

# EXPLORATORY DATA ANALYSIS OF ADVERTISING CAMPAIGNS ON TECH PLATFORMS

Campaign Trends, Performance  
& Audience Behavior



## TOOLS USED

Python, Pandas, Matplotlib, Seaborn & Plotly

## PRESENTED BY

Tejas Jadhav



# TABLE OF CONTENTS

Problem Statement & Objective	3
EDA Workflow	4
Key Questions Explored	5
Data Overview	6
Insights	11
Business/Developer Takeaways	21
Future Work	23
Conclusion	24

# PROBLEM STATEMENT & OBJECTIVE

## Problem Statement:

The Digital Advertising landscape features thousands of campaigns across diverse platforms and formats, making it challenging to identify trends, understand audience engagement and determine which campaigns achieve the greatest impact.

## Objective:

To explore the digital advertising campaign performance dataset and uncover patterns, insights, and trends that can help marketers and businesses optimize their campaigns and improve ROI.



# EDA WORKFLOW

For this analysis, a structured workflow was followed, involving data collection, understanding, cleaning, exploration, and summarization of insights to allow a clear understanding of the dataset and its trends.

01

## Data Collection

- Gathered the Google Play Store dataset from Kaggle

02

## Data Understanding & Anomaly Detection

- Looked at data distributions
- Found missing values, outliers, and unusual patterns

03

## Data Cleaning & Treatment

- Fixed incorrect values
- Standardized formats for consistency

04

## Exploratory Analysis

- Univariate Analysis: eg. Platform, Device\_Type, Audience Category etc.
- Bivariate Analysis: Impressions vs Clicks
- Multivariate Analysis: Conversion Rate vs Creative Format vs Device

05

## Insights & Reporting

- Summarized patterns and trends
- Highlighted key findings for developers and businesses

# KEY QUESTIONS EXPLORED

- Which platforms drive the highest revenue efficiency and ROAS and where should budget be reallocated?
- How does TikTok compare to Google and Facebook in terms of reach, revenue generation & scalability potential?
- Which creative formats (Video, Interactive, Text) generate the strongest engagement, conversions & cost efficiency?
- Which audience segments deliver the highest return on investment?
- How does retargeting impact profitability compared to non-retargeted campaigns?
- Where do the biggest funnel drop-offs occur and how can conversion efficiency be improved?
- Are higher budgets actually driving better performance, or is profitability concentrated among a few high-performing campaigns?
- How do campaign objectives (Conversion, Lead Gen, Brand Awareness) differ in terms of profitability & return efficiency?



# DATA OVERVIEW

The dataset provides key information about Google Play Store apps, including ratings, reviews, installs, price, category, and type. Additional details like content rating, genres, app size, last update, and Android version capture app characteristics and trends.

Data Source: Kaggle

## Dataset Size

10000

Records

41

Features

## Campaign Diversity

70,53,92,230

Impressions

6

Platforms

6

industry Verticals



# DATA OVERVIEW

Below is a detailed description of the feature set:

Dataset Features	Type	Feature Description
campaign_id	String	Unique identifier for each advertising campaign
campaign_objective	Categorical	Primary goal of the campaign
platform	Categorical	Advertising platform where the campaign ran (Google Ads, Facebook, LinkedIn, TikTok, Twitter, Instagram)
ad_placement	Categorical	Specific location of the ad display (Feed, Stories, Search, Display Network, In-Stream Video, Sidebar)
device_type	Categorical	User device used to view the ad (Desktop, Mobile, Tablet)
operating_system	Categorical	Device operating system (iOS, Android, Windows, macOS, Other)
creative_format	Categorical	Format of the ad creative (Video, Image, Carousel, Text, Interactive, Story)
creative_size	Categorical	Dimensions of the ad in pixels (e.g., 1080x1080, 1920x1080, 300x250)
ad_copy_length	Categorical	Length category of the ad text (Short, Medium, Long)
has_call_to_action	Boolean	Indicates whether the ad contains a CTA button (True/False)
creative_emotion	Categorical	Emotional tone conveyed by the ad (Fear, Joy, Urgency, Trust, Curiosity, Neutral)
creative_age_days	Numerical (Discrete)	Number of days since the creative was launched (1-90)
target_audience_age	Categorical	Age bracket of the targeted audience (18-24, 25-34, etc.)

