

# EXPLORATORY DATA ANALYSIS

**On Twitch TV Channels** 

Ishan Manchanda C22011

Navam Pradhan C22016

DS Spring 2022

STS-I Project



### ABSTRACT -

Within the gaming industry, live-streaming is becoming very popular as a form of online entertainment. Especially the so called social live streaming services (SLSSs) as a new type of social media have established in the last few years. One of the most prominent and current examples is Twitch.tv which provides the opportunity for streamers to broadcast a game and react to viewers' comments in real time. The pandemic was a major catalyst for the explosion of gaming and live streaming into the mainstream culture.

The main goal of this study is to analyse the spread of Twitch users among different channels. This Project explores the Top 1000 channels of Twitch.tv platform. The data has been collected from Kaggle.

Moreover, esports as an industry is growing at an alarming rate in India due to the popularity of mobile phone games such as PUBG mobile. The total size of India's esports industry will touch Rs 1100 crore by FY25.

## Introduction:

Twitch is an American video live streaming service that focuses on video game live streaming, including broadcasts of esports competitions, in addition to offering music broadcasts, creative content, and "in real life" streams. It is operated by Twitch Interactive, a subsidiary of Amazon.com, Inc. It was introduced in June 2011 as a spin-off of the general-interest streaming platform Justin.tv. Content on the site can be viewed either live or via video on demand.

The popularity of Twitch eclipsed that of its general-interest counterpart. In October 2013, the website had 45 million unique viewers, and by February 2014, it was considered the fourth-largest source of peak Internet traffic in the United States. At the same time, Justin.tv's parent company was re-branded as Twitch Interactive to represent the shift in focus – Justin.tv was shut down in August 2014. That month, the service was acquired by Amazon for US\$970 million, which later led to the introduction of synergies with the company's subscription service Amazon Prime. Twitch acquired Curse in 2016, an operator of online video gaming communities and introduced means to purchase games through links on streams along with a program allowing streamers to receive commissions on the sales of games that they play.

By 2015, Twitch had more than 100 million viewers per month. In 2017, Twitch remained the leading live streaming video service for video games in the US, and had an advantage over YouTube Gaming, which shut down its standalone app in May 2019. As of February 2020, it had 3 million broadcasters monthly and 15 million daily active users, with 1.4 million average concurrent users. As of May 2018, Twitch had over 27,000 partner channels.

Twitch is designed to be a platform for content, including esports tournaments, personal streams of individual players, and gaming-related talk shows. A number of channels do live speed running. The Twitch homepage currently displays games based on viewership. As of June 2018, some of the most popular games streamed on Twitch are *Fortnite*, *League of Legends*, *Dota 2*, *PlayerUnknown's Battlegrounds*, *Hearthstone*, *Overwatch* and *Counter-Strike: Global Offensive* with a combined total of over 356 million hours watched.

Twitch has also made expansions into non-gaming content; such as in July 2013, the site streamed a performance of 'Fester's Feast' from San Diego Comic-Con, and on July 30, 2014, electronic dance music act Steve Aoki broadcast a live performance from a nightclub in Ibiza. In January 2015, Twitch introduced an official category for music streams, such as radio shows and music production activities, and in March 2015, announced that it would become the new official live streaming partner of the Ultra Music Festival, an electronic music festival in Miami.

#### **Charity**

Broadcasters on Twitch often host streams promoting and raising money towards charity. By 2013, the website has hosted events which, in total, raised over US\$8 million in donations for charitable causes, such as Extra Life 2013. As of 2017, Twitch has raised over US\$75 million in donations for charitable causes. The biggest charity event of Twitch is Zevent, a French project created by Adrien Nougaret and Alexandre Douchary, with more than US\$10 million raised for Action Contre la Faim in October 2021.

#### **Esports**

ESL tournaments have aired on Justin.tv and later Twitch.tv since 2009. The platform has also been a long time broadcaster of the Evolution Championship Series.

Twitch has been the official broadcaster of the League of Legends World Championship since 2012, as well as other League of Legends tournaments organized by Riot Games.

Dota 2's premier tournament The International has been livestreamed on Twitch since 2013.

The platform airs *Rocket League* tournaments organized by Psyonix since 2016. The ELeague also broadcasts events on Twitch since 2016.

Twitch and Blizzard Entertainment signed a two-year deal in June 2017 to make Twitch be the exclusive streaming broadcaster of select Blizzard esports championship events, with viewers under Twitch Prime earning special rewards in various Blizzard games. Twitch also reached a deal in 2018 to be the streaming partner of the Overwatch League, with the site also offering an "All-Access Pass" with exclusive content, emotes, and in-game items for *Overwatch*. Blizzard switched to rival platform YouTube in 2020.

Fortnite Battle Royale competitions have aired on Twitch since its launch in 2017, including the E3 2018 Fortnite Pro-Am and the 2019 Fortnite World Cup.

The NBA 2K League has been livestreamed on Twitch since its inception in 2018.

