



# Scrapping **TikTok**



Ratu Risha Ulfia  
G1501231041



Follow Us +



# OUTLINE



1

Tiktok

2

Collecting Data

3

DATA PROCESSING

4

Analyzing

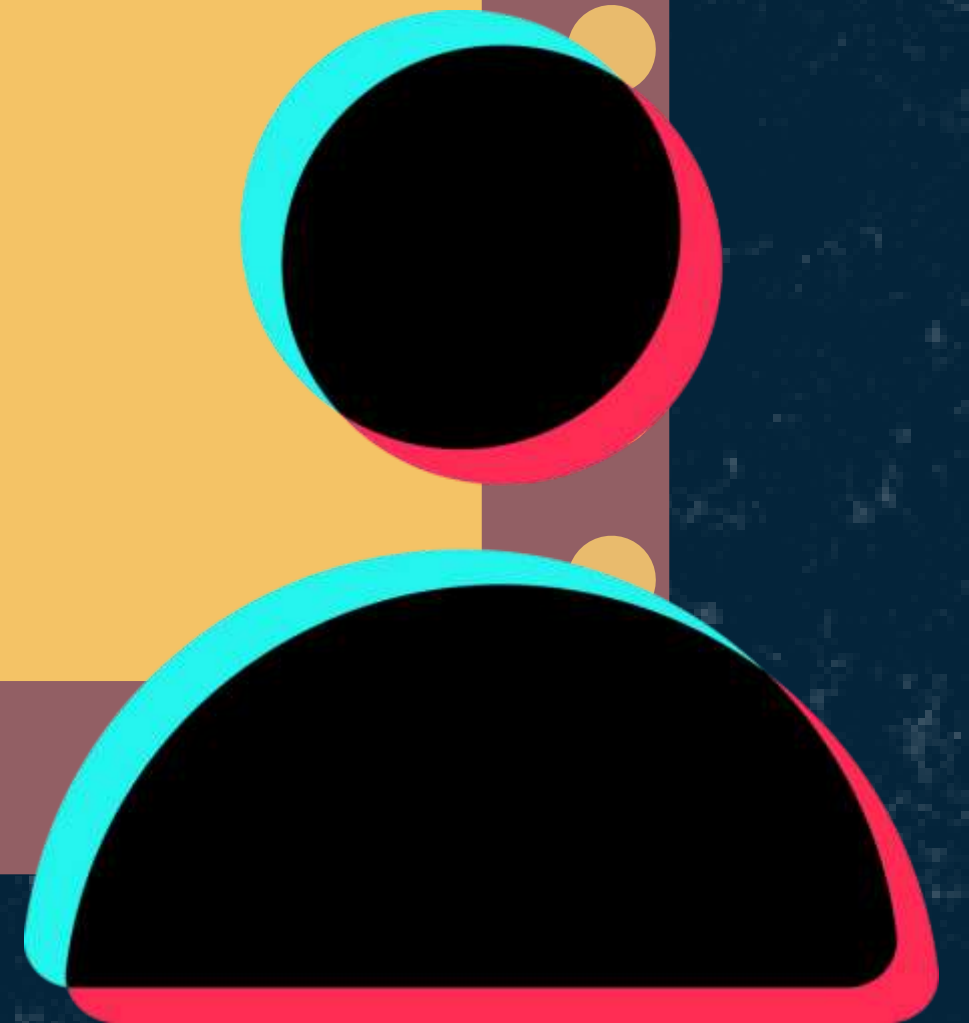
5

Conclusion

4

Insight

# About **TIKTOK**





# DESCRIPTION

**HypeAuditor TikTok** adalah sebuah platform web yang menyediakan analitik dan manajemen kampanye untuk pemasaran influencer di TikTok. Platform ini dirancang untuk membantu merek, pemasar, dan agensi untuk menemukan, menganalisis, dan mengelola influencer TikTok dengan lebih efektif.





# **COLLECTING DATA**

# COLLETING DATA

1

Informasi yang akan di scrapping pada website

<https://hypeauditor.com/top-tiktok/>



2

MongoDB adalah perangkat lunak yang digunakan untuk mengelola dan menyimpan database hasil scraping yang dilakukan.



3

Digunakan untuk menulis, mengembangkan, dan menjalankan skrip R yang melakukan scraping, pembersihan, dan analisis data dari hypeauditor Tiktok.



