

DSA210 Term Project

Fall24-25

Tiktok Engagement Analysis vs Personal Posting

Seher Azem

30511

Introduction	4
Motivation	4
Data Overview	4
1. Engagement Data	4
2. Personal Posts	4
3. Additional Insights	5
Methodology	5
1. Data Cleaning and Preparation:	5
2. Analysis:	5
3. Visualization:	5
Visualization and Analysis of my Tiktok Data	6
Browsing Data	6
Daily TikTok Browsing Activity	6
Weekly TikTok Browsing Activity	7
TikTok Browsing Activity by Day of the Week	8
Hourly TikTok Browsing Activity	9
TikTok Likes Trends	9
Monthly TikTok Likes Trends	9
TikTok Likes by Days of the Week	11
Analysis of Hashtags and Interests in Liked Videos	12
Primary Interests	12
Secondary Interests	13
Word Cloud Insights	13
Favorite Sounds Analysis	14
Analysis of Favorite Sounds Activity by Year and Month:	14
Favorite Sounds Activity by Year	15
Aggregated Favorite Sounds Activity by Month	15
Favorite Sounds Activity by Day of the Week	17
Analysis of Posts	18
Number of Posts Per Year	18
Posts Visibility Per Year	19
Distribution of Likes Per Post by Year and Top Sounds by Median Likes	20
Top Sounds by Median Likes:	20
Posts by Hour of Day vs Activity Hours	22
Hourly Posting Activity by Day of the Week:	23

Followers Vs Following Analysis:-----	24
Hypothesis Testing -----	25
Conclusion -----	28
Limitations and Future Work -----	28

Introduction

This project explores how my engagement on TikTok – including browsing history, likes, favorites, sounds, and follows and followers– shapes my personal posting behavior on the platform. The main hypothesis of my project is that the themes, styles, and creative decisions in my TikTok posts are directly influenced by the content I engage with. By analyzing patterns in engagement and personal posting habits, the final goal is to uncover correlations and derive actionable insights.

Motivation

This study investigates how social media algorithms and personal preferences impact creative output by identifying trends in engagement habits, exploring overlaps between consumed and created content, and providing recommendations to refine personal content strategies. Another motivation is just so I can personally see how online I really am and how influenced I can get from social media consumption. I chose Tiktok because it is my most used social media platform app and I spend the most time on it.

Data Overview

1. Engagement Data

- **Browsing Data:** Tracks the videos watched, providing insights into content consumption habits.
- **Liked Videos:** Highlights trends and preferences in the content that resonates most.
- **Favorite Sounds:** Includes saved audio tracks, reflecting recurring audio themes and popular trends.
- **Hashtags:** Captures recurring tags used or engaged with, showing thematic preferences.

2. Personal Posts

- **Posted Content:** Metadata about captions, themes, hashtags, and sounds used in my posts.

3. Additional Insights

- **Favorite Videos:** Includes deeper preferences as indicated by saved favorites.
- **Profile Data:** Insights from following patterns and follower engagement metrics.
- **Analysis of TikTok Browsing Activity**

This dataset provides a comprehensive foundation for analyzing how engagement habits translate into creative output, focusing on the interplay between consumed and created content.

Methodology

1. Data Cleaning and Preparation:

- Duplicates were removed.
- Metadata was standardized for consistency.

2. Analysis:

- Trends in engagement data were explored using Python scripts.
- Personal posts were categorized by themes, sound usage, and creative styles.
- Comparisons were made between engagement data and posting behavior.
- Hypothesis testing was conducted using Chi-squared tests to evaluate the relationships between engagement and posting patterns.

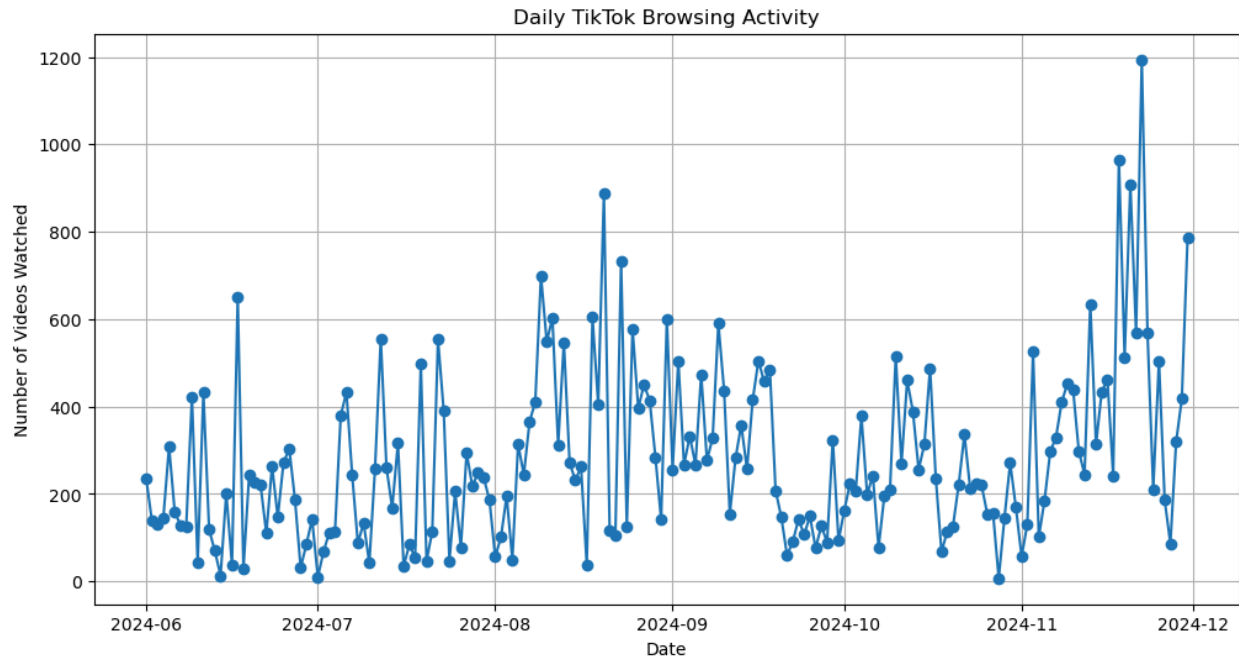
3. Visualization:

- Graphs were created to illustrate trends and relationships.

