Maria Cardillo

Google Ads Dataset - Data Journal

August 16th, 2025

Dataset: Google Ads Sales Dataset (Kaggle)

Dataset

For this project, I'm using a dataset containing daily advertising performance metrics for a single location. The data includes a range of KPIs like clicks, conversions, sale amounts, cost, spend, ROAS, and day-of-week breakdowns. Each row represents aggregated performance for one day, allowing us to analyze both individual metric trends and how those metrics relate to each other over time.

While the dataset is fairly clean, it does have some limitations. Since it only covers one location, we're not able to compare performance geographically. Likewise, device-level performance was included but did not appear to significantly influence results during initial exploration. However, the dataset is strong for identifying temporal patterns, understanding spend efficiency, and investigating relationships between metrics like daily ROAS and daily spend.

Business Scenario

Digital ad campaigns are only as effective as their return on investment—and for this business, ROAS is the key performance indicator. The goal of this analysis is to identify the trends, patterns, and relationships within the data that could guide future ad spend decisions and optimize performance.

My main curiosity is around how sales and ROAS fluctuate over the week, and whether certain days consistently outperform others. I also want to explore how daily spend impacts efficiency, as well as other funnel metrics like clicks, conversions, and sale amounts.

The primary business question here is: *How can we use historical campaign data to make more informed spend decisions that maximize ROAS while maintaining or growing sales volume?*

Data Tools

For this project, I'm using Python in Google Colab. This setup makes it easy to store the dataset in Google Drive and work directly from there without bogging down my local machine. I will rely heavily on pandas for data cleaning and manipulation, matplotlib and seaborn for visualization, and scikit-learn's preprocessing tools for normalizing metrics.

Colab also allows for an iterative approach—testing one visualization or metric relationship, adjusting it, and immediately seeing the results. Since the dataset was relatively small, processing speed wasn't an issue, and working in a notebook format kept the workflow transparent and easy to share.

Data Cleaning

First, I'm checking for null values in the dataset. The conversion rate column has over 600 missing entries, but that's an easy fix—I can simply recalculate it using the conversions and clicks.

Next, I'm looking at the raw metrics with missing data. I start by standardizing the date format and extracting the Day_of_Week so I can see if there are patterns that might help me fill in the blanks. If I find that missing values are tied to certain days, I can use the mean or median for that day to fill them.

After plotting and reviewing the data, I'm not seeing any strong correlation between nulls and specific days. The missing values are random and not linked to a systematic cause. Since the variation in these metrics is high, filling in with averages could skew results. I'm deciding to drop those rows entirely, which removes about 18% of the data.

Now I'm reassigning column data types so the fields are easier to work with. I'm also cleaning the object columns—since they have only a few unique values, I'm manually mapping them to fix typos, inconsistent casing, and other formatting issues.

I'm also creating two new calculated fields to deepen the analysis:

Cost per Conversion = Spend / Conversions

Revenue per Click = Revenue / Clicks

Finally, I'm creating a daily summary dataframe so I can view metrics grouped by day of week. This will make it much easier to spot seasonal or day-based trends.

Data Analysis

I'm starting with a daily comparison of Clicks vs Conversions. Clicks are much higher than conversions, so I'm using a dual-axis plot (twinx) to keep them on separate scales while still showing how they move together. The two lines track closely, which tells me higher traffic generally results in higher conversions.

Next, I'm checking conversion rate by platform. I expect desktop, mobile, and tablet to differ significantly, but the results are surprisingly close—desktop and mobile are almost identical, with only a small dip for tablet users. Platform clearly isn't a big driver of conversion rate in this dataset.

I’m also normalizing funnel metrics—Clicks, Conversions, and Sale Amount—by day of week using MinMaxScaler so I can compare them on the same scale. This helps me see how different parts of the funnel behave relative to each other throughout the week.

These steps are giving me a baseline understanding of how the data behaves, and they’re helping me spot areas that might be worth deeper investigation, like whether certain days are more profitable or whether external factors are influencing sales.

As I step back from the exploratory work, a clear weekly pattern emerges. Friday and Saturday drive the highest clicks and conversions, giving me dependable volume. Wednesday stands out for efficiency — it delivers the strongest conversion rate even without peak traffic. Saturday pairs strong sales with the highest ROAS, making it the best overall return day. In contrast, Sunday underperforms across the board with low clicks, low conversions, and the weakest ROAS.

Device performance doesn’t materially shift the story. Tablets generate fewer clicks and conversions, but their ROAS keeps pace with desktop and edges out mobile. Overall, device type shows minimal variation in efficiency, so I don’t treat it as a major optimization lever in this dataset.

Keywords tell a more actionable story. “Data analytics online” brings the most traffic and conversions, but it lags on ROAS. “Data analytics course” delivers the highest ROAS despite fewer conversions, and “Data analytics training” posts the best cost per conversion with a solid ROAS. Despite those differences in traffic and return, conversion rates across keywords stay fairly tight (~4.9–5.2%), which suggests the main levers are spend mix and pricing/offer alignment rather than on-page conversion mechanics.

Based on these patterns, I’m recommending rebalancing budget by day and keyword. I would increase Saturday and Wednesday spend to maximize ROAS, maintain Friday for reliable volume while tightening targeting to lift return, and limit Sunday to reduce inefficiency. On the keyword side, I would prioritize “Data analytics course” and “Data analytics training,” keep a balanced investment in “learn data analytics,” and reduce “data analytics online” given its lower ROI. The goal is to reallocate dollars from high-cost, low-return segments toward the top performers so I can raise ROAS without sacrificing conversion volume.

Overall, this gives me a simple, testable playbook: shift spend toward the highest-return days and keywords, trim the weak spots, and monitor the lift. If these patterns hold as new data comes in, this approach should help the team maximize ROAS, sustain conversions, and cut wasted budget — a practical roadmap for smarter, more efficient ad spend.