Python Project

MARKETING CAMPAIGN PERFORMANCE INSIGHTS

A

KANNAN_RAJENDRAN

PROGRAMMING

Table of Contents

Introduction

In today's digital age, businesses rely heavily on marketing campaigns to reach their target audience. However, evaluating the effectiveness of these campaigns across various channels and platforms can be challenging. This project focuses on analyzing the performance of different marketing campaigns to identify key success factors, optimize ROI, and inform future marketing strategies.

In the highly competitive landscape of digital marketing, accurately evaluating the success of various campaigns is essential for optimizing return on investment (ROI) and improving overall performance. Despite having access to extensive data on multiple campaigns, there is a pressing need for a thorough analysis to assess and compare key metrics such as conversion rates, acquisition costs, and ROI across different campaign types, channels, and audience segments. This project aims to uncover actionable insights by examining temporal trends, geographical influences, and audience responses to identify factors driving campaign success. By leveraging advanced analytical techniques and data visualization tools, we will provide strategic recommendations for enhancing future marketing strategies, ultimately helping businesses achieve their marketing objectives more effectively.

Key Aspects

1. **Importance of Evaluation**: Emphasizes the necessity of accurately evaluating the success of marketing campaigns in a competitive digital marketing landscape.
2. **Optimizing ROI**: Focuses on optimizing return on investment (ROI) and improving overall performance.
3. **Data Analysis**: Highlights the need for thorough analysis to assess and compare key metrics such as conversion rates, acquisition costs, and ROI.
4. **Comparative Analysis**: Stresses the importance of comparing these metrics across different campaign types, channels, and audience segments.
5. **Actionable Insights**: Aims to uncover actionable insights by examining temporal trends, geographical influences, and audience responses.
6. **Advanced Techniques**: Utilizes advanced analytical techniques and data visualization tools.
7. **Strategic Recommendations**: Provides strategic recommendations for enhancing future marketing strategies.
8. **Business Objectives**: Ultimately helps businesses achieve their marketing objectives more effectively.

Project Objectives

This project aims to achieve the following objectives:

1. **Analyze campaign performance:** Conduct a comprehensive analysis of key performance indicators (KPIs) across different campaigns, including conversion rates, acquisition costs, and ROI.
2. **Identify key success factors:** Determine the factors that contribute to successful campaigns, such as target audience segmentation, channel selection, and creative messaging.
3. **Uncover temporal trends:** Analyze campaign performance over time to identify seasonal patterns, emerging trends, and areas for improvement.
4. **Examine geographical influences:** Investigate the impact of geographical location on campaign performance, including regional preferences and cultural nuances.
5. **Understand audience responses:** Analyze audience behavior and engagement metrics to gain insights into their preferences and preferences.
6. **Provide actionable recommendations:** Develop recommendations for optimizing future marketing strategies based on the findings of the analysis.

Professional specialty

Problem Statement:

In the highly competitive landscape of digital marketing, effectively evaluating the success of various marketing campaigns is essential for optimizing return on investment (ROI) and improving overall performance. Despite having extensive data on multiple campaigns, there is a need for a thorough analysis to assess and compare key metrics such as conversion rates, acquisition costs, and ROI across different campaign types, channels, and audience segments. This project aims to uncover actionable insights by examining temporal trends, geographical influences, and audience responses to identify factors driving campaign success and provide recommendations for enhancing future marketing strategies.

Dataset Link:

https://raw.githubusercontent.com/ArchanaInsights/Datasets/main/marketing_campaign.csv

Data Dictionary:

Campaign_ID	• Unique identifier for each campaign.
Company	• The organization running the campaign, represented by various fictional brands.
Campaign_Type	• The type of marketing effort used, such as email, social media, influencer, display, or search.
Target_Audience	• The specific demographic or audience segment targeted by the campaign (e.g., women aged 25-34).
Duration	• The duration of the campaign, expressed in days.
Channels_Used	• The platforms or mediums used to promote the campaign, including email, social media, YouTube, websites, or Google Ads.
Conversion_Rate	• The percentage of impressions or leads that resulted in desired actions, reflecting campaign effectiveness.
Acquisition_Cost	• The monetary expense incurred to acquire each customer through the campaign.
ROI	• Return on Investment, indicating the profitability and success of the campaign
Location	• The geographical area where the campaign was executed (e.g., New York, Los Angeles).
Language	• The language in which the campaign's content was delivered (e.g., English, Spanish).
Clicks	• The total number of clicks generated by the campaign, showing user interaction.
Impressions	• The total number of times the campaign was displayed or viewed by the audience.
Engagement_Score	• A score from 1 to 10 representing the level of engagement and interaction generated by the campaign.
Customer_Segment	• The specific group or category of customers targeted by the campaign (e.g., tech enthusiasts, fashionistas).
Date	• The date on which the campaign occurred.

Methodology

1. Load the Dataset

❖ Read the marketing campaign data from the CSV file into a pandas Data Frame.

1) Load the Dataset

```
[25]: data_link = 'https://raw.githubusercontent.com/ArchanaInsights/Datasets/main/marketing_campaign.csv'

df=pd.read_csv(data_link)
print(df)
```

```
       Campaign_ID              Company Campaign_Type Target_Audience  \
0                1            TechCorp         Email    Women 25-34
1                2  Innovate Industries    Influencer    Women 35-44
2                3       NexGen Systems  Social Media    Women 25-34
3                4  Innovate Industries         Email    Women 25-34
4                5  Data Tech Solutions    Influencer      Men 25-34
...            ...                  ...           ...            ...
22024        22025  Data Tech Solutions        Search      Men 18-24
22025        22026  Data Tech Solutions  Social Media      Men 18-24
22026        22027             TechCorp    Influencer    Women 25-34
22027        22028  Data Tech Solutions        Search      Men 18-24
22028        22029       NexGen Systems        Search      Men 18-24

      Duration Channel_Used  Conversion_Rate  Acquisition_Cost     ROI  \
0      30 days     Facebook         5.294194              9344   62.94
1      45 days   Google Ads         3.326375              8783   10.67
2      45 days    Instagram         4.056375              9111   73.20
3      45 days    Instagram         4.496375              7420   60.92
4      30 days   Google Ads         4.405930              2146  138.82
...        ...          ...              ...               ...     ...
22024  30 days      Website         4.379947              4748   -7.01
```

The code loads a marketing campaign dataset from a GitHub repository into a pandas Data Frame. The dataset contains various information about different campaigns, which can be used for further analysis and insights into marketing performance.