Visualization and Analysis of my Tiktok Data

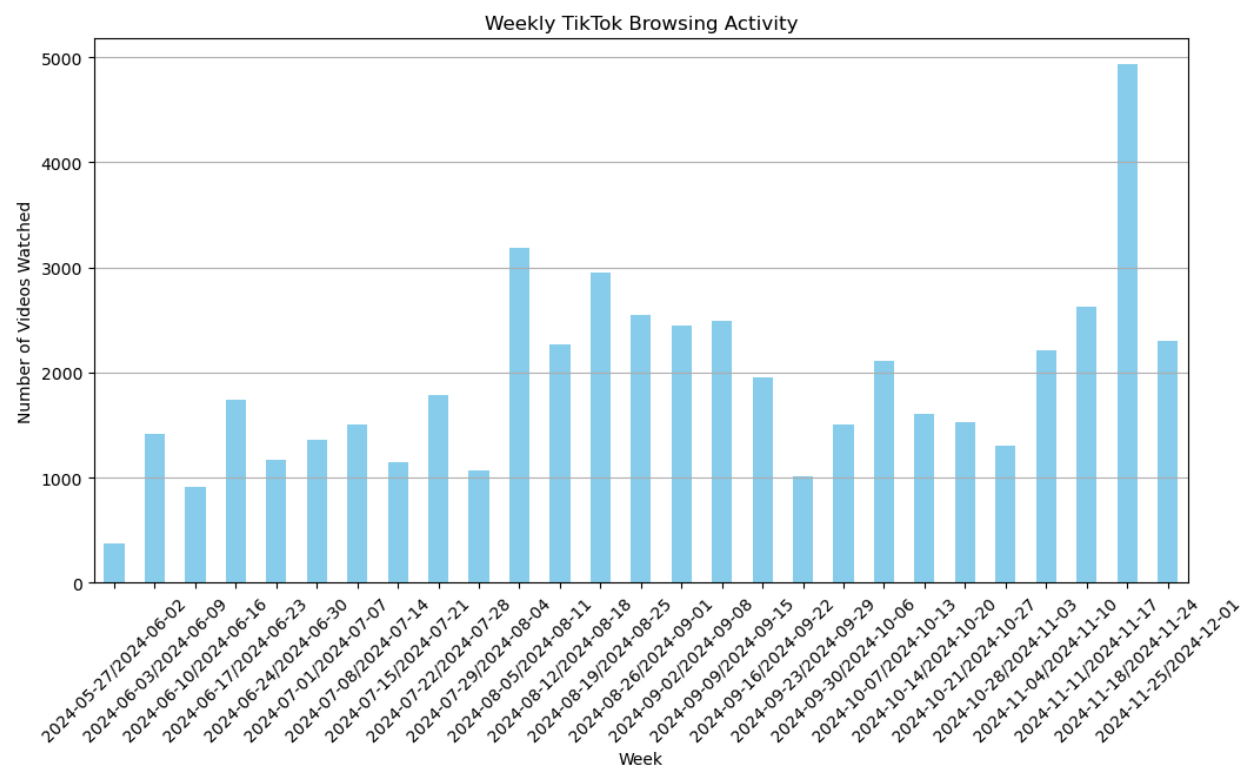
Browsing Data

Daily TikTok Browsing Activity



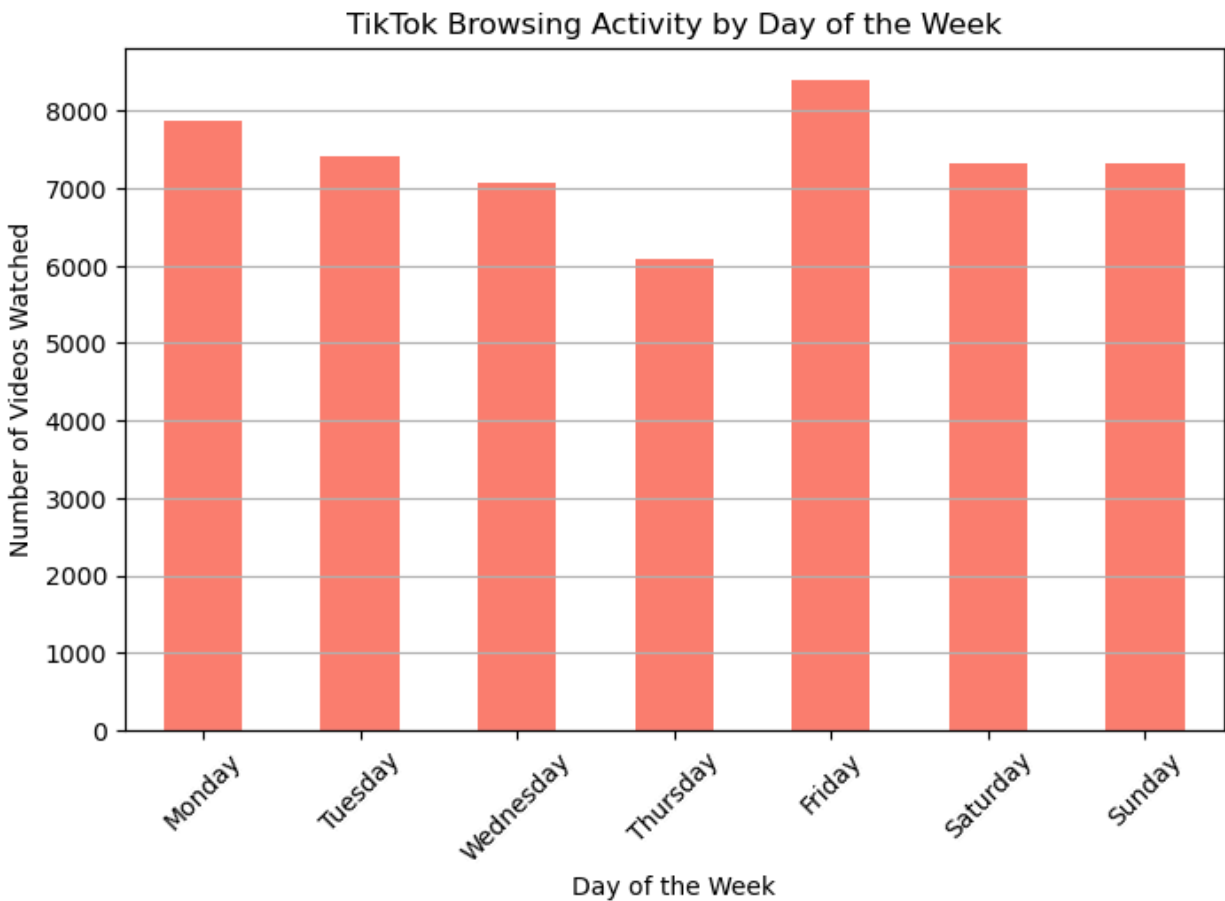
The figure shows fluctuating daily TikTok browsing activity, with notable spikes around late November and early December, suggesting periods of higher engagement. Activity levels generally increase over time, indicating a growing usage pattern. Drops in browsing may align with specific off-platform events or breaks. Peaks could correspond to weekends, holidays, or viral trends capturing attention. So because of this I decided to look at my weekly browsing activity trends

Weekly TikTok Browsing Activity



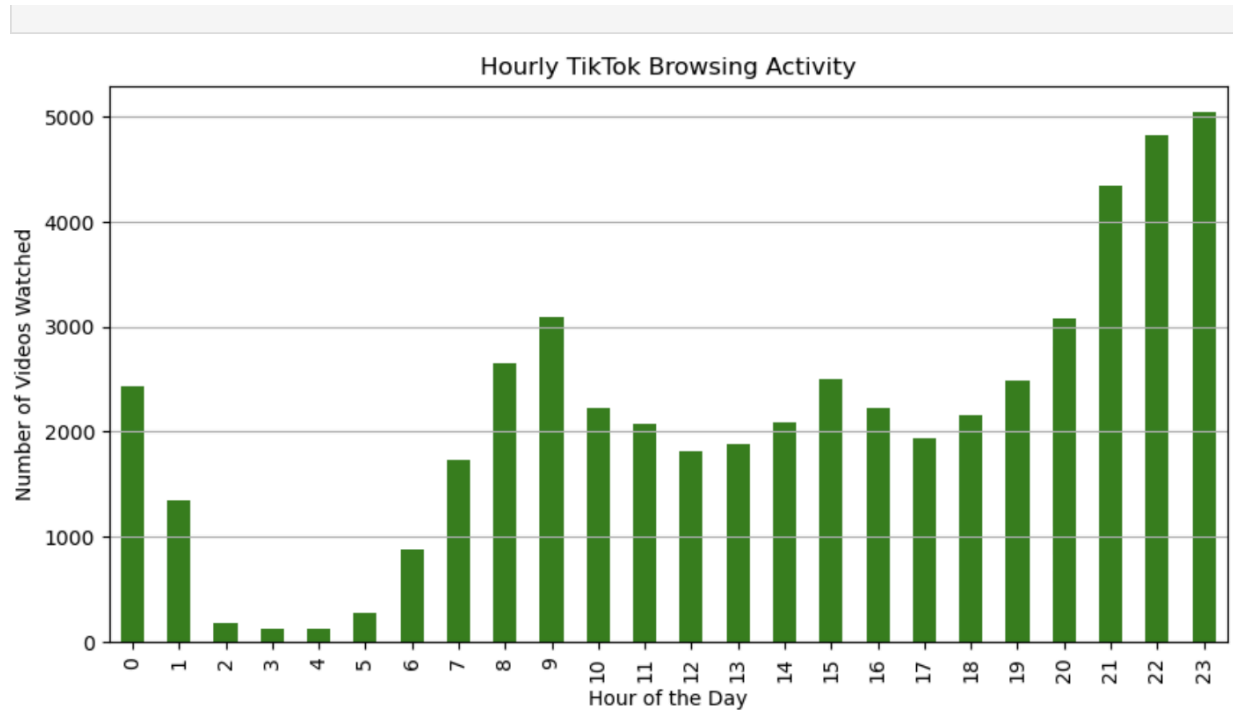
This chart shows the weekly trends in TikTok browsing activity. Activity remains consistent for most weeks, with a significant spike in mid-August and a sharp increase in late November. These peaks may correspond to periods of higher engagement, such as viral trends, seasonal breaks, or personal availability. The end-of-year surge highlights increased browsing activity, potentially due to more leisure time or heightened platform engagement during that period. Occasional dips, such as in early October, suggest reduced activity influenced by external factors like personal schedules or events.

TikTok Browsing Activity by Day of the Week



This chart highlights TikTok browsing patterns across the days of the week. Friday shows the highest activity, suggesting increased usage at the end of the workweek, likely due to more leisure time. Monday also exhibits high engagement, which may reflect a strong start to the week for browsing. Midweek days like Wednesday and Thursday show slightly lower activity levels, indicating possible reductions in free time or shifts in focus toward work or other responsibilities. Weekend activity is consistent but does not surpass Friday's peak, hinting at stable but less spiked engagement during leisure days.

