

# Project Proposal: YouTube Content Analysis

UCLA STATS418  
Stephanie Lu

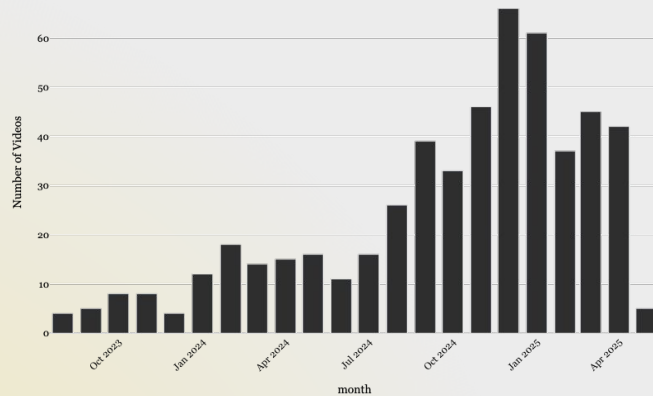
May 2025

# Dataset

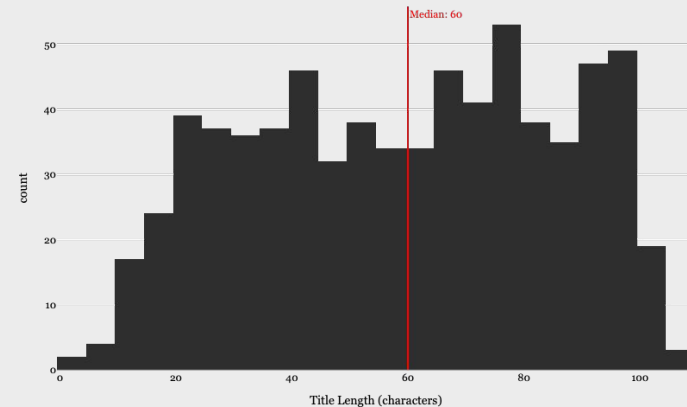
- **YouTube Data API v3** for video metadata from beauty & lifestyle creators
- Manually selected 15\* channels across **macro and micro** tiers
- Queried each creator's channel ID to retrieve their 50 most recent videos, including **titles, publish dates, and video IDs.**
- Additional API calls to collect **view counts, like counts, and comment statistics**
- Parsed and aggregated the results into a structured dataset using pandas for downstream analysis.

# EDA

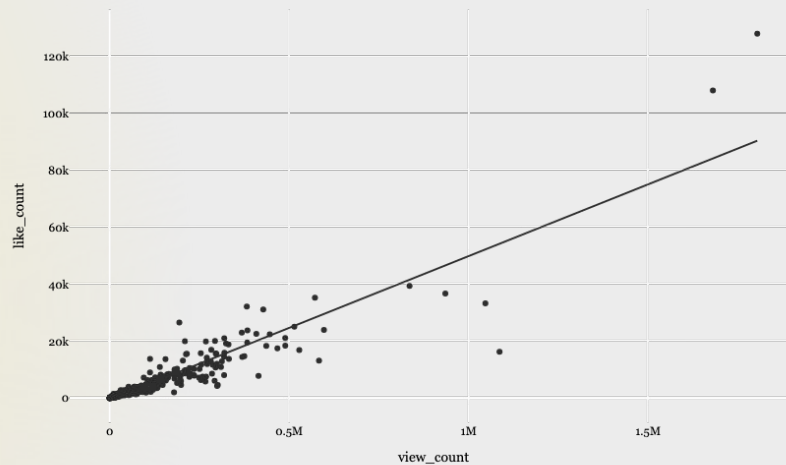
Monthly Video Uploads



Distribution of Video Title Lengths



View Count vs. Like Count



# Proposal

## YouTube Title Rewriter

- Analyze patterns in high-performing YouTube video titles across beauty and lifestyle creators
- Would allow users to input a draft video title and return alternatives to optimize for engagement
- Simple NLP and rule-based heuristics
- To support creators in writing attention-grabbing titles that improve click-through and views