Olivia Pivovar
May 3, 2025

Ensemble Data Engagement Analysis: Adidas Social Media

---

1. Overview of Approach

This project analyzes unstructured, raw social media data from Adidas' TikTok and Instagram accounts, provided via Ensemble's API exports. The goal was to extract meaningful engagement insights and provide actionable recommendations based on content performance. I began by cleaning and structuring the data, which involved parsing multiple nested JSON files, unzipping archives, and normalizing post-level fields such as captions, hashtags, timestamps, and performance metrics (likes, views, comments, shares).

The data was completely unprocessed, requiring careful inspection and design of extraction logic. For TikTok, JSON responses followed a consistent structure and were easily parsed. I was able to collect 20,171 valid TikTok posts. However, Instagram files came in mixed formats—some contained only error messages or empty structures, while others followed a nested data.posts[].node format. After developing logic to identify and extract from the usable files, I successfully collected 15,871 valid Instagram posts.

Once structured, I calculated derived engagement metrics and evaluated several content features to determine what drove audience interaction. These included posting time, caption length, and emoji presence. I prioritized aspects showing early promise during exploratory visualizations, balancing depth with feasibility.

Some elements of this approach (e.g., analyzing emoji use and caption length impact) were inspired by brainstorming supported with generative AI tools. However, all data validation, engineering, plotting, and interpretation were performed independently.

---

2. Key Insights

Platform Comparison

- Instagram significantly outperformed TikTok in raw engagement rate. Instagram's average rate was 137.02, while TikTok was only 0.0716. This disparity stems from how views dilute TikTok's engagement when normalized.
- Visualizations confirmed the disparity, with TikTok engagement curves appearing flat and Instagram showing strong variability.

Timing

- Instagram engagement spiked around 7 AM, and again between 6 PM and 9 PM, suggesting optimal post timing for visibility. Engagement rates were best for posts made on Saturday (186.51) and Thursday (240.08).
- TikTok showed minimal variation by hour; engagement rates remained near 0.07 regardless of post time. However, the most optimal time frame is between 7 AM and 9 AM. Engagement rates performed best on posts made on Monday (0.0873) and Wednesday (0.0889).

Caption Length

- Instagram posts with captions averaging 126.7 characters had the highest engagement, especially in the 100–150 character range.

- TikTok captions averaged 76 characters, and changes in length had negligible effect on engagement.

Emoji Use

- On Instagram, posts with emojis had an average engagement rate of 157.23, while posts without emojis averaged only 84.49 — a nearly 86% increase.
- On TikTok, posts with emojis had slightly lower engagement (0.0706 vs. 0.0772 for non-emoji posts).

Hashtags

- On Instagram, top-performing posts featured campaign-driven hashtags like #YouGotThis (1,600) and #adizero (680), suggesting effective brand messaging through product-specific and motivational tags.
- On TikTok, high-engagement posts frequently used hashtags like #adidas (5,017 times), #football (2,269), and #basketball (625) — reflecting the platform's strong alignment with Adidas' athletic and cultural branding.

---

3. Recommendations

1. Prioritize Instagram for content intended to maximize engagement, given its significantly stronger return per post.

2. Use emojis in Instagram captions to boost performance; they likely add visual/emotional appeal.

3. Post between 6–9 AM or 6–9 PM on Instagram for best reach.

4. Write slightly longer Instagram captions (100–150 characters) for better engagement.

5. On TikTok, since metadata has little impact, focus on trends and visual creativity to drive performance.

6. Use campaign-specific or sport-focused hashtags (e.g., #adizero, #football) to boost engagement, and apply them consistently across platforms.

---

4. Suggestions for Future Research

- Content Type Analysis: Classify posts by video/image format to test engagement by media type.
- Sentiment Analysis: Score captions and test if positive or emotional tone correlates with engagement.
- Hashtag Effectiveness: Analyze top-performing hashtags in both platforms.
- Cross-Platform Strategy: Track engagement on identical posts cross-posted to TikTok and Instagram.

Visual Supplement (Provided Separately)

A zip archive of comparison charts and platform-specific graphs is included with this report submission, supporting all insights discussed above.

Appendix A: Use of Generative AI Tools

Purpose of AI Use

Generative AI (ChatGPT) was used during this project to support ideation, brainstorming, and communication. The objective was to explore possible content features that might influence engagement and to refine the clarity and structure of written insights. All technical tasks, including data extraction, cleaning, transformation, analysis, and visualization, were created personally using Python and the raw Ensemble API data.

AI Prompting Strategy

AI support was primarily used in the early stages of the project to help brainstorm engagement drivers worth testing—such as emoji use, caption length, and time-of-day posting effects. It was also used during the writing phase to suggest ways to rephrase findings, align the tone of different sections, and structure the report more clearly.

Verification of AI Output

Any insights or suggestions provided by AI were critically reviewed and verified through my own analysis. I reviewed and edited all AI-generated output myself, ensuring that any language used aligned with my own interpretation and that any numbers or statistics mentioned accurately reflected the underlying data. Rather than using AI as a replacement for the work, I treated it as a collaborative tool — using its ideas as starting points and expanding on them through hands-on coding, visualization, and interpretation. I ensured final conclusions were shaped through my own processing, not automatically generated responses.

Examples of Prompts Used

Prompt 1: “What social media post attributes are most likely to affect engagement on TikTok and Instagram?”
AI Output: “You could analyze variables such as: caption length, posting time (hour or day), use of emojis, presence of hashtags, or content type (image vs. video). Comparing these features with engagement metrics like likes, comments, or views may reveal valuable patterns.”

Prompt 2: “Rephrase this recommendation to match a professional tone and formatting.”
AI Output: “Use campaign-specific or sport-focused hashtags (e.g., #adizero, #football) to boost engagement, and apply them consistently across platforms.”

Prompt 3: “Summarize a chart comparing engagement by hour across two platforms.”
AI Output: “Describe which hours perform best on each platform, highlight any overlap or contrasts, and mention if one platform shows more variability than the other. For example: ‘Instagram shows strong peaks around 7 AM and 6–9 PM, while TikTok engagement remains relatively flat across hours.’”