# DATA OVERVIEW

Continued...

Dataset Features	Type	Feature Description
target_audience_gender	Categorical	Gender of the targeted audience (Male, Female, All)
audience_interest_category	Categorical	Interest segment of the targeted audience
income_bracket	Categorical	Household income level of the audience (<\$50K, \$50K-\$100K, etc.)
purchase_intent_score	Categorical	Behavioral purchase intent level (Low, Medium, High)
retargeting_flag	Boolean	Indicates whether the campaign is a retargeting effort (True/False)
quarter	Numerical (Discrete)	Calendar quarter of the campaign (1-4)
day_of_week	Categorical	Day of the week the ad was shown (Monday-Sunday)
hour_of_day	Numerical (Continuous)	Hour of the day the ad was displayed (0-23)
campaign_day	Numerical (Discrete)	Day number within the campaign lifecycle (1-90)
quality_score	Numerical (Discrete)	Platform-assigned quality rating of the ad (1-10)
actual_cpc	Numerical (Continuous)	Actual cost per click paid for the ad (\$0.25-\$17.00)
impressions	Numerical (Discrete)	Number of times the ad was displayed (5,000-500,000)
clicks	Numerical (Discrete)	Number of ad clicks ( $\geq 10$ and $\leq$ impressions)



# DATA OVERVIEW

Continued...

Dataset Features	Type	Feature Description
conversions	Numerical (Discrete)	Number of completed actions from the ad (0+ and $\leq$ clicks)
ad_spend	Numerical (Continuous)	Total campaign spend calculated as clicks $\times$ actual_cpc
revenue	Numerical (Continuous)	Revenue generated from the campaign (\$)
bounce_rate	Numerical (Discrete)	Percentage of immediate exits (10–90%)
avg_session_duration_seconds	Numerical (Continuous)	Average time users spend on the site after clicking (10–600 seconds)
pages_per_session	Numerical (Discrete)	Average number of pages viewed per session (1.0–10.0)
industry_vertical	Categorical	Business sector of the advertiser
budget_tier	Categorical	Campaign budget classification (Low, Medium, High)
CTR	Numerical (Continuous)	Click-through rate (%) = (clicks / impressions) $\times$ 100
CPC	Numerical (Continuous)	Cost per click (\$) = ad_spend / clicks
conversion_rate	Numerical (Continuous)	Conversion rate (%) = (conversions / clicks) $\times$ 100
CPA	Numerical (Continuous)	Cost per acquisition (\$) = ad_spend / conversions
ROAS	Numerical (Continuous)	Return on ad spend = revenue / ad_spend
profit	Numerical (Continuous)	Profit (\$) = revenue – ad_spend

---

# DATA QUALITY CHALLENGES & ANOMALIES

Few inconsistencies were found in the dataset, which could have affected the analysis if left unaddressed.

## DATA ANOMALIES

- Columns requiring cleaning or type conversion: Start\_Date.
- Derived metrics (CTR, conversion rate, CPA, ROAS) may not perfectly match base metrics due to pre-computation or rounding, so recalculating them from raw data is recommended for accurate analysis.



# DATA CLEANING & TREATMENT

Inconsistent and missing values were addressed, and key features were cleaned and standardized for analysis.

## DATA CLEANING SUMMARY

- Start\_Date was cleaned and converted to proper datetime format to ensure accurate temporal analysis.
- Derived performance metrics (CTR, conversion rate, CPA, ROAS) were recalculated from base fields to resolve rounding or pre-computation inconsistencies and ensure analytical accuracy.





