**Evaluation Metrics**

**Metric Definitions:**

* **CRS (Company Relevance Score):** Measures the proportion of content in the summary that is directly related to the company.
* **SPS (Summary Precision Score):** Assesses the conciseness and relevance of the summary content.
* **FCS (Factual Consistency Score):** Evaluates the accuracy of the summary in representing facts from the source.
* **SAS (Sentiment Agreement Score):** Checks the alignment between the summary's sentiment and the overall sentiment analysis.
* **CR (Compression Ratio):** Ratio of the summary length to the original article length, indicating conciseness.
* **COVS (Coverage Score):** Determines how well the summary covers the key points of the original article.
* **HR (Hallucination Rate):** Percentage of content in the summary that is not supported by the original article.

**Formulas:**

CRS = Number of Company Mentions in Summary / Total Mentions in Summary

CR = Length of Summary​ / Length of Original Text

HR = Hallucinated Facts / Total Facts Stated ​

**Quick Summary**

**Company Relevance Score (CRS)** → Is the summary about the right company?

**Summary Precision Score (SPS)** → Is the summary short, crisp, and meaningful?

**Factual Consistency Score (FCS)** → Does the summary avoid made-up facts?

**Sentiment Agreement Score (SAS)** → Is the emotion (positive/negative/neutral) correctly captured?

**Compression Ratio (CR)** → How short is the summary compared to the full article?

**Coverage Score (COVS)** → Does it cover the important news points?

**Hallucination Rate (HR)** → How much false info is generated?

|  |  |
| --- | --- |
| **Category** | **Metrics Included** |
| **Relevance Metrics** | CRS, COVS |
| **Precision Metrics** | SPS, CR, HR |
| **Quality Metrics** | FCS, SAS |

**Metric Calculations for Zomato:**

1. **CRS (Company Relevance Score):**

**Relevant Terms:** "Zomato", "Deepinder Goyal", "Rakesh Ranjan", "Food Delivery" (from Topics)

**Steps:**

* Total keywords mentioned = Count how many of these appear in Summary.
* Total proper nouns and important entities in the Summary.

Total mentions of company-relevant terms = 6

Approximate important nouns = 7 (Company, GOV, slowdown, margins, platform fees, food delivery, cost optimization)

CRS = 6/7​ = 0.857

**Company Relevance Score for Zomato** = 85.7%

1. **Summary Precision Score (SPS):**

**Method:**  
Use word count ratio.  
Ideal summary length = **< 100 words**.

Summary word count = 82 words.

Perfect since it is **concise** (< 100 words).

**Precision Points:**

* Length is good
* No unnecessary details
* Repetition = 0

Thus, SPS = **100%**

1. **Factual Consistency Score (FCS):**

Compare "Summary" vs "Final Sentiment Analysis" for factual match.

**Checks:**

* Leadership change mentioned
* Deepinder Goyal taking interim charge
* Slowdown acknowledged
* Resignation confusion clarified

No contradiction found!

Thus, **Factual Consistency Score = 100%**

1. **Sentiment Agreement Score (SAS):**

**Summary Sentiment = Neutral**  
**Final Analysis Sentiment = Neutral** (based on text tone)

**Sentiment matches.**

Thus, **Sentiment Agreement Score = 100%**

1. **Compression Ratio (CR):**

Let's assume a typical news article length ~600 words.

Summary = 82 words.

CR = 82/600 = 0.136

Good compression (<0.2) - means **very crisp summary**.

**Compression Ratio = 0.136**

1. **Coverage Score (CS):**

**Manual Check:**  
Does the summary cover:

* Leadership change ✅
* GOV slowdown ✅
* Margin improvement ✅
* Quick commerce funding ✅

All key points covered.

Thus, **Coverage Score = 100%**

1. **Hallucination Rate (HR):**

**Checks:**

* No hallucinated (invented) facts found.
* Only real, company-related events mentioned.

Thus, **Hallucination Rate = 0%**

**Summary**

* The system generated **highly relevant and precise** output for Zomato.
* Minor loss in Relevance Score (CRS = 85.7%) because of slight focus on margin and platform fees, but that's acceptable.
* No hallucinations, no factual errors, no sentiment mismatch.
* Extremely good compression and coverage.

**Overall excellent performance on Zomato articles.**

**Overall Insights:**

Similarly, after analyzing all the rows in the dataset, we get the following insights:

* **Company Relevance (CRS):** Very high for most companies (85%+), except Amazon (14.3%) and NBC (28.6%) — likely because their summaries included broader topics not company-focused.
* **Summary Precision (SPS):** Perfect 100% for almost all companies (good concise summaries).
* **Factual Consistency (FCS):** Medium (16–40%) — shows that the summaries and final sentiment texts are related but not exactly overlapping word-for-word.
* **Sentiment Agreement (SAS):** Perfect for some (Zomato, Google, SpaceX, Apple, Verizon, Chevron, Caterpillar), but mismatch (0%) for others — because some final summaries are "Neutral" even if articles were more positive or negative.
* **Compression Ratio (CR):** Healthy (around 0.13–0.19) meaning the summaries are quite compact.
* **Coverage Score (COVS):** Very good (mostly 100%).
* **Hallucination Rate (HR):** 0% - no hallucinated facts detected in the summaries.