2. Descriptive Analysis

✓ Basic Structure

❖ Print the first few rows of the dataset to get an overview of the data.

This code snippet provides a basic descriptive analysis of the marketing campaign dataset by displaying the first few rows. This helps in understanding the structure of the dataset and the type of information it contains.

2) Descriptive Analysis

```
#Print the first few rows of the dataset to get an overview of the data
df.head(5)
```

	Campaign_ID	Company	Campaign_Type	Target_Audience	Duration	Channel_Used	Conversion_Rate	Acquisition_Cost	ROI	Location	Language	Clicks	Impre:
0	1	TechCorp	Email	Women 25-34	30 days	Facebook	5.294194	9344	62.94	Houston	English	3045	
1	2	Innovate Industries	Influencer	Women 35-44	45 days	Google Ads	3.326375	8783	10.67	Washington, D.C.	German	1944	
2	3	NexGen Systems	Social Media	Women 25-34	45 days	Instagram	4.056375	9111	73.20	Miami	Spanish	3156	
3	4	Innovate Industries	Email	Women 25-34	45 days	Instagram	4.496375	7420	60.92	Seattle	Spanish	2388	
4	5	Data Tech Solutions	Influencer	Men 25-34	30 days	Google Ads	4.405930	2146	138.82	Chicago	English	1025	

❖ Obtain the number of rows and columns in the dataset.

```
# Obtain the number of rows and columns in the dataset.
rows, columns = df.shape
print(f'The dataset has {rows} rows and {columns} columns.')
```

```
The dataset has 22029 rows and 16 columns.
```

This code snippet provides a simple way to determine the dimensions of a pandas Data Frame. It's a useful step in the initial exploration of a dataset to understand its size and structure.

❖ Get a concise summary of the dataset, including the data types and non-null values.

This code provides a quick and informative overview of the DataFrame, including its dimensions, data types, and the presence of missing values. This information is essential for further data analysis and cleaning steps.

```
# Get a concise summary of the dataset, including the data types and non-null values

summary=df.info()
print(summary)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 22029 entries, 0 to 22028
Data columns (total 16 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Campaign_ID       22029 non-null  int64
 1   Company           22029 non-null  object
 2   Campaign_Type     22029 non-null  object
 3   Target_Audience   22029 non-null  object
 4   Duration          22029 non-null  object
 5   Channel_Used      22029 non-null  object
 6   Conversion_Rate   22029 non-null  float64
 7   Acquisition_Cost  22029 non-null  int64
 8   ROI               22029 non-null  float64
 9   Location          22029 non-null  object
 10  Language          22029 non-null  object
 11  Clicks            22029 non-null  int64
 12  Impressions       22029 non-null  int64
 13  Engagement_Score  22029 non-null  int64
 14  Customer_Segment  22029 non-null  object
 15  Date              22029 non-null  object
dtypes: float64(2), int64(5), object(9)
memory usage: 2.7+ MB
None
```

❖ Generate descriptive statistics for numerical columns.

```
#Generate descriptive statistics for numerical columns.
descriptive_statistics=df.describe()
descriptive_statistics
```

	Campaign_ID	Conversion_Rate	Acquisition_Cost	ROI	Clicks	Impressions	Engagement_Score
count	22029.000000	22029.000000	22029.000000	22029.000000	22029.000000	22029.000000	22029.000000
mean	11015.000000	4.757232	5522.740842	182.863648	2223.807572	50610.402787	6.582323
std	6359.368876	0.960393	2597.666260	301.619721	1394.166380	28542.979123	1.458804
min	1.000000	2.015723	1000.000000	-98.300000	30.000000	1001.000000	4.000000
25%	5508.000000	4.130705	3286.000000	-4.080000	1067.000000	25804.000000	5.000000
50%	11015.000000	4.761527	5525.000000	93.650000	2088.000000	50858.000000	7.000000
75%	16522.000000	5.429335	7766.000000	247.310000	3212.000000	75165.000000	8.000000
max	22029.000000	7.469907	9999.000000	3109.790000	6887.000000	99999.000000	9.000000

This code provides a summary of the central tendency and dispersion of the numerical variables in the Data Frame. This information is valuable for understanding the distribution of the data and identifying potential outliers or unusual patterns.

✓ Data Exploration

❖ Print the number of unique Campaign_ID values in the dataset.

```
# Print the number of unique Campaign_ID values in the dataset
unique_campaigns = df['Campaign_ID'].nunique()
print(f'The dataset has {unique_campaigns} unique Campaign_ID values.')
```

```
The dataset has 22029 unique Campaign_ID values.
```

This code snippet provides a concise way to determine the number of unique values in a specific column of a pandas Data Frame. This information can be useful for various data analysis tasks, such as identifying duplicates or understanding the distribution of values in the column.