# INSIGHTS





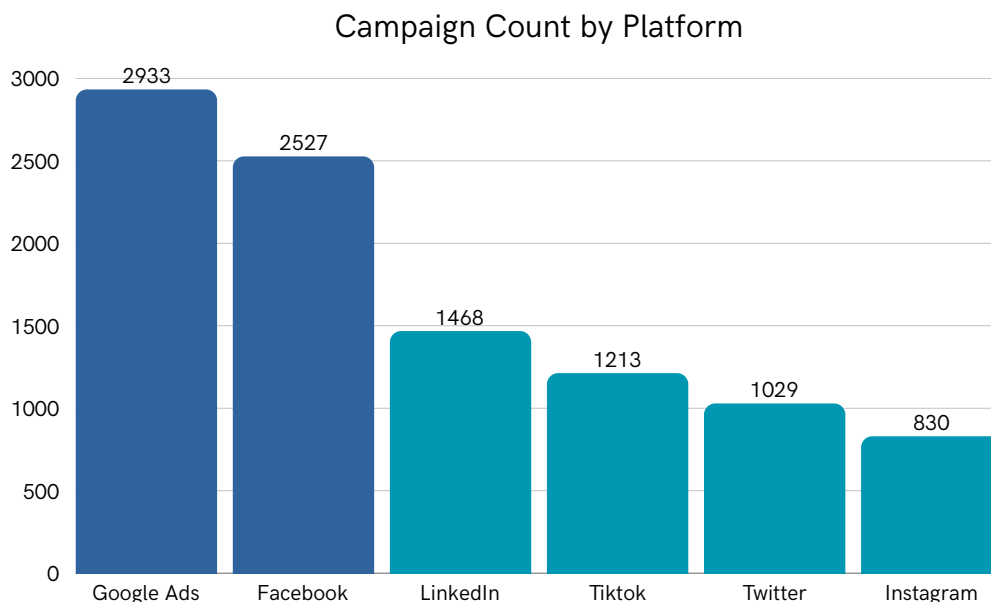
# Platform & Market Landscape Overview



# Google Ads & Facebook dominate campaign allocation, capturing the majority of advertising activity

## 1,667

Average Campaign Volume



## Key observations

- Google Ads has the highest campaign count (~2,933), followed by Facebook (~2,527), significantly ahead of other platforms.
- Instagram has the lowest campaign count (~830), while LinkedIn, TikTok, and Twitter fall in a mid-to-lower tier range.

