

INTRODUCTION TO ANALYSIS

The reason for this analysis is to help aspiring badminton men's singles players choose where they should train abroad. By understanding which countries produce the high quality or volume of BWF players ranked in the top 200, players can make more informed decisions depending on what traits they seek in a training environment. The analysis can also help players understand which country's training systems worked the best and try to take aspects of them to incorporate into their own learning.

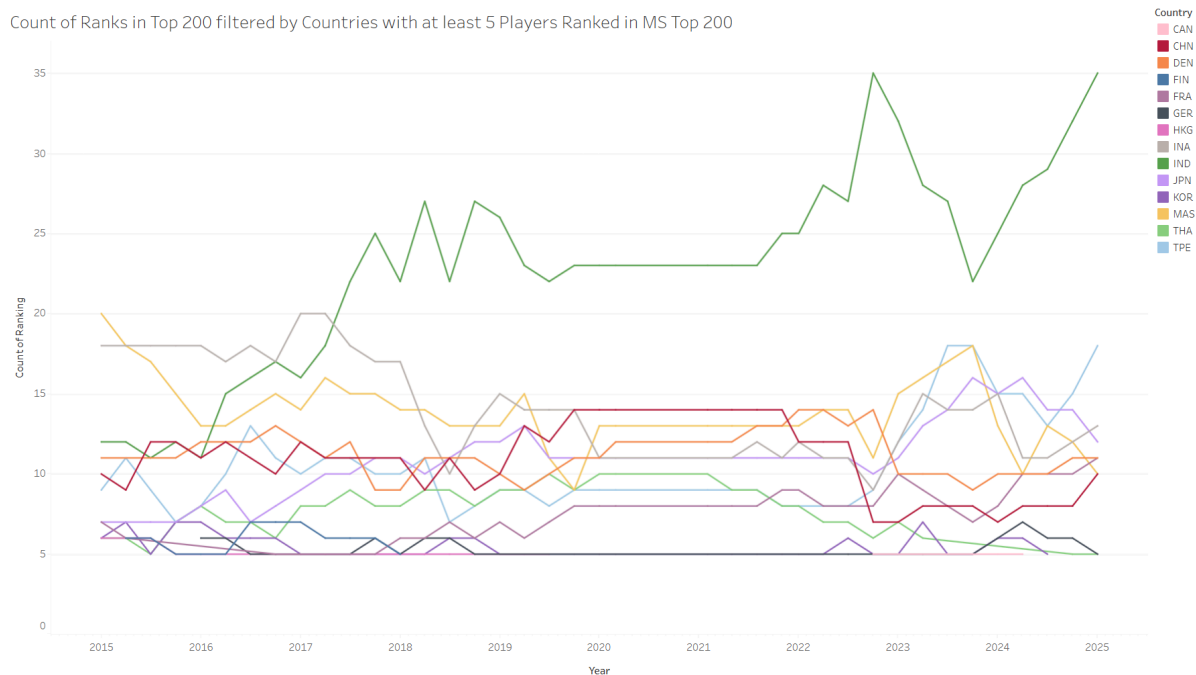
The goal of this analysis is to find which countries have a more successful badminton training system in terms of number of top 200 players and total number of ranking points in comparison to their population size. All analysis is done on men's singles rankings.

[BWF Unique Players in Top 200 Rank Analysis](#) - COUNTD() so only individuals are counted once. Analysis focuses on a country's diversity of talent in the Top 200 rankings.

[BWF Rank Instances in Top 200 Rank Analysis](#) - COUNT() so only individuals are counted multiple times. Analysis focuses on a country's consistency and sustained presence in the Top 200 rankings.

Only visualizations with the players in the top rankings to country population ratio differ. It is denoted as players per million.

ANALYSIS



Countries with at least 5 Players Ranked in MS Top 200 show India has had the highest volume of top 200 players for the majority of the decade. There is an additional spike in players in 2025.

