

The Spike of Global Interest in Online Chess

By WGM Nadya Ortiz

December 12, 2020

1. Introduction

Chess is one of the most ancient board games in history and it has been played by millions of people worldwide. During this 2020-year, online chess has gained popularity that we have not seen before and millions of new fans have joined online chess platforms.

This project aims to illustrate the current global interest in the online chess.com platform, by finding the number of chess players per country, such as grandmasters, college students, kids among other groups who are actively playing online. In addition, visualizing in a map chess players from different countries and clustering the types of chess opening from around the world.

Quantifying how many people are currently playing chess online, may be beneficial for chess organizers that are looking for new sponsors who could profit from having their names associated with online tournaments and events. It may also encourage beginners or amateur chess players to keep learning new techniques and take advantage of the proven educational and cultural diversity benefits from chess.

2. Data

The chess datasets and geological location information were created by:

- Chess.com Python API [1] allows users to download public data from chess.com website.
- The API has different methods such as getting player info, games, countries, clubs, professional players among other functionalities.
- Google Geocoding Python API (geopy) [2] to locate the coordinates of each country.

To reduce the scope of the data, we will explore these 3 main group: Titled Players, Chess University and Chess Kids:



Data Players:

Using chess.com API I was able to download each player record that consists of: Player Info name, country, joined date, last online date, amount other fields.

The data frame for each group will contain: country, location, last_online among others:

df_players.head(n=3)										
	@id	country_name	country	followers	is_streamer	joined	last_online	location	name	
ahachess	https://api.chess.com/pub/player/ahachess		Vietnam	VN	30	False	2011-06-13 22:01:00	2020-11-16 14:33:00	Tp. Hồ Chí Minh	An Nguyen Thi Thanh
ambotsari	https://api.chess.com/pub/player/ambotsari		Greece	GR	3	False	2020-07-13 21:51:00	2020-11-16 10:51:00	NaN	Anna-Maria Botsari
anastasiyakarlovych	https://api.chess.com/pub/player/anastasiyakar...		Greece	GR	66	False	2017-05-05 15:22:00	2020-12-08 14:29:00	NaN	Anastasiya Karlovych

3 rows x 91 columns

2.1 Titled Players

The International Chess Federation (FIDE), awards chess titles to classify the strength of a player. Once awarded, FIDE titles are held for life. Grandmaster (GM) is the highest title a chess player can attain. Any title may be earned by any player. However, there is a separation for women's titles and they are restricted only to female players.

These are the Official FIDE Titles and number of chess players around the world [4]:

Open titles, January 2020 ^[4]			
Title	Men	Women	Total
Grandmaster (GM)	1,655	37	1,692
International Master (IM)	3,738	116	3,854
FIDE Master (FM)	8,067	37	8,104
Candidate Master (CM)	1,708	19	1,727
Total	15,168	209	15,377

Women's titles, January 2020 ^[4]	
Title	Total
Woman Grandmaster (WGM)	458
Woman International Master (WIM)	846
Woman FIDE Master (WFM)	1,737
Woman Candidate Master (WCM)	762
Total	3,803

2.2 University Players

This group has the most number of players in chess.com. It is the world's #1 online academy.[3]

2.3 ChessKids

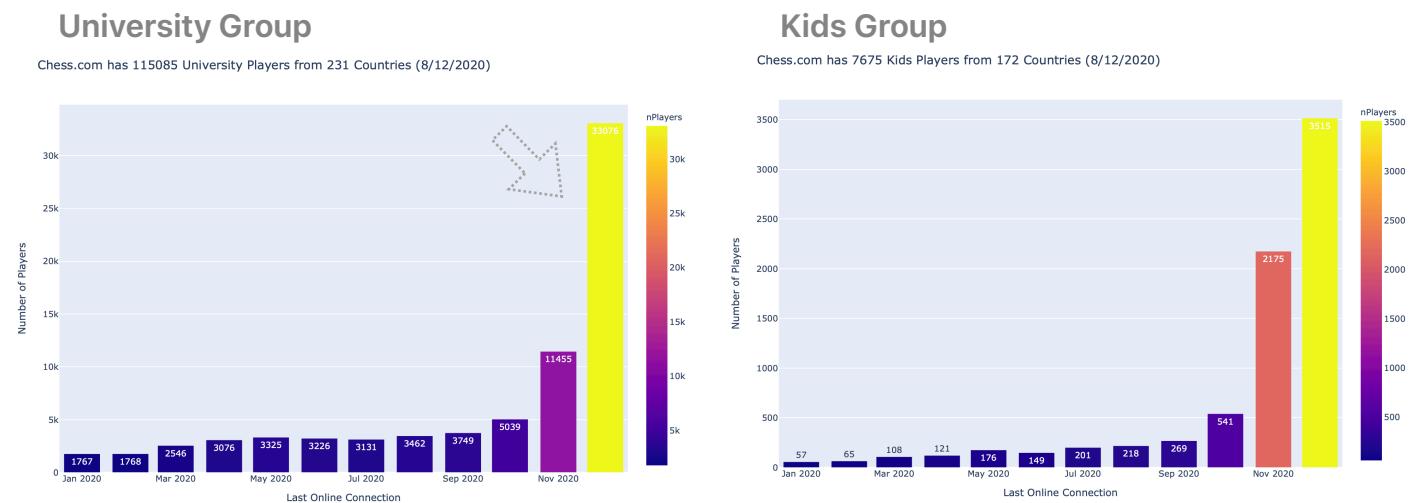
This is the official group for kids, with latest news and updates

3. Methodology

We will describe and analyze each group independently, quantify the number of players for each group, their countries and differences between them.

