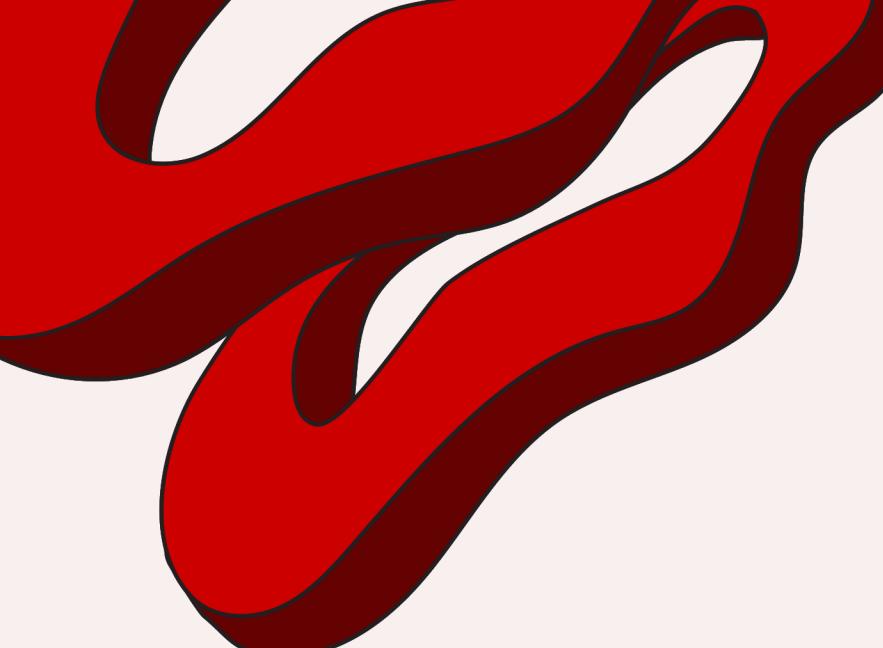
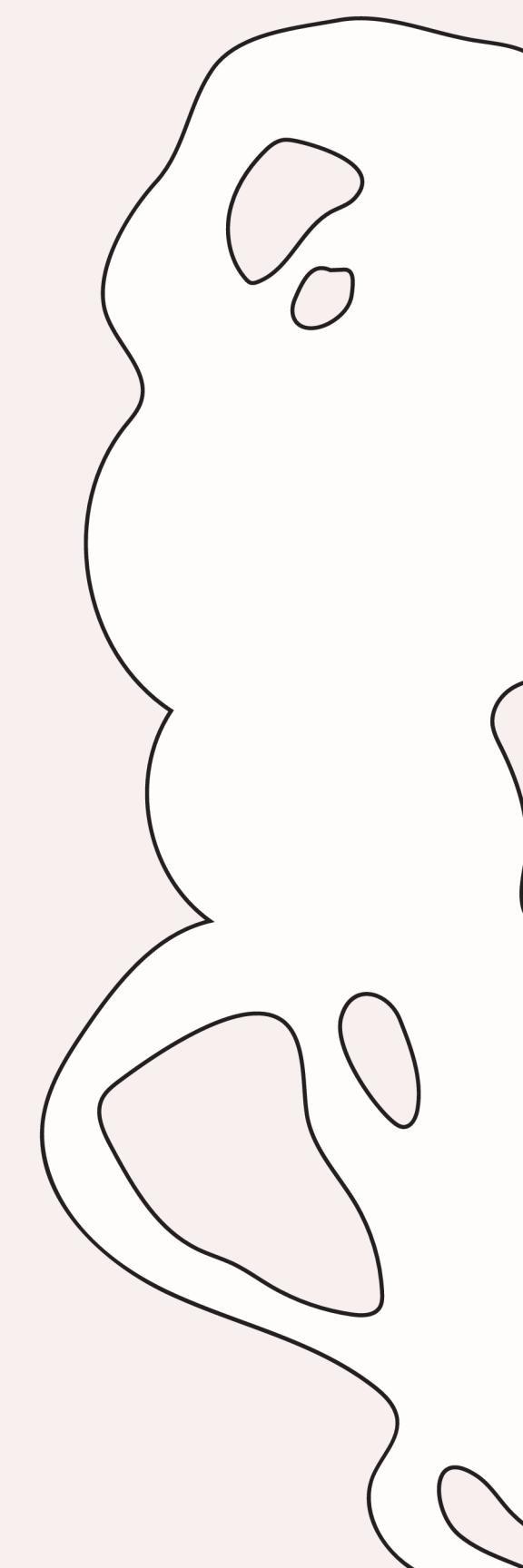