Count of New T200 Players by Country with at least 5 players in T200

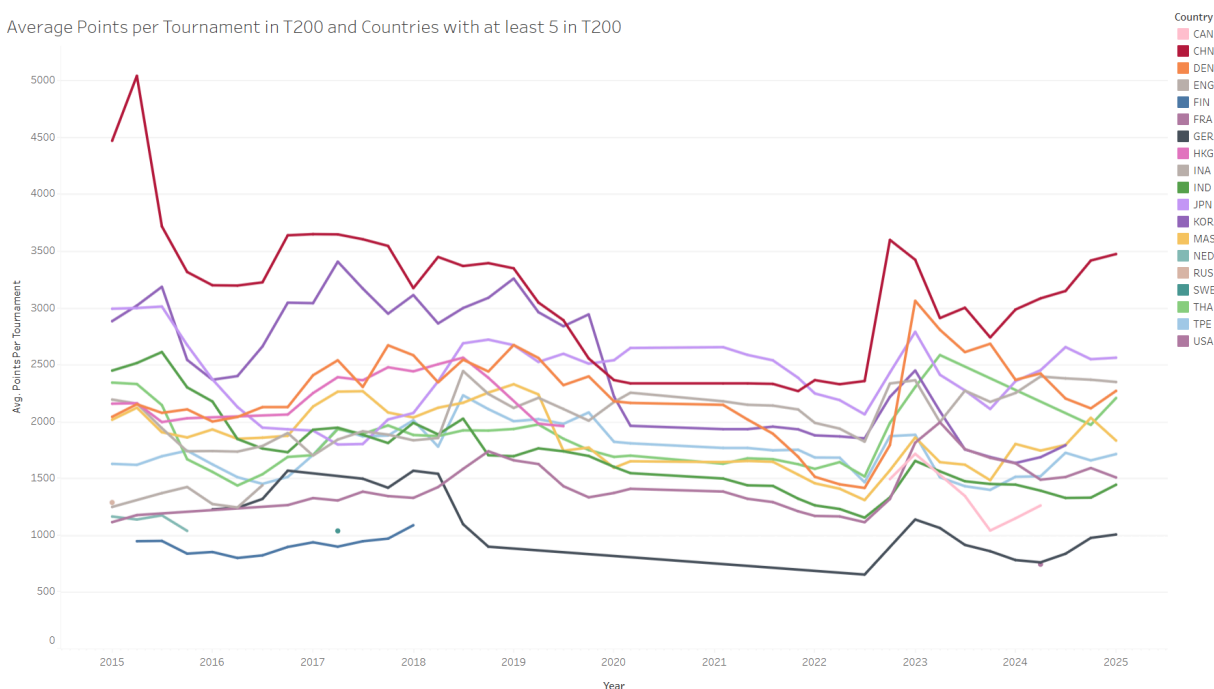


This visualization shows in the last 6 years (10 years if viewed on the Tableau website) new top 200 players primarily come from India. China had many new players in 2019, but Indonesia and Taiwan showed a spike in the recent years of 2023 and 2024. Japan and Malaysia had an influx of elite players in 2022 and 2023, but Japan has greatly slowed. We can conclude here that the Indian badminton system is producing a high volume of top 200 players, and the Indonesian, Malaysian, and Taiwanese badminton systems have produced a high volume of top 200 players in recent years.

Most notable badminton training environments to study:

- India for all years, most notably 2022 and 2024
- China in 2019
- Indonesia, Japan, Malaysia, and Taiwan have produced many new Top 200 players in recent years
- Taiwan and India had the most new Top 200 players in 2024 (most recent year)