4

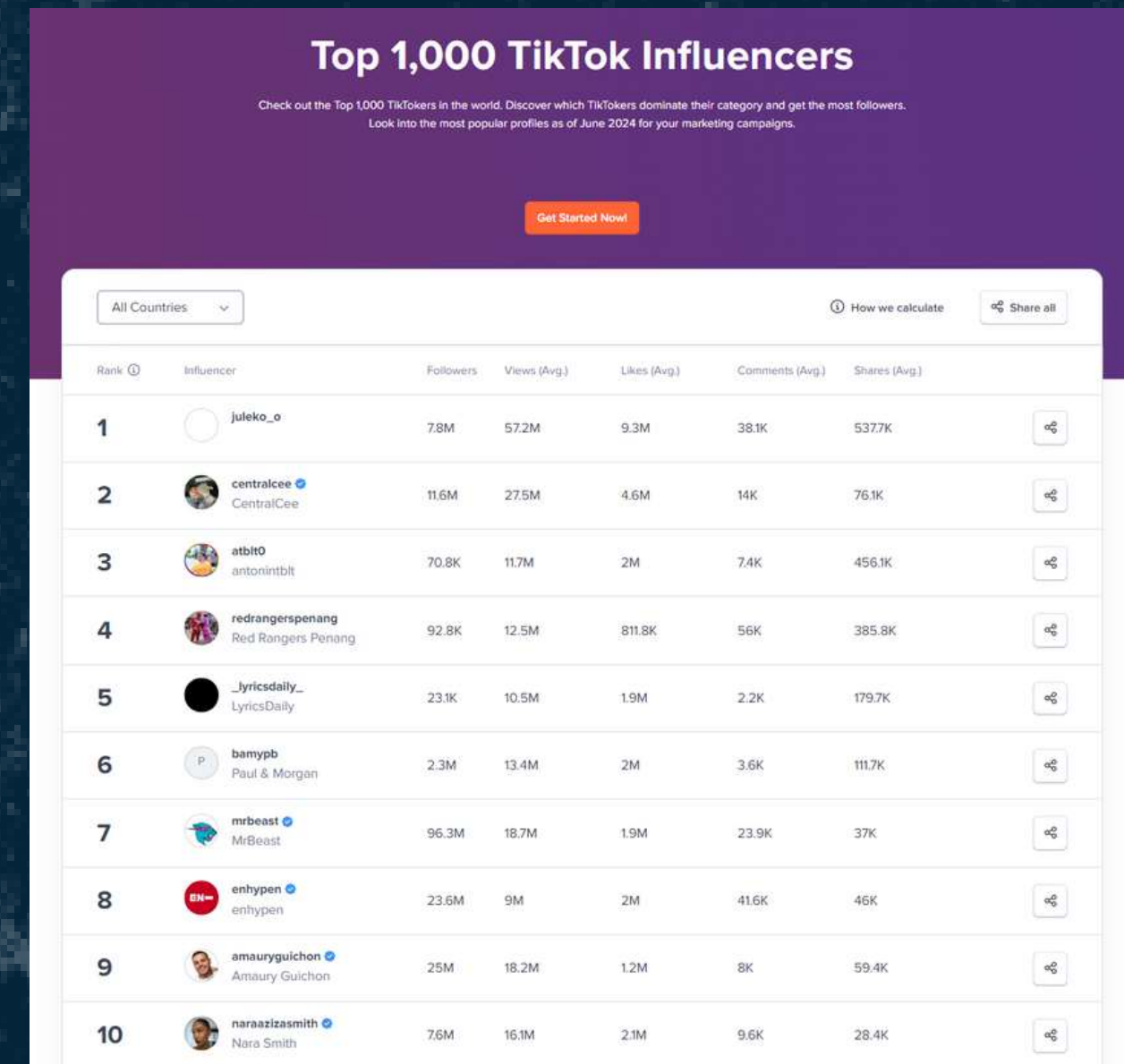
Digunakan untuk mengotomatisasi scraping data setiap hari setiap jam 01.00 WIB dengan menjalankan skrip R secara terjadwal.



# SCRAPPING DATA

## Data yang di Scraping sebagai berikut:

- **\_id:** ID unik yang mungkin dihasilkan oleh MongoDB untuk setiap dokumen dalam koleksi. ID ini digunakan untuk mengidentifikasi setiap entri secara unik dalam database.
- **comments:** Jumlah komentar yang diterima oleh influencer pada TikTok. Angka ini menunjukkan tingkat interaksi pengguna terhadap konten yang diposting oleh influencer.
- **followers:** Jumlah pengikut (followers) yang dimiliki oleh setiap influencer di TikTok. Ini menunjukkan popularitas dan jangkauan audiens dari influencer tersebut.
- **influencer:** Nama pengguna (username) dari influencer TikTok. Ini mengidentifikasi siapa influencer tersebut.
- **likes:** Jumlah likes yang diterima oleh konten yang diposting oleh influencer. Ini adalah indikator lain dari tingkat interaksi dan popularitas konten mereka.
- **rank:** Peringkat influencer dalam suatu kategori atau berdasarkan metrik tertentu. Peringkat ini bisa berdasarkan jumlah pengikut, interaksi, atau metrik lainnya.
- **shares:** Jumlah kali konten dari influencer dibagikan oleh pengguna lain. Ini menunjukkan sejauh mana konten tersebut dianggap menarik dan layak dibagikan oleh audiens.
- **time\_scraped:** Tanggal dan waktu ketika data tersebut diambil (scraped). Ini membantu dalam mengetahui seberapa terbaru data yang digunakan.
- **view:** Jumlah tampilan (views) yang diterima oleh konten influencer. Ini menunjukkan seberapa banyak orang yang telah melihat konten mereka.



**Top 1,000 TikTok Influencers**

Check out the Top 1,000 TikTok influencers in the world. Discover which TikTok influencers dominate their category and get the most followers. Look into the most popular profiles as of June 2024 for your marketing campaigns.

[Get Started Now!](#)

Rank	Influencer	Followers	Views (Avg.)	Likes (Avg.)	Comments (Avg.)	Shares (Avg.)
1	juleko_o	7.8M	57.2M	9.3M	38.1K	537.7K
2	centralcee CentralCee	11.6M	27.5M	4.6M	14K	76.1K
3	atbit0 antonintbit	70.8K	11.7M	2M	7.4K	456.1K
4	redrangerspenang Red Rangers Penang	92.8K	12.5M	811.8K	56K	385.8K
5	_lyricsdaily_ LyricsDaily	23.1K	10.5M	1.9M	2.2K	179.7K
6	bamypb Paul & Morgan	2.3M	13.4M	2M	3.6K	111.7K
7	mrbeast MrBeast	96.3M	18.7M	1.9M	23.9K	37K
8	enhypen enhypen	23.6M	9M	2M	41.6K	46K
9	amauryguichon Amaury Guichon	25M	18.2M	1.2M	8K	59.4K
10	naraazizasmith Nara Smith	7.6M	16.1M	2.1M	9.6K	28.4K

