



Follow

Data Management for Advanced Applications - 16:954:694:01

@Twitter Search Application – Team 08

 Rutgers University

 Presented on April 26

Note: Cohort presentation is completed (Cohort -3)



Team Members:

Anirudh Gaur (216007356)

Prashanth Aripirala (219000924)

Rishik Salver (217001710)

Rohit Macherla (219008045)

Data Summary

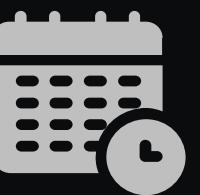
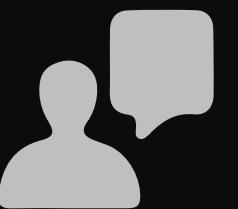


Twitter is one of the most popular social media platforms in the world, with millions of users posting tweets on a daily basis.

The Twitter search application harnesses the power of Twitter's vast database to provide users with a simple yet powerful tool to quickly and easily find the most relevant information on any topic.

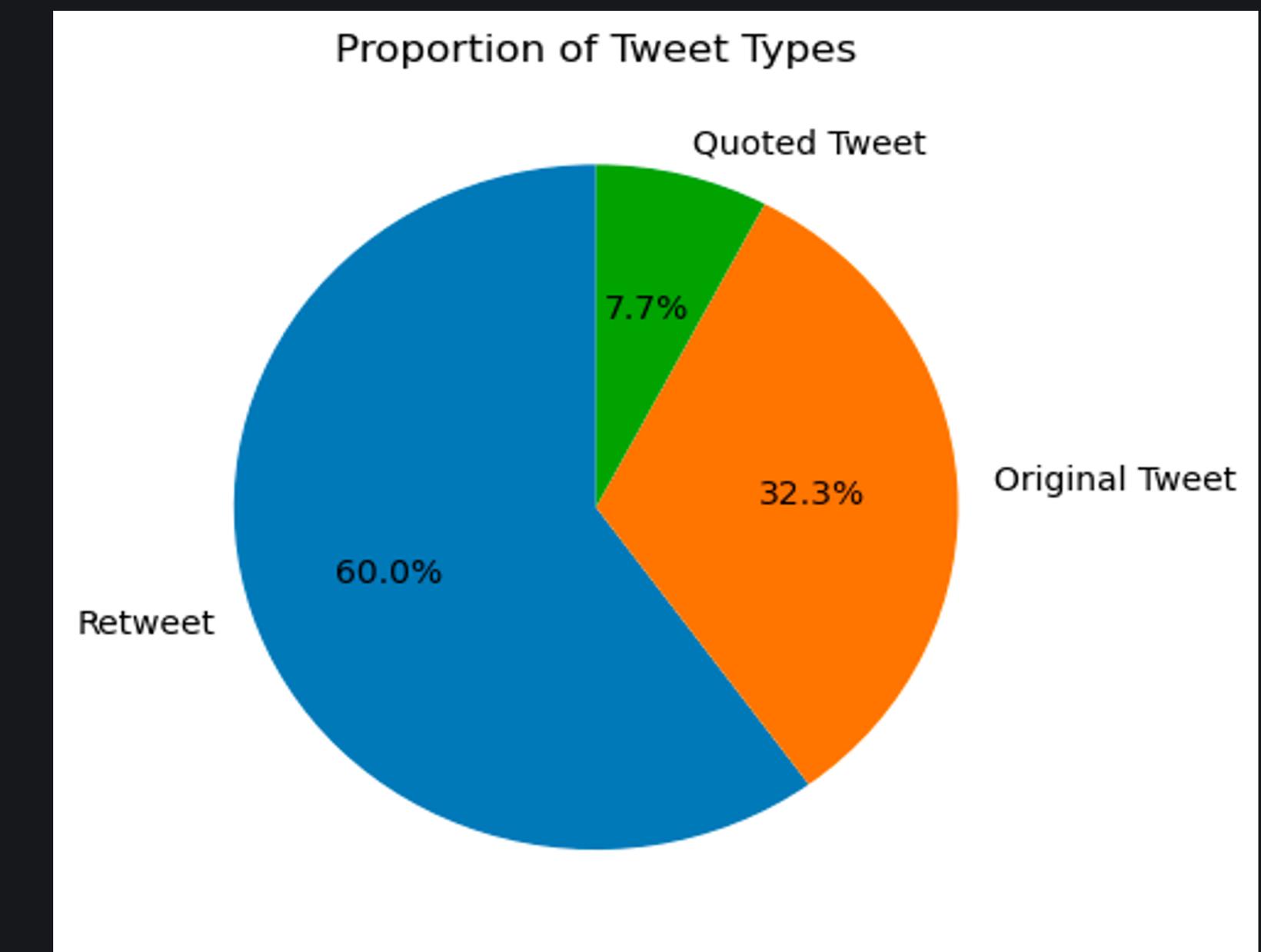
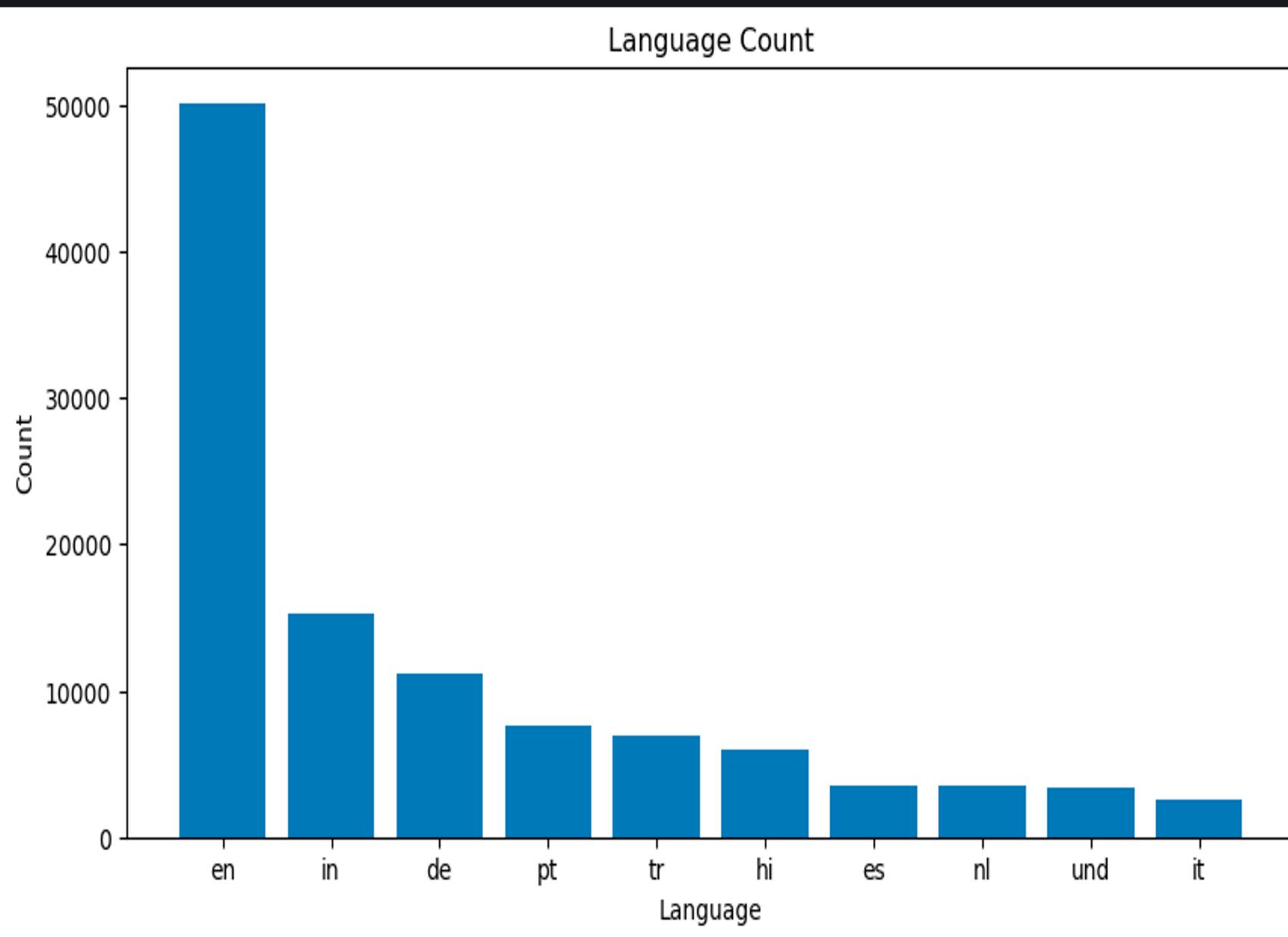
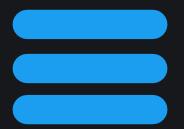