As the COVID-19 pandemic suspended motorsports competitions around the world, several series launched simulated racing competitions with real-life professional drivers. Some series had official broadcasts on Twitch, such as Formula One and IMSA. Many drivers also had their personal live streams on Twitch, as was the case of several eNASCAR iRacing Pro Invitational Series and INDYCAR iRacing Challenge drivers.

Creators and audience

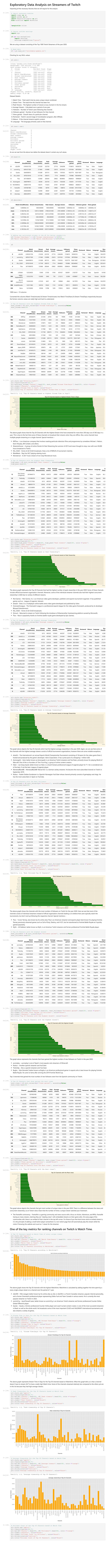
#### **Streamers**

Streamer Ninja had been among Twitch's top personalities, with over 14 million followers. In August 2019, however, Ninja announced that he would move exclusively to a Microsoft-owned competitor, Mixer. After Ninja left, the top three streamers in October 2019 based on follower count were Tfue (7.01 million followers), Shroud (6.45 million followers) and TSM Myth (5.1 million followers). Twitch began signing exclusivity deals with high-profile streamers in December 2019, starting with DrLupo, TimTheTatman, and Lirik, who had a combined 10.36 million followers at the time. Dr DisRespect signed a multi-year deal in March 2020. In May 2020, Twitch signed popular streamers Summit1g, dakotaz and JoshOG to multi-year exclusive deals. On June 26, 2020, Dr DisRespect was banned from Twitch for unexplained reasons and his channel was removed from the site. Following the discontinuation of Mixer in late-July 2020, both Ninja and Shroud (who had also defected to the service) re-signed exclusively with Twitch.

As of March 2022, there have been six streamers to have reached over 100,000 concurrent subscribers. These streamers are Ninja, Shroud, Ranboo, Ludwig Ahgren, Casimito and Ironmouse. In April 2021, *Business Insider* reported that "over the past 31 days, Ahgren has streamed non-stop in an attempt to break the record of 269,154 subscribers held by gaming personality Tyler 'Ninja' Blevins. By the end of the month-long stream, Ahgren had over 282,000 subscribers on his channel. [...] At one point during his sleep cycle, his channel had the most concurrent viewers of any on the platform". In analysis of the October 2021 data leak, multiple news outlets reported that the

three top-earning Twitch content creators are *Critical Role* (\$9,626,712), XQc (\$8,454,427), and Summit1g (\$5,847,541). Sisi Jiang, for *Kotaku*, reported that "excluding streams that are run by multiple people (such as *Critical Role*), there are no women in the top third of top-earning Twitch content creators"; in total, there are only three women in the top 100 and only one is a woman of colour. Jiang highlighted that these streamers are "*Valorant* streamer Pokimane at 39th place, cosplayer Amouranth at 48th, and music streamer Sintica at 71st" and commented that "in spite of the complaints about the 'hot tub meta,' 'titty streamers,' and how some male streamers perceive that female streamers are stealing views from men, the numbers show that only a small percentage of women are among the ranks of Twitch's highest-earning content creators".

In August 2021, DrLupo left Twitch for an exclusivity deal with YouTube; TimTheTatman followed in September 2021, as did Ludwig Ahgren in November 2021. Nathan Grayson, for *The Washington Post*, commented that when streamers moved to Mixer in 2019, Twitch quickly locked down multiple streamers in exclusivity deals; however, streamers who moved to Mixer saw their audiences undergo "a marked downsizing. [...] It demonstrated that many viewers within Twitch's ecosystem, when deprived of their favourite big streamers, will just find other Twitch streamers to take their place. [...] Now Twitch is bargaining from a place of confidence. That allows it to re-evaluate previous deals made when streamers had more leverage". Grayson reported that lower offers from Twitch coupled with Twitch's higher streaming hour requirement ("YouTube's contracts start at 100 hours of streaming time per month while Twitch's start at 200") has made YouTube's exclusivity deals "tantalizing" to some Twitch streamers. Grayson wrote that "Ryan Wyatt, head of YouTube Gaming, said that allowing streamers to have a better work-life balance is a big priority for him"; DrLupo cited work-life balance as part of his decision to leave Twitch.