# HASIL SCRAPING

MongoDB Compass - PrakMDS/scrape\_tiktok.tiktok

Connect Edit View Help

**PrakMDS** Documents scrape\_tiktok.tikt...

My Queries Databases Search

- admin
- config
- local
- sample\_airbnb
- sample\_analytics
- sample\_geospatial
- sample\_guides
- sample\_mflix
- sample\_restaurants
- sample\_supplies
- sample\_training
- sample\_weatherdata
- scrape\_tiktok
- tiktok**

scrape\_tiktok.tiktok 210 DOCUMENTS 1 INDEXES

Documents Aggregations Schema Explain Plan Indexes Validation

Filter Type a query: { field: 'value' } Reset Find More Options

ADD DATA EXPORT COLLECTION 1 - 20 of 210

```
{
  "_id": ObjectId('66581dc55f54c8360e011b72'),
  "time_scraped": "2024-05-30T06:33:38.334+00:00",
  "rank": "2",
  "influencer": "janneksp!ace",
  "followers": "4.2M",
  "view": "44M",
  "likes": "3.7M",
  "comments": "35.6K",
  "shares": "341.3K"
}
```

```
{
  "_id": ObjectId('66581dc55f54c8360e011b71'),
  "time_scraped": "2024-05-30T06:33:38.334+00:00",
  "rank": "1",
  "influencer": "nikolaisavic",
  "followers": "3M",
  "view": "84.3M",
  "likes": "11.7M",
  "comments": "26.1K",
  "shares": "299K"
}
```





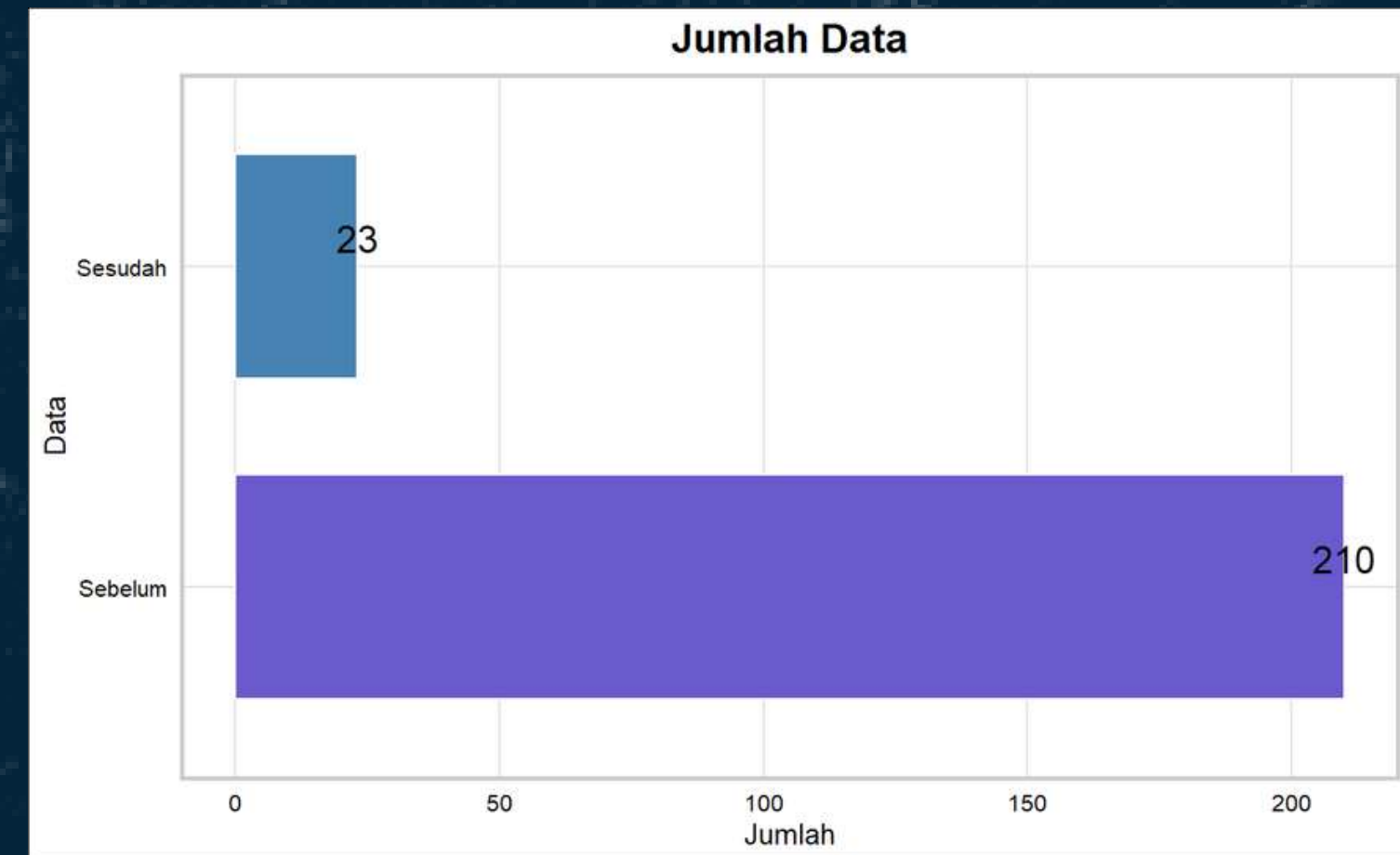
# **DATA PROCESSING**

# DATA PROCESSING



## Menghapus data duplikat

Tujuan menghapus data duplikat adalah untuk memastikan analisis dan visualisasi yang dilakukan didasarkan pada data yang bersih, akurat, dan dapat diandalkan. Hal ini penting agar hasil analisis memberikan wawasan yang valid dan bermanfaat. Menghilangkan duplikasi juga meningkatkan efisiensi pemrosesan data dan mengurangi kesalahan dalam interpretasi.