Impact of Content Category on Like-to- View Ratio

| Faye Le

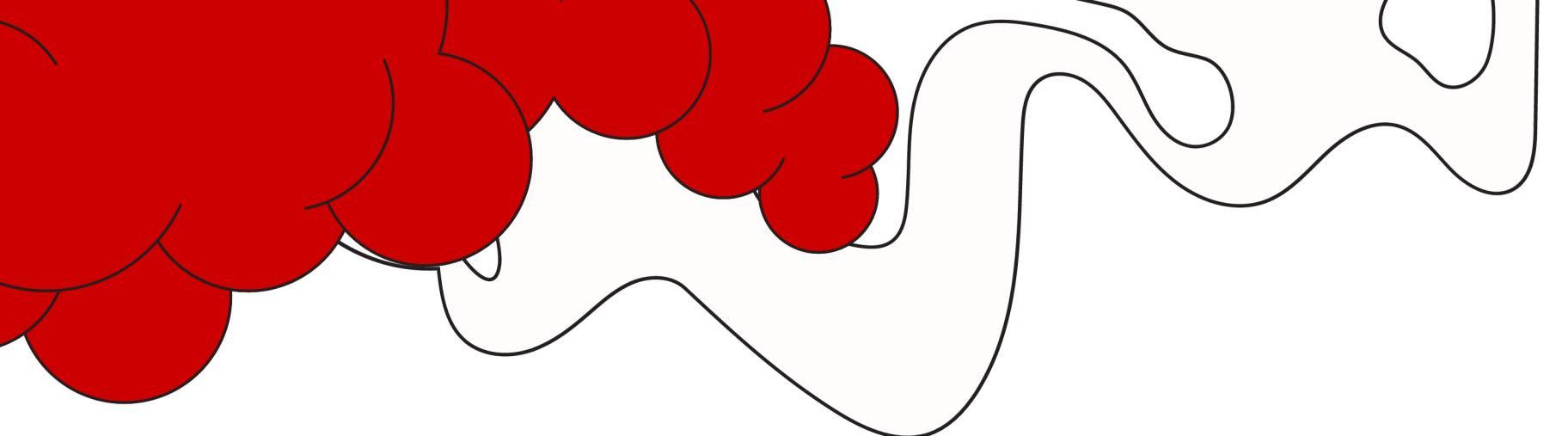


Questions

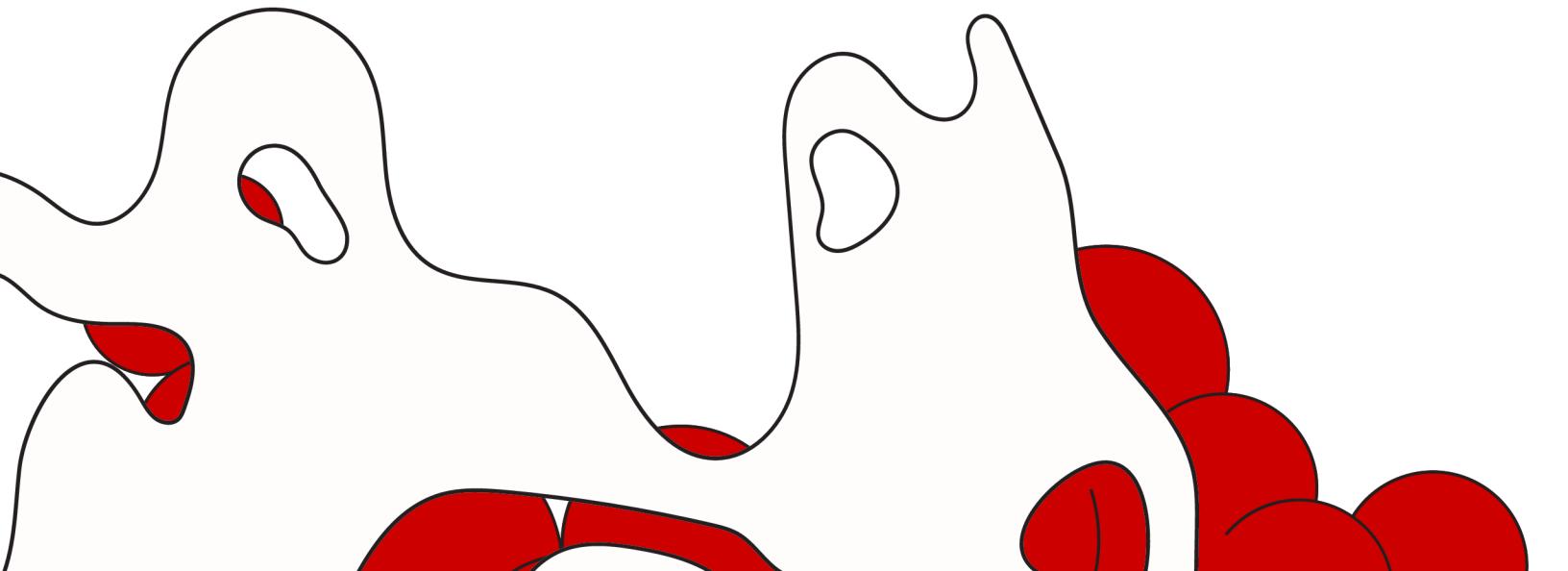


How does the content category affect the like-to-view ratio? Are there any categories where viewers are more likely to interact with the content through likes?

Hypothesis: Educational content might have a lower like-to-view ratio compared to entertainment-focused categories like "Video Games"



Why?



Insight into how actively engaged viewers are with content, beyond simply watching the video

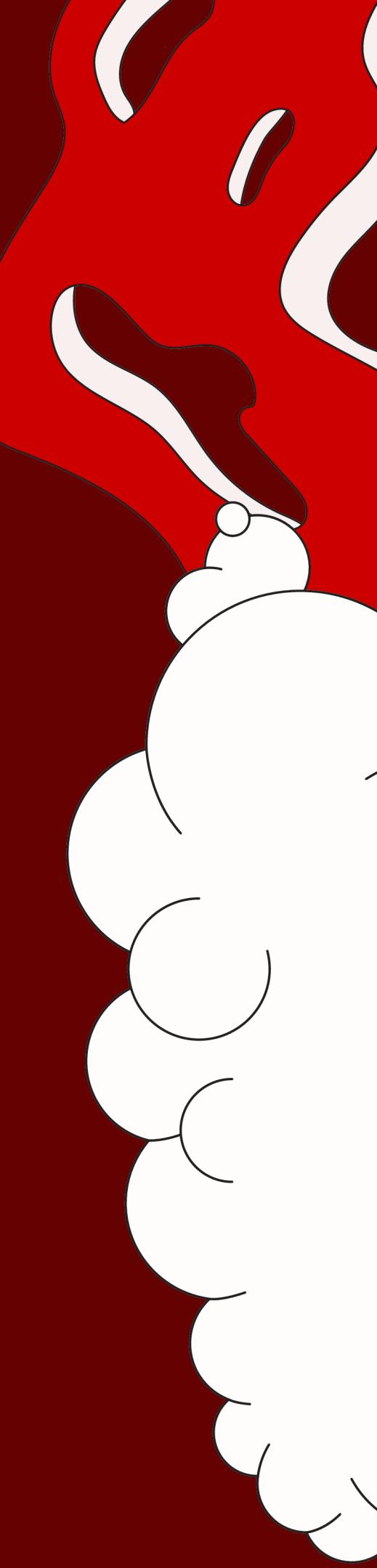
Signal that the content is well-received or of high quality

Recommendation algorithms

Monetization and Sponsorships

Tractable data & Data retrieval

- Kaggle
- Sep to Dec 2022
- Top subscribed Youtuber
- Prep data



```
November 2022 Data:  
   Channel name    Youtuber name      Category Subscribers \\\n0       tseries          T-Series  Music & Dance     229.5M  
2     PewDiePie        PewDiePie  Animation     111.6M  
3     MrBeast6000      MrBeast  Video games     109.1M  
6    WWEFanNation        WWE  Video games     91.8M  
7 zeemusiccompany  Zee Music Company  Music & Dance     89.7M
```

```
      Country Views avg. Likes avg. Comments avg.  
0       India   103.8K     2.8K      158  
2 United States     1.4M    110.7K      7.8K  
3 United States     38.4M     3.1M     56.7K  
6 United States    144.9K     5.1K     207  
7       India    89.6K     2.4K      91
```

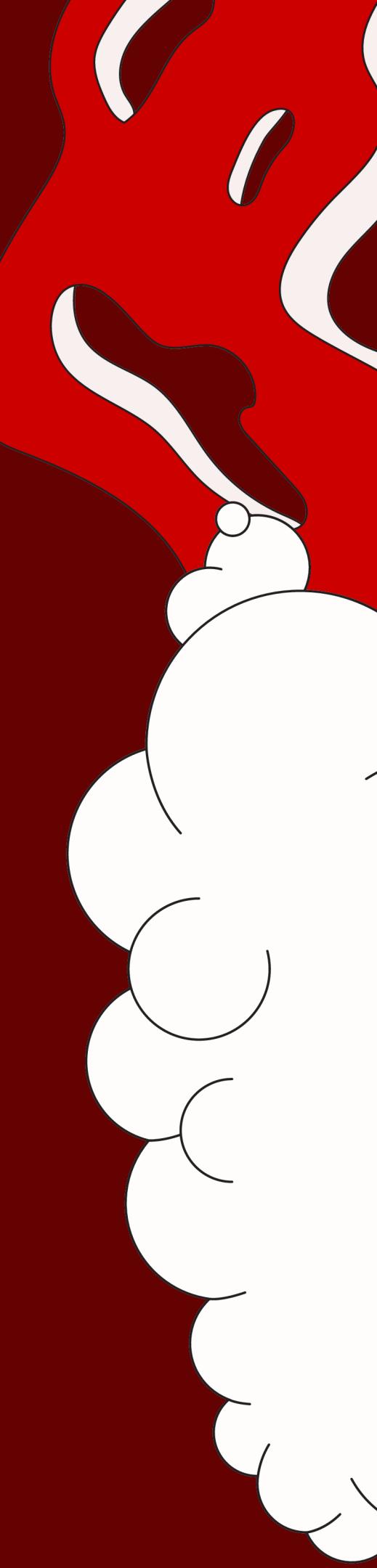
```
September 2022 Data Rows and Columns: (722, 8)  
October 2022 Data Rows and Columns: (572, 8)  
November 2022 Data Rows and Columns: (581, 8)  
December 2022 Data Rows and Columns: (579, 8)
```

```
def convert_to_number(s):  
    if pd.isna(s):  
        return s  
    if s[-1] == 'K':  
        return float(s[:-1]) * 1000  
    elif s[-1] == 'M':  
        return float(s[:-1]) * 1000000  
    else:  
        return float(s)
```