	Text(0.5, 1.0, 'Follower C	ount of Top 50 Chan	Follower Count of T	Top 50 Channels
	7000000 - (\$6000000 - 5000000 -			
	Follower Count (in millions) 2000000 - 20000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 20000000 - 20000000 - 2000000 - 2000000 - 2000000 - 20000000 - 20000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 20000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 20000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 20000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 20000000 - 20000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 20000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 2000000 - 20000000 - 20000000 - 20000000 - 20000000 - 20000000 - 20000000 - 20000000 - 20000000 - 200000000			
	xQcOW	Anomaly  The Tatman  LIRIK  (riotgames)  Rubius  auronplay  sodapoppin  ryang1125)  Lineacki	alanzoka - ohhCamage - Lord_Kebun - LCK_Korea - Castro_1021 - rDisrespect - eRealKnossi - TheGrefin	voba - Pestily - Pestily - Solary - Gotaga - Got
	sum sum Sum G G G G G G G G G G G G G G G G G G G	Anomaly TimTheTatman LIRIK Riot Games (riotgames) Rubius auronplay MontanaBlack88 sodapoppin COD (hanryang1125)	alanzoka alanzoka CohhCarnage Lord_Kebun LCK_Korea Castro_1021 DrDisrespect TheRealKnossi uu	GamesD Mo Mo doi han BeyondTh
In [149	<pre># Calculating which Channel df['Popularity'] = df['Wa' most_popular = df.sort_va.</pre>	tch time(Minutes)']	<pre>/df['Stream tim</pre>	
In [150	plt.figure(figsize=(20,7) plt.barh(most_popular['Chaplt.xlabel('Popularity Ind	) annel'].head(20), m dex', color='Black'		opularity'].head(20), color ='maroon')
Out[150	plt.ylabel('Channels', co.plt.title('Top 20 Most Pop Text(0.5, 1.0, 'Top 20 Mos	pular Channels', co	')	
	Anomaly - RiotGamesBrazil - TheGrefg - LEC - TheRealKnossi - OverwatchLeague -		Top 20 Most F	Popular Channels
	Riot Games (riotgames) - StarLadder_cs_en - MontanaBlack88 - KEEMSTAR - Call of Duty (callofduty) - LCK -			
	RocketLeague - LCK_Korea - Rubius - Asmongold - LCS - auronplay -			
	dota2ti_ru-dota2ti-			larity Index 120000 140000 160000 1800000 180000 180000 180000 180000 180000 180000 180000 180000 1800000 1800000 1800000 1800000 1800000 1800000 1800000 180000 180000000 1800000 1800000 1800000 1800000 1800000 18000000 18000000 1800000 180000000 18000000 18000000 1800000000
In [151	this list had high watch time, average concurrent viewershi	but less stream time rela p when the channel was	atively to other cha	annels. So, We can believe that all of these channels had high
	df_language = df[['Stream		viewers','Avera	age viewers','Followers','Language']]\
Out[151			<b>Followers Followers</b> 5682.20 816537.60	_
	1       Chinese       104.56         2       Czech       97.77         3       English       86.45	7 9918.50	3424.80201714.531827.67247374.835112.81705685.65	
	4       Finnish       145.57         5       French       91.54         6       German       80.81	4 33865.41 1 31274.73	1027.00 255595.00 3506.53 439734.56 4278.18 479367.57	
	7       Greek       85.66         8       Hungarian       77.66         9       Italian       85.51	6 24780.00	1961.00 82289.00 2140.50 347856.50 2205.53 328961.88	
	10       Japanese       99.35         11       Korean       83.25         12       Other       161.54	5 18633.04	4762.60160145.803913.88262173.871149.0030986.00	
	13       Polish       89.65         14       Portuguese       110.07         15       Russian       63.40	7 38463.48	2905.25 397090.58 3800.93 555868.26 6594.23 339730.49	
	16       Slovak       103.01         17       Spanish       50.51         18       Swedish       88.95	1 53186.99	<ul><li>2072.00 252184.00</li><li>6449.69 779686.12</li><li>1027.00 96691.00</li></ul>	
	19       Thai       114.87         20       Turkish       49.54		2011.36 202150.36 4760.86 718936.00	
In [152	channels in their respective la  Top_Streamers_lang1 = df.o	anguage for each measu	re / Total no of cha	lation is done by adding together all the values of different annels in the specific language.  Lime (Years) ': ['mean', 'min', 'max']})
Out[152	Top_Streamers_lang1  W  mean mi	/atch Time(Years)		
	Language         454.814361         266.68598           Chinese         644.893419         245.29700			
	Czech         546.222118         269.07565           English         897.868339         232.70750           Finnish         410.439326         410.43932	1483.766039 11788.740011		
	French 719.319663 232.48259  German 749.545233 236.06963  Greek 438.647175 438.64717	35 4582.307820		
	Hungarian       441.836687       392.78336         Italian       479.640537       246.51375         Japanese       796.532934       264.33727	490.890011 56 1372.809903		
	Korean         686.770193         232.97625           Other         568.546747         568.54674           Polish         682.050404         256.68247	56 4160.316724 47 568.546747		
	Portuguese         860.898585         236.26355           Russian         715.799955         237.46621           Slovak         578.809361         578.80936	56 10739.328225 10 2805.826142		
	Spanish         715.743500         233.11384           Swedish         250.421946         250.42194           Thai         467.387617         270.34788	41 4925.100143 46 250.421946		
In [153	<b>Turkish</b> 541.686064 241.06632	24 1382.000542	pased on Watch 1	Time
Out[153	Top_Streamers_lang		'', 'Watch Time(	(Years)']].groupby('Language', as_index=False).max()