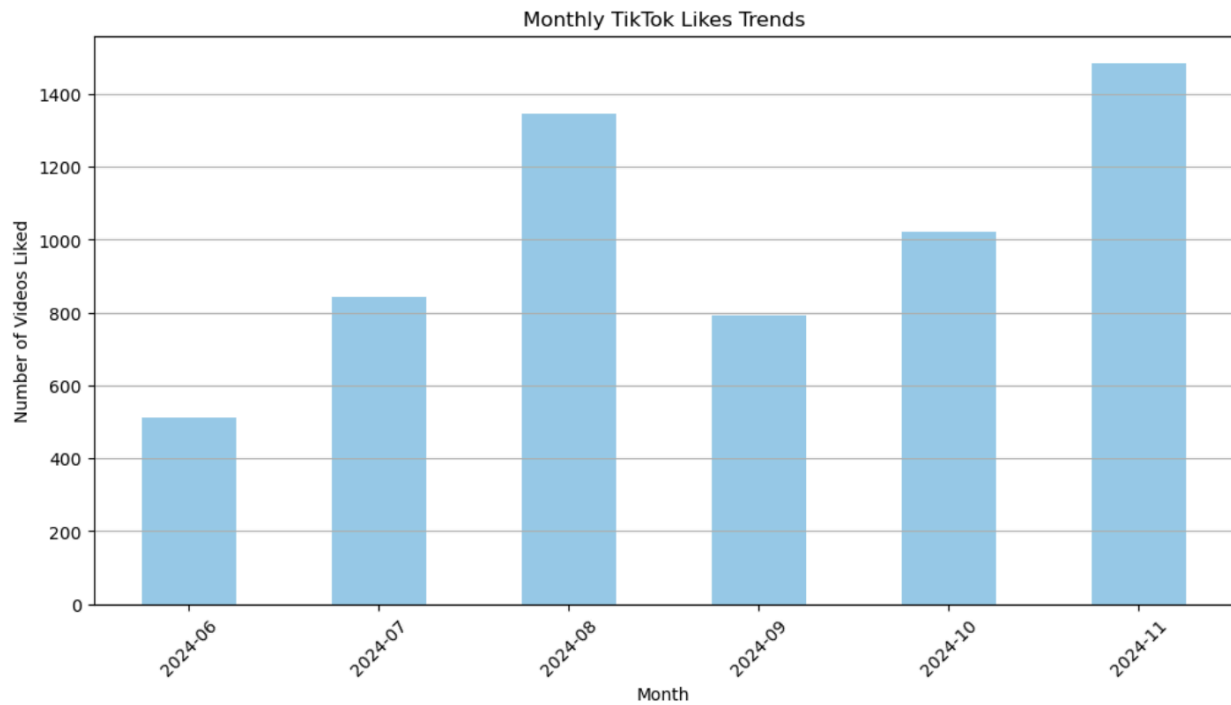
Hourly TikTok Browsing Activity



This chart shows my TikTok browsing activity by hour of the day. I can see that my activity peaks late at night, especially between 9 PM and midnight, which aligns with my habit of using TikTok during my leisure time before bed. Early morning hours, particularly between 2 AM and 6 AM, have the lowest activity, reflecting my usual sleeping schedule. There's also a noticeable spike around 9 AM, which could be during morning breaks, while my afternoon activity remains consistent. The late-night peak clearly highlights how I rely on TikTok as a source of evening entertainment.

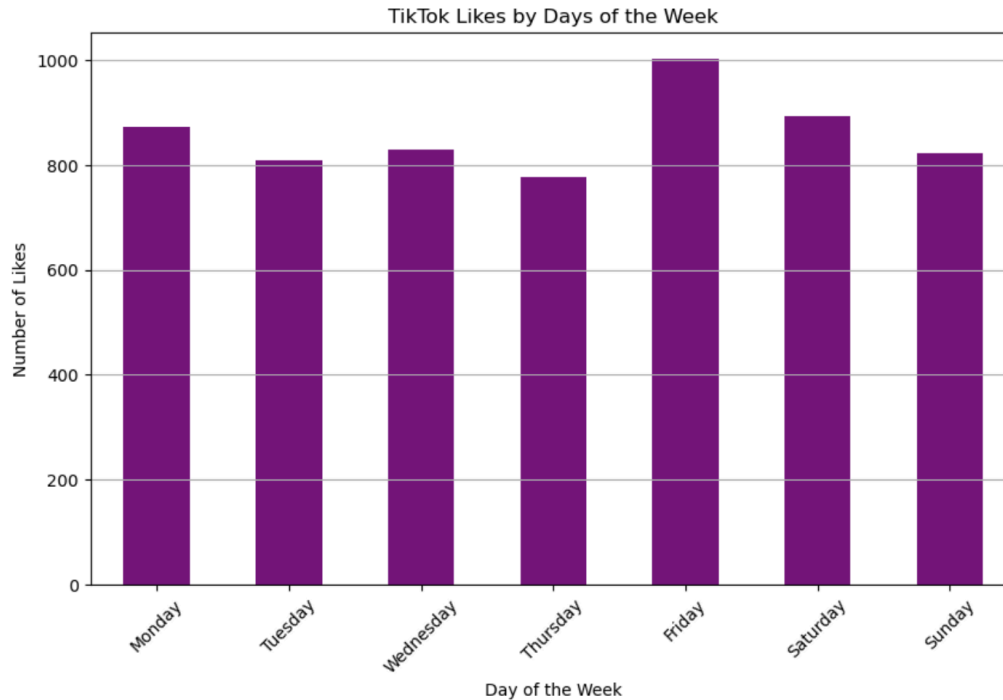
TikTok Likes Trends

Monthly TikTok Likes Trends



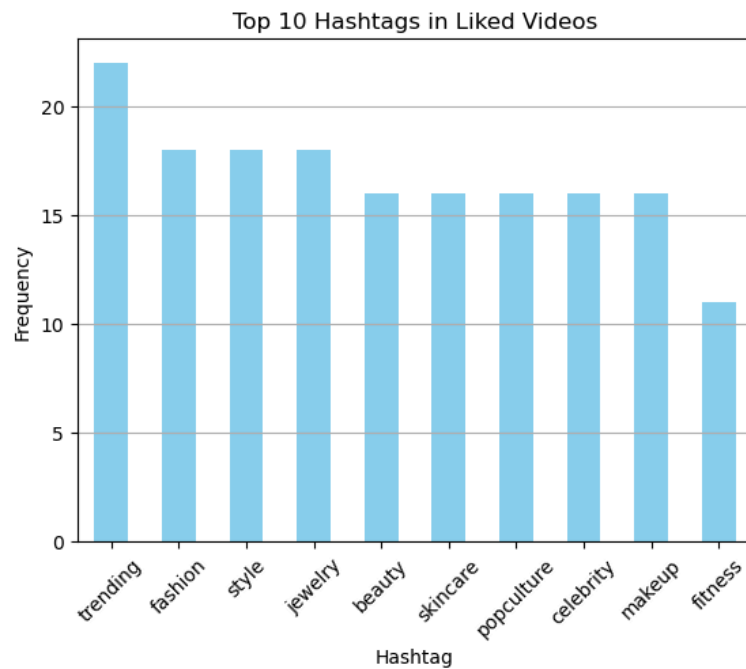
This chart highlights the number of TikTok videos I liked each month. My activity noticeably increased over time, with significant spikes in August and November. August likely reflects higher engagement during summer, while November may coincide with seasonal trends or an increased interest in specific content. June shows the lowest activity, which could align with a period of less frequent app usage or reduced interest in engaging with content due to being busy with summer. These trends indicate how my engagement fluctuates based on personal habits and external factors such as seasons.

TikTok Likes by Days of the Week



This chart shows the number of TikTok likes I gave on each day of the week. Friday stands out as the day with the most likes, which aligns with my increased activity as I wind down the workweek. Likes remain relatively high on Mondays and weekends, reflecting steady engagement during those days. Thursday, however, shows the lowest activity, which might correspond to a busier schedule or reduced browsing time. Overall, my liking behavior suggests more engagement toward the end of the week and during leisure days.

Analysis of Hashtags and Interests in Liked Videos



Primary Interests

The bar chart reveals that my primary interests center around fashion, beauty, and style. Hashtags like "trending," "fashion," "style," and "jewelry" are the most frequently appearing, reflecting my strong engagement with aesthetic content and keeping up with current trends. This dominance

highlights my focus on visual appeal, staying updated with fashion, and exploring creative expressions through style.