Steps

- Standardize column names across all datasets, drop unnecessary columns
- Handle missing or NaN values, particularly in key fields like likes rate, category, and subscribers
- Convert text-based numerical values (e.g., "477.9M") into actual numbers to enable calculation of engagement rates.
- Ensure all datasets have consistent structures for easier combination and comparison

Exploratory Data Analysis





Calculate the like-to-view ratio for each channel

Group the channels by content category

Visualize the distribution

Identify whether certain categories, such as Education, tend to have lower like-to-view ratios compared to categories like Video Games

	Like-to-View Ratio			
	mean	median	std	count
Category				
ASMR	0.042764	0.035227	0.024950	5
Animals & Pets	0.048212	0.044909	0.020867	4
Animation	0.058867	0.055249	0.033963	67
Autos & Vehicles	0.156667	0.156667	NaN	1
Beauty	0.043919	0.043919	0.014890	2
DIY & Life Hacks	0.043035	0.036735	0.018181	8
Daily vlogs	0.059352	0.059186	0.044275	30
Design/art	0.093737	0.093737	0.043753	2
Education	0.055533	0.049124	0.034616	12
Fashion	0.058496	0.045924	0.027668	7
Fitness	0.040678	0.044697	0.007759	3
Food & Drinks	0.051549	0.037524	0.034223	16
Health & Self Help	0.039199	0.039199	0.026838	2
Humor	0.048530	0.044274	0.022137	20
Movies	0.055102	0.050937	0.036636	55
Music & Dance	NaN	0.037092	NaN	180
Mystery	0.037439	0.037439	NaN	1
News & Politics	0.023443	0.020282	0.014534	48
Science & Technology	0.059124	0.048780	0.029664	15
Sports	0.046866	0.033762	0.039244	11
Toys	0.036791	0.041876	0.017869	13
Travel	0.055759	0.055759	NaN	1
Video games	0.057105	0.049246	0.030773	72

Sep

	Like-to-View Ratio			
	mean	median	std	count
Category				
ASMR	0.017165	0.017165	0.003285	2
Animals & Pets	0.051676	0.051676	0.001438	2
Animation	0.049651	0.043050	0.033707	61
Beauty	0.051900	0.051900	NaN	1
DIY & Life Hacks	0.042395	0.033333	0.024991	9
Daily vlogs	0.038152	0.040083	0.029767	27
Design/art	0.057942	0.057942	NaN	1
Education	0.048757	0.045714	0.019335	17
Fashion	0.055115	0.055115	0.055741	2
Fitness	0.043704	0.039000	0.009113	3
Food & Drinks	0.051395	0.044066	0.034905	14
Health & Self Help	0.026202	0.025872	0.009971	3
Humor	0.048284	0.042508	0.024113	11
Movies	0.053569	0.044088	0.037472	90
Music & Dance	0.047419	0.037676	0.041106	194
News & Politics	0.023068	0.017361	0.020228	46
Science & Technology	0.058569	0.045638	0.030247	17
Sports	0.052219	0.047170	0.028150	9
Toys	0.034352	0.030894	0.020930	4
Travel	0.041545	0.041545	NaN	1
Video games	0.051063	0.049171	0.028680	60

