

Trend Identification

Hello, and welcome to your Scrollmark coding project! Please send back your completed project within the day you received it. You're free to spend as much time as you'd like, but you'll be evaluated as per expected work product in line with 90-120 minutes of work.

Problem Statement

You are free to use the entirety of the Internet, and any GPT tools are completely in-scope for usage.

A common problem for social media managers for large brands is listening to what their community wants and acting accordingly. Naturally, given that social media platforms are a noisy place, the data is also noisy.

Your challenge is to create a metrics-driven report of trends and topics of interest, and how they shift over time, for consumers of [@treehut](#), a popular skincare brand that's active on Instagram.

We've prepared a [corpus](#) of ~18,000 comments left on their posts in March 2025 containing the

- `timestamp`
- `media_id`
- `media_caption`
- `comment_text`

Put yourself in the shoes of a social media manager and process this data however you see fit to extract relevant data.

Evaluation Criteria

Here's what we're looking for:

1. **Actionable Insights & Visuals** - This is the most important deliverable. We want to see if you have the product sense to **surface insights beyond surface level** that social media managers would find useful. Quality > Quantity.
2. **Well-Organized Codebase** - Clear folder structure, README, meaningful commit history, no superfluous files
3. **Robust Data Pipeline** - Modular functions or classes for each stage, graceful error handling
4. **Thoughtful Extension Plan** - A good senior engineer makes their own scope. How would you expand the scope of this project if you had a month?

Deliverables

1. **Code Repository** (GitHub or ZIP) with:
 - **README .md** (see checklist below)
 - All source code + pinned dependencies
2. **Report** (PDF, markdown, notebook, or hosted app) containing:
 - **Executive summary** (1–2 paragraphs) of your approach & key findings
 - **Visualizations + insights** with captions
 - **Instructions** for running and extending your work
3. **Extension Proposal** (1 page or section in README):
 - Ranked features you'd build next and why.
4. **AI & Tool Usage Disclosure** - Which tools you leveraged