	<ol> <li>Arabic shongxbong</li> <li>Chinese 龜狗 (gueigotv)</li> <li>Czech Xnapycz</li> </ol>	1545.590040 1483.766039		
	<ul> <li>3 English xQcOW</li> <li>4 Finnish OfficialAndyPyro</li> <li>5 French rhobalas_lol</li> <li>6 Corman xTheSolutionTV</li> </ul>	410.439326 2942.536872		
	<ul><li>6 German xTheSolutionTV</li><li>7 Greek Pun1shers_TV</li><li>8 Hungarian WeAreTheVR</li></ul>	438.647175 490.890011		
	9 Italian veloxey 10 Japanese スタンミ (sutanmi) 11 Korean 홍방장 (game2eye)	2152.795634 4160.316724		
	<ul><li>12 Other LigaPro1</li><li>13 Polish polsatgames</li><li>14 Portuguese zorlaKOKA</li></ul>	1364.338527 10739.328225		
	<ul><li>15 Russian ybicanoooobov</li><li>16 Slovak resttpowered</li><li>17 Spanish zEkO</li></ul>	578.809361 4925.100143		
	<ul><li>Swedish CyrusTWO</li><li>Thai voicetv</li><li>Turkish wtcN</li></ul>	1042.696775		
In [154 In [155	Tanguage_TISC = GI[ Bangua	age'].values		
	<pre>text = "" for word in language_list   text = text + " " + wo</pre>			
	wordcloud = WordCloud(wid:  plt.figure(figsize=(10,10 plt.imshow(wordcloud, into plt.axis("off") plt.margins(x=0, y=0)	))		
		nese	Frei	nch
	Greek	Ru	ssia	
	German			Swedish
	Eng	lis	S N Itali	ian
	Spai	nish	Turki	sh
	Czech		olish	
		Por	tugu	ese
In [156	Portuguese # Top 20 Channels with the	e Mature Stream Tag	Ţ	amers stream are English, Russian, Spanish, German, Korean,
Out[156	<pre>df[df['Mature'] == True][  Channel Watch time()</pre>	['Channel', 'Watch  (Minutes) Stream time(m	time (Minutes)',	
	10TimTheTatman2815MontanaBlack8824	334436990 408460990	108780 14206 67740 18160 115305 10783	67 5265659 00 2911316
	<ul><li>24 TheRealKnossi 18</li><li>32 MOONMOON 15</li></ul>	311696100 527882945	56010 2884! 124680 2489 118995 4584	59 1260160 92 923448
		861024835 252711830	164235 14406 141135 5644	66 746865
		. 50	174270 3674	
	<ul><li>63 Quin69 11</li><li>65 dasMEHDI 11</li><li>70 Vader 11</li></ul>	10952500	174270 3674 231465 4768 184305 1628 109140 3396	42 538532 83 299048 89 424374
	63       Quin69       11         65       dasMEHDI       11         70       Vader       11         71       forsen       11         74       Symfuhny       10         77       Rainbow6       10	10952500 06781045 076179485 031011170	231465 4768 184305 1628	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197
	63       Quin69       11         65       dasMEHDI       11         70       Vader       11         71       forsen       11         74       Symfuhny       10         77       Rainbow6       10         80       Trainwreckstv       10         84       ZeratoR       10         89       AdmiralBulldog       9	10952500 06781045 076179485 031011170 021699920 011013035 072317520	231465 4768 184305 1628 109140 3396 137400 4565 82380 1354 148425 493 84960 2622 154845 1668	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253
	63       Quin69       11         65       dasMEHDI       11         70       Vader       11         71       forsen       11         74       Symfuhny       10         77       Rainbow6       10         80       Trainwreckstv       10         84       ZeratoR       10         89       AdmiralBulldog       9         114       TeePee       7	10952500 06781045 076179485 031011170 021699920 011013035 072317520 789698115 779867430	231465 4768 184305 1628 109140 3396 137400 4567 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951
In [179	63 Quin69 11  65 dasMEHDI 11  70 Vader 11  71 forsen 11  74 Symfuhny 10  77 Rainbow6 10  80 Trainwreckstv 10  84 ZeratoR 10  89 AdmiralBulldog 9  114 TeePee 7  116 Agraelus 7  The table above shows the Top 20  • Mature Tag :- It is for those characteristics.	10952500 106781045 1076179485 1031011170 1021699920 10113035 1072317520 1089698115 1079867430 10 Channels that had the light hannels whose stream contains and the light hannels whose stream	231465 4768 184305 1628 109140 3396 137400 4566 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2358 Mature tag on their	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951  r streams.  ppropriate for younger audience. This doesn't mean sexual
In [179	63 Quin69 11  65 dasMEHDI 11  70 Vader 11  71 forsen 11  74 Symfuhny 10  77 Rainbow6 10  80 Trainwreckstv 10  84 ZeratoR 10  89 AdmiralBulldog 9  114 TeePee 7  116 Agraelus 7  The table above shows the Top 20  • Mature Tag :- It is for those characteristics.	10952500 106781045 1076179485 1031011170 1021699920 1011013035 1072317520 1089698115 1079867430 10 Channels that had the light hannels whose stream contains ages are a counts () ages are a counts () ages ages ages ages ages ages ages ages	231465 4768 184305 1628 109140 3396 137400 4566 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2358 Mature tag on their	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951  r streams.  ppropriate for younger audience. This doesn't mean sexual
	63 Quin69 11  65 dasMEHDI 11  70 Vader 11  71 forsen 11  74 Symfuhny 10  77 Rainbow6 10  80 Trainwreckstv 10  84 ZeratoR 10  89 AdmiralBulldog 9  114 TeePee 7  116 Agraelus 7  The table above shows the Top 20  • Mature Tag :- It is for those change and the standard of t	10952500 106781045 1076179485 1031011170 1021699920 10113035 1072317520 1089698115 1079867430 10 Channels that had the limit hannels whose stream contains and the limit hannels whose stream contains ages 10 counts () 11 counts () 12 counts () 13 counts () 14 counts () 15 counts () 16 counts () 17 counts () 18 counts () 19 counts () 19 counts () 10 counts () 11 counts () 12 counts () 13 counts () 14 counts () 15 counts () 16 counts () 17 counts () 18 counts	231465 4768 184305 1628 109140 3396 137400 4566 82380 13543 148425 4933 84960 26223 154845 1668 170010 7874 169515 2358 Mature tag on their	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951  r streams.  ppropriate for younger audience. This doesn't mean sexual