Nov

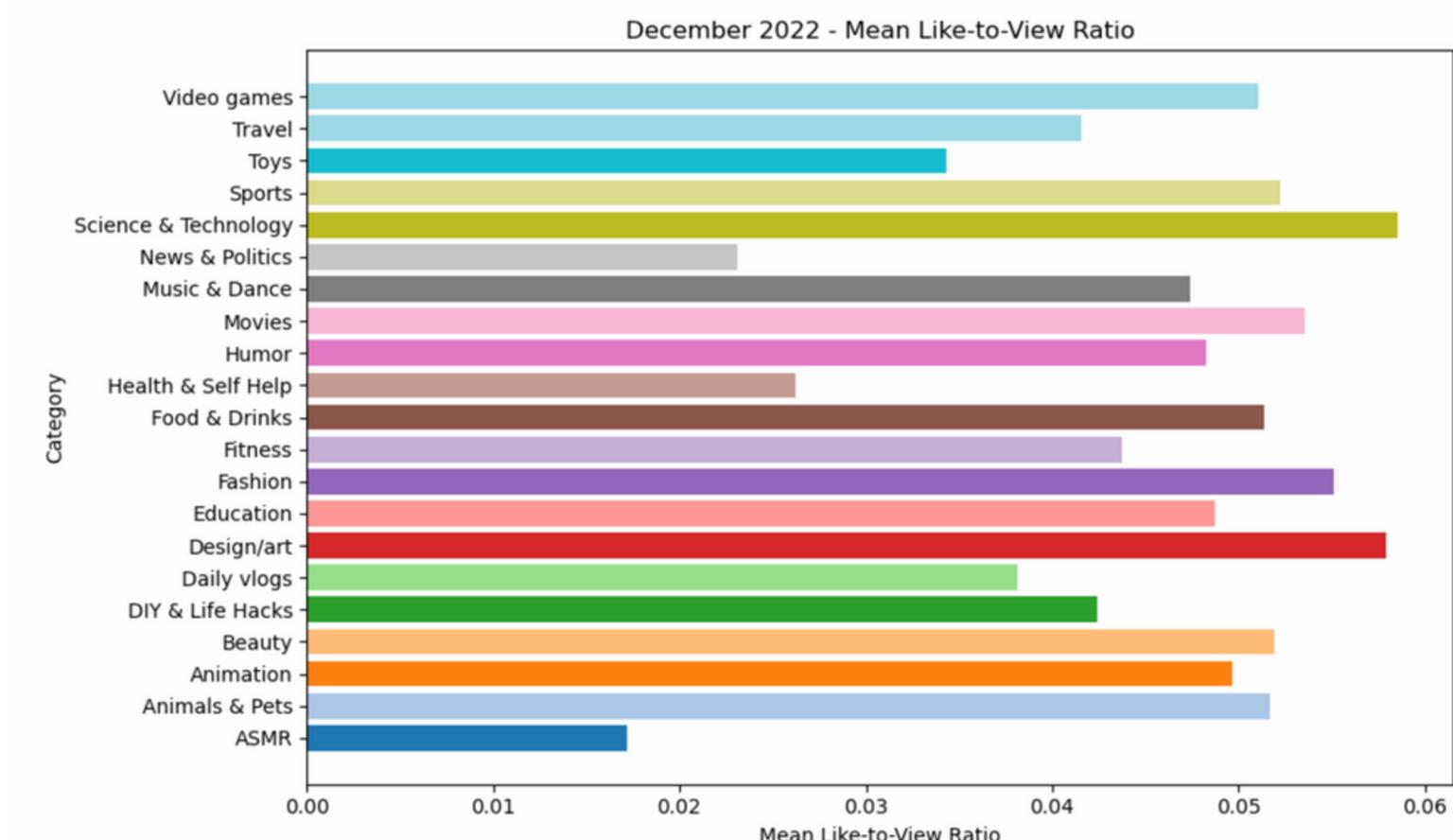
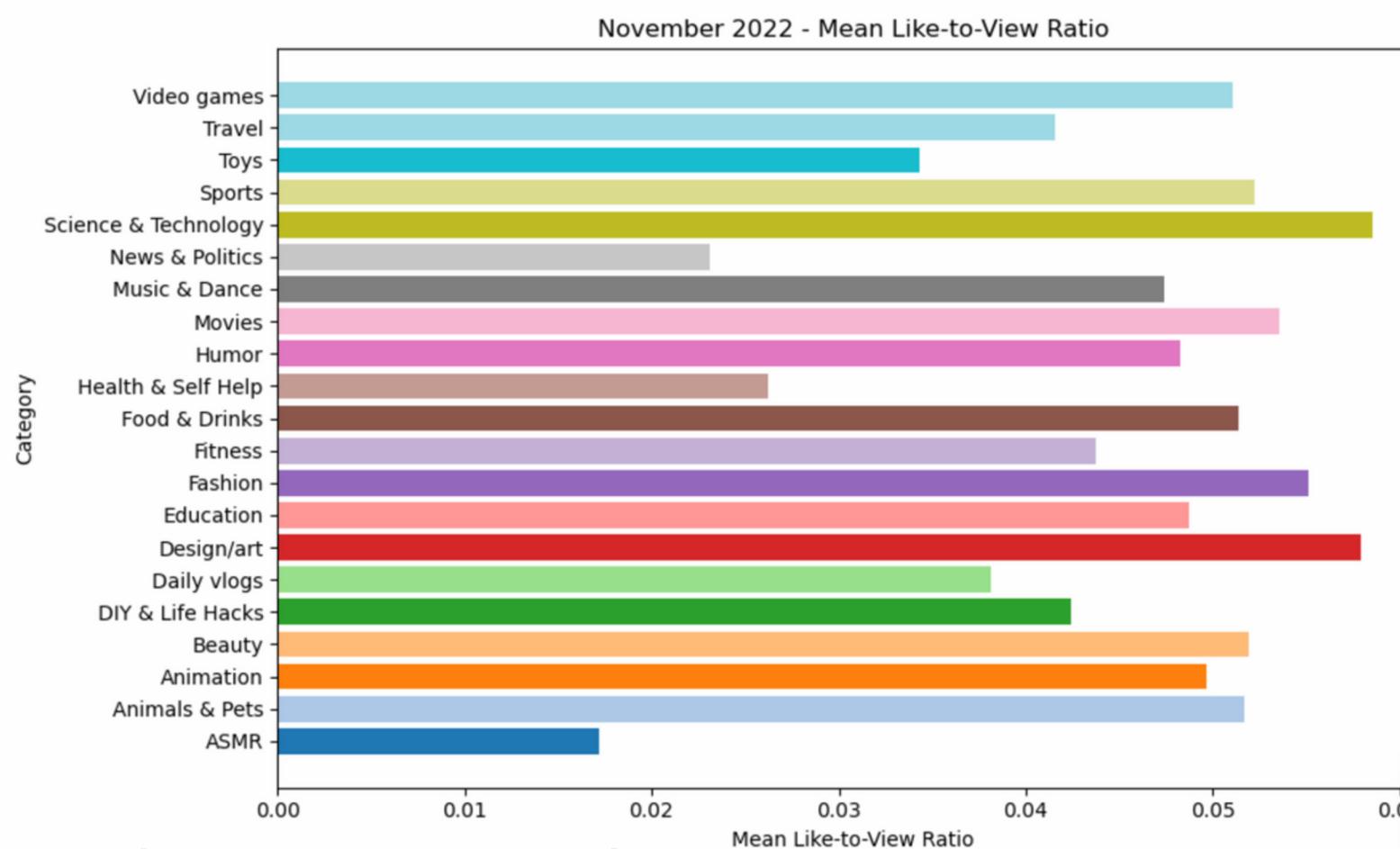
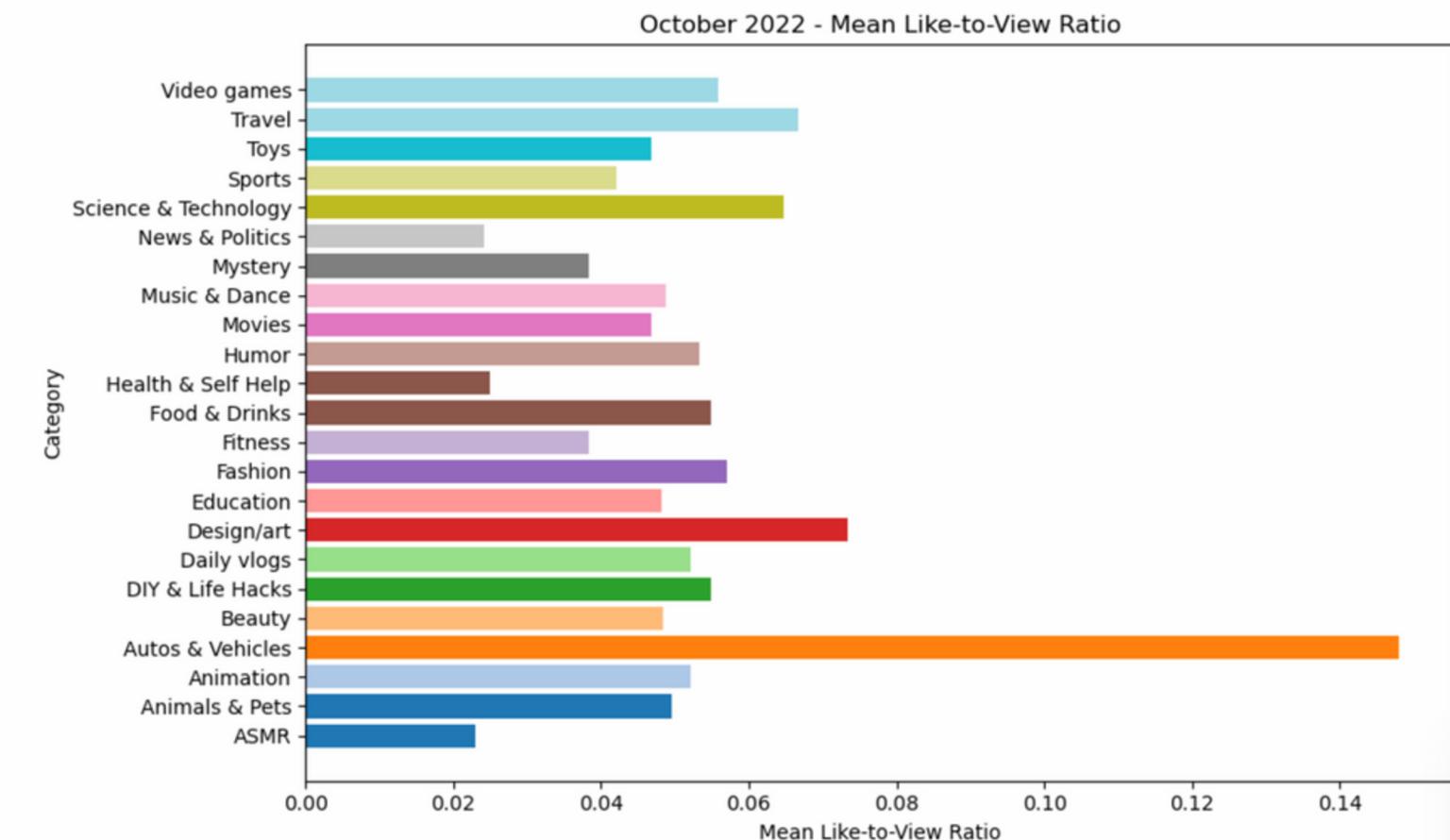
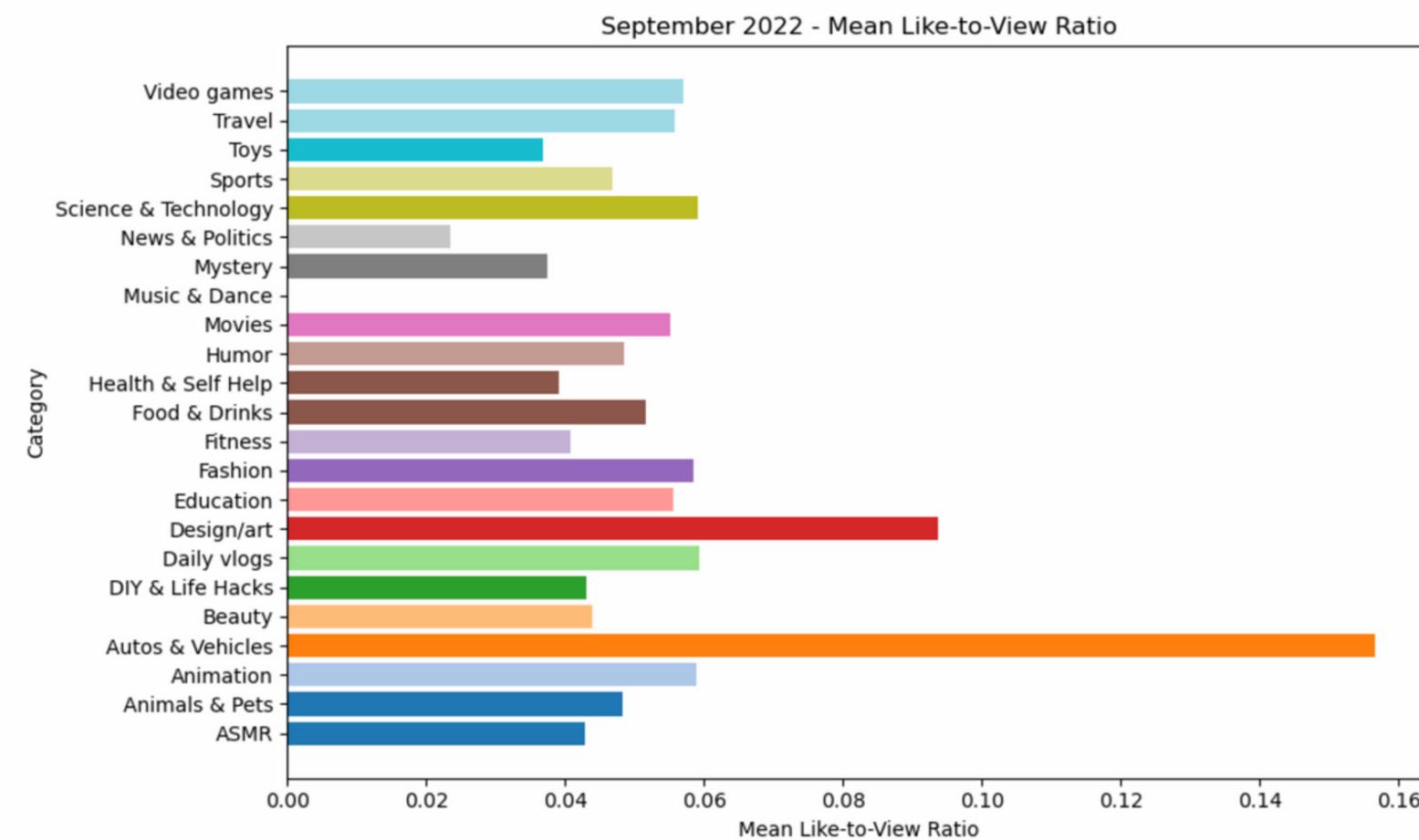
	Like-to-View Ratio			
	mean	median	std	count
Category				
ASMR	0.022934	0.012000	0.019314	3
Animals & Pets	0.049580	0.048260	0.020304	4
Animation	0.052143	0.052016	0.027160	56
Autos & Vehicles	0.148052	0.148052	NaN	1
Beauty	0.048288	0.048288	0.009210	2
DIY & Life Hacks	0.054937	0.042090	0.030517	14
Daily vlogs	0.052135	0.052247	0.033403	27
Design/art	0.073363	0.067836	0.038807	3
Education	0.048213	0.050884	0.020169	14
Fashion	0.056941	0.041024	0.030891	7
Fitness	0.038256	0.046047	0.015242	3
Food & Drinks	0.054791	0.043527	0.031673	15
Health & Self Help	0.024997	0.024997	0.003212	2
Humor	0.053333	0.043165	0.032608	29
Movies	0.046843	0.031607	0.038940	33
Music & Dance	0.048787	0.037205	0.043535	185
Mystery	0.038426	0.038426	NaN	1
News & Politics	0.024173	0.019117	0.015730	45
Science & Technology	0.064758	0.052839	0.037481	17
Sports	0.042158	0.032389	0.028819	11
Toys	0.046715	0.045373	0.031111	20
Travel	0.066629	0.066629	NaN	1
Video games	0.055926	0.047067	0.033272	79