Average Points per Tournament in T200 and Countries with at least 5 in T200

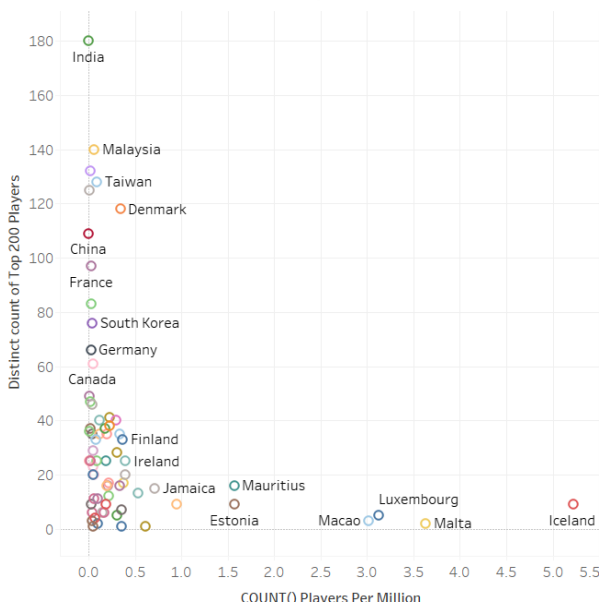


While India has produced a higher volume of players in the top 200, China has on average produced higher quality players in the top 200 rankings as shown by average points per tournament. This visualization shows Chinese players earn more ranking points per tournament as compared to other countries. In recent times, China maintains a significant lead while the next countries like Japan, Taiwan, Denmark, and Thailand cluster slightly below. However, this could also be explained by China only sending their best players to compete whereas other countries allow any citizen to compete. It is noteworthy that South Korea was a prominent challenger to China's quality in 2015-2020, but has not produced the same results since. South Korea's badminton system from 2015-2020 should also be accounted for when discussing the best badminton systems.

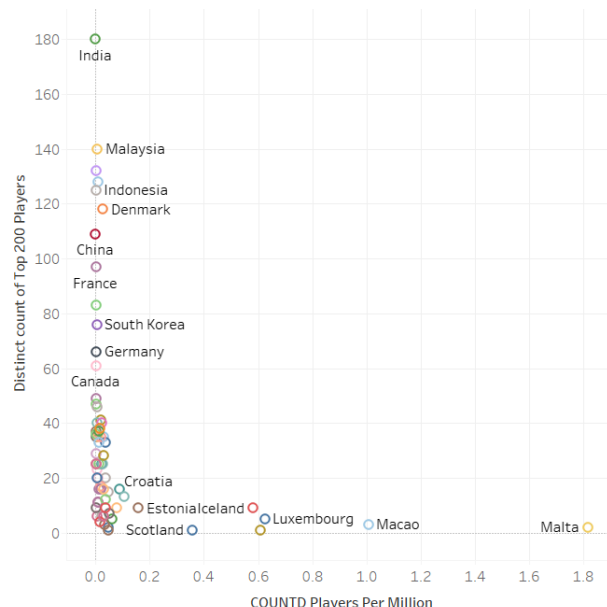
Most notable badminton training environments to study:

- China 2015-2019, 2022-2025
- South Korea 2015-2019
- Japan 2020 (Note Ranking Freeze due to COVID-19)

Rank Instances Per Million Population 2020-2025



Distinct Players Per Million Population 2020-2025



Players_per_million is a field calculated by counting the number of players in the top 200 and then dividing by the sum of that country's population and again by one million. A higher number means that country has produced more players in the top 200 relative to their population compared to another.

In the two scatterplots, we use **COUNT()** in the first and **COUNTD()** in the second. Using **COUNT([BWF ID])** will consider multiple appearances of the same player (rank instances), which is valuable for analyzing a country's consistency and presence in the rankings over time. On the other hand, **COUNTD([BWF ID])** counts unique players, which helps us understand the breadth of a country's talent pool without overemphasizing individuals. Together, these metrics provide a holistic view of both consistency and diversity in country representation in the top 200 rankings.