Secondary Interests

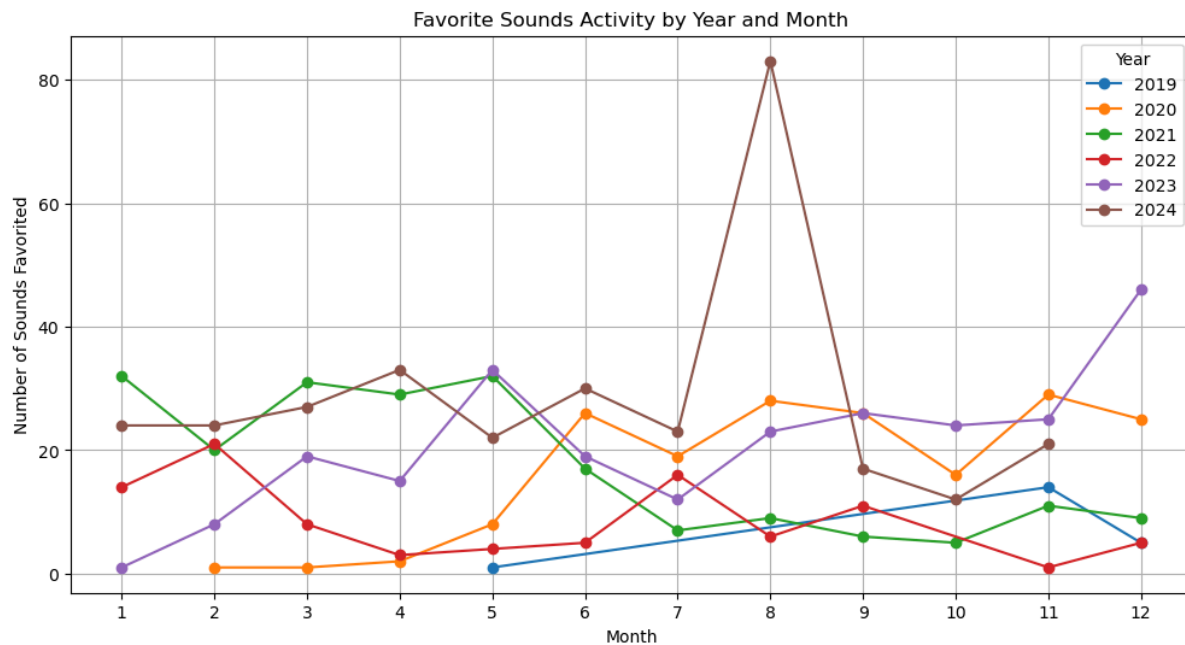
Other frequently used hashtags, such as "skincare," "popculture," "celebrity," and "fitness," point to additional interests in wellness, self-care, and cultural trends. These hashtags suggest that I also engage with content that promotes health, beauty routines, and lifestyle improvements. This broadens my focus beyond just fashion to include a more holistic perspective on personal well-being and popular culture.

Word Cloud Insights

The word cloud complements the bar chart findings by providing a visual representation of recurring themes. Prominent words such as "fashion," "beauty," and "style" underline my primary interests, while smaller words like "fitness," "recipes," "aesthetic," and "positivity" showcase secondary interests in health, food, and motivational content. The diverse themes suggest that I value a combination of aesthetic appeal, wellness, and inspiration. My liked videos clearly demonstrate a strong inclination toward fashion and beauty content, underscored by a secondary focus on wellness, cultural trends, and self-care. These preferences reflect my appreciation for aesthetically pleasing, lifestyle-oriented content that promotes a sense of style, health, and positivity.

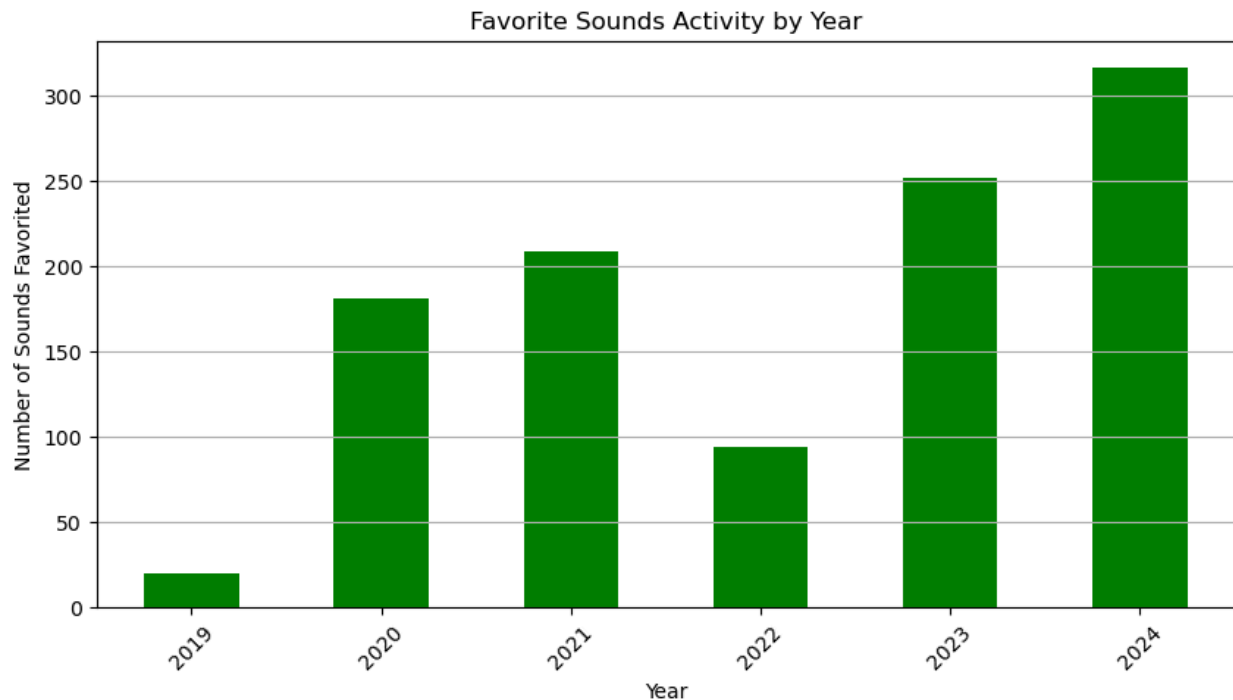
Favorite Sounds Analysis

Analysis of Favorite Sounds Activity by Year and Month:



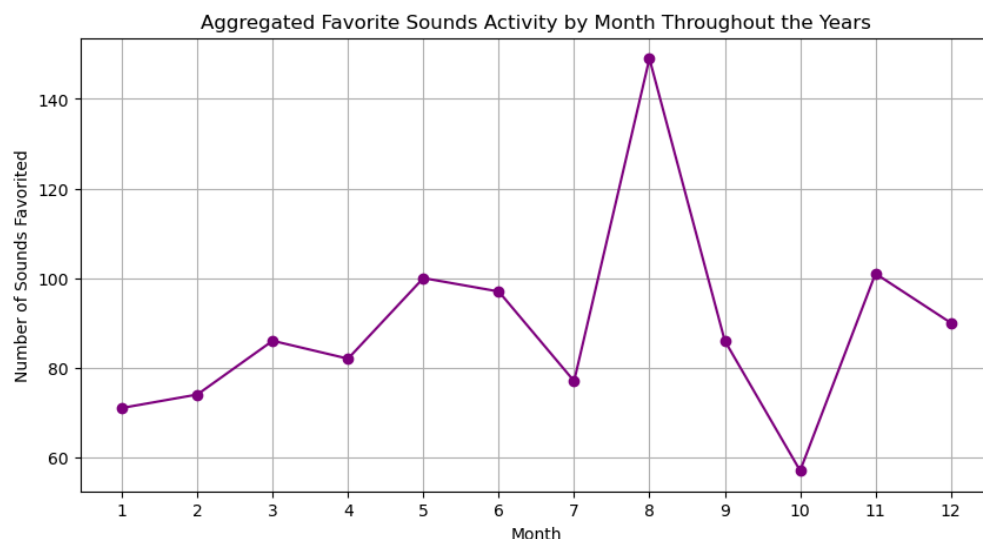
This line graph shows monthly trends for favorite sounds across different years. A sharp spike in July 2023 stands out, with over 80 sounds favorited in that month, reflecting a peak in activity. From 2019 to 2022, monthly engagement appears steady but relatively low. However, 2023 and 2024 demonstrate significant growth in engagement, especially in mid-year months like July and August, indicating a period of heightened interaction with sounds on TikTok.