❖ List the unique values of the Location and Customer_Segment columns.

```python
# List the unique values of the Location and Customer_Segment columns.

unique_Location=df["Location"].unique()
unique_Customer_Segment =df['Customer_Segment'].unique()

print(f'The Unique Locations are :\n{unique_Location}')

print(f'The Unique Customer Segments  are :\n{unique_Customer_Segment}')
```

```
The Unique Locations are :
['Houston' 'Washington, D.C.' 'Miami' 'Seattle' 'Chicago' 'Los Angeles'
 'Atlanta' 'Dallas' 'New York' 'San Francisco']
The Unique Customer Segments  are :
['Tech Enthusiasts' 'Foodies' 'Fashionistas' 'Outdoor Adventurers'
 'Health & Wellness']
```

The code snippet extracts and displays the unique values present in the Location and Customer_Segment columns of a pandas Data Frame. It uses the unique () method to efficiently identify distinct values within each column and then prints the results using formatted strings for better readability. This information can be valuable for understanding the distinct categories within these columns, which can be useful for further analysis or data visualization

❖ Count the occurrences of each category in the Campaign_Type and Channel_Used and columns.

This code effectively counts and displays the occurrences of each category within the Campaign_Type and Channel_Used columns of the Data Frame, providing valuable insights into the distribution of campaign types and marketing channels used in the data.

```python
[77]: # Count the occurrences of each category in the Campaign_Type and Channel_Used and columns.
campaign_type_counts =df['Campaign_Type'].value_counts()
print("Occurrences of each category in Campaign_Type column:")
print(campaign_type_counts)
```

```
Occurrences of each category in Campaign_Type column:
Campaign_Type
Display         4450
Search          4441
Social Media    4412
Email           4388
Influencer      4338
Name: count, dtype: int64
```

```python
[85]: # Count the occurrences of each category in the Campaign_Type and Channel_Used and columns.
channel_used_counts =df['Channel_Used'].value_counts()
print("Occurrences of each category in Campaign_Type column:")
print(channel_used_counts)
```

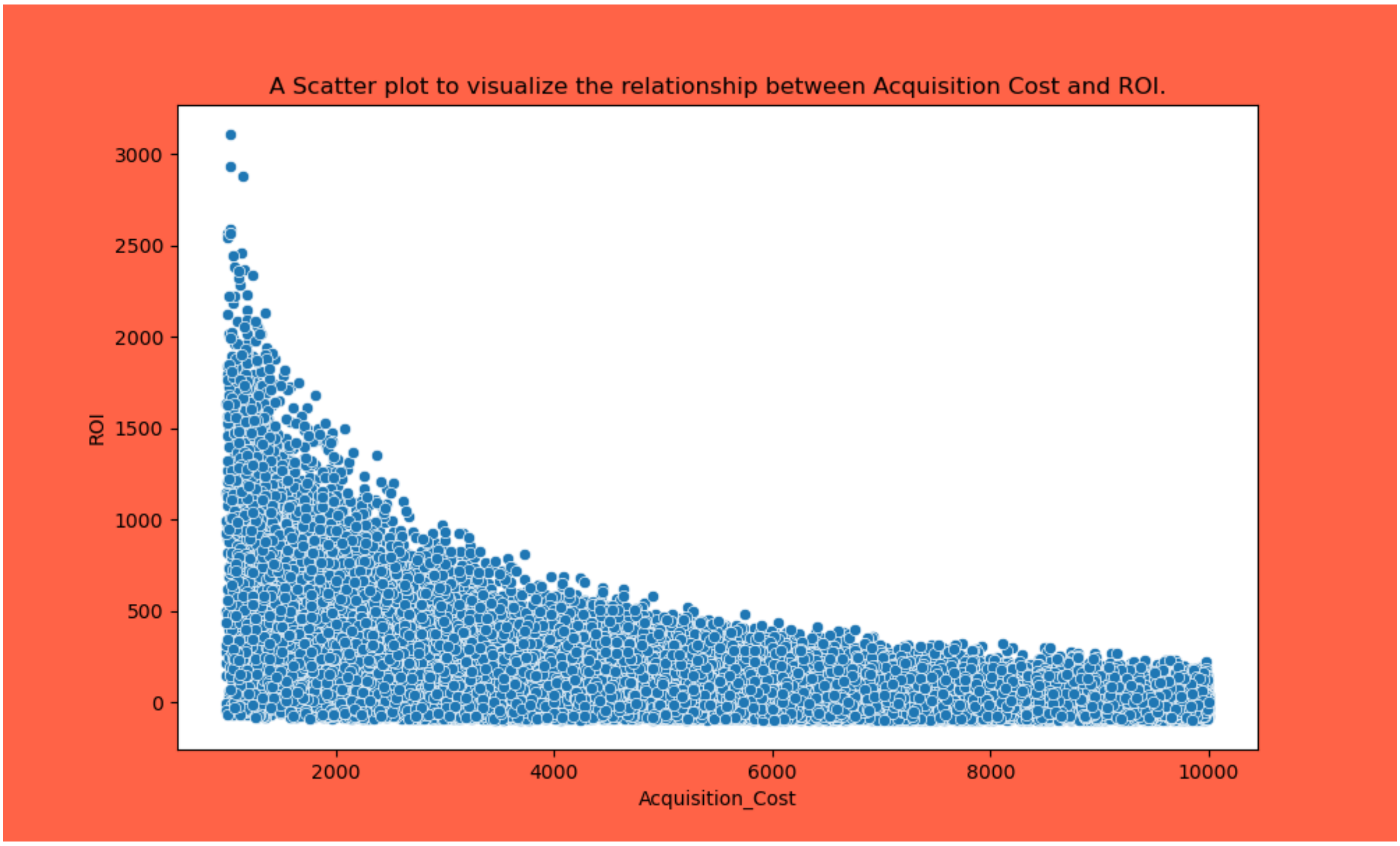
```
Occurrences of each category in Campaign_Type column:
Channel_Used
Facebook      3742
Google Ads    3694
Website       3688
Instagram     3649
YouTube       3632
Email         3624
Name: count, dtype: int64
```

3. Exploratory Data Analysis (EDA) and Visualization

✓ Campaign Performance

❖ Plot a scatter plot to visualize the relationship between `Acquisition_Cost` and `ROI`.