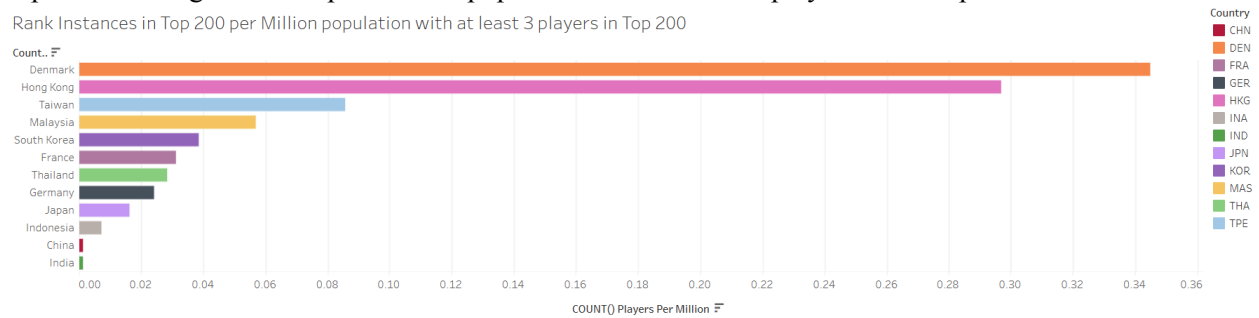
Again, we see India produce the highest volume of top 200 players over 2020-2025. However, we also see countries like Malta, Macao, and Luxembourg produce top 200 rankings despite having relatively small populations. When we look at overall ranking presence, Iceland, Mauritius, and Estonia also show high representation ratios. However, since that chart uses **COUNT()**, the disappearance in the **COUNTD()** chart implies the same players represented these countries and there was no new talent produced.

While the ratio is the star of these visualizations, the actual count of players and rankings is important. For example, in both charts, Macao only has a count of three players in the top 200 while Malta only has two throughout five years. Such few occurrences may skew our analysis and may be outliers. Better cases to study would be Iceland (9 players in the top 200 from 2020-2025) and Luxembourg (5 players in the top 200 from 2020-2025).

These visualizations highlight the potential for smaller countries to achieve high representation ratios due to their population size rather than sheer volume. It suggests that while these countries may not dominate in total player count, their rankings carry more significance relative to their population. This could indicate either better badminton systems in talent development or the presence of a few exceptionally skilled individuals driving their ratios upward. However, we must be cautious to not overanalyze and draw conclusions on outliers.

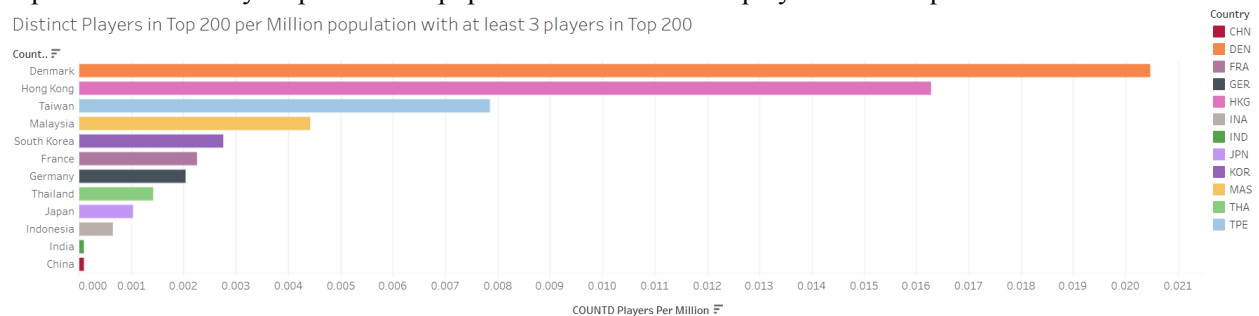
Top 200 Ranking Instances per Million population with at least 3 players in the top 200

Rank Instances in Top 200 per Million population with at least 3 players in Top 200



Top 200 Distinct Players per Million population with at least 3 players in the top 200

Distinct Players in Top 200 per Million population with at least 3 players in Top 200



The top 200 ranking count per million and top 200 distinct players per million charts are both filtered by only showing countries with at least three players in the top 200 in order to limit the effect of outliers.

Both charts show Denmark and Hong Kong far ahead of other countries in terms of population ratio, but are not as far ahead when it comes to distinct top 200 players. This suggests that Denmark and Hong Kong rely on a few consistent players dominating rankings over time, while countries like Malaysia and Taiwan, though slightly lower in ranking count per million, may produce a broader pool of high level players. This may also imply Denmark and Hong Kong focus their elite-level training environments to sustain individual performance, whereas Malaysia and Taiwan emphasize building a wider talent base.