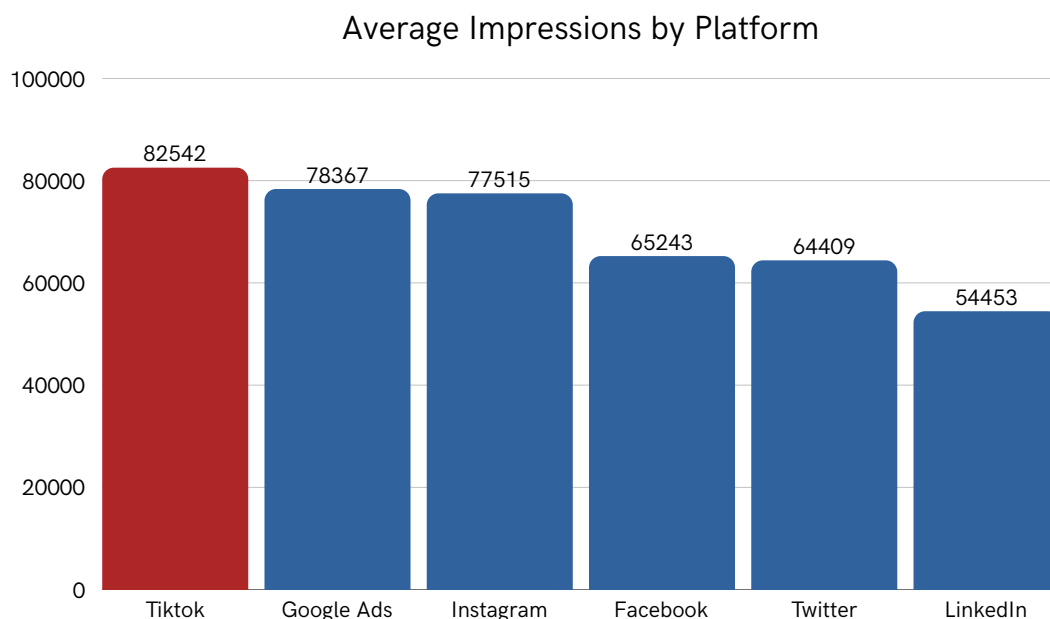
## Business Insights

- Budget concentration appears heavily skewed toward Google Ads and Facebook, indicating these platforms are primary acquisition channels.
- Performance optimization efforts should prioritize Google and Facebook first, while smaller platforms may require selective or experimental investment rather than equal budget distribution.

# TikTok delivers the highest average reach despite not leading in campaign volume

## 70,539

Average Impression Count



## Key observations

- TikTok has the highest average impressions (~82,542), outperforming all other platforms.
- LinkedIn records the lowest average impressions (~54,453), while Google Ads and Instagram are close behind TikTok but still lower.

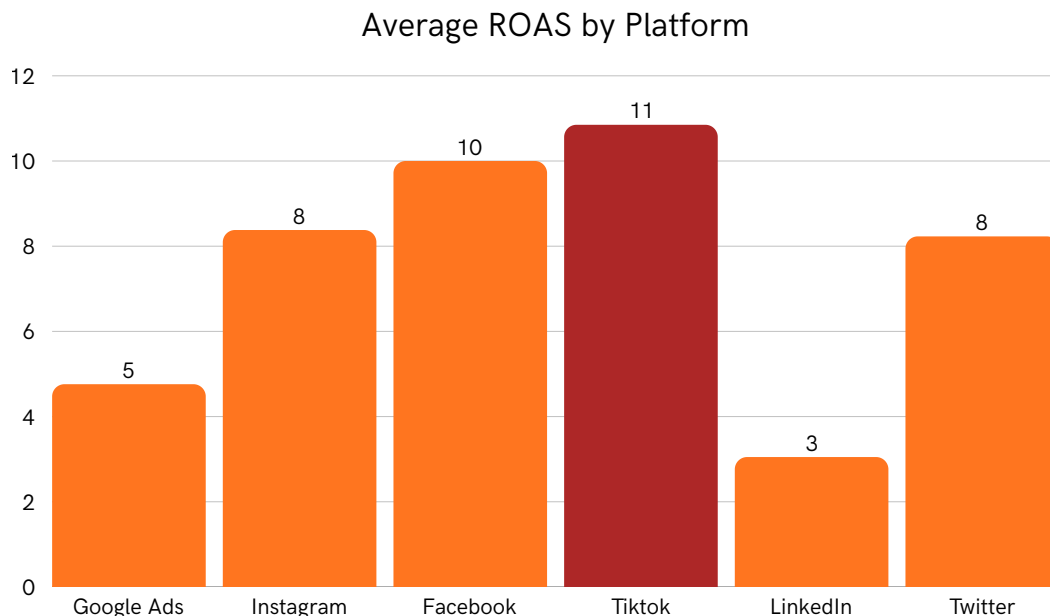
## Business Insights

- TikTok appears to generate stronger reach per campaign, indicating higher exposure efficiency at the campaign level.
- If awareness and reach are primary objectives, reallocating incremental budget toward TikTok could increase total visibility without proportionally increasing campaign count.

# TikTok and Facebook generate the strongest return efficiency, while LinkedIn significantly underperforms

## 7.52

Average Advertising ROAS



## Key observations

- TikTok has the highest average ROAS (~10.85), followed closely by Facebook (~10.00).
- LinkedIn records the lowest average ROAS (~3.05), substantially below all other platforms.