Oct

	Like-to-View Ratio			
	mean	median	std	count
Category				
ASMR	0.017165	0.017165	0.003285	2
Animals & Pets	0.051676	0.051676	0.001438	2
Animation	0.049651	0.043050	0.033707	61
Beauty	0.051900	0.051900	NaN	1
DIY & Life Hacks	0.042395	0.033333	0.024991	9
Daily vlogs	0.038152	0.040083	0.029767	27
Design/art	0.057942	0.057942	NaN	1
Education	0.048757	0.045714	0.019335	17
Fashion	0.055115	0.055115	0.055741	2
Fitness	0.043704	0.039000	0.009113	3
Food & Drinks	0.051395	0.044066	0.034905	14
Health & Self Help	0.026202	0.025872	0.009971	3
Humor	0.048284	0.042508	0.024113	11
Movies	0.053569	0.044088	0.037472	90
Music & Dance	0.047419	0.037676	0.041106	194
News & Politics	0.023068	0.017361	0.020228	46
Science & Technology	0.058569	0.045638	0.030247	17
Sports	0.052219	0.047170	0.028150	9
Toys	0.034352	0.030894	0.020930	4
Travel	0.041545	0.041545	NaN	1
Video games	0.051063	0.049171	0.028680	60

Dec

	Like-to-View Ratio			
	mean	median	std	count
Category				
ASMR	0.017165	0.017165	0.003285	2
Animals & Pets	0.051676	0.051676	0.001438	2
Animation	0.049651	0.043050	0.033707	61
Beauty	0.051900	0.051900	NaN	1
DIY & Life Hacks	0.042395	0.033333	0.024991	9
Daily vlogs	0.038152	0.040083	0.029767	27
Design/art	0.057942	0.057942	NaN	1
Education	0.048757	0.045714	0.019335	17
Fashion	0.055115	0.055115	0.055741	2
Fitness	0.043704	0.039000	0.009113	3
Food & Drinks	0.051395	0.044066	0.034905	14
Health & Self Help	0.026202	0.025872	0.009971	3
Humor	0.048284	0.042508		



Conclusion

Educational content tends to have a slightly lower like-to-view ratio compared to **Video Games**.