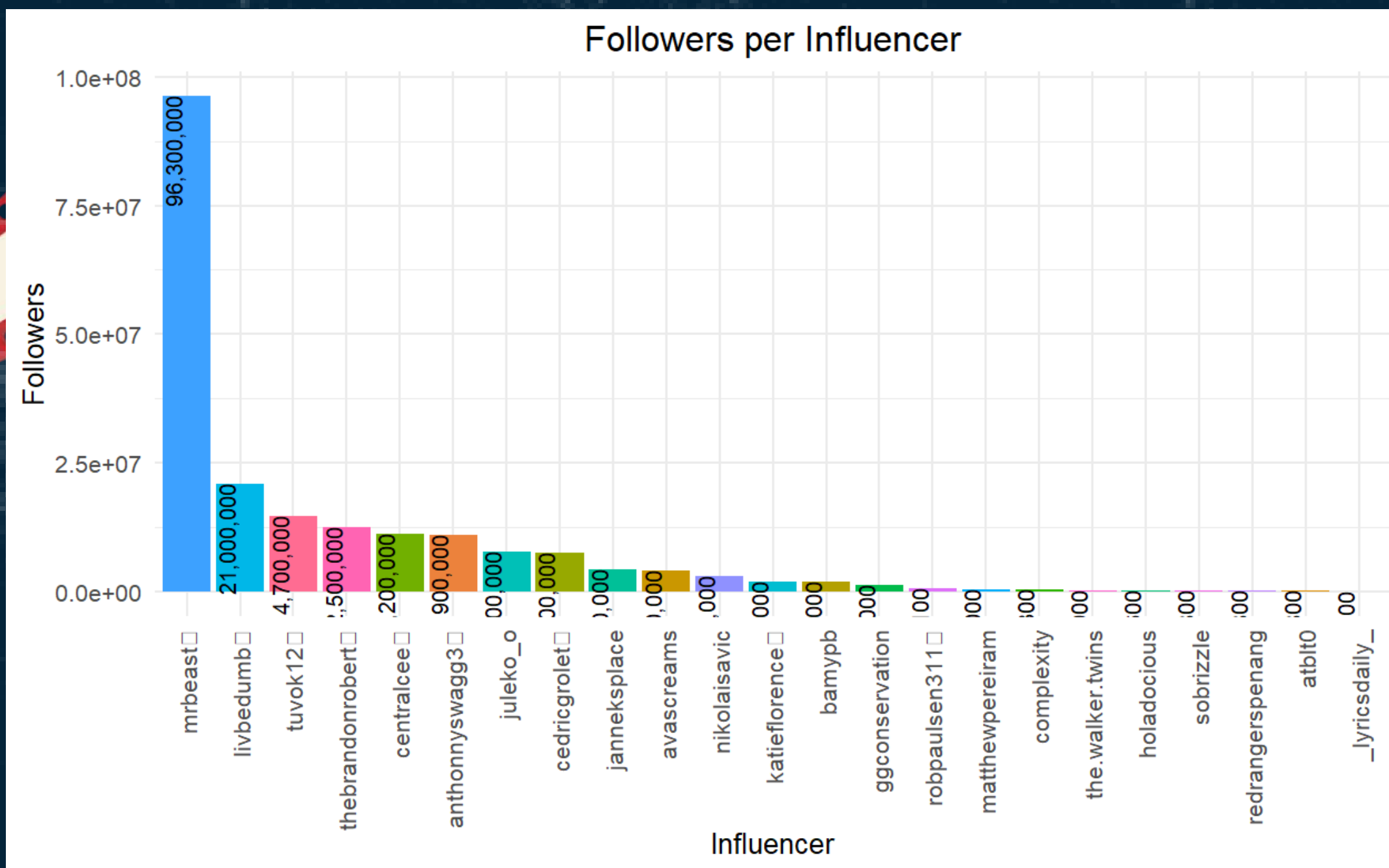
Data yang awalnya berjumlah 210 setelah dilakukan penghapusan data duplikat menjadi 23 data. Hal ini terjadi karena banyak data yang terambil sama setiap harinya karena tidak adanya update per hari.