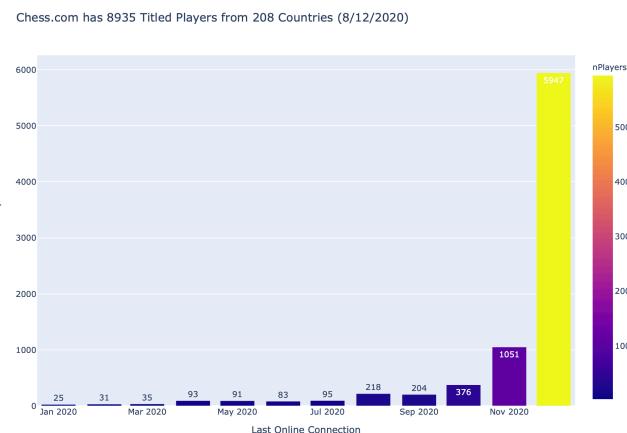
3.1 Number of Players:

All data was collected before December 8, 2020. As we can observed from the plots below, all groups show a spike of new connections during the last month:

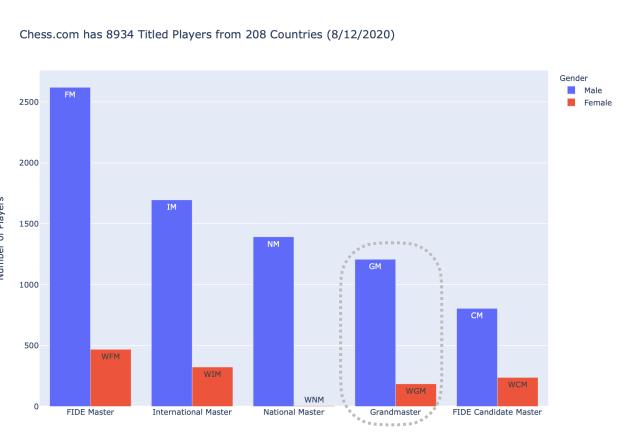


The only group we can identify women is in the Titled group. Based on the data analyzed we observe a significant gap between women and men. For example, there are 1207 Grand masters compared to 185 Women Grand masters. This gap holds for all other titled players such as IM. Vs WIM, MF vs WIM and so on.

Titled Group



Women vs Men



3.2 Number of Countries:

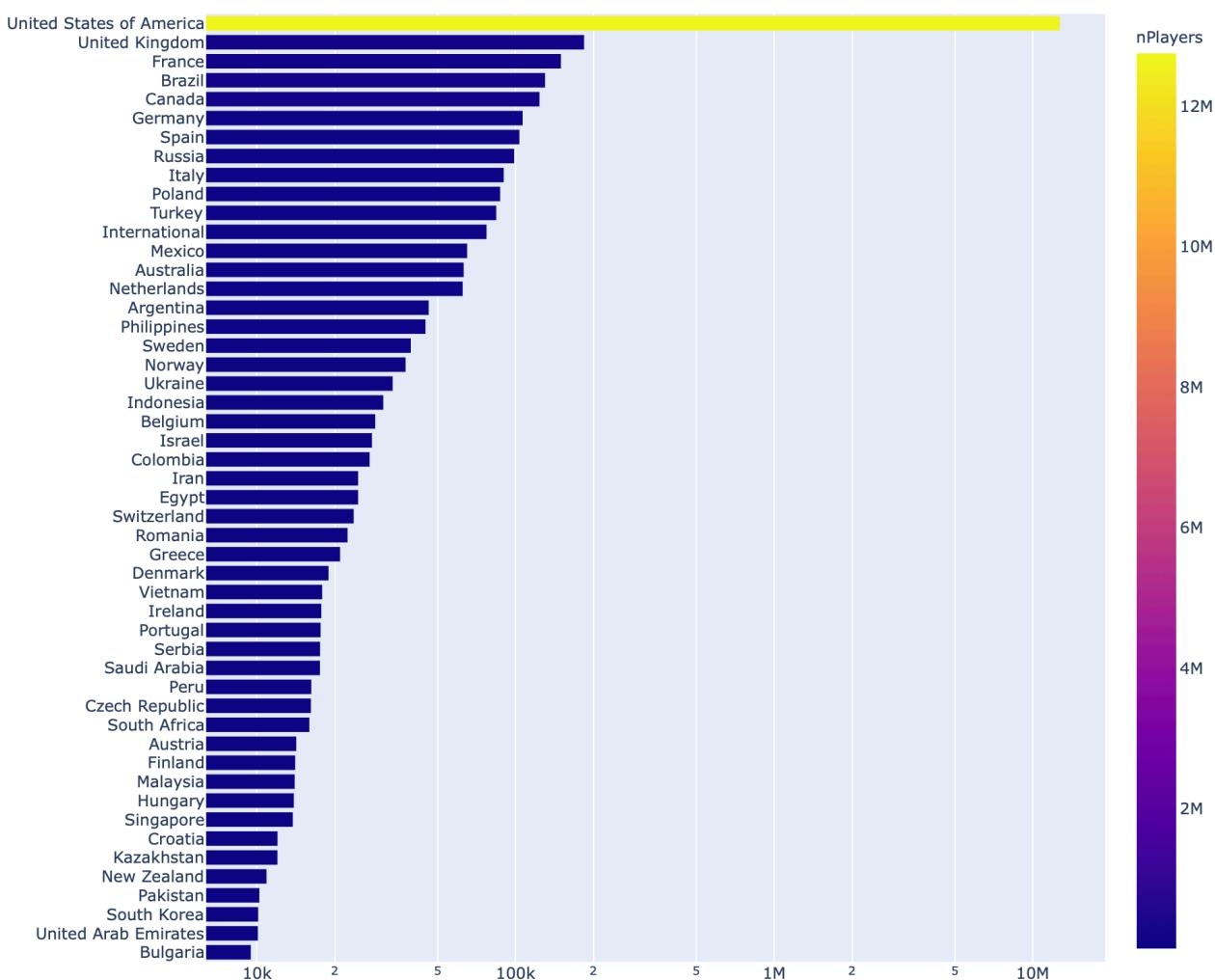
All data was collected before December 12, 2020. chess.com online platform. It has 228 locations from around the globe with more than 15 million players.