This Twitter search application is designed for fast and efficient data retrieval from large database and cache. From the JSON file, around 106,000 unique users and 134,000 tweets were extracted.



The tweets in the provided JSON file were posted on only 2 different dates in 2020, during the COVID-19 pandemic. As a result, sentiment analysis of the tweets was skewed towards negative due to the global uncertainty and anxiety surrounding the pandemic during that time period.

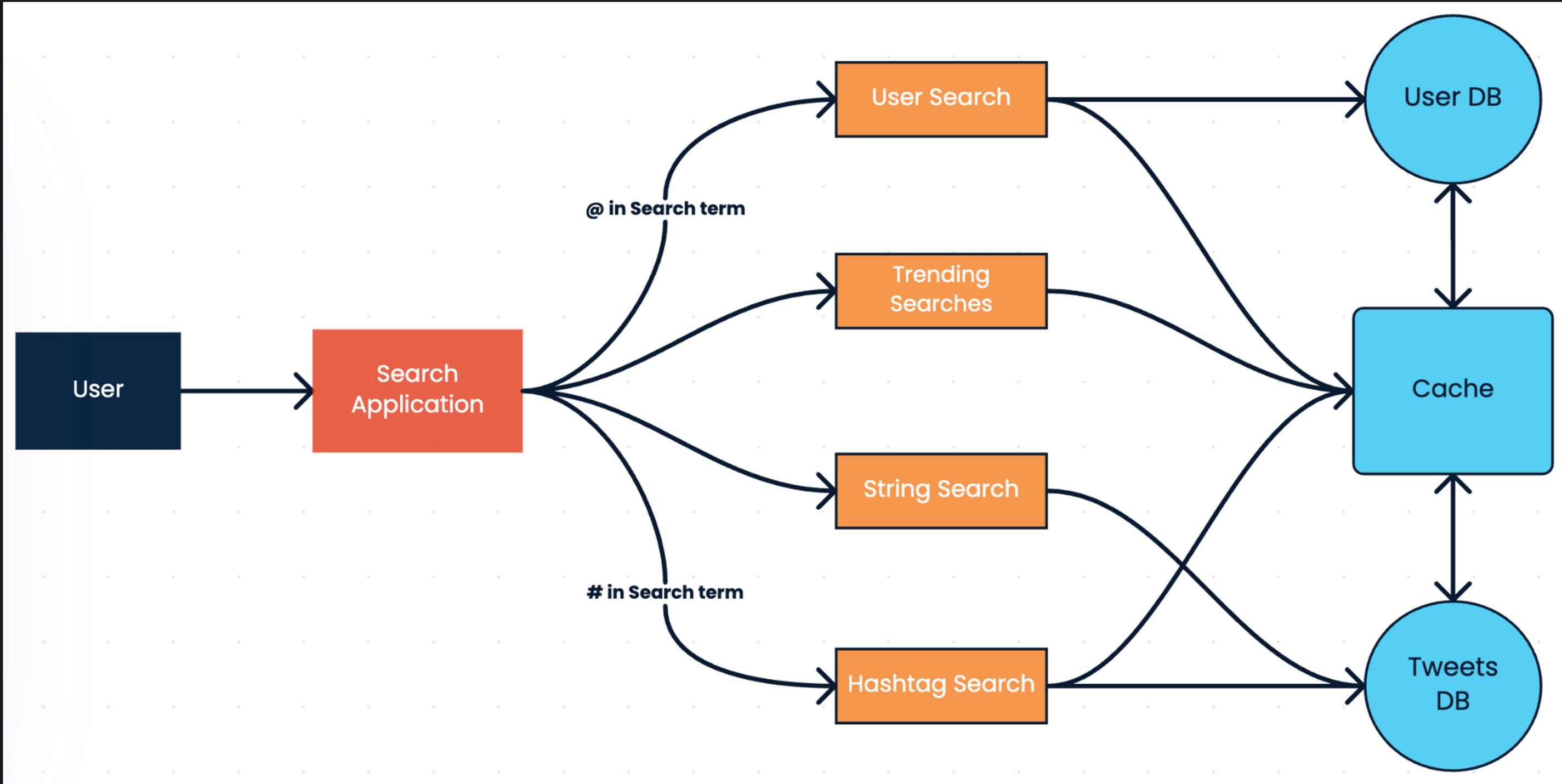
Tweets Analysis



NEXT ➞

SYSTEM
ARCHITECTURE

System Architecture



NEXT ➞

USER DATA LOAD

User Data Load



Local instance 3306
mydatabase.users

Table Details

Engine: **InnoDB**
Row format: **Dynamic**
Column count: **10**
Table rows: **106375**
AVG row length: **301**
Data length: **30.6 MiB**

Column	Type	Default Value	Nullable	Character Set
id	varchar(255)		NO	utf8mb4
name	varchar(255)		YES	utf8mb4
screen_name	varchar(255)		YES	utf8mb4
verified	tinyint(1)		YES	
followers_count	bigint		YES	
friends_count	bigint		YES	
created_at	datetime		YES	
location	varchar(255)		YES	utf8mb4
tweets_count	bigint		YES	
Description	text		YES	utf8mb4

Indexes in Table

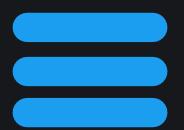
Visible	Key	Type	Uniq...	Columns
<input checked="" type="checkbox"/>	PRIMARY	BTREE	YES	id
<input checked="" type="checkbox"/>	idx_users_name	BTREE	NO	name



