

FULL DATA CLEANING & ENHANCEMENT SUMMARY

◆ 1. Initial Cleanup of Raw Data


The dataset contained social media campaign data with fields like ad performance, demographics, and conversions. The initial steps ensured the data was usable, clean, and structured for analysis.

✔ Steps Taken:

- **Date Conversion:** Converted `reporting_start` and `reporting_end` from text to proper date format for filtering and time-based insights.
- **Null Handling:** Fields like `total_conversion` and `approved_conversion` had missing values — these were handled using conditional logic to prevent calculation errors.
- **Data Types Standardization:** Ensured numeric columns (`impressions`, `clicks`, `spent`) were stored as proper numerical types.
- **Duplicate Checks:** Checked for duplicates in `ad_id` and ensured it remained a unique identifier.
- **Categorical Cleanup:** Cleaned `gender` and `age` values for consistency in filters (e.g., standardizing case, formatting).

+ 2. Derived KPI Columns Created

To make the dataset **dashboard-ready** and replicate professional marketing insights, new columns were added using calculated logic.

 Column	Formula	Benefit/Purpose
CTR (Click Rate)	<code>clicks / impressions</code>	Measures how often people clicked after viewing ad
CPC (Cost/Click)	<code>spent / clicks</code> (only if clicks > 0)	Shows cost-efficiency per user click
Conversion Rate	<code>total_conversion / clicks</code> (only if clicks > 0)	Helps evaluate how well clicks converted to actions
Approval Rate	<code>approved_conversion / total_conversion</code>	Measures quality of conversions (e.g., valid leads)
Campaign Day	Date part of <code>reporting_start</code>	Useful for filtering/reporting by campaign date
Engagement Score	<code>clicks + total_conversion + approved_conversion</code>	Composite score to identify top-performing ads

Each metric helps **visualize campaign effectiveness**, segment performance by age/gender/region, and compare campaigns across time.

🏆 Outcome

After cleaning and enhancement, your dataset is now:

- Fully structured for **Power BI dashboarding**
- Rich with **actionable KPIs**
- Ready for use in **visual storytelling** to support marketing decisions