As we can observe from the plot, USA has more than 10 millions players which is a remarkable difference compared to other countries. It's worth mentioning that there are other online chess platforms like chess24 or lichess that are more popular in Europe and Asia compared to Chess.com platform that is used more in USA and America continent.

*In Chess.com:

- The country name "international" is when a player does not want to disclose his/her nationality.
- Islands are classified as countries, therefore there are more than 195 official country names.

Top 50 Countries out of 228 from 15.21 million Players @ Chess.com (12/12/20)

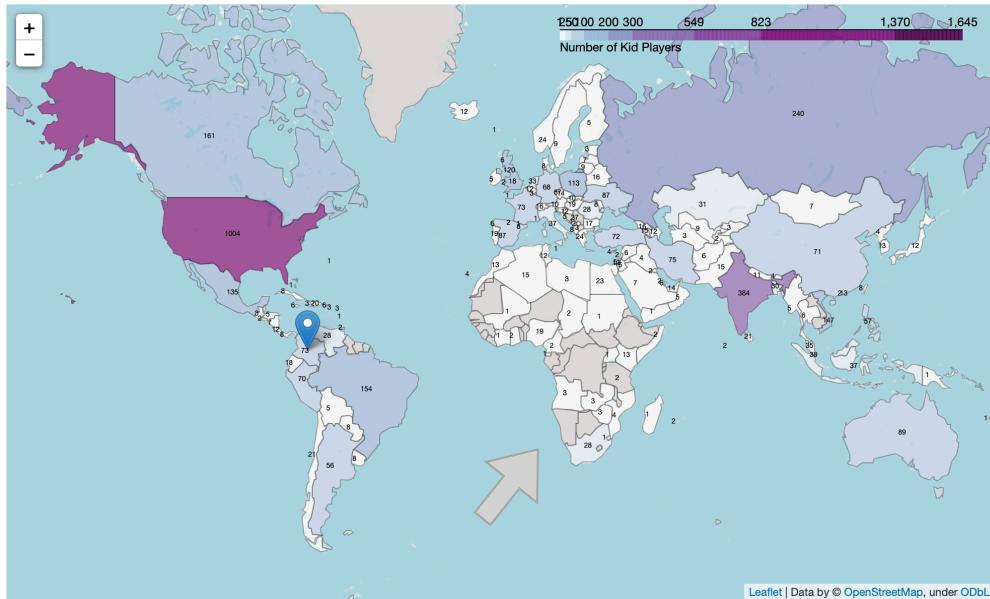


3.3 Chess Players World Map:

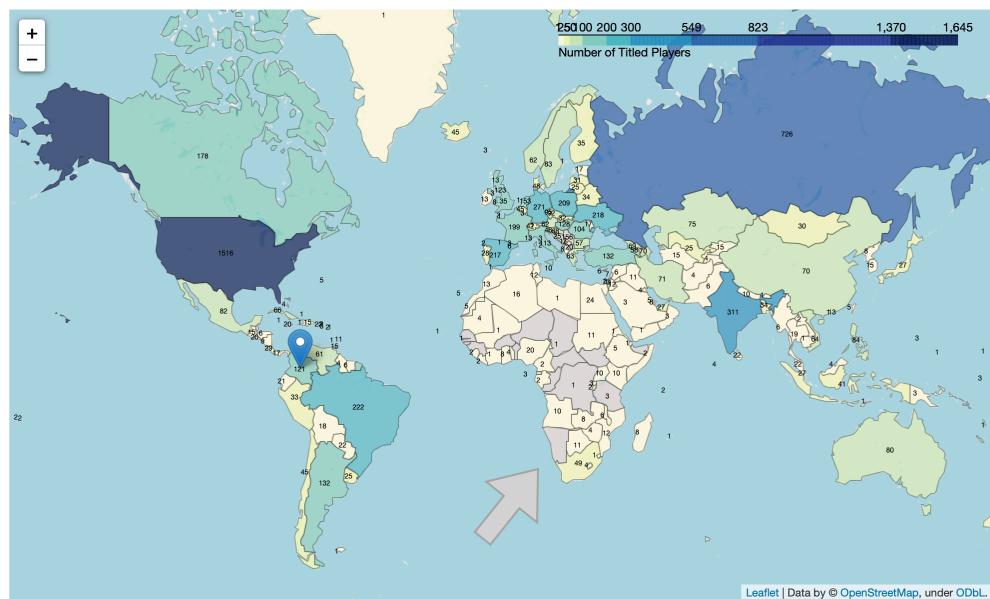
As we can observe from both ChessKids and Titled players maps there are 3 main dominant countries USA, Russia and India.

Emerging countries in Africa show a very low number of players, although I do not know the numbers from other platforms (like chess24 and lichess) this is a reflection of our society where countries with social-economic disadvantage have less access to have a computer and be connected online.