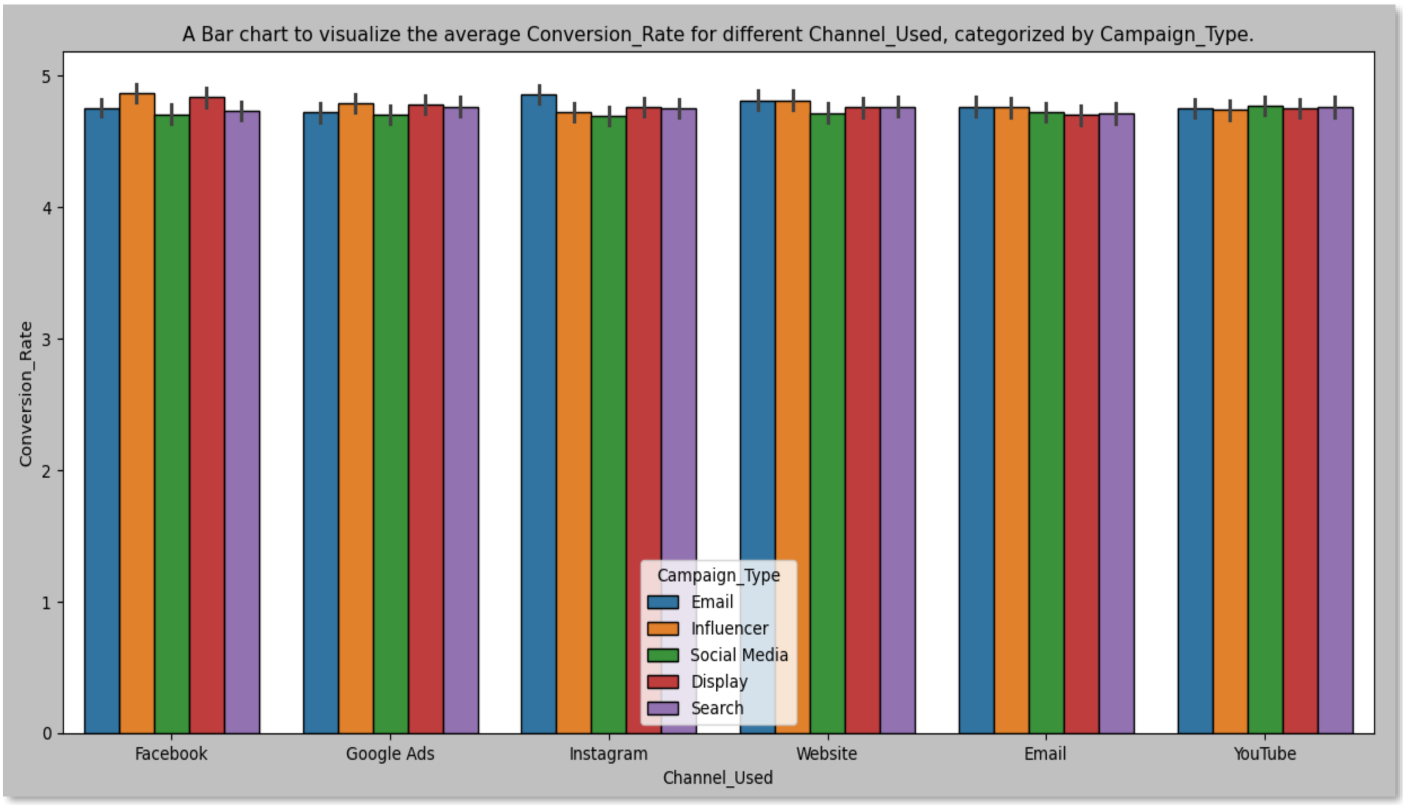
```
import matplotlib.pyplot as plt
import seaborn as sns
# Plot a scatter plot to visualize the relationship between Acquisition_Cost and ROI.
plt.figure(figsize=(10,6),facecolor='tomato')
sns.scatterplot( data= df,x='Acquisition_Cost',y='ROI')
plt.title("A Scatter plot to visualize the relationship between Acquisition Cost and ROI.")
plt.savefig('A Scatter plot to visualize the relationship between Acquisition_Cost and ROI.')
plt.show()
```



The scatter plot visualizes the relationship between Acquisition Cost and Return on Investment (ROI) for a set of marketing campaigns. It shows that as the Acquisition Cost increases, the ROI generally decreases. This suggests that while spending more on acquiring customers may lead to some increase in ROI, it is not a linear relationship and there may be diminishing returns at higher acquisition costs.

- Create a bar chart to visualize the average Conversion_Rate for different Channel_Used, categorized by Campaign_Type.

```
#Create a bar chart to visualize the average Conversion_Rate for different Channel_Used, categorized by Campaign_Type.
import seaborn as sns
import matplotlib.pyplot as plt
plt.figure(figsize=(15,7),facecolor='silver')
sns.barplot(data= df,x='Channel_Used',y='Conversion_Rate', hue ='Campaign_Type',estimator='mean',edgecolor='black')
plt.title('A Bar chart to visualize the average Conversion_Rate for different Channel_Used, categorized by Campaign_Type.')
plt.savefig('A Bar chart to visualize the average Conversion_Rate for different Channel_Used, categorized by Campaign_Type.')
plt.show()
```