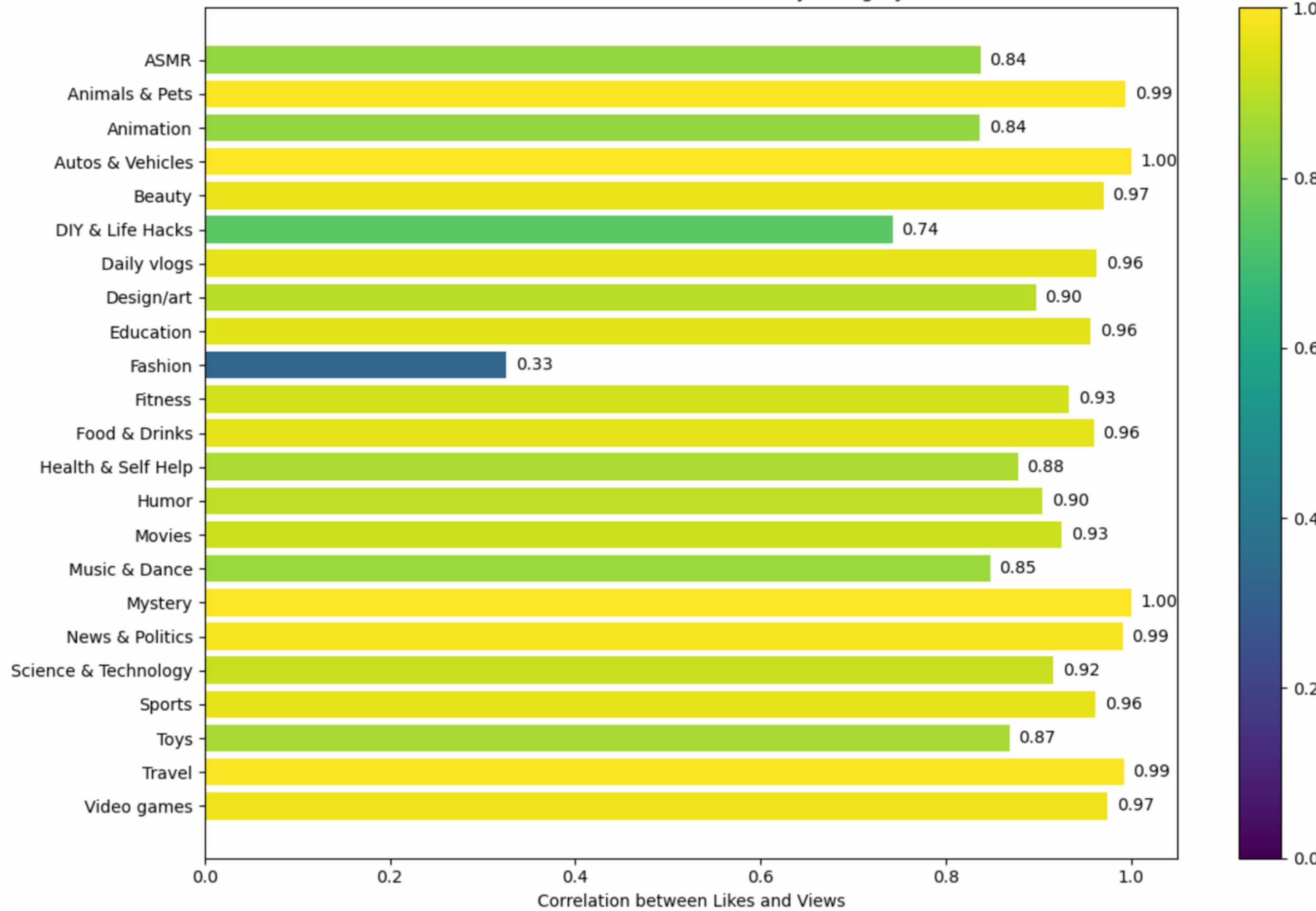
Video Games, Movies, and Music & Dance are more likely to interact with content by giving likes.

Categories with high engagement (e.g., **Video Games, Science & Technology, Autos & Vehicles**) tend to involve **more interactive and emotionally engaging content**, which leads to more viewer interaction.

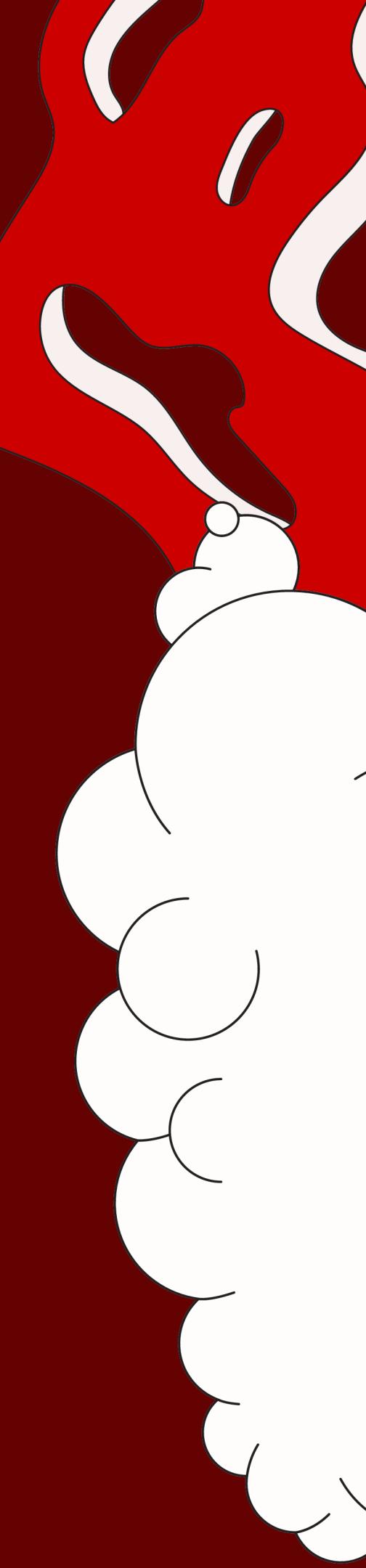
On the other hand, content in categories like **News & Politics and Health & Self Help** tends to attract fewer likes, suggesting lower engagement despite potentially high view counts.