In [179	63 Quin69 11  65 dasMEHDI 11  70 Vader 11  71 forsen 11  74 Symfuhny 10  77 Rainbow6 10  80 Trainwreckstv 10  84 ZeratoR 10  89 AdmiralBulldog 9  114 TeePee 7  116 Agraelus 7  The table above shows the Top 20  • Mature Tag: It is for those characteristic content or violence activity. E  # Breakup of Top 10 Language   language = df['Language'] language = danguage[:7] lang2 ['Others'] = 120  plt.figure(figsize=(25,8) plt.style.use('Solarize_Language') language   language, labels=language, labels=	10952500 106781045 1076179485 1031011170 1021699920 10113035 1072317520 1089698115 1079867430 10 Channels that had the limit hannels whose stream contains and the limit hannels whose stream contains ages 10 counts () 11 counts () 12 counts () 13 counts () 14 counts () 15 counts () 16 counts () 17 counts () 18 counts () 19 counts () 19 counts () 10 counts () 11 counts () 12 counts () 13 counts () 14 counts () 15 counts () 16 counts () 17 counts () 18 counts	231465 4768 184305 1628 109140 3396 137400 4566 82380 13543 148425 4933 84960 26223 154845 1668 170010 7874 169515 2353 Mature tag on their	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951  r streams.  ppropriate for younger audience. This doesn't mean sexual
	63 Quin69 11  65 dasMEHDI 11  70 Vader 11  71 forsen 11  74 Symfuhny 10  77 Rainbow6 10  80 Trainwreckstv 10  84 ZeratoR 10  89 AdmiralBulldog 9  114 TeePee 7  116 Agraelus 7  The table above shows the Top 20  • Mature Tag: It is for those characteristic content or violence activity. E  # Breakup of Top 10 Language   language = df['Language'] language = danguage[:7] lang2 ['Others'] = 120  plt.figure(figsize=(25,8) plt.style.use('Solarize_Language') language   language, labels=language, labels=	10952500 106781045 1076179485 1031011170 1021699920 10113035 1072317520 1089698115 1079867430 10 Channels that had the limit hannels whose stream contains ages 10 counts () 11 counts () 12 counts () 13 counts () 14 counts () 15 counts () 16 counts () 17 counts () 18 counts () 19 counts () 19 counts () 10 counts () 10 counts () 11 counts () 12 counts () 13 counts () 14 counts () 15 counts () 16 counts () 17 counts () 18 counts () 1	231465 4768 184305 1628 109140 3396 137400 4566 82380 13543 148425 4933 84960 26223 154845 1668 170010 7874 169515 2353 Mature tag on their	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951  r streams.  ppropriate for younger audience. This doesn't mean sexual
	63 Quin69 11  65 dasMEHDI 11  70 Vader 11  71 forsen 11  74 Symfuhny 10  77 Rainbow6 10  80 Trainwreckstv 10  84 ZeratoR 10  89 AdmiralBulldog 9  114 TeePee 7  116 Agraelus 7  The table above shows the Top 20  • Mature Tag: It is for those characteristic content or violence activity. E  # Breakup of Top 10 Language   language = df['Language'] language = danguage[:7] lang2 ['Others'] = 120  plt.figure(figsize=(25,8) plt.style.use('Solarize_Language') language   language, labels=language, labels=	10952500 106781045 1076179485 1031011170 1021699920 10113035 1072317520 1089698115 1079867430 10 Channels that had the limit hannels whose stream contains ages 10 value_counts() 10 sum() 10 ight2') 10 g2.index, autopct=' 11 Languages') 12 of Languages') 13 entage of Languages 14 English 15 English	231465 4768 184305 1628 109140 3396 137400 4566 82380 13543 148425 4933 84960 26223 154845 1668 170010 7874 169515 2353 Mature tag on their	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951  r streams.  ppropriate for younger audience. This doesn't mean sexual
	63 Quin69 11  65 dasMEHDI 11  70 Vader 11  71 forsen 11  74 Symfuhny 10  77 Rainbow6 10  80 Trainwreckstv 10  84 ZeratoR 10  89 AdmiralBulldog 9  114 TeePee 7  116 Agraelus 7  The table above shows the Top 20  • Mature Tag: It is for those change of the stay	10952500 106781045 1076179485 1031011170 1021699920 10113035 1072317520 1089698115 1079867430 10 Channels that had the limit hannels whose stream contains ages 10 value_counts() 10 sum() 10 ight2') 10 g2.index, autopct=' 11 Languages') 12 of Languages') 13 entage of Languages 14 English 15 English	231465 4768 184305 1628 109140 3396 137400 4566 82380 13543 148425 4933 84960 26223 154845 1668 170010 7874 169515 2353 Mature tag on their	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951  r streams.  ppropriate for younger audience. This doesn't mean sexual
	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 9 114 TeePee 7 116 Agraelus 7  The table above shows the Top 20  • Mature Tag :- It is for those changuage = df ['Language'] language = df ['Language'] language = df ['Language'] lang2 = language[7: lang2 = language[7: lang2 = language[7: lang2 = language] language = df ['Language'] 12	10952500 106781045 1076179485 1031011170 1021699920 1011013035 1072317520 1089698115 109867430 10 Channels that had the liminate whose stream contains ages 10 channels whose stream contains ages 11 channels whose stream contains ages 12 channels whose stream contains ages 13 channels whose stream contains ages 14 channels whose stream contains ages 15 channels whose stream contains ages 16 channels whose stream contains ages 17 channels whose stream contains ages 18 channels whose stream contains ages 19 channels whose stream contains ages 10 channels whose stream contains ages 11 channels whose stream contains ages 12 channels whose stream contains ages 13 channels whose stream contains ages 14 channels whose stream contains ages 15 channels whose stream contains ages 16 channels whose stream contains ages 17 channels whose stream contains ages 18 channels whose	231465 4768 184305 1628 109140 3396 137400 4566 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2359 Mature tag on their content may be inapportent or abusive language.	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951  r streams.  ppropriate for younger audience. This doesn't mean sexual