Favorite Sounds Activity by Year



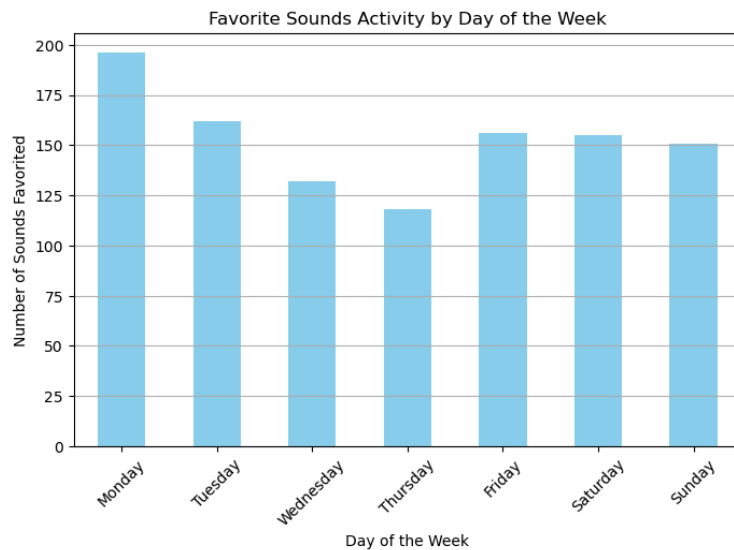
This bar chart highlights the annual growth in favorite sound activity. In the early years, such as **2019**, activity is minimal, reflecting TikTok's relatively limited popularity at the time. However, engagement jumps significantly in **2020 and 2021**, coinciding with the COVID-19 pandemic, when TikTok experienced a surge in global popularity as people turned to the platform during lockdowns. After a slight decline in 2022, activity boomed in **2023 and 2024**, with 2024 seeing the highest number of sounds favorited, marking TikTok's growing dominance and influence in recent years.

Aggregated Favorite Sounds Activity by Month



This line graph aggregates favorite sound activity by month across all years. **August** emerges as the most active month, with consistent engagement peaks, suggesting a seasonal trend. Activity during March through June remains moderate, while months like October and December show lower engagement. The spike in August may correlate with platform trends, seasonal breaks, or other factors driving higher content engagement during that period

Favorite Sounds Activity by Day of the Week

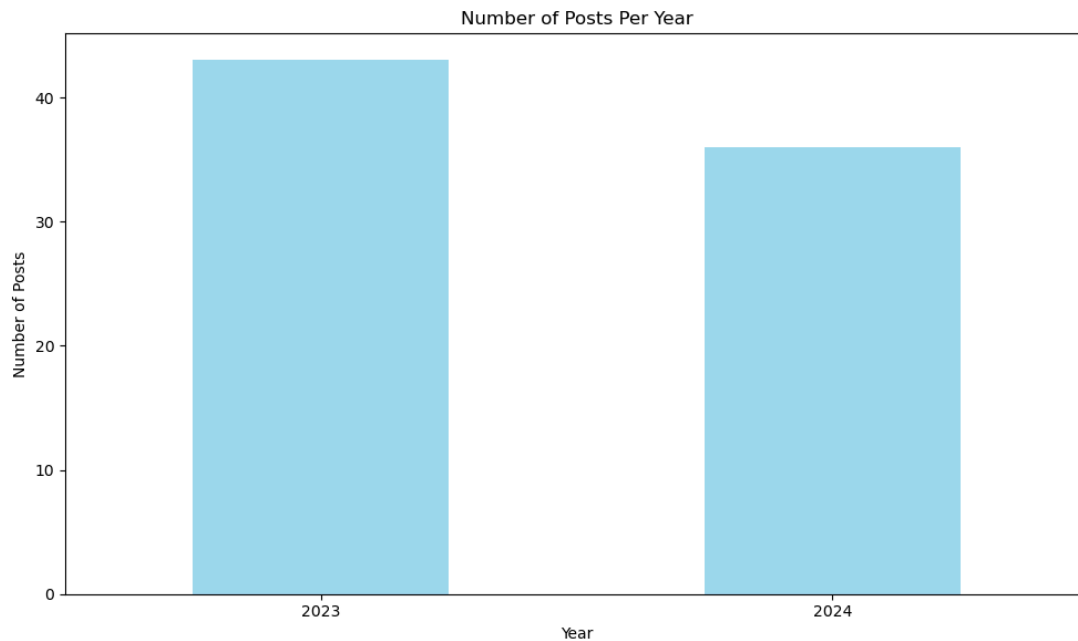


This bar chart demonstrates weekly engagement patterns. **Monday** shows the highest activity, with close to 200 sounds favorited, indicating a strong start to the week. **Friday**, **Saturday**, and **Sunday** also show higher engagement levels, possibly reflecting increased leisure time during weekends. In contrast, midweek days like **Wednesday** and **Thursday** see the lowest activity, suggesting that engagement might be influenced by routine work or school schedules.

The data reflects TikTok's journey from a relatively niche platform in **2019** to a global sensation by **2020 and 2021**, driven by the COVID-19 pandemic. As TikTok continued to expand its user base and influence, engagement surged further in **2023 and 2024**, with 2024 marking a peak in activity. Seasonal patterns show August as the most active month, while weekly trends highlight Mondays and weekends as key days for engagement. This growth trajectory underscores TikTok's increasing cultural relevance and its ability to captivate users with engaging content and trends

Analysis of Posts

Number of Posts Per Year

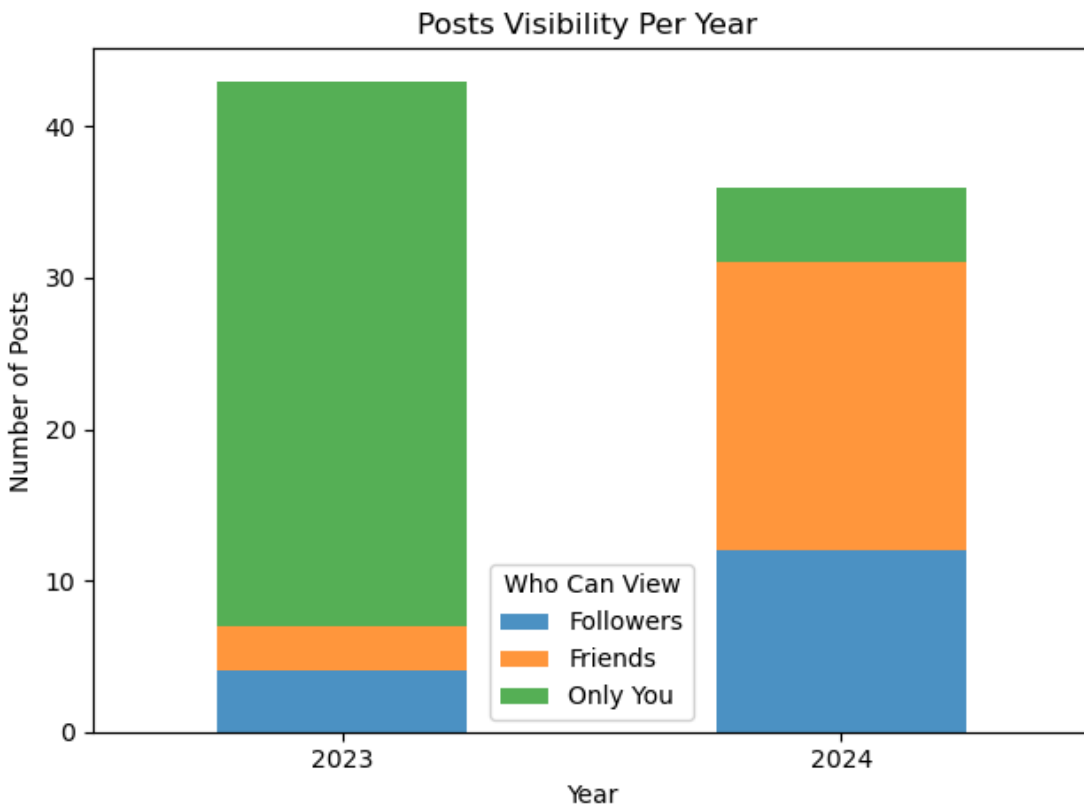


Looking at this graph, my posting activity was at its peak in 2023, with over 40 posts. In 2024, I posted slightly less—just over 30 posts—but it still reflects a consistent effort in sharing content. 2023 seemed to be a year where I was highly engaged in creating and sharing on TikTok, perhaps driven by more inspiration or a desire to experiment with different trends and sounds.

In 2024, the slight dip in posting might reflect a shift in my focus. Maybe I started prioritizing quality over quantity, spending more time curating or engaging with content rather than posting as frequently. It could also mean that I was exploring other creative outlets or simply enjoying the platform more as a consumer rather than a creator.