ChessKids Map



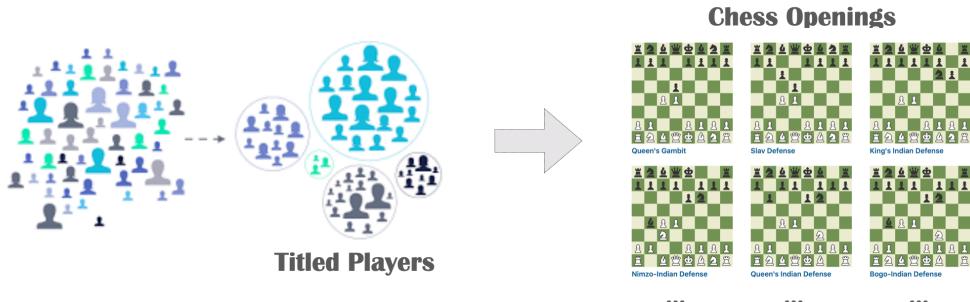
Titled Players Map



* I added a ping in Colombia. I feel proud of my country and it always gives me joy to see many Colombians playing online. Hopefully the numbers will keep growing

3.4 Clustering Chess Openings

A chess opening refers to the initial moves played during a chess game. There are thousands of openings and millions of variants played around the world. We are going to explore and cluster the different kinds of openings played in chess.com by professional chess players during the 2020 year.



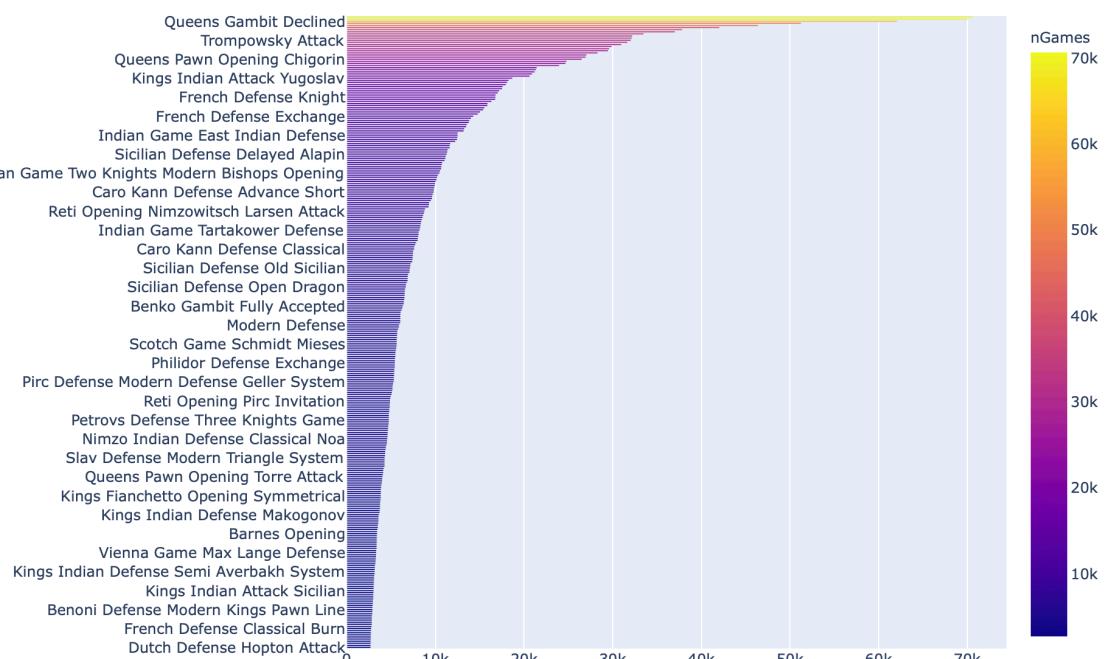
Time Control:

It refers to how much time each player has during a game. To reduce the dataset, we will focus on Blitz (3 mins per player) time control. The Blitz data frame, contains information for each titled user from all the games played during 2020. It has result, ECO, Opening name, TimeControl in seconds.

df_blitz.head()										
	user	Date	White	Black	Result	ECO	Opening	WhiteElo	BlackElo	TimeControl
0	ahachess	2020.01.07	ahachess	quesfr	1-0	D30	Queens Gambit Declined	2252	2200	180
1	ahachess	2020.01.07	Olga_Zhuravleva	ahachess	0-1	B34	Sicilian Defense Open Accelerated Dragon	2322	2231	180
2	ahachess	2020.01.07	FearNoEvil12	ahachess	1-0	B23	Sicilian Defense Closed Grand Prix	2054	2199	180

There are more than 3 millions games in total. Here are the 300 most common Opening Names