LIMITATIONS

- Population Metrics:

The analysis uses the total country population rather than the population actively participating in badminton. Even if data on badminton-playing populations were available, it would still require filtering for individuals who aspire to reach elite levels of competition. Additionally, age demographics, especially for youth players, play a significant role in talent development but are even less likely to exist. However, population was still included since a larger population still increases the likelihood of a stronger top 1% gene pool, both for the general population and the population of badminton players.

- Potential Misrepresentation of Talent by Ranks

Reaching the top 200 is influenced by many factors that cannot be captured in ranking data. For instance, Denmark may focus on sending only their absolute best players to international competitions, while other countries might send more players in general. Countries with high player volume may also include individuals with access to more personal funding, allowing them to participate in ranking tournaments and earn points independently. Additionally, regions with more badminton likely have more strong players competing in the same tournaments. This leads to many strong players not getting the ranking points their

level deserves. Players in weaker regions are able to gain ranking points more easily due to limited competition.

- Potential Allegiance Changes and Location Changes

Some top players (such as [Zhang Beiwen](#) [China -> Singapore -> USA] or [Ade Resky Dwicahyo](#) [Indonesia -> Azerbaijan]) may choose to represent a different country. This skews results because these players were developed in other countries and shift the ranking points from their origin country to the new country of allegiance. Additionally, some players like [Ygor Coelho](#) (represents Brazil but trains and resides in Denmark), may still represent their country of origin, but move to a different country to train.

- Nature of Sport

Badminton, like many sports, is highly individual and talent-dependent. Success at the elite level often depends on the exceptional qualities of individuals. Top athletes in any sport are natural outliers in terms of skill, physical conditioning, and performance. This makes it difficult to draw conclusions since our sample size is restricted by 200 players per date.

CONCLUSION

The analysis shows the following:

- India consistently produces the highest volume of top 200 players
 - Implies players should go to India if they want more abundant practice partners
- Indonesia, Japan, Malaysia, and Taiwan have produced many new top 200 players in 2022-2025
 - Implies players should go to India if they want more abundant practice partners today
- India and Taiwan have produced significantly more new top 200 players in 2024 compared to other countries
 - Implies players should go to India if they want more abundant practice partners today
- China's players earn significantly more ranking points per tournament compared to other countries' players
 - Implies players should train in China if they want the highest level of practice partners
- South Korea was a strong contender for points per tournament, especially from 2017-2019
 - Implies players should study how players for South Korea trained in 2017-2019 and incorporate those training methods
- Relative to population size, Iceland and Luxembourg produced the most top 200 players
 - Implies players should study how players for Iceland and Luxembourg trained in and incorporate those training methods
- Denmark and Hong Kong's ranking presence is stronger than their distinct top 200 player count, implying they may be focusing on their elite players
 - Implies training in Denmark or Hong Kong may not be as effective since only players at the highest level are focused on
- Taiwan and Malaysia especially see a relative increase when looking at distinct player count instead of ranking instances, implying they focus on building a strong player pool
 - Implies training in Taiwan or Malaysia may be the most effective since the focus is more on all players and less on the elite players

TABLEAU VISUALIZATION LINKS

[BWF Unique Players in Top 200 Rank Analysis](#) - COUNTD() so only individuals are counted once. Analysis focuses on a country's diversity of talent in the Top 200 rankings.

[BWF Rank Instances in Top 200 Rank Analysis](#) - COUNT() so only individuals are counted multiple times. Analysis focuses on a country's consistency and sustained presence in the Top 200 rankings.

DATA SOURCES

Badminton World Federation (BWF) [ranking fansite](#) web scraper: blocked

BWF [corporate ranking website](#) web scraper: blocked

Solution: Manually download 2015-2025 ranking files.

