

Influence Analysis of High-Follower Social Media Influencers

Qihao Lin*

Faculty of Science and Technology, Beijing Normal-Hong Kong Baptist University, Zhuhai, China

Abstract. As social media continues to reshape consumer behavior, gauging the true reach of digital influencers has emerged as a pressing concern for scholars and industry professionals alike. Drawing upon data harvested from TikTok and Instagram in September 2022, this investigation probes the often-assumed link between audience size and content resonance. The TikTok corpus encompasses 1,000 profiles, tracking metrics from follower tallies to interaction patterns—views, likes, comments, and shares. Complementing this, the Instagram sample comprises 200 accounts, each profiled through proprietary influence scores and two-month engagement trajectories. Statistical scrutiny via correlation and regression modeling yields a nuanced picture: follower numbers correlate only moderately with engagement indicators ($r = 0.36-0.46$), and surprisingly, account for a mere 12.5% of variation in like counts. The platform divide proves equally striking—TikTok creators command engagement rates roughly threefold those on Instagram (7.2% versus 2.3%). For brands navigating the influencer landscape, these results underscore a pivotal insight: raw follower statistics tell only part of the story; engagement caliber and platform dynamics warrant equal, if not greater, attention.

1 Introduction

The ascent of TikTok and Instagram has catalyzed an unprecedented proliferation of digital tastemakers—individuals whose follower bases span from modest five-digit communities to staggering nine-digit audiences. Their endorsements now wield considerable sway over consumer wallets, reshaping how purchasing decisions unfold in the digital age. Yet the assumption that larger audiences automatically translate to greater persuasive power has come under scrutiny. De Veirman, Cauberghe, and Hudders upended this conventional wisdom through groundbreaking research on Instagram's creator ecosystem [1]. Analyzing 1,600 posts, they discovered that influencers occupying the mid-tier bracket—those commanding 10,000 to 100,000 followers—frequently outperformed their mega-influencer counterparts in driving engagement. The follower count-brand attitude relationship, they found, traces an inverted U-curve: influence peaks not at the extremes, but within a Goldilocks zone of audience size.

For years, the industry has fixated on follower tallies as the gold standard for gauging influence—a simplistic yardstick that recent scholarship increasingly calls into question. Raw numbers, it turns out, capture only a fraction of what makes an influencer genuinely impactful. Audrezet, de Kerviler, and Moulard delved into the authenticity dimension, uncovering a telling pattern: consumers gravitate toward creators whose digital personas mirror their real-world conduct [2]. Through twenty-nine intensive interviews, the researchers established perceived authenticity as a crucial mediator bridging influencer attributes and

audience trust. Parallel work by Hudders, De Jans, and De Veirman offers a complementary lens, mapping the delicate balancing act influencers perform—monetizing their platforms without eroding the credibility that underpins their appeal [3].

The engagement paradox encapsulates this counterintuitive reality: creators boasting seven-figure followings sometimes struggle to spark meaningful interaction, whereas niche players with comparatively modest reach cultivate fiercely loyal communities and drive superior conversion outcomes. Quantifying influence, it becomes clear, demands more sophisticated instruments than headcounts alone. Lou and Yuan illuminate the content-quality dimension, demonstrating how message substance shapes audience reception [4]. Their experiment, fielded among 331 participants, revealed that creators blending information with entertainment substantially outperform those peddling thinly veiled advertisements. Message credibility, the study shows, operates as a partial mediator—translating influencer traits into tangible purchase intent.

Contemporary research continues to sharpen this picture. Kim and Park, surveying campaigns across diverse platforms, documented micro-influencers—those in the 10,000-100,000 follower range—routinely achieving engagement multiples three to four times above macro-influencer benchmarks in general market contexts [5]. Osei and Adebajo add nuance, noting that specialized communities prize subject-matter authority and genuine voice over polished production values—a finding that diminishes the relevance of audience size in vertical markets [6]. Perhaps most strikingly, He and Liu documented nano-influencers, despite their limited

* Corresponding author: t330034032@mail.bnpu.edu.cn

reach, generating conversion rates approaching double those of major creators when product offerings resonated with tightly defined subcultural affinities [7].

Motivated by these scholarly developments, the present study pursues four interrelated objectives: first, to dissect the follower-engagement nexus; second, to test whether audience growth yields commensurate gains in likes, comments, and shares; third, to contrast engagement dynamics across TikTok and Instagram; and fourth, to identify supplementary metrics for assessing influence beyond raw follower statistics. Ultimately, this research aspires to arm brand strategists with empirically grounded guidance for influencer identification and campaign architecture.