# **ANALYZING DATA**



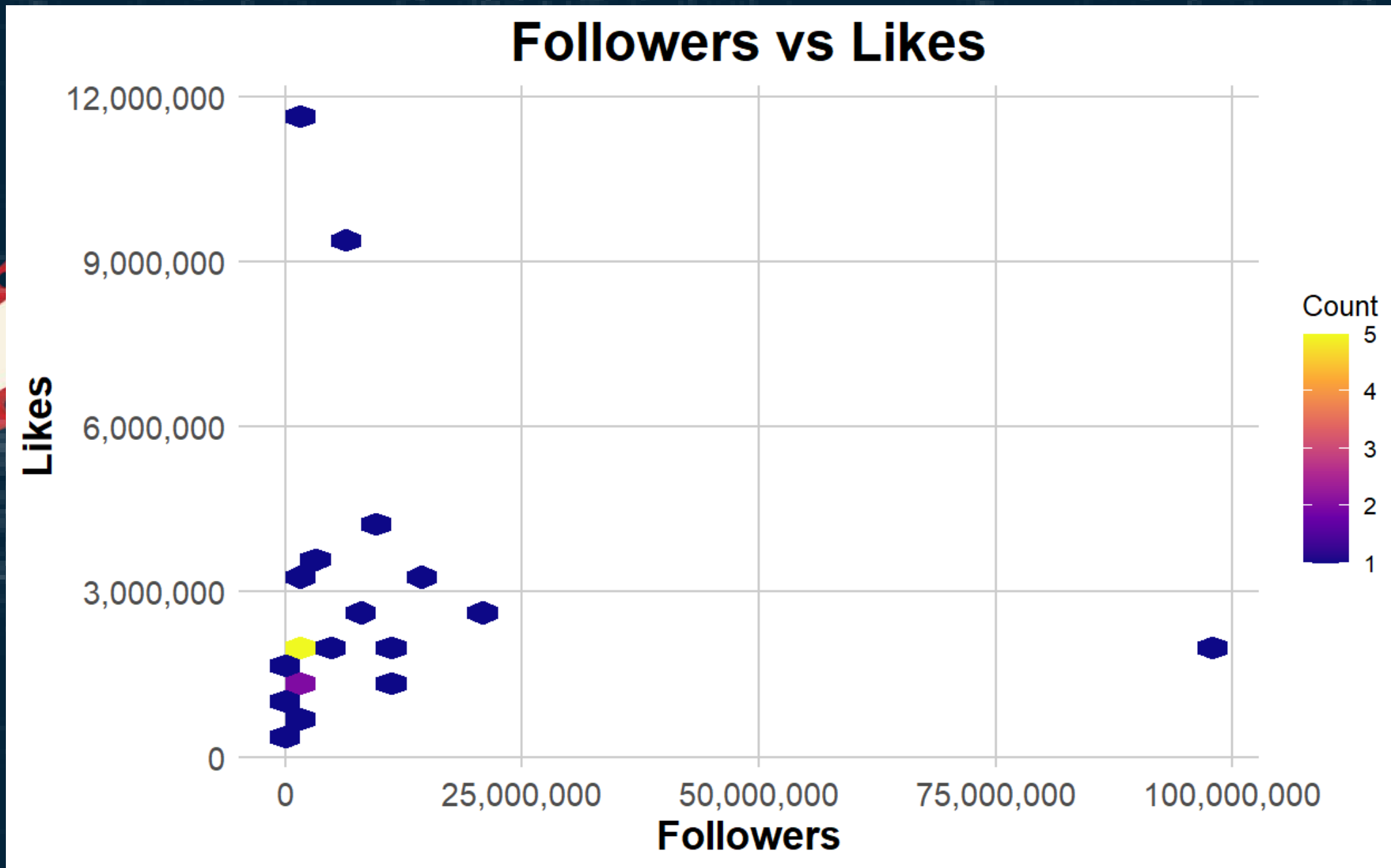
# FOLLOWERS PER INFLUENCER



Visualisasi ini menunjukkan bahwa **MrBeast adalah influencer yang dominan di TikTok berdasarkan jumlah pengikut**. Terdapat beberapa influencer lainnya yang juga sangat populer, namun dengan jumlah pengikut yang jauh lebih sedikit. Ini memberikan wawasan tentang siapa saja yang memiliki pengaruh besar di platform dan bagaimana pengikut tersebar di antara berbagai influencer.