This trend shows how my relationship with TikTok has evolved. While I've stayed active, my approach to posting seems to have matured, balancing creativity with other ways of engaging on the platform. This balance highlights how TikTok continues to be an important part of how I express myself online.

Posts Visibility Per Year

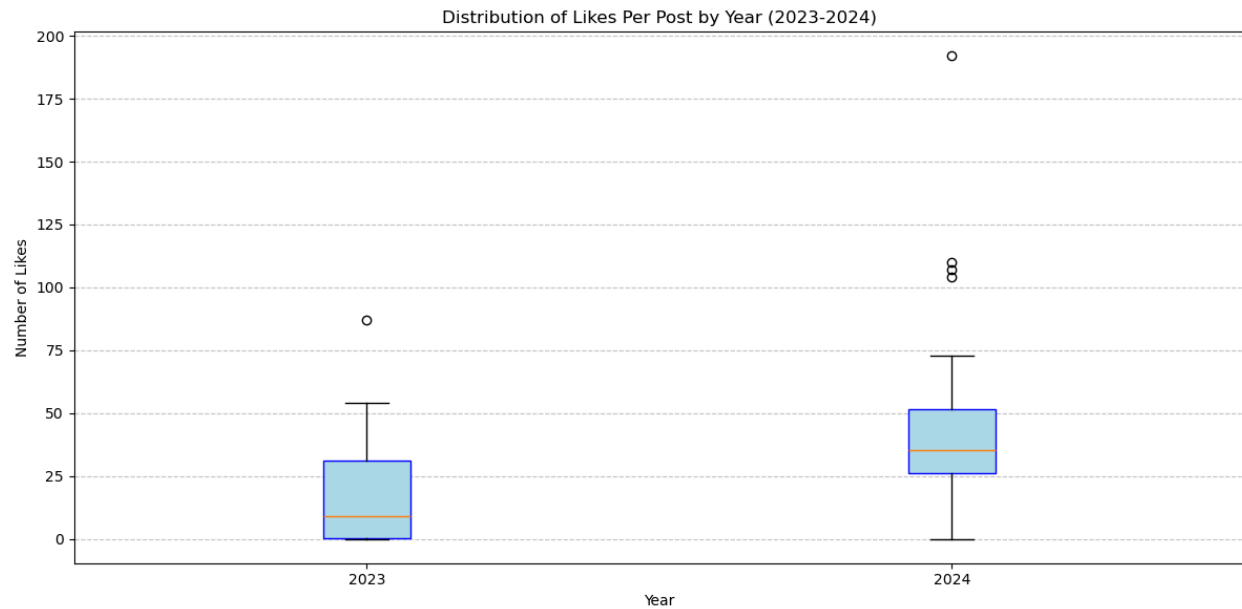


Here, it's clear that in **2023**, almost all of my posts were set to "**Only You**" visibility, meaning I likely used TikTok as a more private space for experimenting with content or saving posts for myself. A very small portion was visible to **friends** or **followers**, which shows I wasn't as focused on sharing publicly during that time.

In **2024**, this trend shifted. There's a noticeable increase in posts visible to **friends** and **followers**, suggesting that I became more comfortable sharing my content with others. This change could mean I gained more confidence in my posts or wanted to connect more with my TikTok audience. However, I still kept a significant number of posts private, which indicates that I value TikTok as a personal creative space.

This means my posting habits evolved from being mostly private to a mix of private and shared content. It shows how my engagement with TikTok has grown from being personal to slightly more social and outward-facing, reflecting a shift in how I use the platform over time.

Distribution of Likes Per Post by Year and Top Sounds by Median Likes



Top Sounds by Median Likes:

Sound

original sound - youssifm.gameel 192.0

Cola (Sped Up Version) - sped up nightcore 110.0

L'amore dice ciao (From "La matriarca - The Libertine") - Main Titles - Armando Trovajoli
107.0

Originalton - theresa_vders

104.0

tumblr girls - rattyfishsquad

87.0

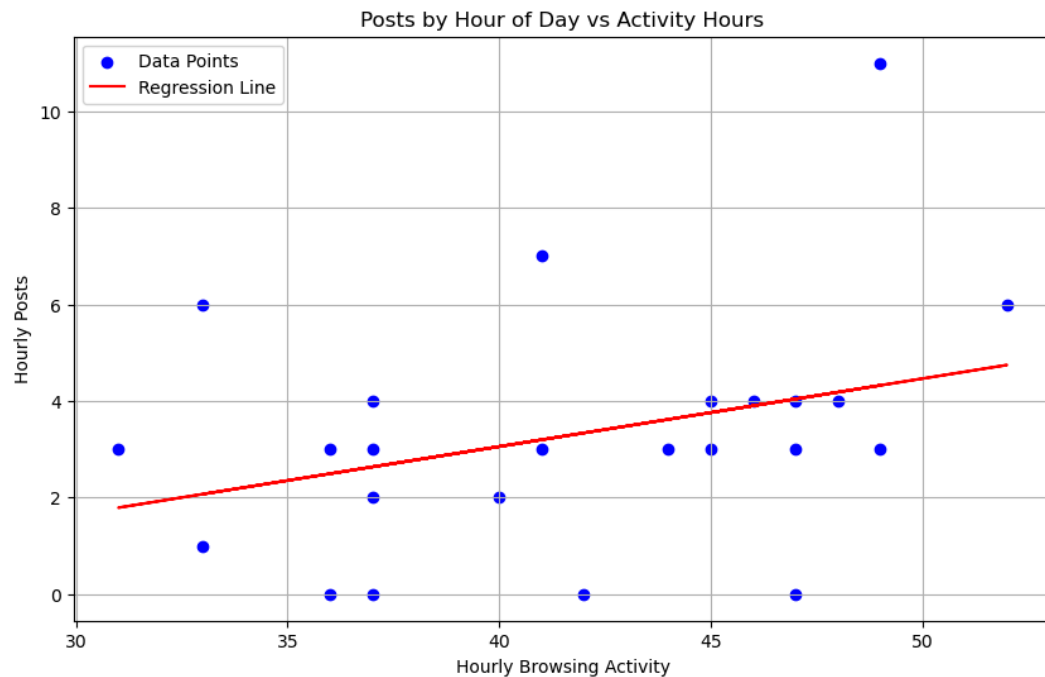
The box plot shows how my posts performed in terms of likes across 2023 and 2024. In **2023**, the median number of likes is lower, with most posts receiving less engagement and only a few outliers getting higher likes. However, in **2024**, there's a noticeable improvement. The median number of likes increased, and the overall distribution of likes shifted upward, indicating that my posts were generally more engaging in 2024. This suggests I might have refined my content strategy or aligned better with my audience's preferences.

When looking at the "Top Sounds by Median Likes," it's clear that the sounds associated with my most liked posts are diverse. For example, the **original sound by youssifm.gameel** led to the highest median likes (192), followed by the sped-up version of **Cola** and iconic sounds like **L'amore dice ciao** and **tumblr girls**. This tells me that while some of my most engaging posts may have used trending sounds at those times, others gained popularity through unique or niche audio choices.

Understanding this helps me see what resonates most with my audience. Are my most-liked posts using trending sounds or more personalized ones? This data reveals that while trends (like **Cola**) play a role, originality and using sounds with a distinct vibe (like **L'amore dice ciao**) also attract significant engagement. This insight is valuable because it can guide me in deciding whether to lean more into trends or focus on developing unique content with niche sounds.

From this, I can conclude that my most-liked posts are not solely dependent on trending sounds—they show a mix of trendy and unique choices. This means I've successfully balanced following TikTok trends while also exploring my personal preferences. Going forward, I can experiment more with niche sounds and creative content while strategically incorporating popular trends to maintain and boost engagement.

