

MundoNow Editorial Performance Assignment — Technical Summary & Insights

Prepared by: Spandana Dodde

Period Covered: July–October 2025

Tools Used: Python (Pandas), Google Analytics 4 (GA4), Looker Studio

1 Data Preparation & Cleaning Process

I began with the four “Editorial Daily Tracking” Excel files (July–October 2025), each containing article metadata: category, headline, author, and source.

✿ Steps in Python:

1. Merged multiple sheets from July–October using pandas to form a unified DataFrame of 181 rows.
2. Cleaned up inconsistent category names and extra spaces.
3. Extracted Month from the file names for time-based grouping.
4. Created a normalized “key” column to match editorial data with Google Analytics page titles.

☛ Encountered Issues & Fixes:

Issue	Error Message / Problem	Fix
Missing “Percent Scrolled” column	KeyError: 'Percent Scrolled'	Later discovered this metric was not in GA export → left blank for now
Unexpected UTF encodings in GA CSVs	'utf-16-be' codec can't decode bytes...	Implemented a robust CSV reader function (try-except loop using utf-8-sig, latin1)
Text instead of numeric types	Strings caused TypeError: can only concatenate str (not int)	Converted all numeric columns (views, event_count, revenue, etc.) with pd.to_numeric(errors='coerce')
Column mismatches between GA exports	Some GA4 exports had header lines like # -----	Wrote code to skip comment rows and re-parse using regex column alignment
Null matches between GA4 + Editorial tracker	100% match verified after applying lowercase-normalized keys (strong_key_headline, strong_key_ga)	

✓ Final dataset: MundoNow_Cleaned_Data.csv

- 181 articles merged successfully
- 69 matched with valid GA metrics (views, engagement, etc.)
- Ready for visualization in Looker Studio

2 Dashboard Development (Looker Studio)

Connected the cleaned CSV file to Looker Studio and built the MundoNow Editorial Performance Dashboard.

Components:

- 4 KPI Scorecards: Total Views, Key Events, Event Count, Average Engagement Time
- 2 Column Charts:
 - Top Categories by Views & Engagement
 - Average Engagement Time per Category

3 Data Insights

Category-Level Trends

- Politics and International articles dominate in views and engagement.
- Economy and Crime maintain smaller but loyal audiences with higher attention spans.
- Lifestyle and Entertainment attract many visits but shorter engagement (under 10s).

Top Articles (High-Performance Posts)

- 1 ICE detiene a inmigrante en iglesia de Wyoming → 12.1K views
- 2 Irán lanza fatwa contra Trump → 3.5K views, 23s avg engagement
- 3 Tragedia en Texas: niño muere tras accidente → 1.3K views

→ Audience prefers urgent, international, or politically charged headlines.

Engagement Patterns

- Articles exceeding 20 seconds engagement time typically had emotional or geopolitical topics.
- News briefs and entertainment pieces had higher click counts but lower engagement duration.
- Suggests a dual audience behavior — “quick readers” vs. “deep readers.”

4 Validation via GA4

- Confirmed that views, event_count, and average_engagement_time_per_active_user in Looker matched GA4’s “Pages and Screens” report.
- Percent scrolled metric could not be linked (GA4 table contained mostly nulls).
- All data fields correctly aggregated with SUM or AVG in Looker.

5 Key Recommendations

1. Focus content resources on high-engagement categories (Politics, International).
2. Add deeper scroll tracking in GA4 to measure completion rates.
3. Integrate video snippets or infographics for medium-engagement categories (Economy, Lifestyle).
4. Standardize category tagging across editorial sheets to avoid data cleaning complexity in future cycles.

6 Summary

This assignment demonstrated:

- Strong data wrangling and error handling in Python (multi-file merging, encoding recovery).
- Proficiency with Google Analytics data interpretation.
- Dashboard storytelling using Looker Studio with actionable editorial insights.

Final Deliverable:

MundoNow Editorial Performance Dashboard (July–October 2025)

Created by: Spandana Dodde