FOLLOW

# FOLLOWERS VS LIKES

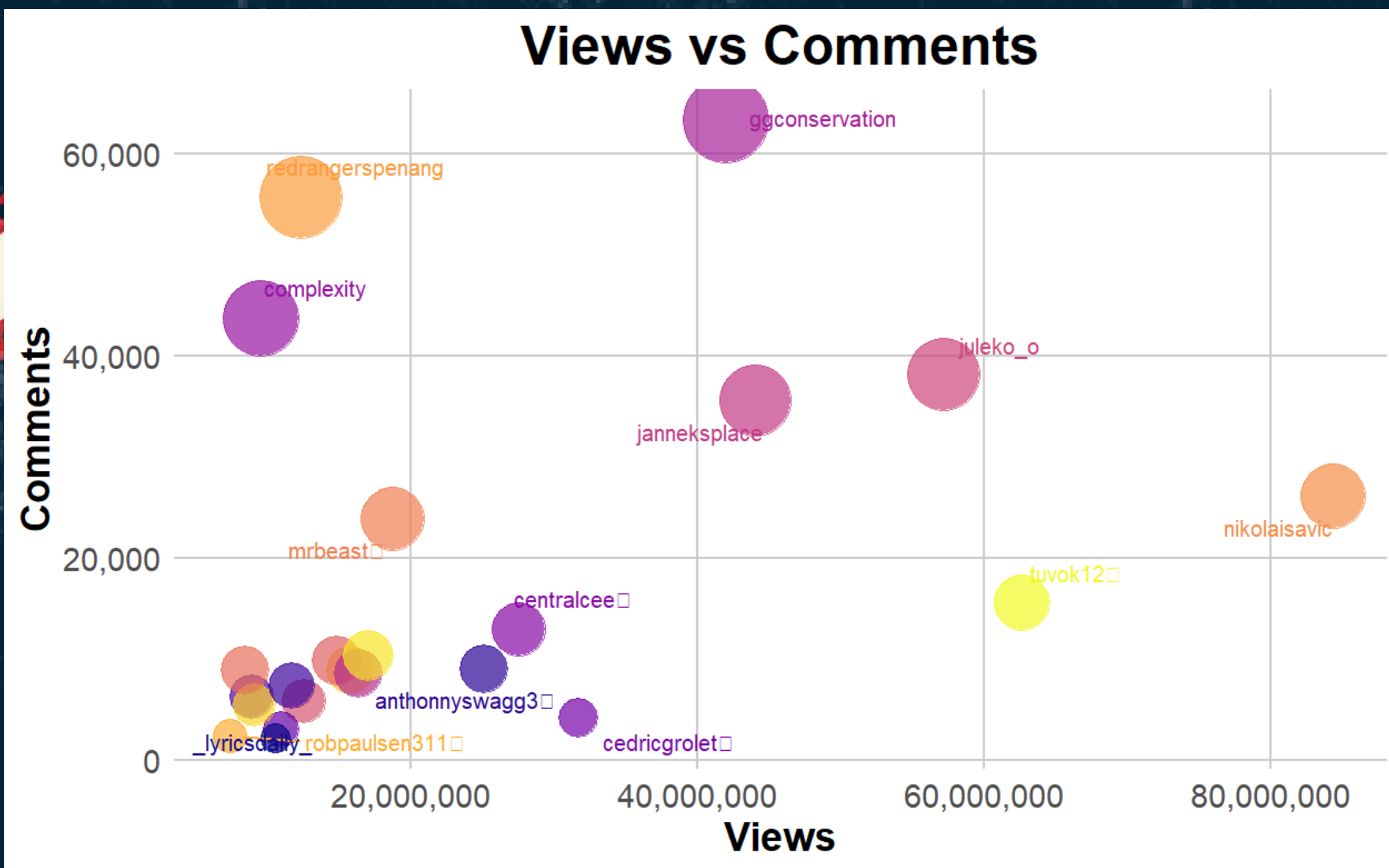


Visualisasi ini mengungkapkan bahwa sebagian besar influencer TikTok memiliki jumlah pengikut dan likes yang relatif rendah, dengan beberapa outliers yang memiliki jangkauan dan interaksi yang sangat tinggi. Ada korelasi positif antara jumlah pengikut dan likes, menunjukkan bahwa **peningkatan pengikut biasanya diikuti oleh peningkatan likes**. Influencer dengan pengikut yang sangat tinggi, seperti MrBeast, sangat menonjol dan menunjukkan perbedaan signifikan dalam popularitas dan keterlibatan dibandingkan dengan yang lainnya.





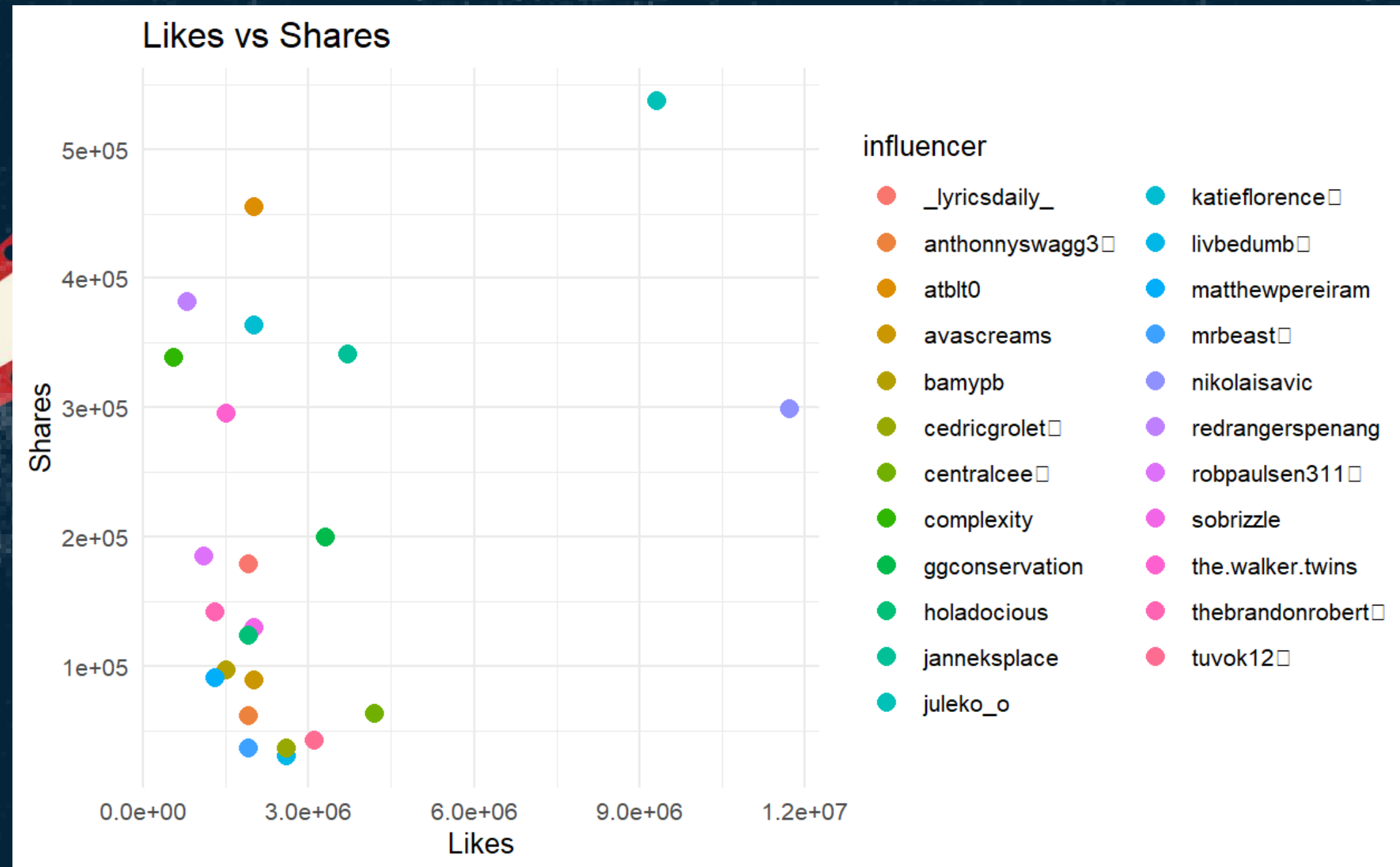
# Views vs Comments



- Ada korelasi positif antara jumlah tampilan dan jumlah komentar: influencer dengan lebih banyak tampilan cenderung mendapatkan lebih banyak komentar.
- Beberapa influencer menonjol dengan tingkat interaksi yang sangat tinggi, sementara sebagian besar influencer memiliki tingkat interaksi yang lebih rendah.
- Variasi dalam jumlah komentar di antara influencer dengan jumlah tampilan yang sama menunjukkan perbedaan dalam kualitas atau jenis konten yang mereka bagikan.