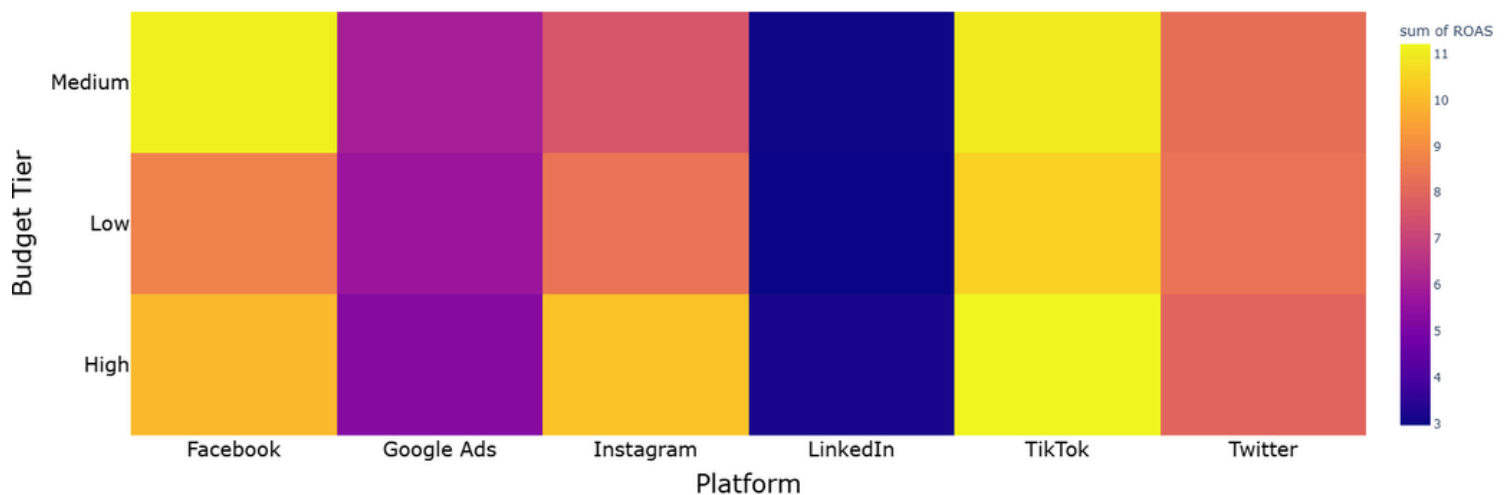
## Business Insights

- Revenue efficiency is concentrated on TikTok and Facebook, indicating these platforms convert spend into revenue more effectively.
- Budget reallocation from lower-ROAS platforms (particularly LinkedIn) toward higher-performing ones could improve overall marketing profitability and blended ROAS.



# Higher budgets amplify ROAS on TikTok & Instagram, while LinkedIn consistently underperforms regardless of spend

Average ROAS by Platform and Budget Tier



## Key observations

- TikTok and Instagram show the highest ROAS at the High budget tier, outperforming other platforms at the same spend level.
- LinkedIn records the lowest ROAS across all budget tiers, with minimal improvement even as budget increases.

## Business Insights

- Scaling budgets on TikTok and Instagram is likely to generate stronger revenue efficiency compared to other platforms.
- Budget allocation toward LinkedIn should be reassessed or optimized strategically, as increased spend does not translate into proportional ROAS gains.



