DATA ANALYTICS REPORT

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

GITHUB

TABLE OF CONTENTS

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 missing or 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 categories dominate the Play Store, and where is competition highest?
- What is the distribution of Free vs Paid apps across categories and ratings?
- Who are the target audiences of apps (content rating distribution)?
- How do app sizes vary, and what is the optimal size range?
- How do app ratings behave across categories, review volumes, and time?
- What monetization models work best by category and audience?

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 (≥10 and ≤ impressions)

DATA OVERVIEW

Continued...

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