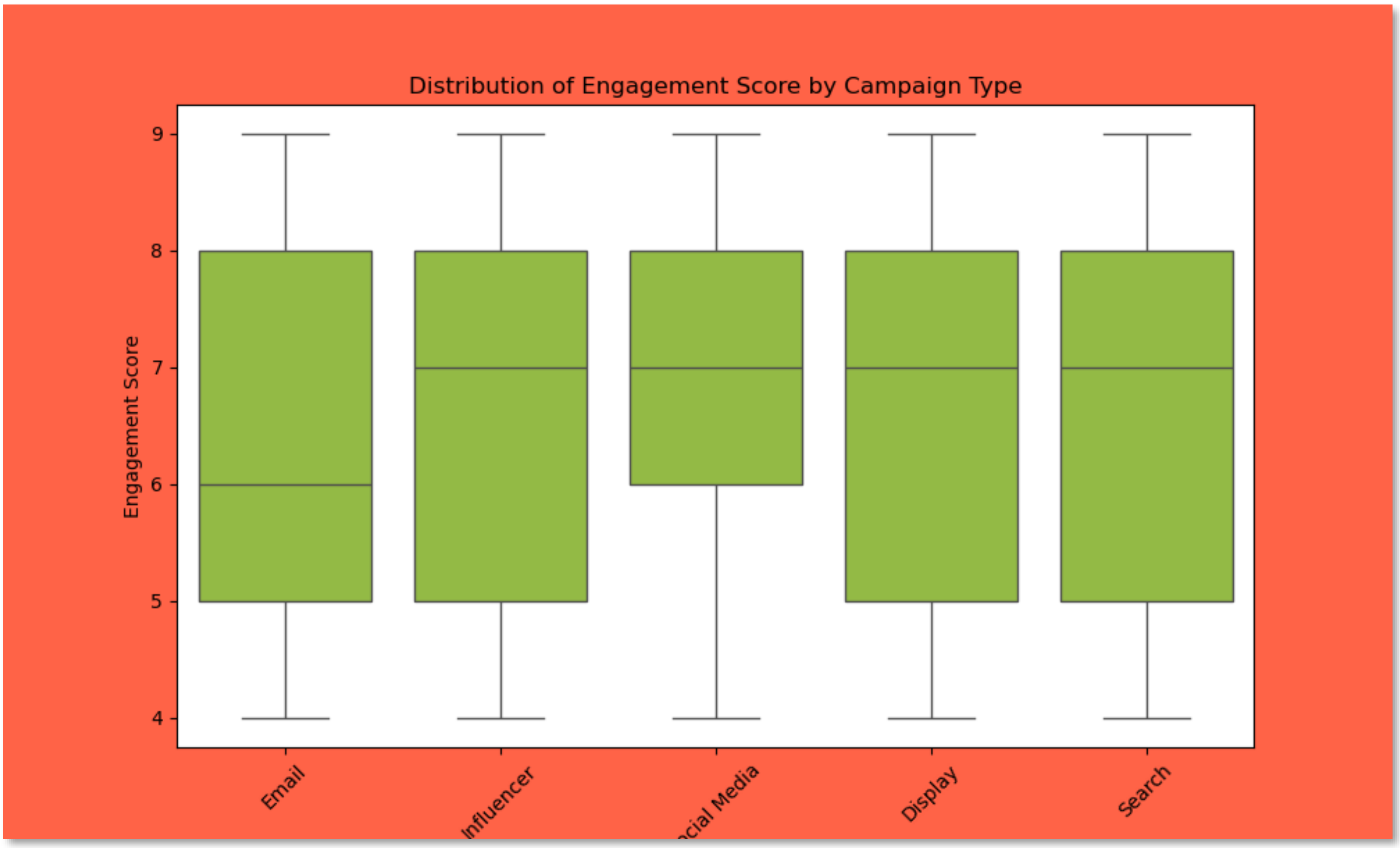
The bar chart presents the average conversion rates for different marketing channels, categorized by campaign type. Each group of bars represents a specific campaign type (Email, Influencer, social media, Display, Search), and within each group, the bars show the average conversion rate for each channel (Facebook, Google Ads, Instagram, Website, YouTube, Email).

Overall, the chart suggests that some channels consistently perform better than others across all campaign types. For example, Facebook and Google Ads generally have higher average conversion rates compared to channels like Instagram and YouTube. However, there are also interesting variations within each campaign type. For instance, while Facebook performs well for most campaign types, its performance might be slightly lower in Influencer campaigns.

This visual representation provides valuable insights into the relationship between campaign type, channel used, and conversion rates. It can help marketers optimize their strategies by selecting the most effective channels for each type of campaign.

❖ Visualize the distribution of Engagement_Score across different Campaign_Type using a box plot.