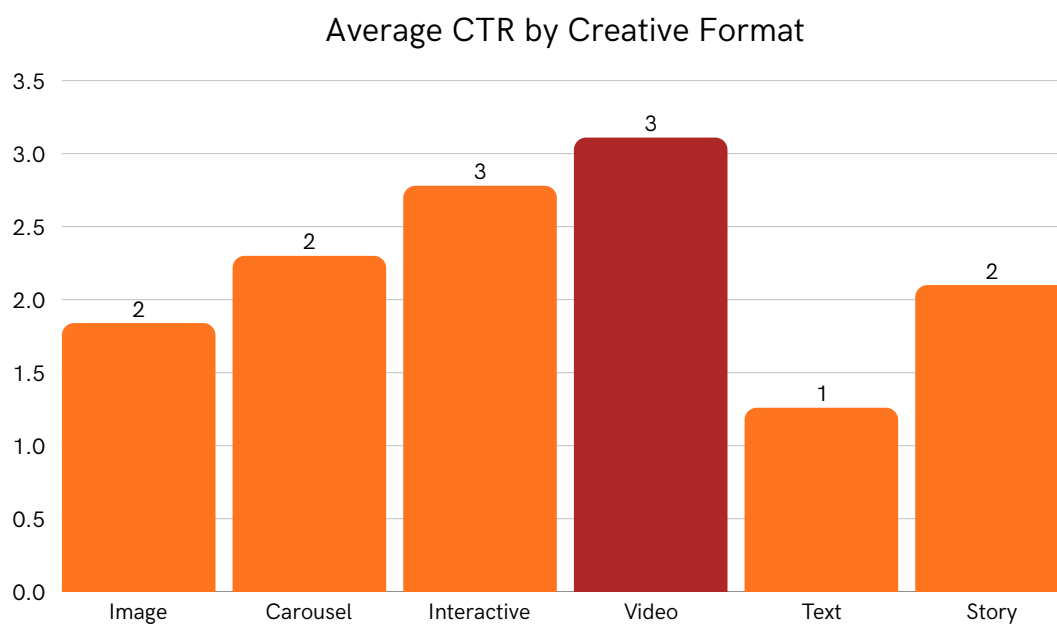
# Creative Strategy & Engagement Analysis



# Video & interactive creatives drive the strongest engagement, while text-based formats significantly underperform

## 2.3

Average CTR by Format



## Key observations

- Video has the highest average CTR (~3.11), followed by Interactive (~2.78), outperforming all other creative formats.
- Text ads record the lowest CTR (~1.26), substantially below visual and dynamic formats like Carousel and Story.

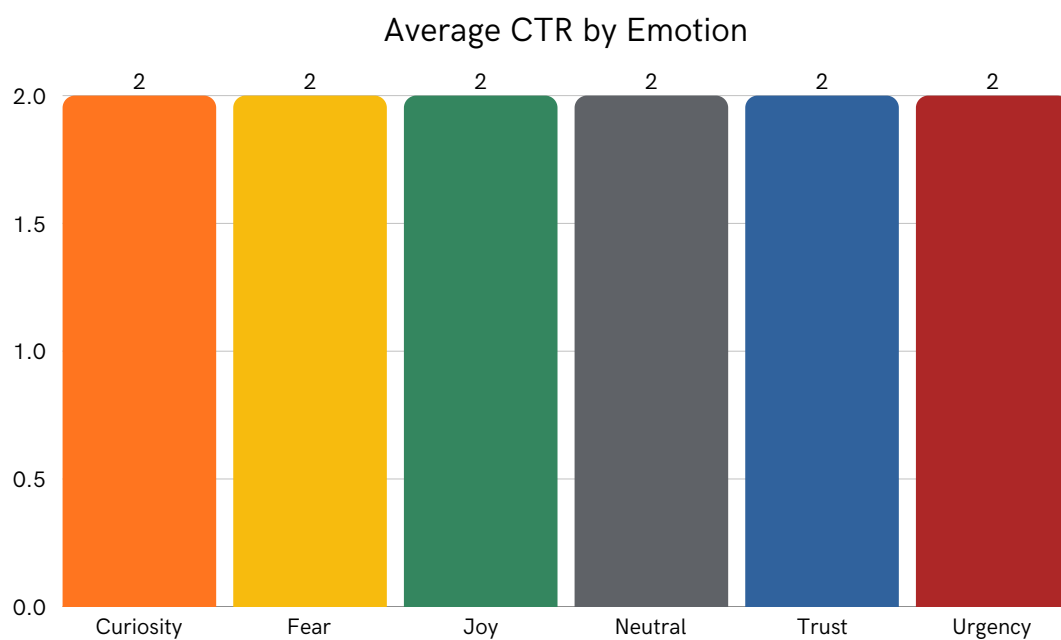
## Business Insights

- Shifting creative investment toward Video and Interactive formats can meaningfully improve engagement and traffic efficiency.
- Text-heavy formats should be deprioritized or redesigned with stronger visual elements to prevent wasted impressions and low engagement rates.