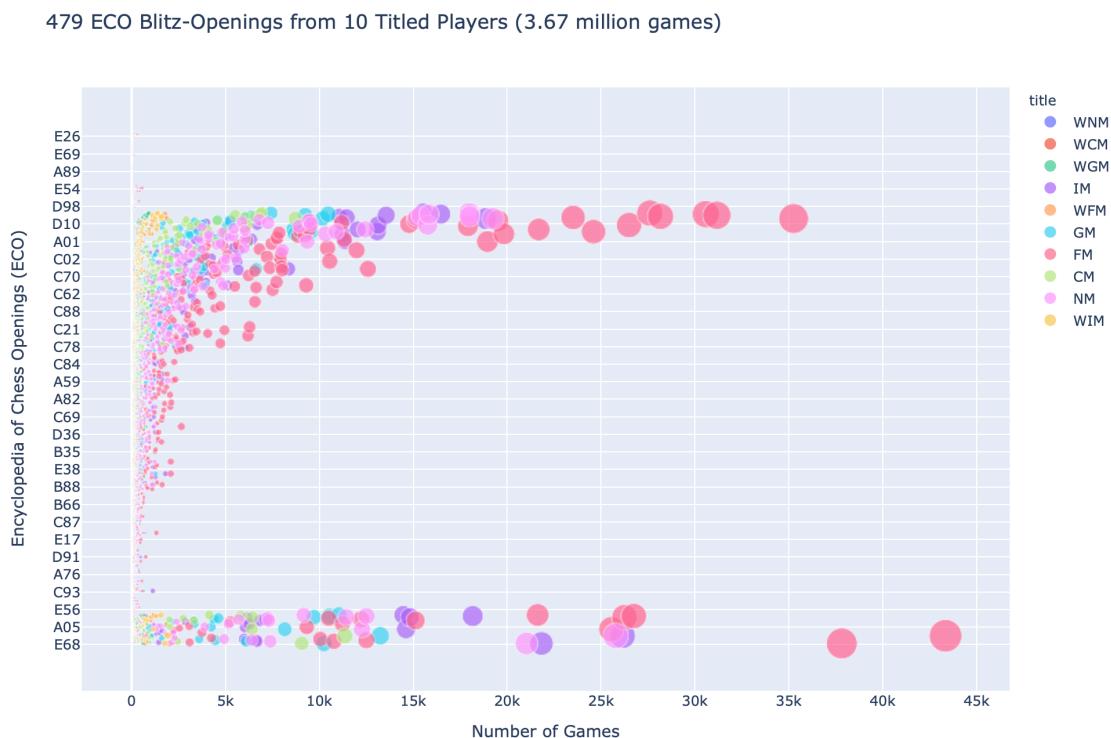
Top 300 Blitz-Openings names from 7194 Titled Players (3.67 million games)



Cluster Blitz Openings (ECO):

The Encyclopedia of Chess Openings (ECO) use a coding system to classify chess openings in five categories: A,B,C,D,E. Each category has 100 subcategories (e.g A00-A99) that represents multiple opening variations for each category. Due to the immense number of chess opening that exists, most players specialize in a limited set of openings. At a high level, a titled player usually selects a repertoire (for white and black openings) based on his/her chess style and preferences to play an open or closed game.

In more than 3 millions games there are 479 unique ECO openings played by 10 titled players:



One-Hot Encoding

Before training, I grouped the openings with the mean of the frequency of occurrence of each ECO.

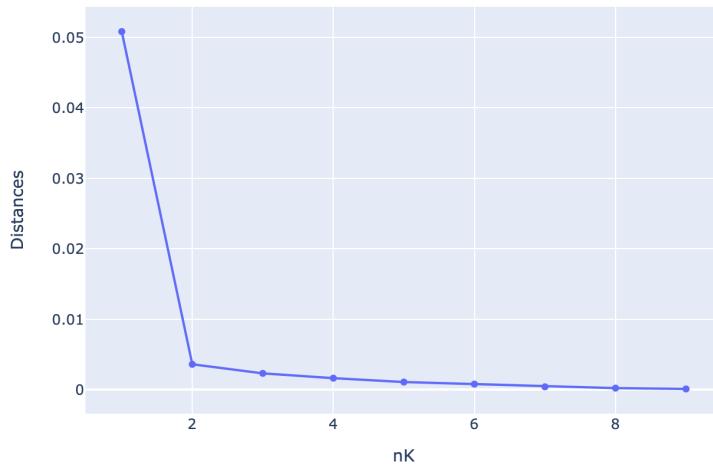
openings_grouped																
	title	A00	A01	A02	A03	A04	A05	A06	A07	A08	...	E90	E91	E92	E93	E94
0	CM	0.029717	0.015202	0.002642	0.001971	0.021871	0.006946	0.006149	0.008691	0.001068	...	0.001861	0.001304	0.000527	0.000038	0.000595
1	FM	0.026866	0.014430	0.002490	0.001689	0.019993	0.007088	0.005476	0.009575	0.001097	...	0.001986	0.001522	0.000561	0.000021	0.000549
2	GM	0.022754	0.013825	0.001583	0.001413	0.021721	0.009450	0.007716	0.014815	0.000979	...	0.002558	0.001064	0.000579	0.000007	0.000741
3	IM	0.026362	0.015855	0.002049	0.001505	0.020704	0.008416	0.005291	0.011650	0.001024	...	0.002389	0.001402	0.000616	0.000013	0.000567
4	NM	0.026913	0.013048	0.003515	0.002299	0.015972	0.005410	0.003878	0.006127	0.000846	...	0.001991	0.001150	0.000687	0.000031	0.000694
5	WCM	0.021518	0.006889	0.002477	0.002554	0.007702	0.003483	0.005379	0.004141	0.000735	...	0.001355	0.001819	0.000851	0.000077	0.000387
6	WFM	0.028096	0.009051	0.002641	0.002116	0.016033	0.005221	0.004078	0.005993	0.000850	...	0.002039	0.001282	0.000293	0.000046	0.000479
7	WGM	0.018832	0.008158	0.001591	0.002285	0.017762	0.006682	0.007232	0.012613	0.001389	...	0.001851	0.001996	0.000174	0.000029	0.000579
8	WIM	0.023945	0.007576	0.002058	0.002198	0.014701	0.006423	0.003975	0.007795	0.001559	...	0.001497	0.001808	0.000343	0.000031	0.000702
9	WNM	0.000000	0.000000	0.000000	0.000000	0.060000	0.040000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000

10 rows x 480 columns

K-means

After training with 10 different k-values, we can observe on the figure below the distances are small due to the limited number of classes to classify. There is a big drop when k=2 meaning there are two major groups as we saw on the ECO-scatter plot above. To get more details between the titled players we will use k=4.