NEXT ➞

TWEETS DATA LOAD

Tweets Data Load



Tweets_data

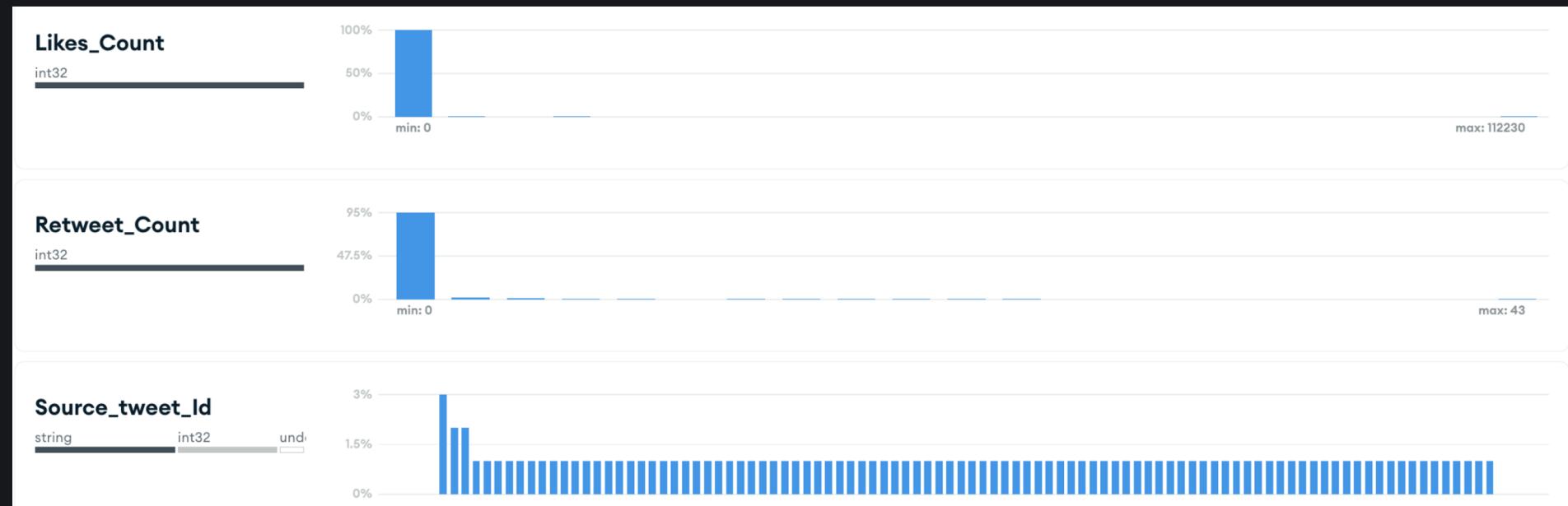
Storage size: 27.66 MB

Documents: 134 K

Avg. document size: 381.00 B

Indexes: 3

Total index size: 33.97 MB



```
_id: ObjectId('64471c7f65ef7286884c8b7b')
created_at: "2020-04-25 09:23:42"
Tweet_Id: "1253977978620518400"
Text: "#25Aprile, il Presidente #Mattarella si è voluto recare all'#Altaredel...
▶ Hashtag: Array
User_Id: "732819391"
User_Name: "Quirinale"
Retweet_Count: 798
Likes_Count: 9524
```



NEXT ➞

UI

UI



TWITTER SEARCH APPLICATION : TEAM 8

Trending Users

Trending Tweets

Trending Hashtags

Enter username, hashtag or part of a string:

Submit



NEXT ➞

SEARCH BY STRING

Search by String



To Support this query,
indexing was done on text
field of all tweets.

Standalone stop words were
not considered for this query

Tweets relevant to "death on"

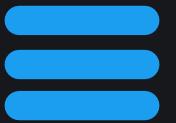
User	Retweets	Likes	Created At	Text	Hashtags
Ahmed Siddique	0	0	2020-04-25 14:47:47	https://t.co/owsm135LeO Watch corona virus patient death on youtube😢 #SaturdayThoughts #KIMJONGUNDEAD... https://t.co/tguyucufZmG	['SaturdayThoughts', 'KIMJONGUNDEAD']
Krishnendu Ghosh	0	0	2020-04-25 14:33:27	Latest news of Bengal, 57 Death in Corona but death only for Corona is 18, I can't understand what its means stand for.#CoronaPoliticsHurts	['CoronaPoliticsHurts']
Michelle Meno	0	0	2020-04-25 12:35:46	Truth traveling for work everyday.China to Italy.. Maybe we should count all the good people of Italy's death on Ch... https://t.co/ms9wrSMwo0	[]
nat❤️	0	0	2020-04-12 18:29:17	maybe it's just me , but I ain't ever wish death on nobody	[]
juwelz v	0	0	2020-04-12 18:27:25	RT @nuffsaidny: wishing death on people is weirdo behavior.	[]



NEXT ➞

SEARCH BY @USERNAME

Search by @Username



Search Results were limited to 5.