	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 9 114 TeePee 7 116 Agraelus 7 The table above shows the Top 20  • Mature Tag :- It is for those change and the second activity. E  # Breakup of Top 10 Languary language = df['Language'] lang_others = language[7: lang2 = language[7: lang2 = language[7: lang2 = language[7: lang2   language   language = df['Language']  Text(0.5, 1.0, 'Percentage of :  Perce  Perce  Perce	10952500 106781045 1076179485 1031011170 1021699920 101013035 1072317520 189698115 1079867430 10 Channels that had the light hannels whose stream contains ages 10 value_counts() 10 sum() 10 sum() 10 sum() 10 sum() 10 sum() 11 sum() 12 sum() 13 sum() 14 sum() 15 sum() 16 sum() 17 sum() 18 sum() 18 sum() 19 sum() 19 sum() 10 su	231465 4768 184305 1628 109140 3396 137400 4566 82380 13548 148425 4938 84960 26228 154845 1668 170010 7874 169515 2358 Mature tag on their content may be inapplied the formula of the second of the	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951 r streams. Oppropriate for younger audience. This doesn't mean sexual guage.  Others
	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 99 114 TeePee 7 116 Agraelus 7  The table above shows the Top 20  • Mature Tag :- It is for those change of the content or violence activity. E  # Breakup of Top 10 Language   language = df ['Language'] lang_others = language[7: lang2 = language[7: lang2 = language[7: lang2 = language[7: lang2 = language] language   displayed   language   language   language   displayed   language   languag	10952500 106781045 176179485 131011170 121699920 111013035 172317520 189698115 179867430 10 Channels that had the languages  .value_counts()  .sum() 1	231465 4768 184305 1628 109140 3396 137400 4567 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2359 Mature tag on their content may be inappretent or abusive land entent or abusive land	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951 r streams. propriate for younger audience. This doesn't mean sexual guage.  Others
	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 9 114 TeePee 7 116 Agraelus 7  The table above shows the Top 20  • Mature Tag:- It is for those changed and the second of the second	10952500 106781045 1076179485 1031011170 1021699920 101013035 1072317520 1089698115 1079867430  10 Channels that had the light hannels whose stream contages 10 channels counts () 10 channels whose stream contages 10 channels whose stream contag	231465 4768 184305 1628 109140 3398 137400 4568 82380 13548 148425 4938 84960 26228 154845 1668 170010 7874 169515 2358 Mature tag on their content may be inapportent or abusive land element or abus	42 538532 83 299048 89 424374 66 1308165 71 2855063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951 r streams.  propriate for younger audience. This doesn't mean sexual guage.  Others  the Top 1000 channels.  similar share.
	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 99 114 TeePee 77 116 Agraelus 77 The table above shows the Top 20  • Mature Tag :- It is for those change of the content or violence activity. E  # Breakup of Top 10 Langual language = df['Language'] lang_others = language[7: lang2 = language[7: lang2 = language[7: lang2 = language] language = (25, 8) plt.style.use('Solarize_L plt.pie(lang2, labels=language] title('Percentage of item of the popularity of Spanish has spanish)  The above Pie Chart illustrates the part of the popularity of Spanish has spanish.	10952500 106781045 1076179485 1031011170 1021699920 1011013035 1072317520 1089698115 1079867430 10 Channels that had the light hannels whose stream contains ages 10 value_counts() 10 sum() 10 sum() 11 sum() 12 sum() 13 sum() 14 september of Languages stream contains and portugues stream contains and portu	231465 4768 184305 1628 109140 3398 137400 4567 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2359 Mature tag on their content may be inapportent or abusive land element or abus	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951 r streams. propriate for younger audience. This doesn't mean sexual guage.  Others  Others  the Top 1000 channels. similar share. nplay & rubius, as seen before.
