HNG

2021 Case Study

An Exploratory Data Analysis of HNG Limited - Marketing Campaign Data

Introduction

Overview

The dataset contains key columns like; Campaign_ID, Company, Company_Type, Target_Audience, Duration, Channel_Used, Conversion_Rate, Acquisition_Cost, ROI, Location, Date, Clicks, Impressions, Engagement_Score, Customer_Segment.

Data types ; The data types available- object,float64, Date(had to convert to datetime64 using Python) and int64

Dataset size: The dataset has a total count of 3,000,009 and 15 Columns with count of 200,006 inclusive of Headers for each Column.

Analysis Objective

Perform exploratory data analysis (EDA) on HNG Limited Marketing Campaign dataset to understand the dataset better and uncover key insights that can guide strategic decision-making.

Key Metrics;

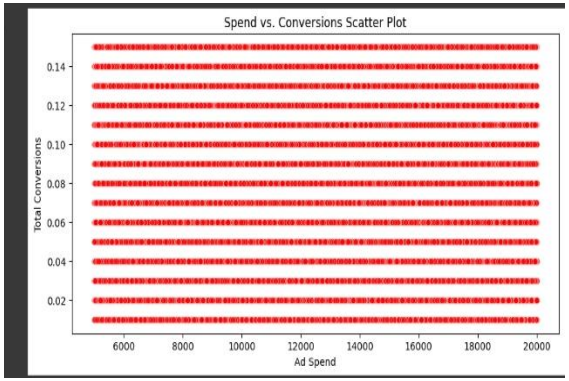
- **Cost per Click(CPC)**
- **Click Through Rate(CTR)**
- **ROI**
- **Conversion Rate**
- **Campaign performance across diff channels**

Insights

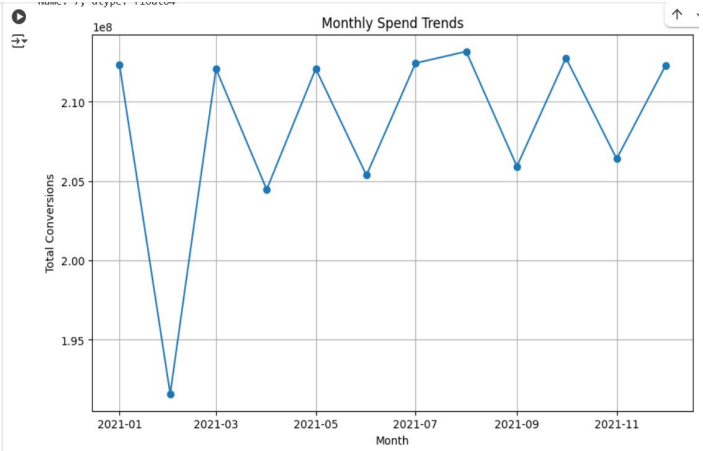
- Email Marketing channel drives the highest Conversion rate of 2,697.38 followed by Google ads with 2,681.24
- Facebook Marketing channel has the lowest total Conversion rate of 2,625.27

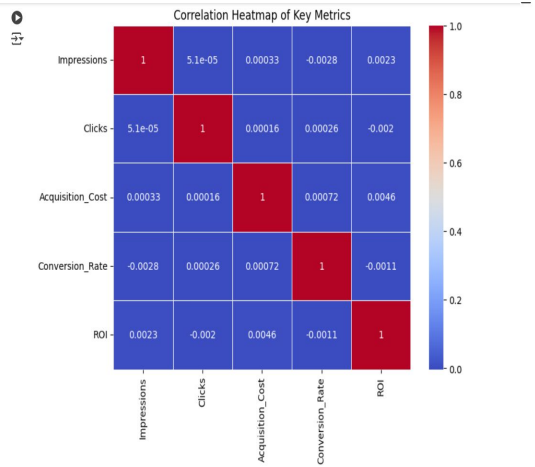
- Google ads has the highest Cost per click with value 108,039,487.6 USD
- Facebook has the lowest cpc with value 105,448,579.3 USD
-

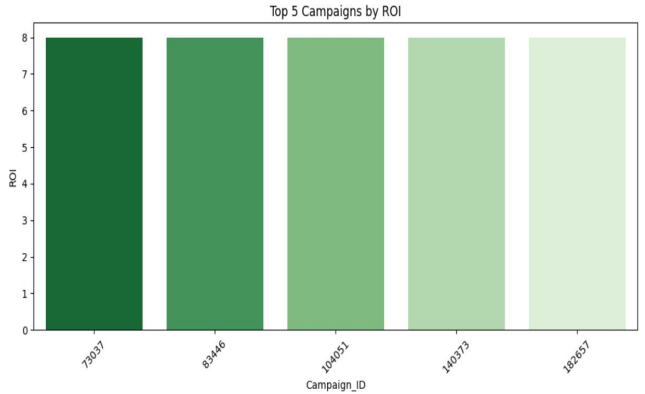
- Email Marketing channel brought in the highest ROI of 167,876.95
- Facebook Channel brought in the lowest ROI of 164,712.81