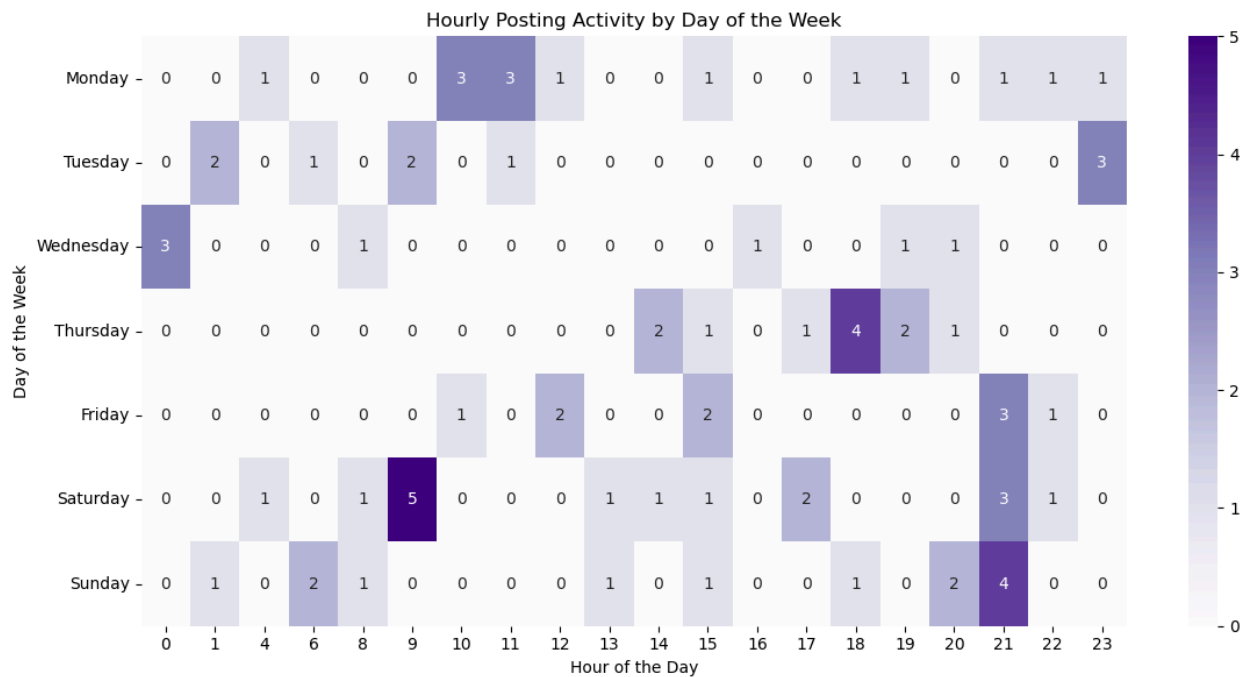
Posts by Hour of Day vs Activity Hours



This scatter plot with a regression line shows the relationship between **hourly browsing activity** and **hourly posting activity**. The positive slope of the regression line suggests a weak but noticeable trend: as my browsing activity increases during certain hours, I'm slightly more likely to post. This means that my posting habits are somewhat influenced by how much time I spend browsing TikTok. For example, if I'm actively engaging with content, I might feel inspired to create and share my own posts.

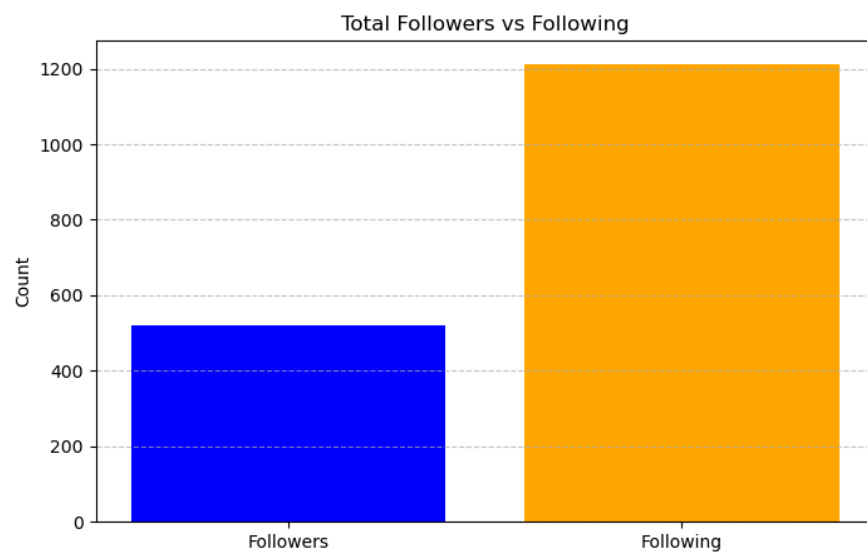
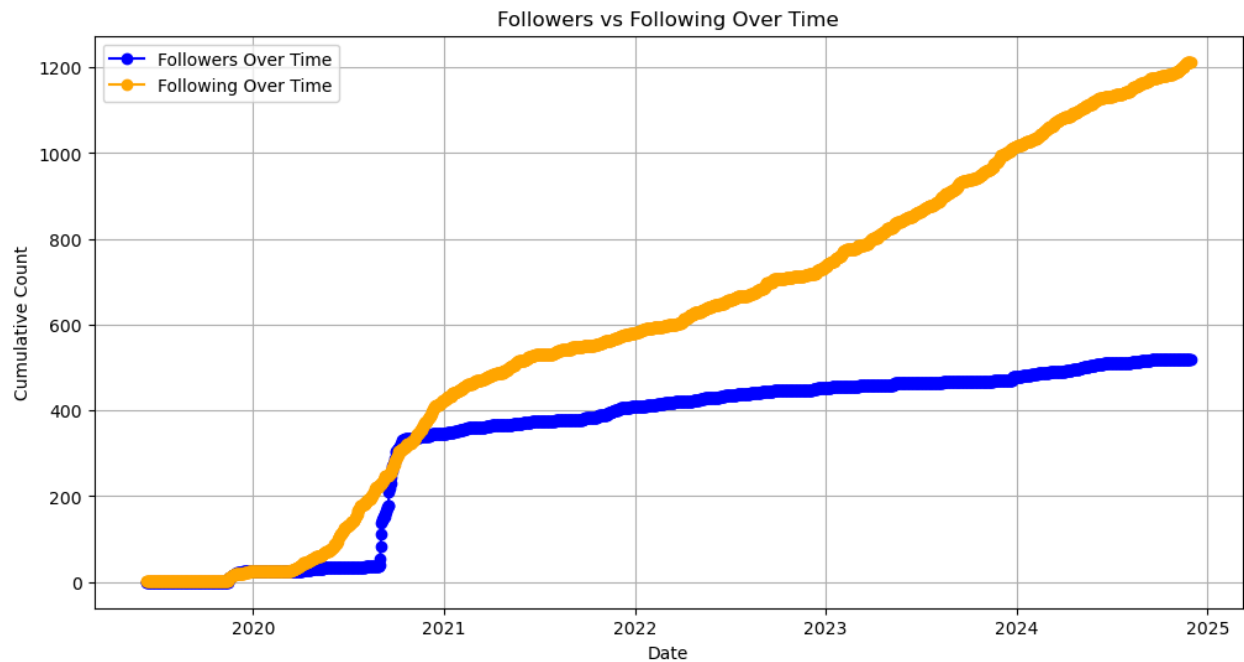
The scatter plot shows that browsing activity can nudge me toward posting more. If I want to increase my posting, I could experiment with browsing during high-engagement hours to gather inspiration.

Hourly Posting Activity by Day of the Week:



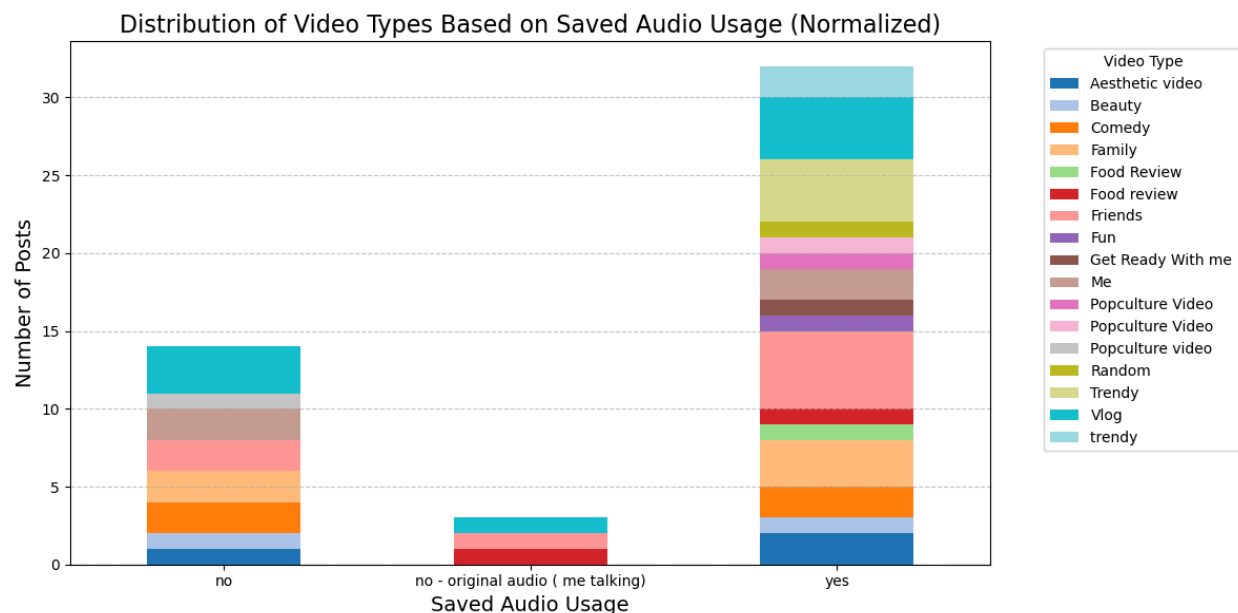
The heatmap reveals my TikTok posting activity across different days and hours, with the darkest shades indicating the highest frequency of posts. My peak posting times are clearly concentrated on Saturday and Sunday evenings around 9 PM, reflecting a strong preference for weekend activity. These times likely align with my availability and higher audience engagement during leisure periods. Sporadic surges can also be seen on Monday and Thursday evenings, though these are less frequent compared to the weekend peaks. Early mornings (12 AM to 6 AM) and midday hours (11 AM to 3 PM) consistently show minimal activity, suggesting these times are less favorable for posting or are occupied by other responsibilities. Overall, my posting patterns emphasize a strategic focus on evening and weekend engagement, balancing personal availability with audience activity trends.