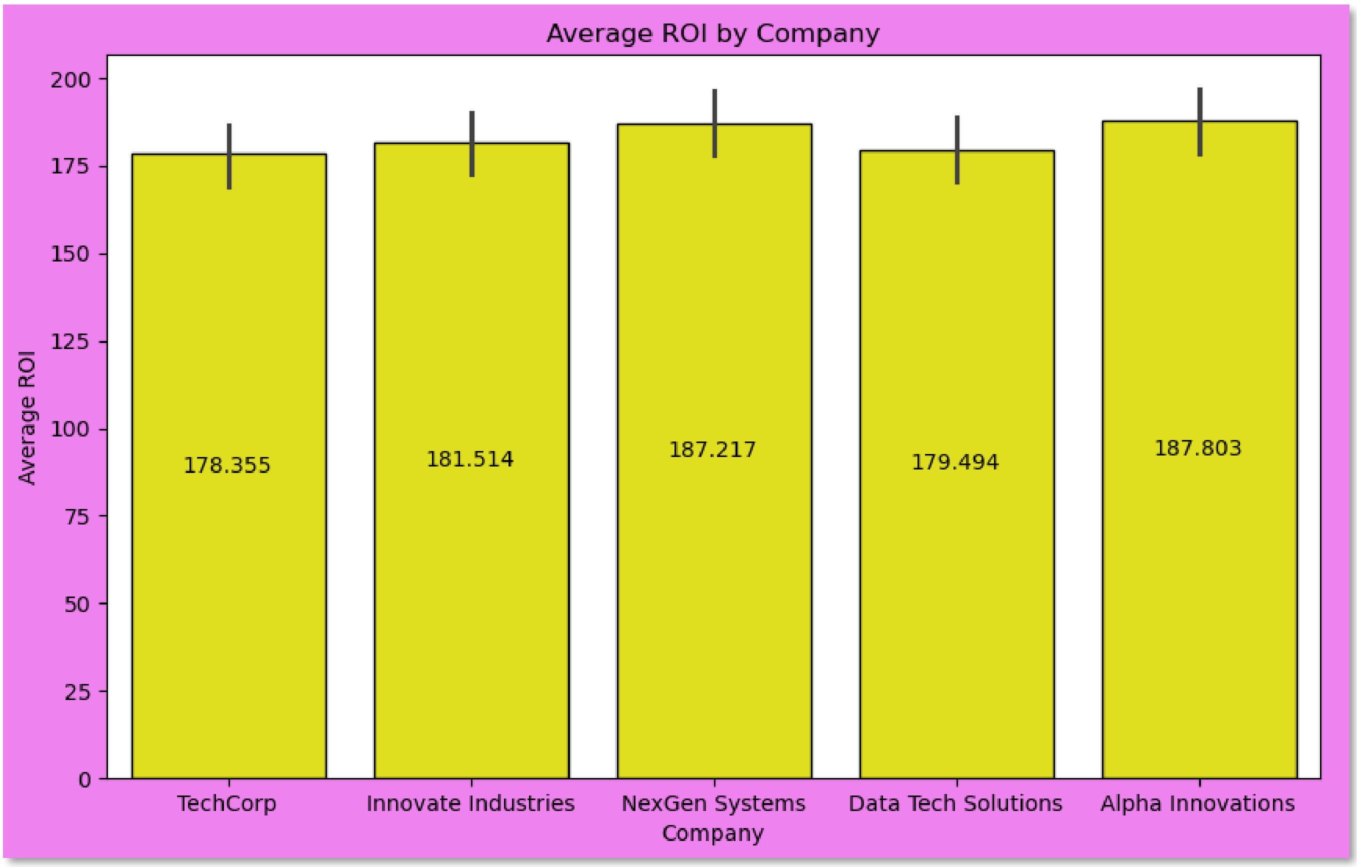
```
#Visualize the distribution of Engagement_Score across different Campaign_Type using a box plot.
plt.figure(figsize=(10, 6),facecolor='tomato')
sns.boxplot(data=df, x='Campaign_Type', y='Engagement_Score',color='yellowgreen')
plt.title('Distribution of Engagement Score by Campaign Type')
plt.xticks(rotation=45)
plt.xlabel('Campaign Type')
plt.ylabel('Engagement Score')
plt.savefig('Distribution of Engagement Score by Campaign Type')
plt.show()
```



The box plot illustrates the distribution of Engagement Scores across different Campaign Types. We can observe that the median Engagement Score is roughly the same for all campaign types, hovering around the 7 marks. The boxes, representing the interquartile range, also show a similar spread for most types. However, the Search campaigns appear to have a slightly wider spread compared to the others. Overall, the plot suggests that while the median Engagement Score is consistent across campaign types, there's some variation in the spread of scores, with Search campaigns exhibiting slightly more variability.

❖ Analyze the average ROI by Company using a bar chart to compare the profitability of campaigns conducted by different companies.

```
#Analyze the average ROI by Company using a bar chart to compare the profitability of campaigns conducted by different companies.
plt.figure(figsize=(10, 6),facecolor='violet')
ax=sns.barplot(data=df , x ='Company',y='ROI',estimator='mean',color='yellow',edgecolor='black')
ax.bar_label(ax.containers[0],label_type='center')
plt.title('Average ROI by Company')
plt.xlabel('Company')
plt.ylabel('Average ROI')
plt.savefig('Average ROI by Company')
plt.show()
```