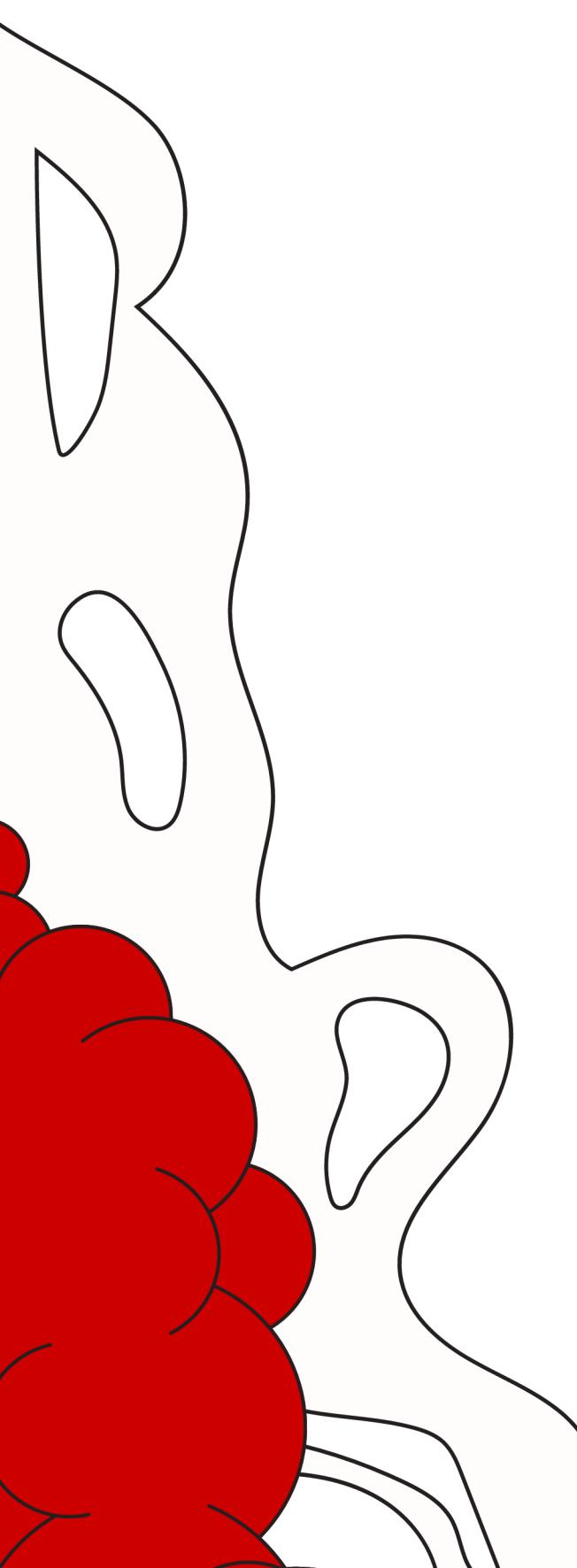
Correlation

Correlation of Likes and Views by Category



Implications for Stakeholders





1. Content Creators and Producers

Implication: Understanding which content drives engagement can help creators refine their content to better suit audience preferences, potentially increasing viewer satisfaction and retention

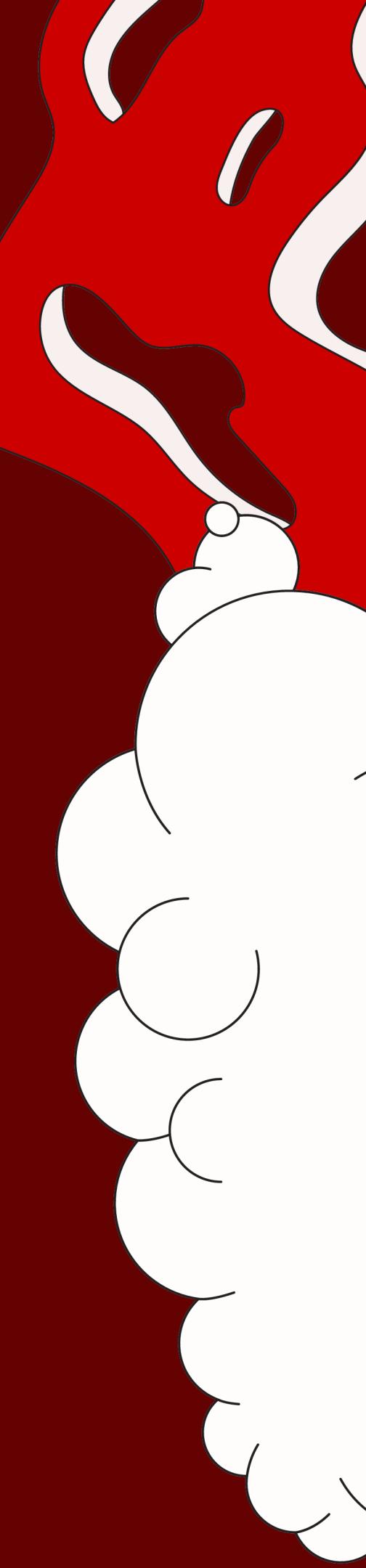
2. Advertisers and Marketers

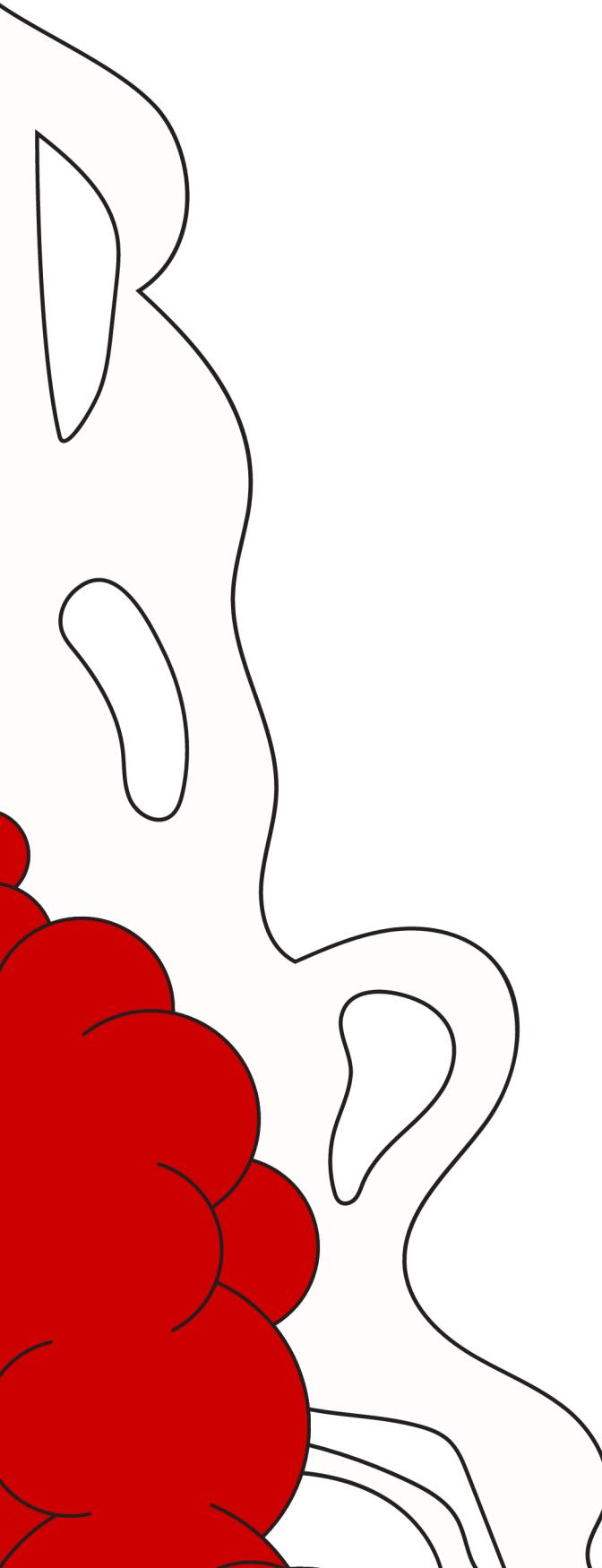
Implication: Marketers can target advertising more effectively by placing ads in content that not only draws views but also engages viewers, as evidenced by high like-to-view ratios.