Each file had rows in the form:

Ranking	BWF ID	Last Name	First Name	Gender	Country	Points	Tour
1	50152	LEE	CHONG WEI	M	MAS	83353	9

Ranking: World Ranking

BWF ID: Unique ID representing a player in the BWF database

Last Name: Last Name

First Name: First and Middle names, can have 2-3 names

Gender: All male because this analysis focuses only on Men's Singles

Country: Country code designated by BWF, not the internationally recognized ISO alpha3 country codes

Points: Total points earned by a player

Tour: Total count of tournaments played

Population was joined to the data through multiple python files (not the best method sorry)

country codes join population.py uses SQL to get data on what country codes are in the data. ISO alpha 3 country codes were attempted, but could not be used because BWF uses their own country coding system. The file was manually edited.

create country name column.py uses the manually edited file from the above python script to create a new column for country name. "SIN" and "SGP" both represent Singapore, so all "SIN" was replaced with "SGP".

input population by country and year.py uses world population data from [Our World in Data](#) and joins it to our file of all dates, rankings, and country names. England, Scotland, and Wales were inputted with United Kingdom ratios sourced from [Wikipedia](#) and the [Office for National Statistics UK](#).

FINAL DATASET:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Date	Ranking	BWF ID	Player Name	Gender	Country	Points	Tournaments Played	Points Per Tournament	Country_Name	Year	Population
2	2015-01-02 00:00:00	1	75787	CHEN Long	M	CHN	90001.3806	12	7500.11505	China	2015	714291690
3	2015-01-02 00:00:00	2	50152	LEE Chong Wei	M	MAS	84630.5918	9	9403.399089	Malaysia	2015	16350474
4	2015-01-02 00:00:00	3	54431	JORGENSEN Jan O	M	DEN	71002.5208	15	4733.501387	Denmark	2015	2824537
5	2015-01-02 00:00:00	4	93627	K. Srikanth	M	IND	62787.0765	17	3693.357441	India	2015	686298734
6	2015-01-02 00:00:00	5	54346	SON Wan Ho	M	KOR	62522.4795	18	3473.471083	South Korea	2015	25576748
7	2015-01-02 00:00:00	6	50906	LIN Dan	M	CHN	57759.5667	8	7219.945838	China	2015	714291690
8	2015-01-02 00:00:00	7	34810	CHOU Tien Chen	M	TPE	57433	20	2871.65	Taiwan	2015	11769049
9	2015-01-02 00:00:00	8	63707	TAGO Kenichi	M	JPN	56264.0466	10	5626.40466	Japan	2015	62204438
10	2015-01-02 00:00:00	9	14587	SUGIARTO Tommy	M	INA	56143.395	13	4318.722692	Indonesia	2015	131513677

Date: Date

Ranking: World Ranking

BWF ID: Unique ID representing a player in the BWF database

Last Name: Last Name

First Name: First and Middle names, can have 2-3 names

Gender: All male because this analysis focuses only on Men's Singles

Country: Country code designated by BWF, not the internationally recognized ISO alpha3 country codes

Points: Total points earned by a player

Tournaments Played: Total count of tournaments played

Points Per Tournament: Points / Tournaments played

Country Name: Full name of the country represented

Year: Year of data

Population: Population of the country during that year

IMPORTANT DATA PROCESSING NOTES:

1. The delimiter used is a space so reverse list indexing is needed since First Name can have 1-3 names.

quarterly rankings compiler.py lines 32-37

```
bwf_id = parts[1]
gender = parts[-4]
country = parts[-3]
points = parts[-2]
tournaments_played = parts[-1]
player_name = " ".join(parts[2:-4])
```

2. file_name == "WR 2019-10-01 (Week 40).xlsx" was returning rows with 4 extra nan values at the end of each list. List indexing was used to remove them.

```
if file_name == "WR 2019-10-01 (Week 40).xlsx": ['200', '86104', 'LOW',
'Pit', 'Seng', 'M', 'AUS', '8460', '17', 'nan', 'nan', 'nan', 'nan']
    parts = parts[0:-4]
```

3. Initial files used to build the test scraper were pdfs. 2015-2023 date were excel sheets. Function excel_to_pdf converts excel files to pdfs so it can be processed in the pdf plumber. There are a total of 39 individual ranking files.