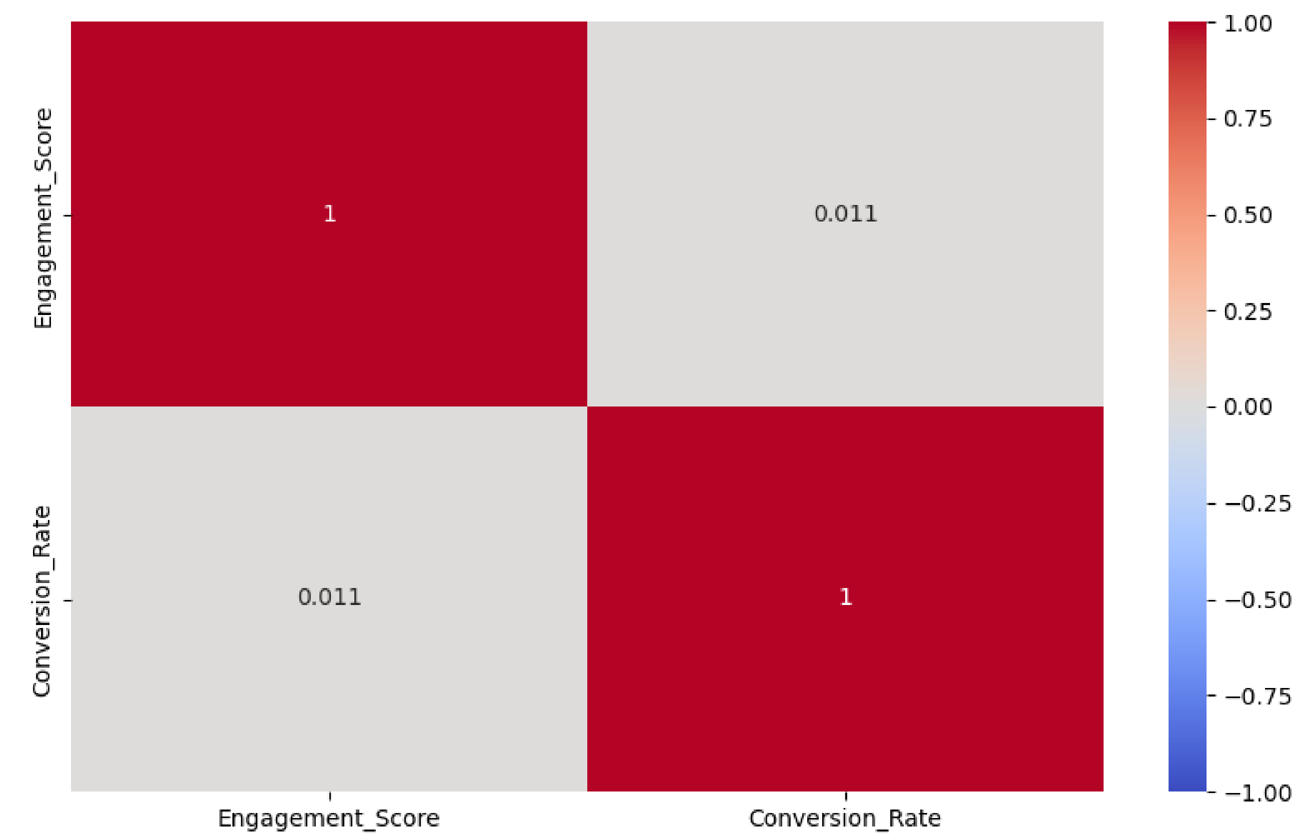


The bar chart presents the average ROI (Return on Investment) achieved by different companies.

- **Overall Performance:** All companies appear to be performing relatively well, with average ROIs ranging from around 178 to 187.
- **Company-Specific Performance:** While all companies are in the same ballpark, NexGen Systems seems to have the highest average ROI, followed closely by Alpha Innovations. TechCorp and Data Tech Solutions have slightly lower average ROIs.

❖ Examine the correlation between Engagement_Score and Conversion_Rate using a heatmap.

```
# Examine the correlation between Engagement_Score and Conversion_Rate using a heatmap.
plt.figure(figsize=(10,6))
sns.heatmap(correlation_matrix,annot=True,cmap='coolwarm', vmin=-1, vmax=1)
plt.savefig('Heatmap for Correlation for data ')
plt.show()
```