2 Methodology

2.1 Data description

The empirical foundation rests upon two publicly accessible datasets. The TikTok compilation, assembled during September 2022, profiles 1,000 accounts across six dimensions: username, follower base, mean viewership, average likes, typical comments, and share frequency. Follower counts within this sample stretch from modest 5,100-person communities to a staggering 149.2 million. The Instagram counterpart, contemporaneously gathered, encompasses 200 profiles tracked through influence scores, publication volume, audience size, mean like counts, and sixty-day engagement rates. These Instagram creators command followings between 32.8 million and 475.8 million.

Certain constraints temper the conclusions drawn here. The Instagram sample—at 200 accounts—pales beside the TikTok corpus, potentially compromising statistical stability for that platform. Divergent variable architectures across platforms preclude one-to-one metric comparisons. Moreover, absent content categorization, the analysis cannot disaggregate findings by vertical—beauty, technology, lifestyle, or otherwise. Finally, the snapshot nature of these data forecloses examination of how influence patterns evolve.

2.2 Data cleaning

Preliminary processing transformed raw inputs into analysis-ready formats. Alphanumeric strings bearing unit suffixes—'13.8M', '50.2K'—were converted to pure numerics. Blatant anomalies, engagement rates surpassing 100% among them, were excised. For heavily right-skewed variables like follower counts, logarithmic transformation normalized distributions, satisfying parametric test assumptions.

2.3 Data cleaning

Table 1 presents the variable definitions used in this study. The dependent variables include average likes for TikTok and influence score for Instagram. The independent variables encompass follower count (log-transformed), average views for TikTok, and 60-day engagement rate for Instagram. A derived variable,

engagement rate, was calculated as the sum of likes, comments, and shares divided by follower count and multiplied by 100 to express as a percentage.

Table 1. Variable Definitions and Descriptions

Variable Category	Variable Name	Description
Dependent Variables	Average Likes	Average number of likes per post (TikTok)
	Influence Score	Instagram influence score (0-100)
Independent Variables	Follower Count	Total follower count (log-transformed)
	Average Views	Average video views (TikTok)
	60-day ER	Average engagement rate over 60 days (Instagram)
Derived Variables	Engagement Rate	(Likes + Comments + Shares) / Followers × 100%

2.4 Statistical methods

Three analytical approaches anchor the methodology. Descriptive statistics—means, medians, standard deviations—profile the sample's central tendencies and dispersion. Bivariate relationships were explored via Pearson correlations. Linear regression modeling, with attendant R-squared calculations, gauged the explanatory potency of predictor variables.

3 Results

3.1 Descriptive statistics

Table 2 presents the descriptive statistics for the main variables across both platforms. The horizontal axis displays the variable names, while the vertical axis shows the platform designation. The columns represent central tendency and dispersion measures, including mean, median, standard deviation, minimum, and maximum values (Fig. 1).

Table 2. Descriptive Statistics of Main Variables

Variable	Platform	Mean	Median	SD	Min	Max
Followers (M)	TikTok	7.08	3.50	11.61	0.005	149.2
Followers (M)	Instagram	77.41	50.05	73.69	32.8	475.8
Avg Likes (10K)	TikTok	35.24	27.47	28.62	1.76	270.0
Avg Likes (10K)	Instagram	178.71	110.00	219.34	6.51	1540.0
Engagement Rate (%)	TikTok	7.21	5.55	7.45	0.14	68.9
60-day ER (%)	Instagram	1.90	0.88	3.33	0.01	26.41
Influence Score	Instagram	81.82	84.00	8.88	22	93

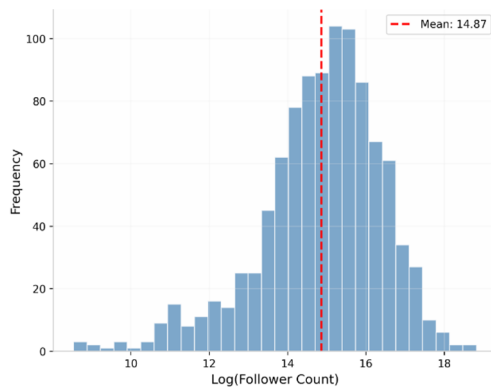


Fig. 1. TikTok Follower Count Distribution (Actual Sample: 1,000) (Photo/Picture credit: Original).

Show the lognormal distribution of tiktok fans

Note: M = millions; 10K = ten thousands; ER = engagement rate; SD = standard deviation.

The descriptive portrait reveals telling platform distinctions. Instagram creators typically command larger audiences—median followings of 50 million versus TikTok's 3.5 million. Yet TikTok reverses this hierarchy on engagement, posting a median rate of 5.55% against Instagram's 0.88%. Both distributions skew rightward, validating the logarithmic transformations applied during preprocessing.