# LIKES VS SHARES



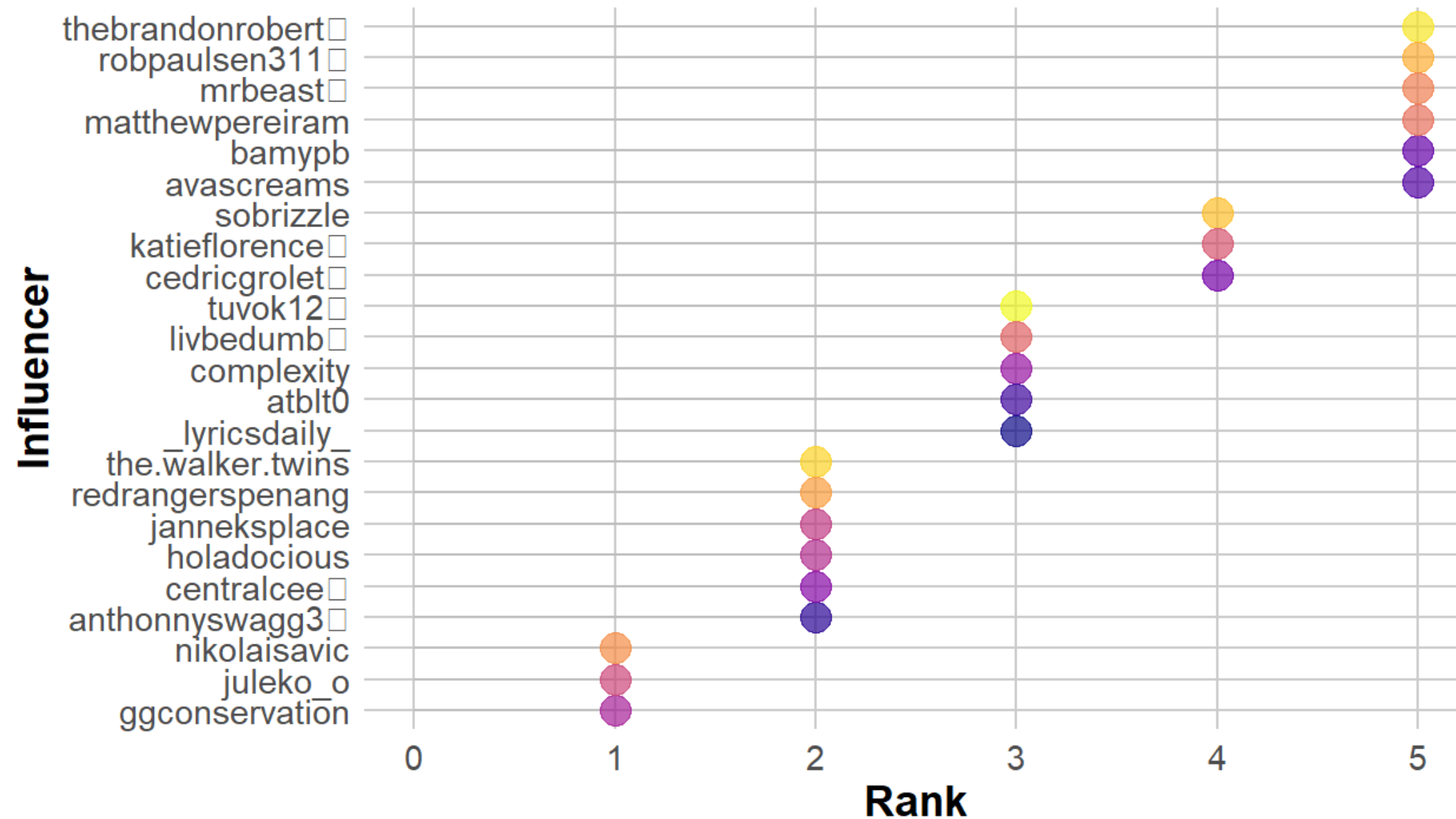
- Popularitas dan Keterlibatan: Influencer seperti Janneksplace dan Nikolaisavic menunjukkan bahwa mereka tidak hanya memiliki banyak pengikut, tetapi juga mampu menghasilkan tingkat keterlibatan yang tinggi melalui likes dan shares.
- Korelasi Antara Likes dan Shares: Ada korelasi positif antara jumlah likes dan shares, tetapi variasi di antara influencer menunjukkan bahwa jenis konten dan cara audiens berinteraksi dapat sangat bervariasi.
- Distribusi Tidak Merata: Sebagian besar influencer memiliki jumlah likes dan shares yang lebih rendah, dengan hanya beberapa yang menonjol dengan jumlah interaksi yang sangat tinggi.