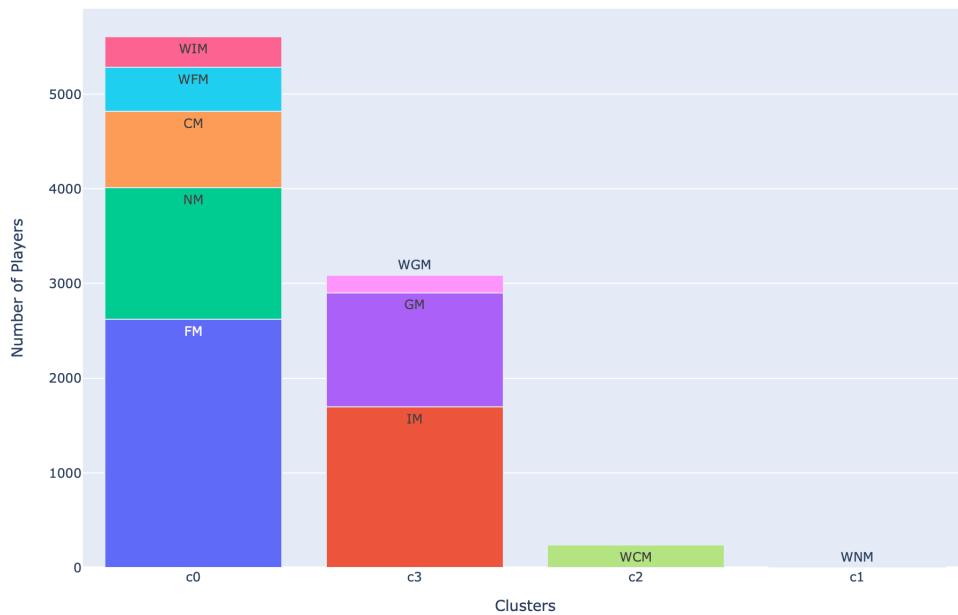
Elbow Method for Optimal k



Clusters:

- Cluster c0: Most of the titled players are in this group.
- Cluster c3: It has the strongest players GM, IM, WGM. With similar openings repertoire.
- Cluster c1,c2: Few titled women in these two groups.

8935 Titled Players Clustered in 4 groups



Most Common Blitz Openings

With more than 3 millions games there were 479 unique ECO blitz-openings. Here are the 10 most common ECO from each titled player.

	title	nPlayers	Chess Title	Gender	Cluster	1st Common	2nd Common	3rd Common	4th Common	5th Common	6th Common	7th Common	8th Common	9th Common	10th Common
0	FM	2619	FIDE Master	Male	c0	A45	A40	A00	B23	D02	B01	D00	B06	B07	A04
1	IM	1695	International Master	Male	c3	A45	A40	A00	B06	D02	D00	A04	A46	D30	B23
2	NM	1392	National Master	Male	c0	A45	A40	A00	B01	D00	D02	B07	B23	C00	B06
3	GM	1207	Grandmaster	Male	c3	A45	D30	B06	D02	A40	A00	A04	A15	B23	B07
4	CM	804	FIDE Candidate Master	Male	c0	A45	A40	A00	D00	D02	B01	A04	A46	B07	C00
5	WFM	468	FIDE Master	Female	c0	A45	A40	A00	D02	B06	D00	B23	A46	B01	B22
6	WIM	323	International Master	Female	c0	A45	A40	B01	A00	D30	D00	B07	D02	B12	B06
7	WCM	237	FIDE Candidate Master	Female	c2	D00	D02	B06	B23	A40	A45	B01	C00	A00	B40
8	WGM	185	Grandmaster	Female	c3	A45	D00	B12	D02	B01	B07	A00	B23	D30	A04
9	WNM	5	National Master	Female	c1	D30	B06	A04	B50	D70	E60	B20	D85	D11	B80

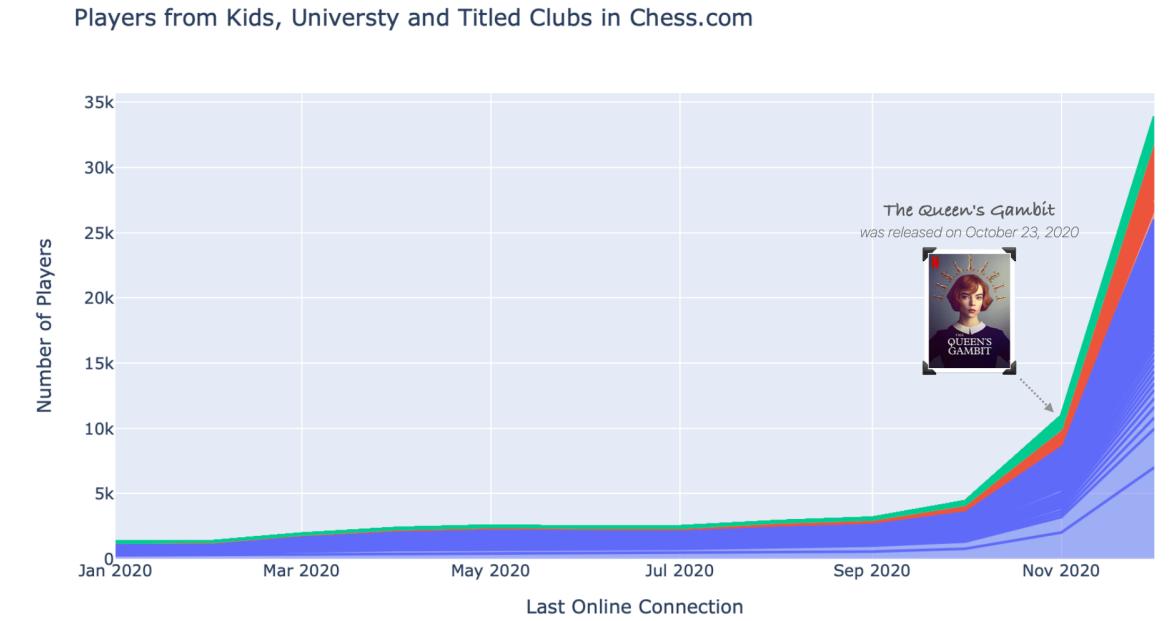
Based on the results ECO-A45 is the most common Opening among Titled Players. From the Grandmaster category (GM: is the highest title a chess player can achieve), the two most common blitz-opening are Indian Game ECO-A45 and The Queen's Gambit ECO-D30.