User Search Results

#	Name	Handle	Verified	Followers	Tweets	Description	Location	Creation Date
1	Rohit Kansal	kansalrohit69	✗	118385	0	Civil Servant IAS; Pr Secy Planning J&K LSE, IIM Kolkata; RTs not endorsement	Jammu, Jammu And Kashmir	2016-05-13 13:43:06
2	Rohitt Jaiswal	rohitjswl01	✓	38921	1	Film Critic Trade Analyst Film Marketing Consultant, For Campaign collaborations & film marketing related query mail me ↪cinematicfreedom@gmail.com	Kolkata, India	2014-04-28 03:47:09
3	Pinky Rajpurohit (ABP News) 🇮🇳	Madrassan_Pinky	✗	23026	1	Marwari chorи brought up in Madras. South India Correspondent with @ABPNews. Focus @DefenceMinIndia @ISRO @DRDO_India and Politics. Ex @NewsNationTV @ANI	Bengaluru, India	2012-10-02 06:24:54
4	Dr.Sanjeev Rajpurohit	DrSanjeevRajp4	✗	12545	5	Assistant Director (Research) J.R.Nagar Rajasthan Vidyapeeth University, Spokesperson, District Congress, Udaipur. RT≠ Endorsement Tweets=ByLifeExperiences	Udaipur, India	2015-08-14 16:09:43
5	Mersal Rohit வி	Mersal_Rohit	✗	8586	1	VIJAY My Silence Is My Attitude Photography♥ Crazy on editing😊 மத்தவங்கள் வேதனை படுத்துற ஒரு சின்ன ஸ்மைல் கூட தப்ப தான்! Enjoy Every Moment @ImagixCreations♥	vijay's ❤	2017-06-21 17:20:01

Enter the user number whose tweets you want to be displayed:

4

Submit



NEXT ➞

USER TWEETS



Tweets for each User were limited to 3

User Tweets

User: Dr.Sanjeev Rajpurohit

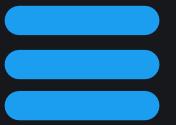
#	Date	Text	Hashtags	Retweets	Likes
1	2020-04-25 14:19:43	RT @ashokgehlot51: A separate ward has been set up at the Mahila Chikitsalaya, Sanganeri Gate, #Jaipur for pregnant women coming from #coro...	['Jaipur']	0	0
2	2020-04-25 13:08:30	@RahulGandhi : Problem is you are building your #CentralVista project at same time. So you not sort of doing just t... https://t.co/7IjWN4aIe3	['CentralVista']	1	0
3	2020-04-25 12:23:19	RT @The_ManishSood: Want to know what we've done wrong in this country under modi govt about Corona pandemic.. Watch this video explaining...	-	0	0



NEXT ►►

SEARCH BY #HASHTAG

Search by #Hashtag



Hashtags are used to categorize and organize tweets around a particular topic, event, or theme.

For this use case, the hashtags were obtained by relevance to the search term and the number of times appeared.

Top 5 Hashtags Matching Your Search for "#corona"

Hashtag	Tweets Count
Corona	5920
corona	1928
coronavirus	825
COVID19	572
Covid_19	503

Enter the hashtag related to which the relevant tweets you want to be displayed:

Submit



NEXT ➞

HASHTAG TWEETS



Tweets for each Hashtag were limited to 3

Tweets of #Covid_19

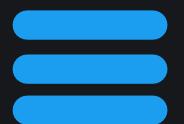
Username	Retweets	Likes	Created At	Text	Hashtags
HARPAL SINGH	0	0	2020-04-25 14:48:25	#Covid_19 #Who_Is_Real_SadGuru #Why_Need_TrueWorship the benefits of true Bhakti which availed by Mirabai / Guru... https://t.co/EPsazvXEOK	['Covid_19', 'Who_Is_Real_SadGuru', 'Why_Need_TrueWorship']
SWR Sport	0	0	2020-04-25 14:48:24	Traurige Nachrichten aus Mannheim. @svw07 #corona #Covid_19 https://t.co/6JRZRimlZG	['corona', 'Covid_19']
Divyanshu Nidhi	0	0	2020-04-25 14:48:22	RT @pranitasubhash: Gratitude 🙏 Preparation and packing of meals #Corona #Covid_19 https://t.co/LvlUlracSy	['Corona', 'Covid_19']