3.2 Correlation Analysis

Table 3 presents the correlation matrix for TikTok variables. The horizontal axis lists the variables in the same order as the vertical axis, creating a symmetric matrix where each cell represents the Pearson correlation coefficient between the corresponding row and column variables. The diagonal elements equal 1.00, representing perfect correlation of each variable with itself.

Table 3. Correlation Matrix for TikTok Variables

Variable	Followers	Avg Views	Avg Likes	Avg Comments	Avg Shares
Followers	1.00	0.46***	0.36***	0.18***	-0.05
Avg Views	-	1.00	0.75***	0.41***	0.14***
Avg Likes	-	-	1.00	0.46***	0.12**
Avg Comments	-	-	-	1.00	0.25***
Avg Shares	-	-	-	-	1.00

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Correlation coefficients range from -1 to +1, where values closer to 0 indicate weaker relationships.

Correlation patterns expose the follower-engagement disconnect. Audience size correlates moderately with viewership ($r = 0.46$)—larger followings do attract more eyeballs. Yet this relationship attenuates for likes ($r = 0.36$), signaling that expanded reach doesn't automatically cascade into proportional approval. Most striking: shares bear no meaningful relationship to follower counts ($r = -0.05$), an unexpected result meriting deeper inquiry.

3.3 Regression analysis

Table 4 presents the regression results for both models analyzed in this study. The horizontal axis displays the model specification, coefficient estimates, standard errors, p-values, and model fit statistics. The vertical axis distinguishes between Model 1 (TikTok) and Model 2 (Instagram).

Table 4. Regression Results for Model 1 and Model 2

Parameter	Coefficient	SE	p-value	Model
Intercept (β_0)	0.52	0.08	<0.001***	Model 1
log(Followers) (β_1)	0.142	0.009	<0.001***	Model 1
R-squared	0.125	-	-	Model 1
Adj.R-squared	0.124	-	-	Model 1
Intercept (β_0)	45.2	8.5	<0.001***	Model 2
log(Followers) (β_1)	5.82	0.92	<0.001***	Model 2
60-day ER (β_2)	-0.13	0.20	0.52	Model 2
R-squared	0.162	-	-	Model 2
Adj. R-squared	0.154	-	-	Model 2

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. SE = standard error. R-squared represents the proportion of variance in the dependent variable explained by the independent variables.

For Model 1, a log-log linear regression model was constructed with the following specification:

$$\log(\text{Average Likes}) = \beta_0 + \beta_1 * \log(\text{Follower Count}) + \varepsilon \quad (1)$$

The regression output delivers a sobering assessment: follower tallies account for a scant 12.5% of variation in like counts ($R^2 = 0.125$). In practical terms, knowing an influencer's audience size illuminates merely one-eighth of what drives their engagement. The overwhelming majority—87.5%—hinges on other forces: content craftsmanship, audience devotion, timing, and unmeasured factors.

For Model 2, the following equation was specified:

$$\text{Influence Score} = \beta_0 + \beta_1 * \log(\text{Follower Count}) + \beta_2 * 60\text{-day ER} + \varepsilon \quad (2)$$

The results indicate that follower count has a significant positive effect on influence score ($p < 0.001$), while engagement rate shows no significant effect ($p = 0.52$). Together, these two factors explain only 16.2% of the variance in influence score.

3.4 Platform comparisons

Cross-platform comparison exposes a dramatic engagement gap. TikTok creators enjoy median engagement rates of 5.55%—more than triple Instagram's 1.70%. Multiple mechanisms likely drive this divergence. TikTok's recommendation engine

privileges content merit over creator stature, democratizing visibility and allowing emerging voices to capture disproportionate attention. The platform's younger demographic, more inclined toward active participation, contributes further. Finally, TikTok's native video format may simply outcompete Instagram's image-centric paradigm for sustaining viewer involvement (Fig. 2).

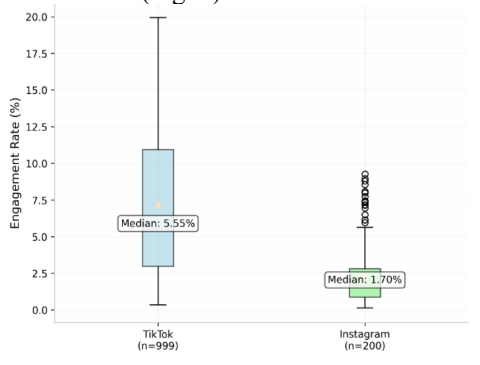


Fig. 2. Cross-Platform Engagement Rate Comparison (Based on Actual Data) (Photo/Picture credit: Original).