For chess curiosity, these are the first moves from the most common openings between GMs.



4. Results

From the 3 major clubs we analyzed: Titled Players, University and ChessKids. We observed a big spike of new users during the last two months.



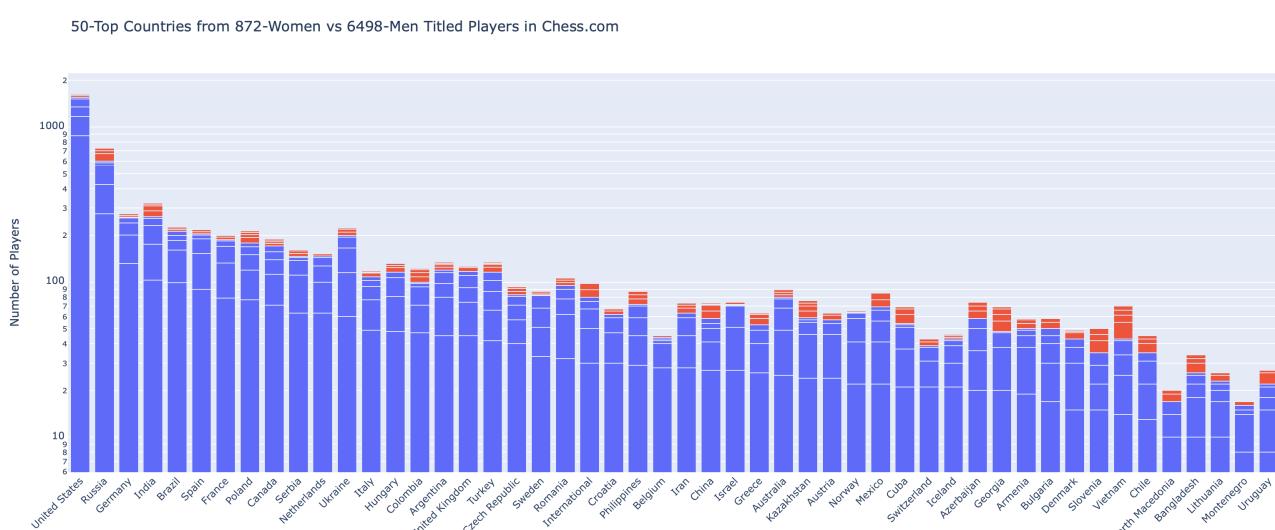
Queen's Gambit Effect

This significant spike of new users is attributed to the new series from Netflix "The Queen's Gambit" [5] released on October 23. The series became the most-watched scripted limited series, attracting 62 million account viewers during its first month. After the show's debut, multiple chess websites have gained thousands of new users.

Women in Chess

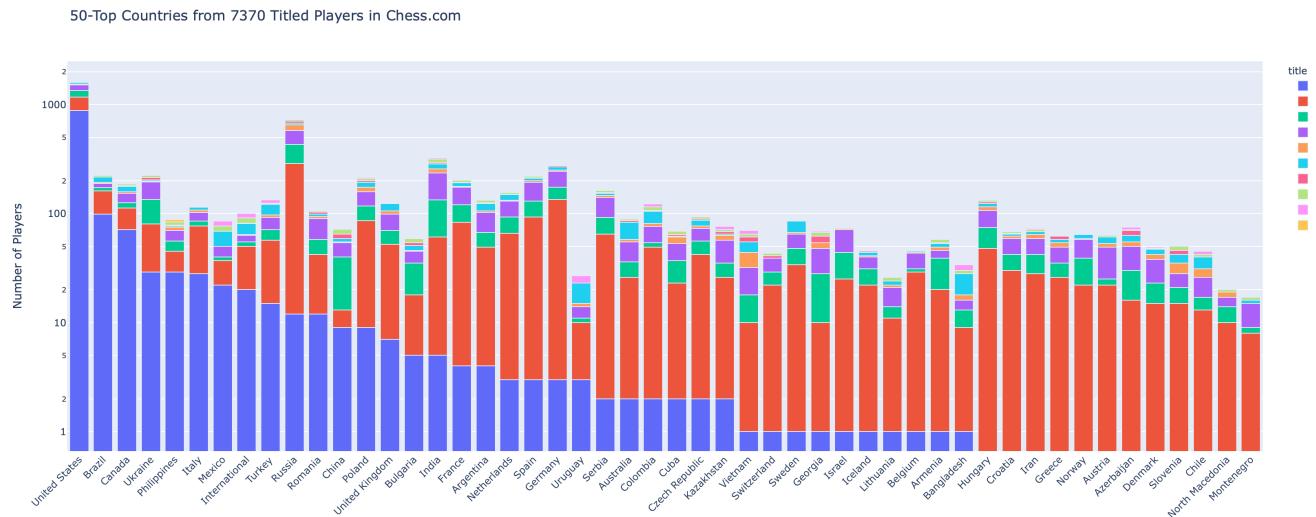
As we can observe from the data, chess remains a male dominated sport. Based on the results there is a significant gap between the number of Titled women vs men from around the world.