Followers Vs Following Analysis:



From the figures above , I observe that I currently have a total of 520 followers on TikTok, while I follow 1,213 accounts. This disparity reflects my casual use of the platform, where my primary focus is on consuming content rather than building a large follower base or establishing a professional presence. Most of my followers are friends, highlighting that I use TikTok as a space for entertainment, lighthearted engagement, and personal connection. Having under 1,000 followers, despite being on the app since 2019, further reinforces that my approach to TikTok is purely for fun, without any ambitions of turning it into a professional or influencer platform. It remains a casual and enjoyable outlet for sharing moments and exploring trends.

Hypothesis Testing

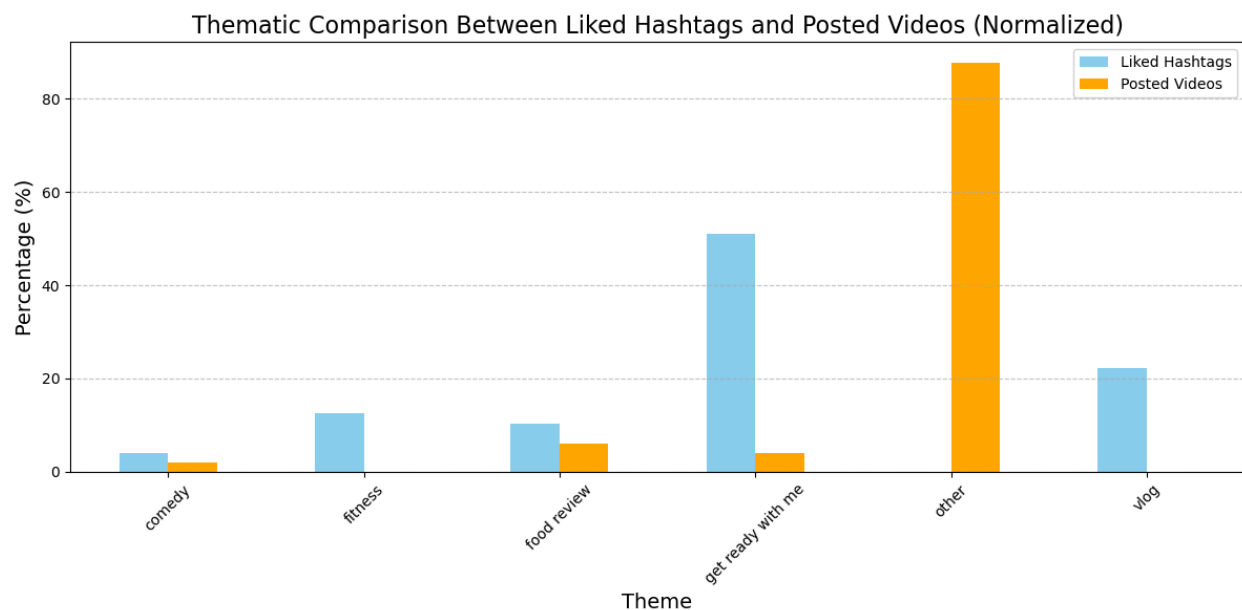


Chi-Square Statistic: 20.77, P-value: 0.9366, Degrees of Freedom: 32

The initial hypothesis stated that the content I engage with on TikTok (e.g., liked videos, saved sounds) directly influences the themes, styles, and creative decisions in my personal posts. The analysis examines the relationship between saved audio usage and video types in my posts. The chart shows that saved audio is used in a majority of posts, covering a wide variety of video

types, while posts without saved audio are less frequent and limited in diversity. However, the Chi-square test results (Chi-square statistic = 20.77, p-value = 0.9366, degrees of freedom = 32) reveal no statistically significant association between video types and saved audio usage.

As the p-value is much greater than 0.05, I fail to reject the null hypothesis, which means there is insufficient evidence to support the claim that saved audio usage influences the choice of video types. This suggests that while saved audio is a common element in my posts, the themes and styles of my content are not solely determined by my engagement with saved sounds. Therefore, my initial hypothesis is only partially supported, as other factors likely play a role in shaping my creative decisions.



Chi-Square Statistic: 193.28

P-value: 0.0000

Degrees of Freedom: 5

The Chi-Square Test indicates a significant relationship. Your posted video themes are influenced by your liked hashtags.

	Theme	Liked Hashtags (%)	Posted Videos (%)
0	comedy	3.977273	2.040816
1	fitness	12.500000	0.000000
2	food review	10.227273	6.122449
3	get ready with me	51.136364	4.081633
4	other	0.000000	87.755102
5	vlog	22.159091	0.000000

This analysis investigates the relationship between the themes of liked hashtags and posted videos to test whether my engagement influences my creative decisions. The chart and table show thematic comparisons, with noticeable differences in the distribution of themes between liked hashtags and posted videos. For instance, themes like "get ready with me" and "fitness"

dominate in liked hashtags, while the "other" category overwhelmingly dominates my posted videos.

The Chi-square test results (Chi-square statistic = 193.28, p-value = 0.0000, degrees of freedom = 5) indicate a statistically significant relationship between the themes of liked hashtags and posted videos. Since the p-value is below 0.05, I reject the null hypothesis, confirming that the themes of my posted videos are influenced by the themes of my liked hashtags.

However, it is important to note that this result could be affected by potential errors. Type I or Type II errors may arise because the hashtags used in the analysis may not exactly match the true content of the videos but rather reflect general trends. Despite this, the general thematic alignment supports my initial hypothesis that the content I engage with influences my creative decisions on TikTok.

Conclusion

This project allowed me to explore how my TikTok engagement habits shape my creative output on the platform. By analyzing trends in browsing, likes, hashtags, and favorite sounds, I discovered clear connections between the content I engage with and my personal posting behavior. The Chi-square tests supported the hypothesis that my engagement influences my creative decisions, particularly in the themes and styles of my posts. However, this influence is not absolute, as other factors also play a role, such as spontaneity, trends, and personal preferences.

Ultimately, as a 21-year-old girl living in the digital age, my relationship with TikTok reflects a broader cultural phenomenon of keeping up with trends and using social media as a creative outlet. TikTok is not just a platform for entertainment but also a space where I explore my interests, experiment with content, and connect with a digital audience. The findings from this analysis have given me a deeper understanding of my creative habits and how social media algorithms subtly guide my decisions.

Limitations and Future Work

This project faced several challenges that highlight areas for improvement in future analyses:

1. **API Limitations:** Due to restrictions with TikTok's API, I had to rely on manual extraction for some data, such as favorite sounds and hashtags. While I was still able to identify general trends, this method introduced potential biases and limited the granularity of the data. Future work could focus on finding more robust tools or methods for automated data collection to enhance accuracy and efficiency.
2. **Content Matching:** The hashtags and themes used in the analysis do not perfectly match the actual content of the videos, which may have introduced errors in the analysis. Future studies could incorporate advanced text or image analysis tools to categorize content more accurately.
3. **Dataset Scope:** The analysis was limited to personal TikTok data. Expanding the dataset to include interactions with peers, comments, and duets could provide a richer understanding of creative influence.
4. **Cultural and Psychological Factors:** While this project focused on engagement and content, future work could explore how cultural trends and personal motivations drive posting behavior. Incorporating psychological surveys or interviews could add depth to the analysis.

Despite these limitations, I was able to uncover meaningful insights and identify trends in my engagement and creative habits. This project highlights the potential for further exploration into how social media platforms influence creativity and self-expression in the digital age.