Out[179	G3 Quin69 11  G5 dasMEHDI 11  T0 Vader 11  T1 forsen 11  T4 Symfuhny 10  T7 Rainbow6 10  80 Trainwreckstv 10  84 ZeratoR 10  89 AdmiralBulldog 9  114 TeePee 7  116 Agraelus 7  The table above shows the Top 20  • Mature Tag:- It is for those of content or violence activity. E  # Breakup of Top 10 Langua language = df['Language'] lang_others = language[7: lang2 = language] lothers'] = 120  plt.figure (figsize=(25, 8) plt.style.use ('Solarize_L plt.pie (lang2, labels=langult.title ('Percentage of : Text(0.5, 1.0, 'Percentage  Perce  Perce  Text(0.5, 1.0, 'Percentage of : Text(0.5, 1.0,	10952500  06781045  076179485  031011170  021699920  011013035  072317520  089698115  079867430  0 Channels that had the light hannels whose stream contains ages  value_counts()  1. sum()  1. sum()  1. sum()  2. index, autopct='Languages')  2. index, autopct='Languages')  2. index autopct='Languages')  2. index autopct='Languages')  2. index autopct='Languages'  3. sum()  48.5%  48.5%	231465 4768 184305 1628 109140 3398 137400 4567 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2359 Mature tag on their content may be inapportent or abusive land element or abus	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951 r streams.  proprpriate for younger audience. This doesn't mean sexual guage.  Others  the Top 1000 channels.  similar share. inplay & rubius, as seen before.
Out[179	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 9 114 TeePee 7 116 Agraelus 7  The table above shows the Top 20  Mature Tag: It is for those change of the content or violence activity. E  # Breakup of Top 10 Language   language = df ['Language'] lang_others = language[7: lang2 = language]   language   lan	10952500  06781045  076179485  031011170  021699920  011013035  072317520  089698115  079867430  0 Channels that had the light hannels whose stream contains ages  value_counts()  1. sum()  1. sum()  1. sum()  2. index, autopct='Languages')  2. index, autopct='Languages')  2. index autopct='Languages')  2. index autopct='Languages')  2. index autopct='Languages'  3. sum()  48.5%  48.5%	231465 4768 184305 1628 109140 3398 137400 4567 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2359 Mature tag on their content may be inapportent or abusive land element or abus	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951 r streams. propriate for younger audience. This doesn't mean sexual guage.  Others  Others  the Top 1000 channels. similar share. nplay & rubius, as seen before.
Out[179	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 9 114 TeePee 7 116 Agraelus 7  The table above shows the Top 20  Mature Tag: It is for those change of the content or violence activity. E  # Breakup of Top 10 Language   language = df ['Language'] lang_others = language[7: lang2 = language]   language   lan	10952500  06781045  076179485  031011170  021699920  011013035  072317520  089698115  079867430  0 Channels that had the light hannels whose stream contains ages  value_counts()  1. sum()  1. sum()  1. sum()  2. index, autopct='Languages')  2. index, autopct='Languages')  2. index autopct='Languages')  2. index autopct='Languages')  2. index autopct='Languages'  3. sum()  48.5%  48.5%	231465 4768 184305 1628 109140 3398 137400 4567 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2359 Mature tag on their content may be inapportent or abusive land element or abus	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951 r streams.  proprpriate for younger audience. This doesn't mean sexual guage.  Others  the Top 1000 channels.  similar share. inplay & rubius, as seen before.
Out[179	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 99 114 TeePee 7 116 Agraelus 7 The table above shows the Top 20  • Mature Tag:- It is for those of content or violence activity. E  # Breakup of Top 10 Langual language = df['Language'] lang others = language[7: lang2 = language[7: lang2 = language[7: lang2 = language[7: lang2 = language] left.style.use('Solarize_L) plt.pie(lang2, labels=langlt.title('Percentage of :  Text(0.5, 1.0, 'Percentage  Percentage of Percentage of :  Percentage of Percentage of :  Percentage of Percenta	10952500 106781045 1076179485 1031011170 101699920 101013035 1072317520 189698115 1079867430 10 Channels that had the lamanels whose stream contains the stream contains the lamanels whose str	231465 4768 184305 1628 109140 3398 137400 4567 82380 13547 148425 4937 84960 26227 154845 1668 170010 7874 169515 2359 Mature tag on their content may be inapportent or abusive land element or abus	42 538532 83 299048 89 424374 66 1308165 71 2355063 71 1501197 79 728097 73 880728 81 694253 41 520519 55 414951 r streams. propriate for younger audience. This doesn't mean sexual guage.  Others  Others  the Top 1000 channels. similar share. nplay & rubius, as seen before.
Out[179	63 Quin69 11 65 dasMEHDI 11 70 Vader 11 71 forsen 11 74 Symfuhny 10 77 Rainbow6 10 80 Trainwreckstv 10 84 ZeratoR 10 89 AdmiralBulldog 99 114 TeePee 7 116 Agraelus 7 The table above shows the Top 20  • Mature Tag:- It is for those of content or violence activity. E  # Breakup of Top 10 Langual language = df['Language'] lang others = language[7: lang2 = language[7: lang2 = language[7: lang2 = language[7: lang2 = language] left.style.use('Solarize_L) plt.pie(lang2, labels=langlt.title('Percentage of :  Text(0.5, 1.0, 'Percentage  Percentage of Percentage of :  Percentage of Percentage of :  Percentage of Percenta	10952500 106781045 1076179485 1031011170 1021699920 101013035 1072317520 1089698115 1079867430 10 Channels that had the lab hannels whose stream contamels whose	231465 4768 184305 1628 109140 3399 137400 4567 82380 13547 148425 4933 84960 26227 154845 1668 170010 7874 169515 2359 Mature tag on their content may be inapplied to a position of the posi	42 538532 83 299048 83 424374 66 1308165 71 235563 77 1250197 73 880728 81 694253 41 520519 55 414951 r streams.  propriate for younger audience. This doesn't mean sexual guage.  Others  the Top 1000 channels.  similar share. riplay & rubius, as seen before.