- Los Angeles has the highest Click through rate(CTR) of 10.000601
- Miami has the lowest CTR of 9.964758
- Campaign_ID 122375 has the highest CTR of 99.20239282
- Campaign_ID 44685 has the lowest CTR of 1.0054293
- Men 18-24 has the highest CTR of 565563.5104
- Women 35-44 has the lowest CTR of 557,928.1631
- Foodies segment have the highest CTR of 568,619.9628
- Fashionistas has the lowest CTR of 556,200.5961







Code and Methodology

```
## TO change date format
def convert_date(date_str):
    formats_to_try = ["%d/%m/%Y", "%m/%d/%Y", "%Y-%m-%d",  # Add all your formats here!
                      "%d-%m-%Y", "%m-%d-%Y"] # Add more if needed
    for fmt in formats_to_try:
        try:
            return pd.to_datetime(date_str, format=fmt, errors='coerce')
        except ValueError:
            pass # Try the next format
    return pd.NaT  # Return NaT if none of the formats match


df['Date'] = df['Date'].apply(convert_date)
df['Date'] = df['Date'].dt.strftime('%d/%m/%Y')
```

```
## TO change date format from object to datetime64
df['Date'] = pd.to_datetime(df['Date'], format='mixed')
```

```
## convert Acquisition_Cost to float
df['Acquisition_Cost'] = pd.to_numeric(df['Acquisition_Cost'].astype(str).str.replace(r'[$,]', '', regex=True), errors='coerce'
```

```
## Calculate Key Metrics
df['CTR'] = (df['Clicks'] / df['Impressions']) * 100  # Click-Through Rate (%)
df['CPC'] = df['Acquisition_Cost'] / df['Clicks']  # Cost Per Click
##df['Conversion Rate'] = (df['Conversion_Rate'] * 100)  # Conversion Rate

print("\nKey Metrics (CTR, CPC):\n", df[['Campaign_ID', 'CTR', 'CPC', 'Conversion_Rate']].head(15))
```

```
# Create a 'Date' column by combining year and month
monthly_ad_spend['Date'] = pd.to_datetime(monthly_ad_spend[['year', 'month']].assign(DAY=1))

# Now you can plot using the new 'Date' column
plt.plot(monthly_ad_spend['Date'], monthly_ad_spend['Acquisition_Cost'], marker='o')
plt.title('Monthly Spend Trends')
plt.xlabel('Month')
plt.ylabel('Total Conversions')
plt.grid(True)
plt.show()
```

```
## LINE GRAPH: Ad Spend Trend per month
# Group by month and sum the acquisition cost(ad spend) for each month
monthly_ad_spend = df.groupby(['year' , 'month'])['Acquisition_Cost'].sum().reset_index()

# Find the month with the highest total sales
max_spend_month = monthly_ad_spend.loc[monthly_ad_spend['Acquisition_Cost'].idxmax()]

# Print the monthly sales data and the month with the highest sales
print("Monthly Spend Data:")
print(monthly_ad_spend)

print("\nMonth with the Highest Sales:")
print(max_spend_month)

# Plotting
plt.figure(figsize=(10, 6))
```

Recommendations

- Tailor future campaigns to the unique needs and preferences of each segment(focus more on Men 18-24) to increase conversion rates.
- Shift budget towards the most effective channels(e.g.Email and Google ads) and reallocate from underperforming ones(e.g. Facebook)
- Leverage Customer Feedback and Insights from surveys, reviews, personal interactions,etc

Actionable Insights

📌Given the success of email marketing, future campaigns should prioritize personalized email strategies for the Men 18-24 age group, while experimenting with additional social media platforms to boost engagement.

📌 The spend trend is both increasing and decreasing It was at its lowest in February 2021 but thereafter increased the following month. It was at an all high in August 2021 followed by January 2021 which had ethe highest ROI. More analysis needs to be done on why spends were that low in February.

📌 Higher spending leads to higher Conversion rate and ROI e.g. Email marketing channel

📌 Priority should be given to Foodies segment as it has the highest CTR and ROI.

Conclusion

Based on these findings, we recommend a shift in focus towards optimizing email strategies while exploring new social media channels. Future analyses should consider seasonality to refine targeting efforts. By implementing these recommendations, we expect to improve future campaign outcomes