# Creative emotion has minimal impact on CTR, with performance remaining nearly uniform across tones

## 2.3

Typical CTR by Emotion (Mean)



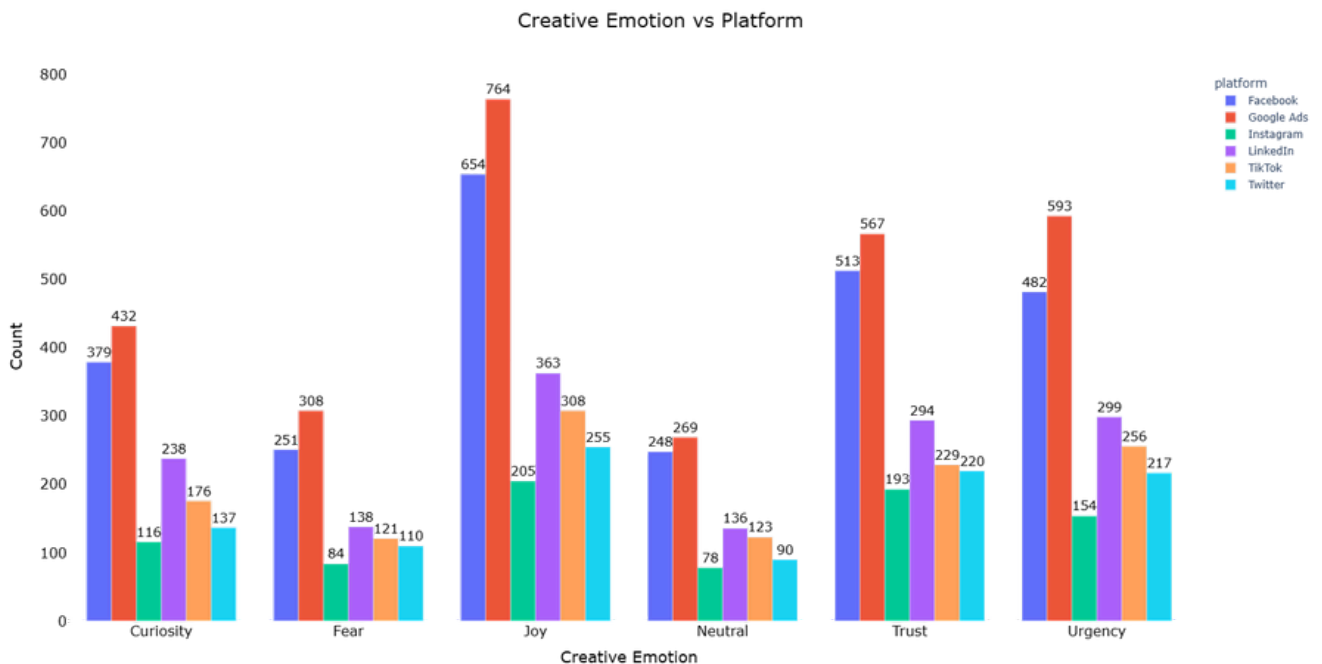
## Key observations

- All creative emotions cluster tightly between ~2.28 and ~2.34 CTR, indicating very small performance variation.
- Trust (~2.34) and Neutral (~2.33) are marginally higher, but the gap versus the lowest (Fear ~2.28) is negligible.