• Bits (virtual goods viewers can buy and cheer on a channel) • Ads (Partners earn a share of revenue generated from any ads on their channel) In [185... df2 = df[['Watch Time(Years)','Stream Time(Days)','Average viewers','Peak viewers','Followers']] plt.style.use('Solarize\_Light2') plt.figure(figsize=(20,5)) plt.boxplot(df2['Watch Time(Years)'], vert=False) plt.xlabel('Watch Time in Years', color='Black') Text(0.5, 0, 'Watch Time in Years') Out[185...

2000

plt.boxplot(df2['Stream Time(Days)'], vert=False) plt.xlabel('Stream Time in Days', color='Black')

plt.figure(figsize=(20,5))

plt.style.use('Solarize\_Light2')

Text(0.5, 0, 'Stream Time in Days')

The median value is 75.16 and IQR is 47.28

plt.style.use('Solarize\_Light2')

Text(0.5, 0, 'Average viewers')

The median value is 2425 and IQR is 3328.5

plt.style.use('Solarize\_Light2')

Text(0.5, 0, 'Peak viewership')

100000

The median value is 16676 and IQR is 28456

plt.style.use('Solarize\_Light2')

Out[167... Text(0.5, 0, 'Follower Count(in millions)')

plt.boxplot(df2['Followers'], vert=False)

plt.figure(figsize=(20,5))

1000000

The median value is 318063 and IQR is 453786

200000

plt.xlabel('Follower Count(in millions)', color='Black')

2000000 3000000

plt.boxplot(df2['Peak viewers'], vert=False) plt.xlabel('Peak viewership', color='Black')

plt.figure(figsize=(20,5))

plt.boxplot(df2['Average viewers'], vert=False) plt.xlabel('Average viewers', color='Black')

plt.figure(figsize=(20,5))

In [168...

Out[168...

In [165...

Out[165...

In [164...

Out[164...

In [167...

In [ ]:

4000

8000

6000

Watch Time in Years

Stream Time in Days

Average viewers

Peak viewership

Follower Count(in millions)

500000

6000000

7000000

8000000

600000

700000

The above box plot has so many outliers because most of the top channels are very popular. The median value is 447.09 and IQR is 514.74

10000

12000

# **Conclusion**

All of the different channels used in this study for analysis have something unique to offer to the viewer in terms of entertainment that's why they are in the Top 1000 ranking out of millions of channels. Using the insights generated in our report, we can see how Exploratory Data Analysis is useful and exposes the trends, patterns, and relationships in the data.

In our report, We are able to analyze which channels are the most popular on the website i.e on the basis of Watch Time, Peak viewership or Average viewership. Also, Which channels have gained the most number of followers and unique viewers over a period of 1 year. Moreover, We are able to analyze the spread of users among different language speaking communities.

To conclude, this report gives us a really unique perspective into the interaction of users across different channels.