3. Audience Groups (Viewers)

Implication: As primary consumers of content, viewers benefit indirectly when content creators and platforms use engagement data to improve content quality and relevance.

Ethical, Legal, and Societal Implications





1. Engagement Metrics Manipulation

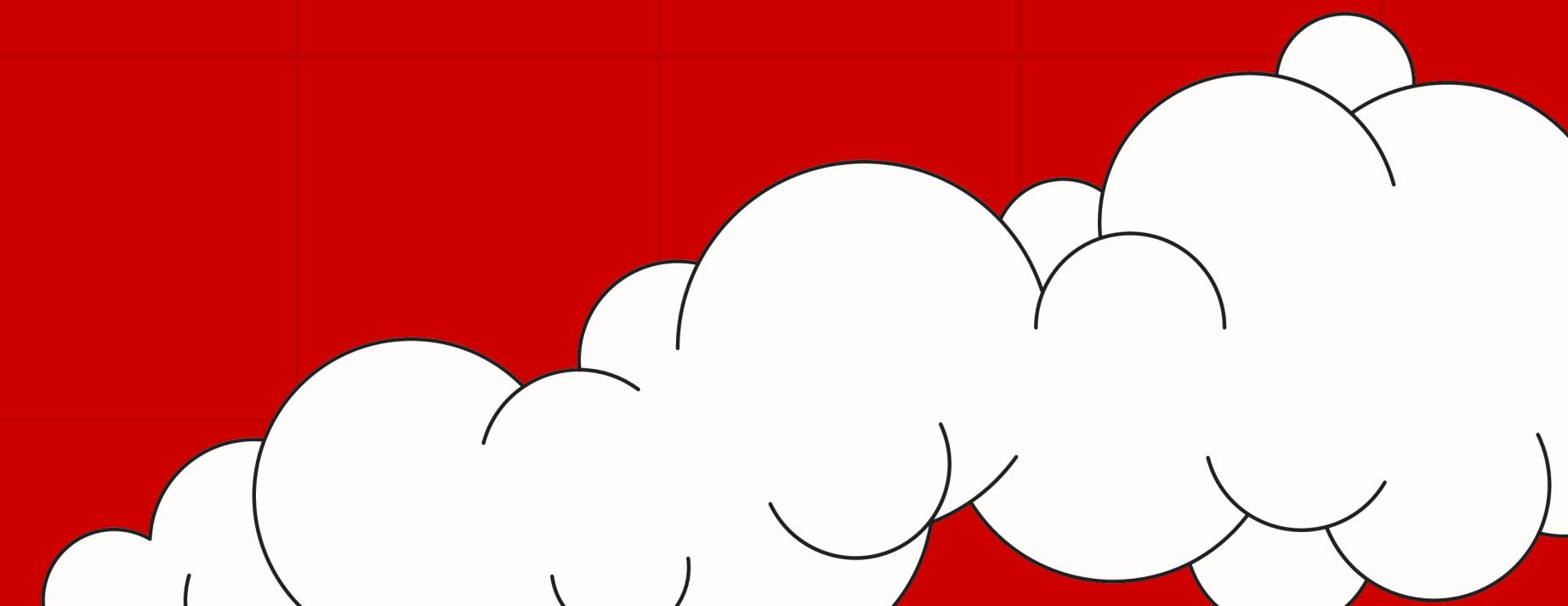
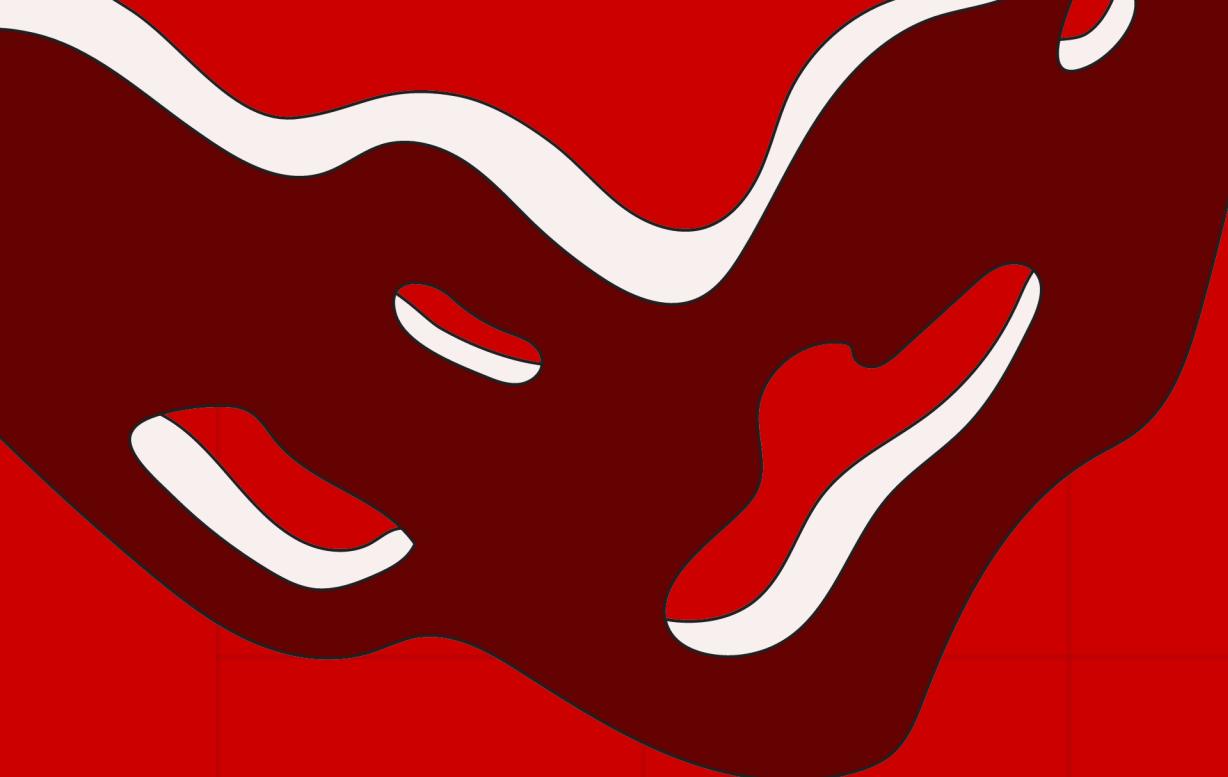
There's a risk of people using unfair methods to boost their likes and views. Platforms need to set rules to prevent such practices to maintain trust.

2. Transparency in Content Promotion

It's legally important for platforms to be clear about how they use data to recommend videos or posts to users.

3. Misinformation Spread

High-engagement content isn't always truthful or helpful. Platforms need to actively manage and moderate content to prevent the spread of false information.



Thank you
for listening