Proj notes

* Attempted to utilize OpenAI, Hugging Face, and Llama.
  + Ran into immediate issues with pricing and API calls. No great free tier options that I was able to access and query.
* Brought up the above issue to Dr. Khan and team during weekly meeting, advised to pursue building AI agent model with Gemini as the backing statistical model.
  + 6/14/2025 pursued this, was able to create a python script that contacted Gemini’s API endpoint and parsed TikTok to collect trends.
    - Utilized a “popularity score” metric to track what was trending
    - Was able to get a JSON output entirely reliant on Gemini’s “gemini-1.5-pro-latest” model

BELOW IS A GREAT EXPLANATION ON WHAT I’VE CURRENTLY ACHIEVED

**🔍 How the popularity\_score is Created**

Your code does **not compute** a popularity\_score through any API, numerical dataset, or real-time scrape. Instead, it **relies entirely on Gemini** to **hallucinate or simulate** that score based on:

1. **Your Prompt Design**:
   * You instruct Gemini to act like a “social media trend analyst.”
   * You ask for output in a **very specific JSON format**, including a "popularity\_score" field from 0–100.
2. **Gemini's Internal Knowledge & Pattern Matching**:
   * Gemini generates that score based on **training data** (e.g. millions of documents, web pages, trend recaps, and statistical approximations from its model training).
   * It recognizes that, for example, something like “Coastal Cowgirl Chic” might be surging and assigns it a higher score.
   * The number is **made up** in a plausible way based on what it "knows" from pretraining and any reinforcement tuning.
3. **No Real Data Connection**:
   * There’s no call to TikTok’s API, Google Trends, Trendhunter, etc.
   * It is **not live scraping** or consulting up-to-date quantitative signals unless you build that in separately.

**⚠️ Implications**

* **Pros**:
  + Super fast, zero external dependencies.
  + Reasonably aligned with public discourse—can help identify directionally plausible trends.
* **Cons**:
  + Scores are synthetic: not based on quantifiable metrics like views, mentions, or sentiment volume.
  + May fabricate trends if they sound realistic.
  + Scores aren’t comparable across queries unless you tightly control prompt format.

**✅ What You Could Do to Improve It**

If you want **actual popularity data**, consider:

* Integrating TikTok Trend APIs via 3rd parties (e.g., Trendpop, BuzzGuru).
* Scraping Google Trends or Twitter/X trending topics with timestamps.
* Combining Gemini's generated trend names with **search volume** or **engagement metrics** for hybrid scoring.