## Business Insights

- Emotional tone alone is unlikely to materially improve engagement; optimization efforts should prioritize format, placement, or targeting instead.
- Creative testing budgets can be consolidated rather than fragmented across emotional variants, as returns appear relatively consistent across tones.

# Google Ads and Facebook campaigns leverage joyful and urgent emotions most effectively.



## Key observations

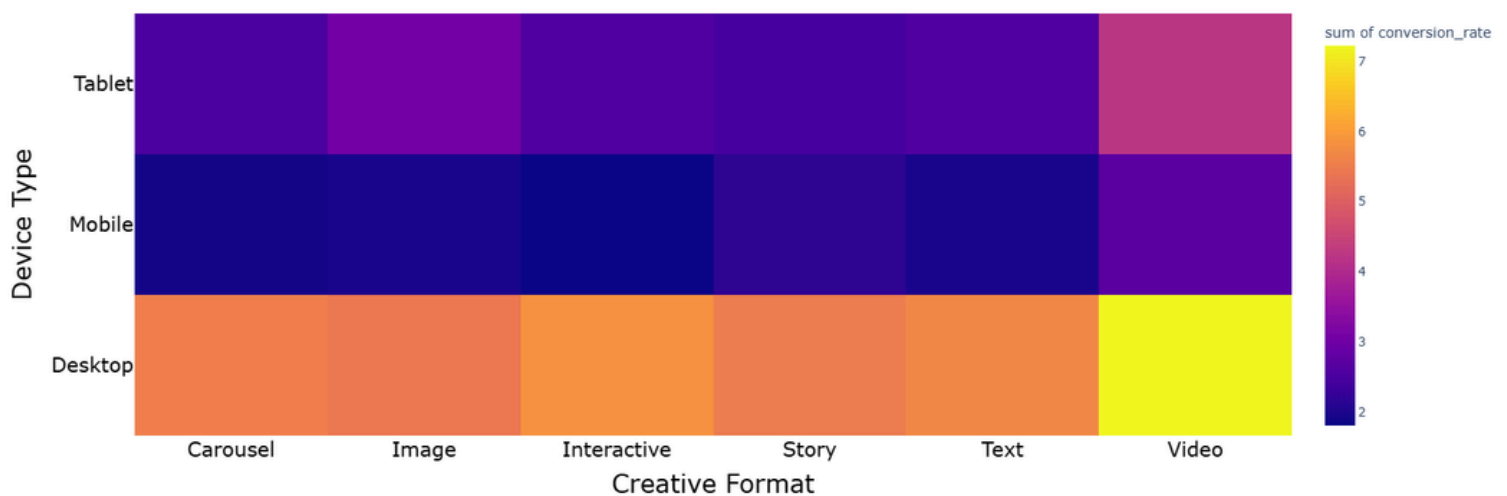
- "Joy" and "Urgency" are the most frequently used emotions on Google Ads and Facebook, with counts of 764 and 593 respectively for Google Ads.
- Platforms like Instagram, TikTok, and Twitter show lower overall usage across all emotions, with comparatively higher counts for "Joy" and "Trust" than "Fear" or "Neutral."

## Business Insights

- Prioritize "Joy" and "Urgency"-driven creatives on Google Ads and Facebook to maximize engagement, given their strong adoption and apparent platform alignment.
- For platforms with lower volume (Instagram, TikTok, Twitter), focus on emotionally resonant content like "Trust" and "Joy" to stand out, rather than spreading effort across less effective emotions.

# Desktop video creatives drive the highest conversion rates across devices.

Conversion Rate by Creative Format and Device Type



## Key observations

- Desktop devices consistently outperform Mobile and Tablet in conversion rate for all creative formats, with Video creatives reaching the peak.
- Mobile and Tablet show much lower conversion rates, with minor variation across creative formats.

## Business Insights

- Prioritize video creatives for Desktop campaigns to maximize conversions and ROAS.
- Consider device-specific optimization: focus higher budgets on Desktop while tailoring creative formats for Mobile and Tablet to improve their performance.