# RANK INFLUENCER



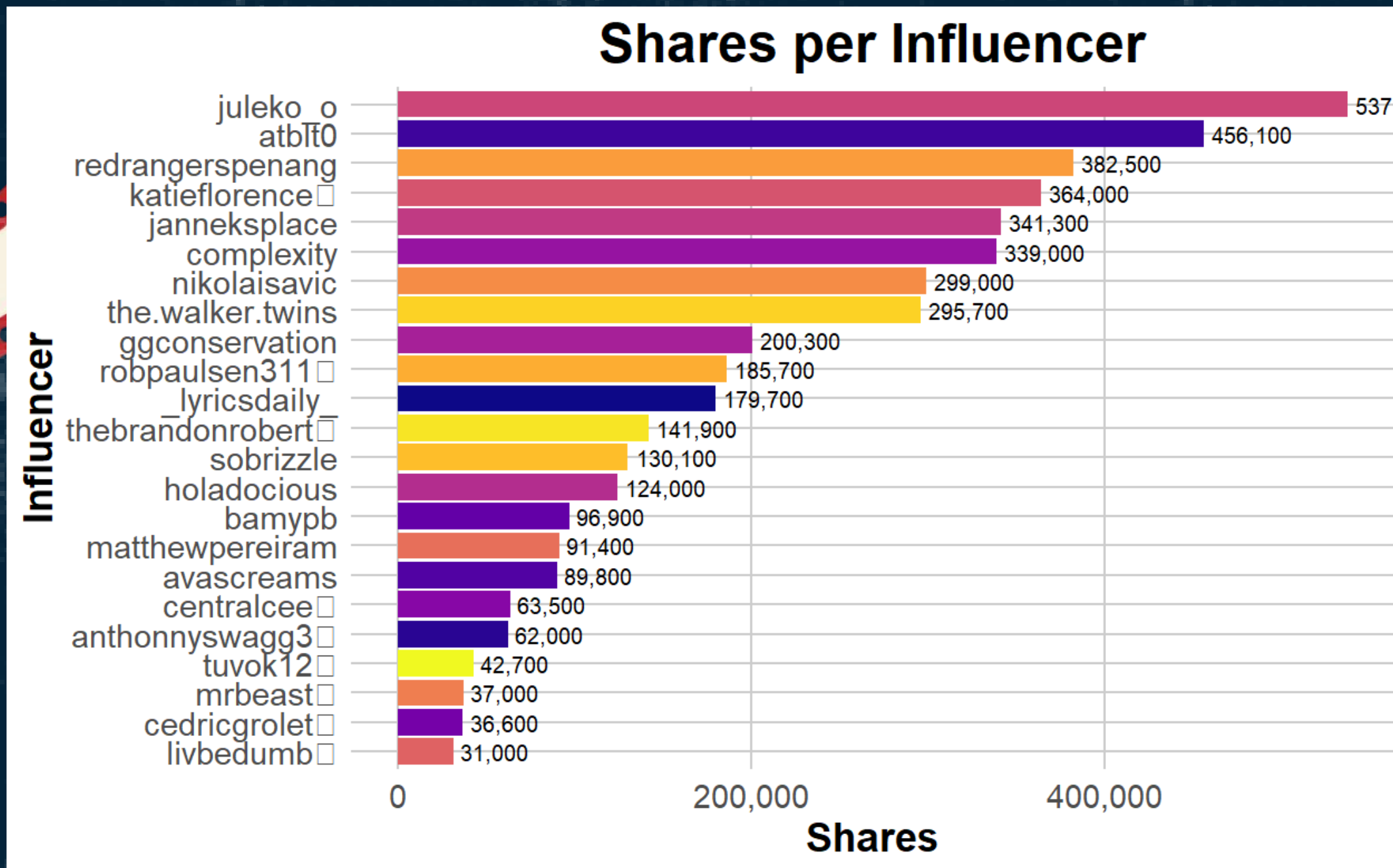
Rank of Each Influencer



Visualisasi ini memberikan gambaran yang komprehensif mengenai distribusi peringkat influencer di TikTok. Dari peringkat 1 hingga 5, kita dapat melihat bagaimana influencer **terdistribusi berdasarkan tingkat pengaruh** mereka. Influencer di peringkat 1 adalah yang paling berpengaruh, sementara peringkat lainnya menunjukkan variasi yang signifikan dalam hal popularitas dan keterlibatan.



# JUMLAH SHARES UNTUK SETIAP INFLUENCER



- Jumlah shares yang tinggi menunjukkan bahwa konten dari influencer seperti Juleko\_o dan Atblt0 sangat menarik dan cenderung dibagikan oleh pengguna, menandakan tingkat keterlibatan yang sangat tinggi.
- Visualisasi ini memberikan wawasan yang jelas tentang bagaimana **konten dari berbagai influencer TikTok diterima dan dibagikan oleh audiens** mereka, menyoroti siapa yang paling berpengaruh dalam hal keterlibatan dan shares.





# CONCLUSION





# CONCLUSION

Dari serangkaian visualisasi ini, kita dapat menyimpulkan bahwa terdapat variasi yang signifikan dalam popularitas dan tingkat keterlibatan di antara influencer TikTok. Beberapa influencer, seperti MrBeast dan Juleko\_o, menonjol dengan jumlah pengikut, likes, views, dan shares yang sangat tinggi, menunjukkan pengaruh dan popularitas yang luar biasa. Namun, sebagian besar influencer memiliki tingkat keterlibatan yang lebih rendah. Korelasi positif antara berbagai metrik (pengikut, likes, views, comments, shares) menunjukkan bahwa influencer yang lebih populer cenderung mendapatkan lebih banyak interaksi di semua aspek. Distribusi yang tidak merata ini menyoroti perbedaan dalam daya tarik dan pengaruh konten di platform TikTok.





**INSIGHT**



# INSIGHT

## Insight Dalam Dunia Bisnis:

### 1. Strategi Pemasaran:

Merek dapat memanfaatkan data ini untuk memilih influencer dengan tingkat keterlibatan tinggi untuk kampanye mereka, memastikan bahwa mereka mendapatkan interaksi maksimal untuk setiap konten yang dipromosikan.

### 2. Pengembangan Konten:

Influencer dapat menggunakan insight ini untuk mengembangkan strategi konten yang lebih efektif, fokus pada pembuatan konten yang tidak hanya menarik pengikut tetapi juga mendorong interaksi tinggi seperti likes dan shares.

### 3. Analisis Kompetitif:

Merek dan influencer dapat menggunakan data ini untuk menganalisis kompetisi dan mengidentifikasi area di mana mereka dapat meningkatkan keterlibatan atau menjangkau audiens baru.

### 4. Penargetan Audiens:

Dengan memahami pola interaksi audiens, merek dapat lebih baik menargetkan audiens yang paling mungkin berinteraksi dengan konten mereka, meningkatkan efisiensi kampanye pemasaran mereka.





# THANK YOU

Saya sangat tertarik dengan data dan sedang menempuh Pascasarjana di IPB University. Berpengalaman menggunakan berbagai software statistik seperti R Studio, Python, minitab dan lainnya untuk menganalisis data.

Konten ini dibuat untuk memenuhi tugas Praktikum Manajemen Data Statistika dan juga sebagai sumber belajar bagi siapa saja yang tertarik. Selamat belajar dan semoga bermanfaat!



Raturisha



ratu.rishaa



Ratu Risha Ulfia