0	1
van Pierre	0	0	1

2020 vs. 2024 Average Athlete Counts

Summary of Athlete Counts per Country - 2020 Tokyo Olympics



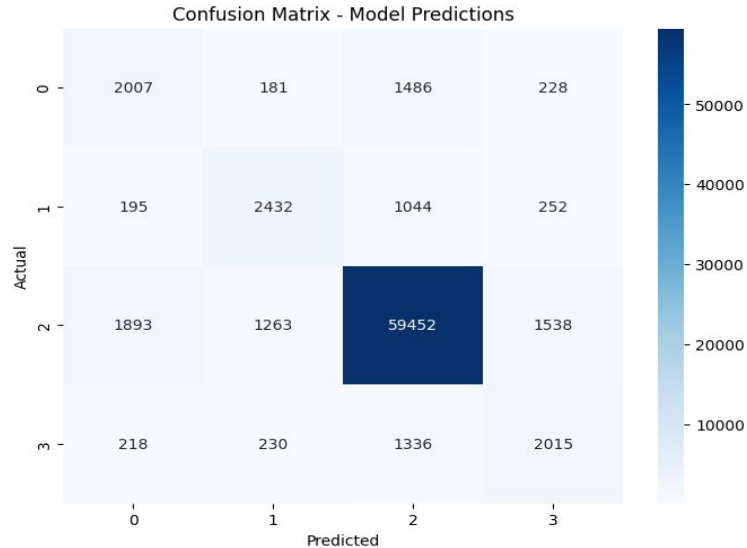
Summary of Athlete Counts per Country - 2024 Olympics



Predicting Future Medal Wins for Countries

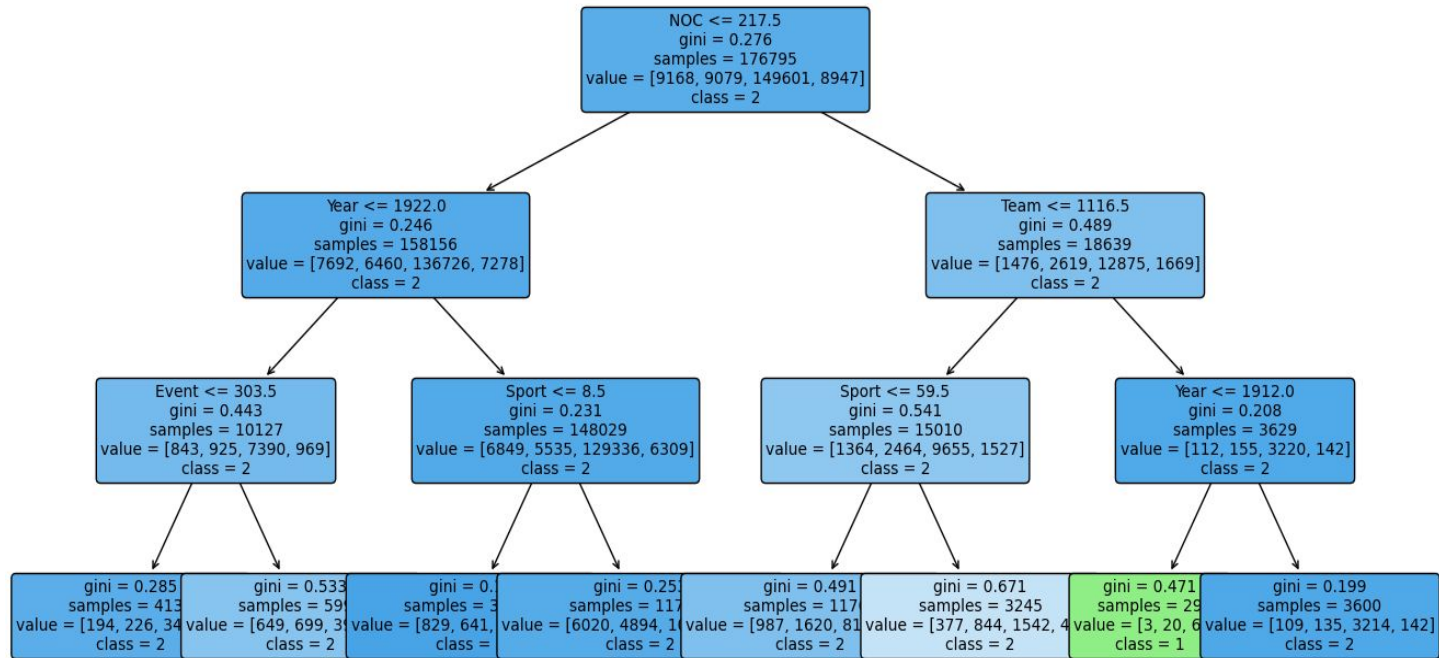
87% overall accuracy

Classification Report:					
	precision	recall	f1-score	support	
0	0.47	0.51	0.49	3902	
1	0.59	0.62	0.61	3923	
2	0.94	0.93	0.93	64146	
3	0.50	0.53	0.51	3799	
accuracy			0.87	75770	
macro avg	0.62	0.65	0.64	75770	
weighted avg	0.87	0.87	0.87	75770	



- Decision Tree model
- Using historical data
- Classify instances into distinct classes (0: No medals, 1: Bronze, 2: Silver, 3: Gold)

Decision Tree





Thank you!
Questions and Comments?