NEXT ➞

TRENDING TWEETS

Trending Tweets



The Tweets were ranked based on a calculated field involving Retweets and Likes.

Higher weightage(60%) was given to retweets as they contribute directly to the spread of tweets compared to likes which was given 40%.

Trending Tweets

User	Retweets	Likes	Created At	Text	Hashtags
Quirinale	798	9524	2020-04-25 09:23:42	#25Aprile, il Presidente #Mattarella si è voluto recare all'#AltaredellaPatria dove ha deposto una corona d'alloro... https://t.co/ev4b4DlqDf	[25Aprile, 'Mattarella', 'AltaredellaPatria']
dr. Shela Putri Sundawa	784	5498	2020-04-25 05:24:38	Ngomongin teori konspirasi corona sama orang yg percaya kalo bumi itu datar, I'm not a smart people, but please dud... https://t.co/aCIWBZg41m	[]
Quirinale	654	4062	2020-04-25 09:56:54	#25Aprile, nel 75° anniversario della #Liberazione il Presidente #Mattarella ha deposto una corona all'Altare della... https://t.co/vDcVcwEMQy	[25Aprile, 'Liberazione', 'Mattarella']
Brit Hume	1474	1831	2020-04-25 11:27:12	And where is the evidence that Covid 19 is easily spread outdoors? https://t.co/YPcJXU1uqw	[]
Only In Russia	586	1352	2020-04-25 11:26:44	Corona disinfecting in Russia. https://t.co/AEF5ccDmM3	[]
♣i§aç♣	445	1532	2020-04-23 19:58:13	Quando eu ligo a televisão e fica falando só de corona vírus	[]
Ben Wikler	986	0	2020-04-25 12:51:03	Milwaukee's health commissioner has now tied 40 coronavirus infections to the April 7 election. https://t.co/fGIzLKzTkm	[]
Funda	732	0	2020-04-25 12:43:26	Gözün çıkışın corona 😊 Ülkece asabii oldu! 🙏 muhtemel psikoloji ektedir 😂😂👉 #PideAlmayaDiyeÇıkıp https://t.co/KoGSbVAMxZ	[PideAlmayaDiyeÇıkıp']
Laura Ingraham	500	0	2020-04-25 13:53:37	A MUST READ...Coronavirus Restrictions: Government Bears the Burden of Proof Before Denying Freedoms National Rev... https://t.co/RcaAK9nwDs	[]
Gu Ru Thalaiva	422	0	2020-04-25 13:01:20	Thalapathy fans from Sivakasi Helped the Poor Family who are affected by this corona Crisis ! They have supplied th... https://t.co/ozoG6d9Ax8	[]