*International country means the player did not disclose the nationality.



Titled Players in Chess.com

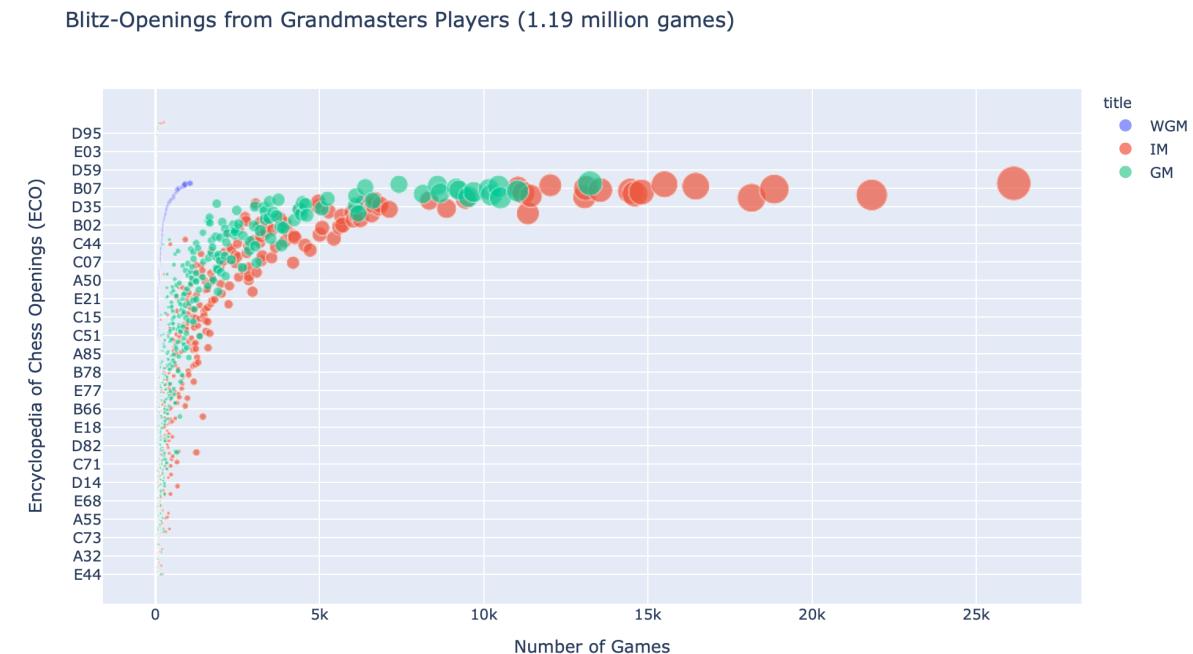
We analyzed 10 Titled players from National (NM), Candidate (CM), FIDE (FM), International (IM) and Grandmasters (GM) for both Women and Men. As we can observe from the plot below, the National and FIDE Masters are the biggest groups among titled players and the most common among different countries.



Clustering Chess Openings

The K-means clustering algorithm was used to classify different chess Blitz-Openings among titled players. Based on the results, there were two main clusters with similar openings repertoire. Cluster c0 that had most of the titled players and Cluster c3 with the strongest chess players.

The picture below represents cluster-c3, with Grandmasters (GM), International Masters (IM) and Women Grandmasters (WGM):



5. Conclusion

On December 12, 2020, the online chess.com platform has around 15million players. To reduce the scope of the project, we analyzed three main groups: Titled players, University and ChessKids with more than 130k players, 228 locations from around the globe.

The data shows a significant gap between women and men. Among all players, National, Candidate, FIDE, International and Grandmasters, the number of titled men players is almost 7 times more than women. This small number of women in chess, is a resemblance of the disparity we see in other fields such as Science, Technology, Engineering, and Math (STEM).

From the thee three groups, we observed a remarkable growth in the number of new users during the last two months. The spike of interest in chess from around the world is due to The Queen's Gambit series from Netflix released on October 23. The series had a transcendental impact that many of us would have not predicted.

* Note:

As a woman grandmaster from Colombia, (although "retired" from competitive chess), I have never seen the popularity that Chess is having nowadays, such as reports from New York Times, Washington Post, Bloomberg among others. I believe chess is having a successful momentum, and this is a great opportunity to look for new sponsors, promote chess in the schools, encourage new girls to learn this beautiful game and achieve master levels. Chess could be portrayed not only as sport but as an educational tool, especially in low income areas with sociocultural disadvantages, chess could bring opportunities and change the course of many lives. -like it did with me.

References:

- ◆ [\[1\] Python Chess.com Wrapper](#)
- ◆ [\[2\] Google Geocoding Python API \(geopy\)](#)
- ◆ [\[3\] University and ChessKids Players](#)
- ◆ [\[4\] FIDE Title Players](#)
- ◆ [\[5\] The Queen's Gambit series from Netflix](#)

Latest News: up to Dec 12,2020

- ◆ [The Queen's Gambit Chess Boom Moves Online - Bloomberg report](#)
- ◆ [The Queen's Gambit' Sends Chess Set Sales Soaring - The New York Times](#)
- ◆ [Five myths about chess - The Washington Post](#)