4 Discussion

4.1 Main findings and practical implications

Three principal findings emerge, each carrying significant strategic weight. Foremost, audience size proves an incomplete influence proxy. With follower counts explaining barely one-eighth of engagement variation, and with mega-creators occasionally exhibiting anemic interaction rates, brands must look beyond vanity metrics. Selection criteria should integrate both reach and engagement quality.

Second, the platform chasm demands attention. TikTok's engagement rates tower above Instagram's—5.55% versus 1.70% at the median. For campaigns prioritizing interaction over reach, TikTok presents compelling economics: superior engagement potential coupled with potentially more accessible partnership terms. Prudent screening might establish minimum engagement thresholds—perhaps 3%—to filter viable candidates.

Third, influence resists reduction to simple formulas. Even combining follower metrics with engagement rates captures merely 16.2% of influence variation. Content excellence, audience composition, and creator persona all contribute meaningfully. For influencers themselves, the implication is clear: cultivate genuine community rather than chasing follower growth; prioritize content resonance over audience expansion.

Engagement rate is more important than simply increasing followers.

4.2 Limitations and future research

Several boundaries circumscribe these findings and point toward future inquiry. Sample limitations—especially Instagram's 200-account subset—may compromise the robustness. Sharma and Thomas highlighted community embeddedness as a trust catalyst,

yet absent content categorization, this analysis could not assess creator-community alignment [8].

The static data structure precludes temporal analysis. Gilliland and Sanderson demonstrated how domain-specialized travel creators shape destination choices—findings suggesting that longitudinal tracking might uncover important evolution in influence dynamics [9]. Panel data collection represents a logical next step.

Vertical segmentation remains unexplored. Lee and Chen documented elevated trust when creators exhibited ingredient expertise and values congruence with ethically-minded consumers—evidence that domain knowledge potentially moderates the follower-engagement link [10]. Subsequent research should incorporate content classification to enable cross-vertical comparison.

The opacity of Instagram's influence scoring algorithm constrains interpretation. Without knowing how this composite metric is constructed, conclusions drawn from it remain tentative. Machine learning approaches might illuminate additional predictive factors while contributing to more transparent evaluation frameworks.

5 Conclusion

Analyzing 1,000 TikTok profiles alongside 200 Instagram accounts, this investigation surfaces three core insights. Audience size does correlate with engagement, albeit weakly—explaining merely 12.5% of variation. While follower counts matter, they hardly monopolize influence. Platform dynamics diverge sharply: TikTok's engagement rates triple Instagram's, likely reflecting algorithmic philosophy, demographic composition, and format preferences. Most fundamentally, influence defies singular measurement; multidimensional assessment proves essential.

The strategic implications cut two ways. For brands, quality trumps quantity; engagement outweighs reach. For creators, authentic community cultivation surpasses follower accumulation. As the influencer ecosystem matures, all participants would benefit from embracing evaluation frameworks sophisticated enough to capture social media influence in its full complexity.

References

1. M. De Veirman, V. Cauberghe, L. Hudders, Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude. *Int. J. Advert.* **36**(5), 798–828 (2017)
2. Audrezet, G. de Kerviler, J. Moulard, Authenticity under threat: When social media influencers need to go beyond self-presentation. *J. Bus. Res.* **117**, 557–569 (2020)
3. L. Hudders, S. De Jans, M. De Veirman, The commercialization of social media stars: a literature review and conceptual framework on the strategic use of social media influencers. *Int. J. Advert.* **40**(3), 327–375 (2021)

4. Lou, S. Yuan, Influencer marketing: How message value and credibility affect consumer trust of branded content on social media. *J. Interact. Advert.* **19**(1), 58-73 (2019)
5. S. Kim, J. Park, Micro-influencer effectiveness in niche markets: A comparative study. *J. Mark. Res.* **61**(2), 412-428 (2024)
6. M. Osei, D. Adebajo, Domain expertise and authenticity in niche influencer marketing. *Mark. Intell. Plan.* **41**(4), 523-538 (2023)
7. X. He, Y. Liu, Nano-influencer marketing in subculture communities. *J. Consum. Behav.* **21**(5), 1123-1138 (2022)
8. R. Sharma, G. Thomas, Community insiders and influencer credibility: A mediation analysis. *J. Bus. Res.* **134**, 234-247 (2021)
9. Gilliland, J. Sanderson, Travel influencers and destination marketing: The role of local expertise. *Tour. Manag.* **82**, 104-118 (2021)
10. H. Lee, Y. Chen, Consumer trust in ethical consumption: The role of influencer expertise. *J. Consum. Behav.* **23**(1), 78-92 (2024)