NEXT ➞

TRENDING USERS,
HASHTAGS

Trending User and Hashtags



Trending Users

Name	Handle	Verified	Followers	Following	Location	Tweets	Bio
Barack Obama	BarackObama	Yes	116518121	607194	Washington, DC	1	Dad, husband, President, citizen.
Donald J. Trump	realDonaldTrump	Yes	78467254	46	Washington, DC	0	45th President of the United States of America
CNN Breaking News	cnnbrk	Yes	57529057	120	Everywhere	0	Breaking news from CNN Digital. Now 56M strong. Check @cnn for all things CNN, breaking and more. Download the app for custom alerts: http://cnn.com/apps
Narendra Modi	narendramodi	Yes	55781248	2364	India	6	Prime Minister of India
Shakira	shakira	Yes	52250613	212	Barranquilla	0	ME GUSTA Shakira & Anuel AA Nuevo Sencillo / New Single
CNN	CNN	Yes	47567385	1106	None	0	It's our job to #GoThere & tell the most difficult stories. Join us! For more breaking news updates follow @CNNSERIALIZED & Download our app http://cnn.com/apps
The New York Times	nytimes	Yes	46359985	904	New York City	1	News tips? Share them here: http://nyti.ms/2FVHq9v
BBC Breaking News	BBCBreaking	Yes	43014510	3	London, UK	0	Breaking news alerts and updates from the BBC. For news, features, analysis follow @BBCWorld (international) or @BBCNews (UK). Latest sport news @BBCSport.
Amitabh Bachchan	SrBachchan	Yes	41596464	1833	Mumbai, India	1	"तुमने हमें पूज पूज कर पत्थर कर डाला ; वे जो हमपर जुमले कसते हैं हमें ज़िंदा तो समझते हैं " ~ हरिवंश राय बच्चन
Salman Khan	BeingSalmanKhan	Yes	40094611	26	MUMBAI	0	Film actor, artist, painter, humanitarian

Trending Hashtags

Hashtag	Tweets Count
Corona	5920
corona	1928
Mattarella	1516
25Aprile	1476
Covid_19	1119
COVID19	1003
coronavirus	825
AltaredellaPatria	806
PideAlmayaDiyeÇıkıp	777
Liberazione	700



NEXT ➞

CACHE IMPLEMENTATION



Cache Implementation

Caching is an essential technique that helps in improving the performance of applications by storing frequently accessed data in a temporary storage area. This temporary storage area, known as cache, enables faster access to that data in the future, which reduces the number of requests made to the database or other external resources. The result is significant improvements in application performance, and a reduction in the load on the application and database servers.

The cache class uses a least-accessed eviction strategy to ensure that frequently accessed keys are retained in the cache. This helps optimize the cache's memory usage and ensures that it can store a large amount of data within the defined limit of 10 MB.



The cache class includes a checkpoint interval of 5 minutes, which allows the cache to regularly save its contents to disk. This ensures that in case of any failures, the cache can quickly recover its contents and continue serving data to users without any disruptions.

The cache class provides a TTL feature that allows the cache to automatically remove stale data from the cache. With a TTL of one week, any data that hasn't been accessed within a week will be automatically removed from the cache, ensuring that only the latest and relevant data is stored.

Cache Performance



Search by @username

From Database

```
*****  
Search term @jimmy  
Time for getting user info: 0.09813242599999938  
Time for getting tweets info: 0.22909174800000187
```

From Cache

```
*****  
Search term @jimmy  
Time for getting user info: 0.00022843799999350267  
Time for getting tweets info: 0.0004839059999994788
```

Search by #hashtag

From Database

```
*****  
Search Term #run  
Time for getting tweets info: 0.347066746000003
```

From Cache

```
*****  
Search Term #run  
Time for getting tweets info: 0.0006296309999953564
```

Incorporating cache into the application significantly enhances the speed of data retrieval, surpassing traditional database fetching by over 500 times.



NEXT ➞

CONCLUSION



Conclusion

- The Twitter search app offers multiple search options, including username, hashtag, and string, providing users with a versatile and efficient tool to retrieve relevant information from Twitter's vast database of tweets.
- With the ability to retrieve and store data in cache, the app performs fast data retrieval times, improving user experience and efficiency.
- The app includes additional features such as the ability to retrieve trending tweets, users, and hashtags with a single click, giving users quick access to the most relevant and popular content on the platform.
- Overall, the search app is a powerful tool that enables users to easily search and retrieve valuable information from Twitter, making it an invaluable resource for businesses, journalists, researchers, and anyone looking to stay informed and up-to-date on the latest trends and conversations on the platform.