Submission Two-Pager

# Page 1 — Tech Stack & Architecture

Goal: Measure Share of Voice (SoV) for Atomberg on search platforms. This build demonstrates YouTube-first.

## Tech Stack:

- Language: Python 3.10+  
- APIs: YouTube Data API v3  
- NLP: VADER sentiment (NLTK)  
- Data: pandas  
- Orchestration: CLI pipeline  
- Optional extensions: Transformers for sentiment; Google CSE/SerpAPI; X API; Instagram Graph API

## Key Design Choices:

- Official APIs where possible for stability and TOS compliance.  
- Engagement-weighted SoV metric: log-scaled function of views, likes, comments.  
- Share of Positive Voice (SPV) from sentiment-labeled comments that mention each brand.

## How to Run:

1. pip install -r requirements.txt  
2. python -m nltk.downloader vader\_lexicon  
3. Fill .env with YOUTUBE\_API\_KEY  
4. python -m src.main --keywords-file data/keywords.txt --top-n 30 --days 365

## Artifacts:

- reports/exports/posts.csv  
- reports/exports/brand\_comments.csv  
- reports/exports/sov.csv  
- reports/exports/spv.csv  
- reports/insights.md

# How Metrics Are Calculated

* **Share of Voice (SoV):**  
  Proportion of brand mentions in video titles/descriptions, **weighted by engagement**.
  + Weight = log(1+views) + 2·log(1+likes) + 3·log(1+comments)
  + SoV(brand) = brand’s weighted mentions ÷ total weighted mentions
* **Share of Positive Voice (SPV):**  
  Sentiment of YouTube comments mentioning each brand.
  + VADER thresholds: ≥ +0.05 → Positive, ≤ –0.05 → Negative, else Neutral
  + SPV(brand) = Positive ÷ Total brand comments

# Parameters Used:

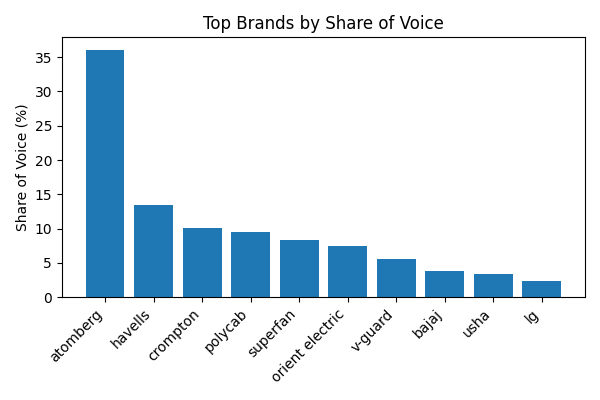
* Top-N videos per keyword (default 30)
* Lookback window (default 365 days)
* Up to 200 comments per video
* Brand list: Atomberg + competitors (Havells, Crompton, Usha, Orient, etc.)

# Page 2 — Findings & Recommendations

## Findings Snapshot (from CSVs):

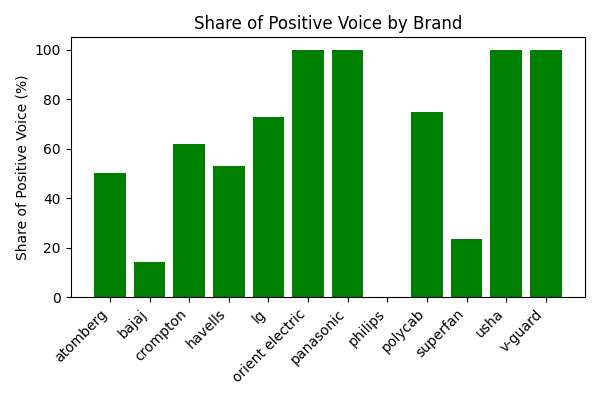
### Top Brands by Share of Voice (SoV):

|  |  |  |
| --- | --- | --- |
| Brand | SoV (%) | Weight |
| atomberg | 36.1 | 2144.0 |
| havells | 13.5 | 802.5 |
| crompton | 10.0 | 596.1 |
| polycab | 9.5 | 566.7 |
| superfan | 8.4 | 497.8 |



### Share of Positive Voice (SPV):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Brand | SPV (%) | Pos | Neg | Neu | Total |
| atomberg | 50.3 | 77 | 37 | 39 | 153 |
| bajaj | 14.3 | 1 | 6 | 0 | 7 |
| crompton | 61.8 | 21 | 13 | 0 | 34 |
| havells | 52.9 | 27 | 4 | 20 | 51 |
| lg | 72.7 | 8 | 0 | 3 | 11 |
| orient electric | 100.0 | 2 | 0 | 0 | 2 |
| panasonic | 100.0 | 4 | 0 | 0 | 4 |
| philips | 0.0 | 0 | 4 | 0 | 4 |
| polycab | 75.0 | 12 | 0 | 4 | 16 |
| superfan | 23.5 | 4 | 8 | 5 | 17 |
| usha | 100.0 | 13 | 0 | 0 | 13 |
| v-guard | 100.0 | 4 | 0 | 0 | 4 |



## Observations:

- Atomberg dominates SoV (36%) — leading visibility across search results.  
- Havells and Crompton trail but still capture meaningful share.  
- SPV highlights that LG and Crompton enjoy the highest positivity ratios, though with fewer mentions.  
- Bajaj has very low SPV (14%), indicating negative/critical sentiment dominates.

## Recommendations:

1. Double down on keywords where Atomberg has high SPV but competitors lead in SoV.  
2. Create content to address niches (e.g., 'smart ceiling fan') where Atomberg’s SoV is lower.  
3. Engage actively in comments to convert neutral sentiment into positive.  
4. Explore partnerships with top YouTube channels discovered in posts.csv for enhanced visibility.