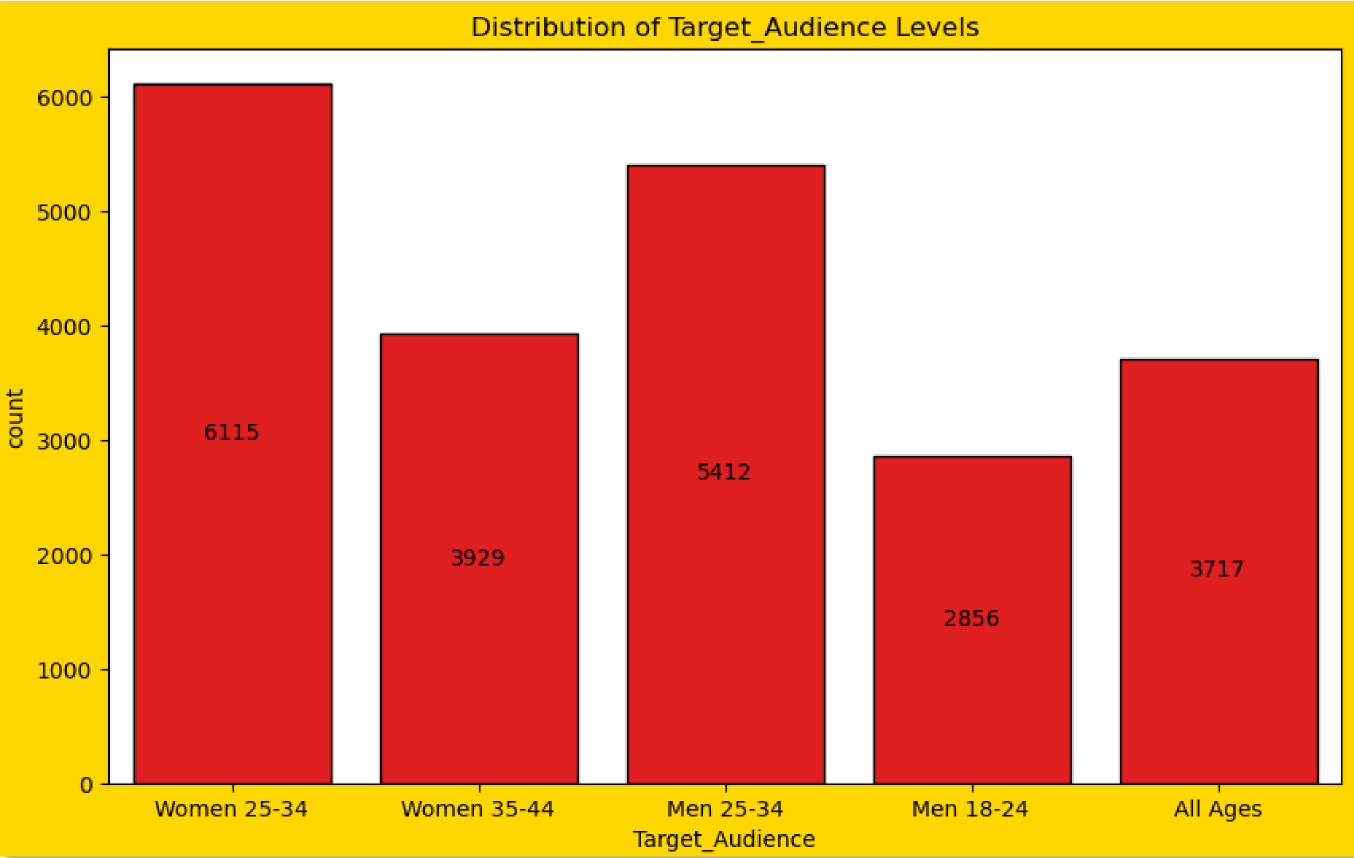


The heatmap depicts the correlation between Engagement Score and Conversion Rate. The diagonal cells, showing a perfect correlation of 1, are expected as they represent the correlation of a variable with itself. The off-diagonal cells reveal a very weak positive correlation of 0.011 between Engagement Score and Conversion Rate. This suggests that while a slight positive association exists, the relationship between these two variables is minimal. It's important to remember that correlation does not imply causation, and other factors might be influencing both variables. Further analysis, such as examining the correlation with other variables like Campaign Type or Channel Used, could provide deeper insights.

✓ Customer Segmentation:

❖ Create a count plot to visualize the distribution of Target_Audience.

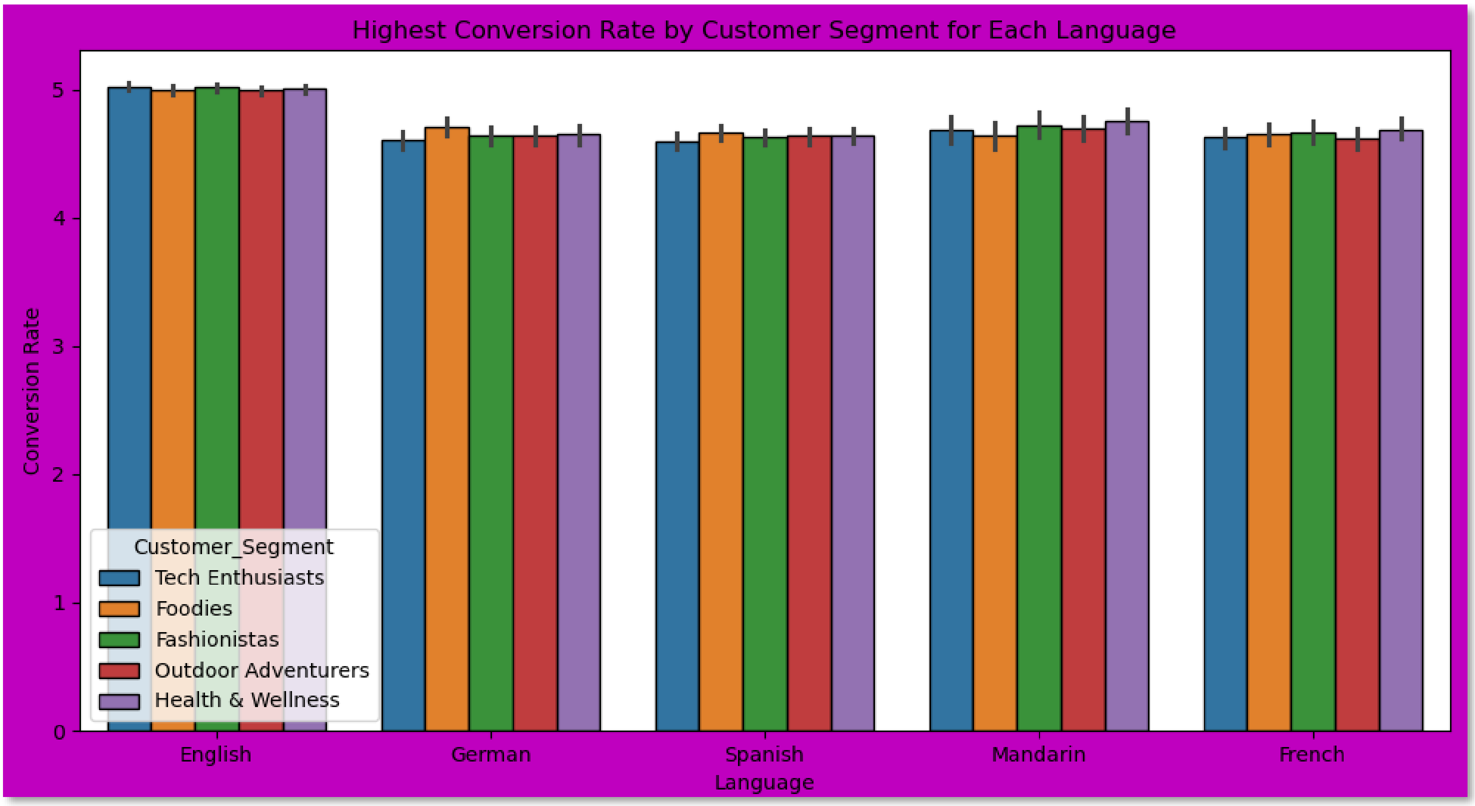
```
#create a count plot to visualize the distribution of Target_Audience.
plt.figure(figsize=(10,6),facecolor='gold')
ax=sns.countplot(data = df, x='Target_Audience',color='r',edgecolor='black')
ax.bar_label(ax.containers[0],label_type='center')
plt.title('Distribution of Target_Audience Levels')
plt.savefig('Distribution of Target_Audience Levels')
plt.show()
```



The plot presents the distribution of Target Audience Levels in the dataset. The x-axis represents the different Target Audience Levels, and the y-axis represents the count of instances for each level. The plot reveals that one Target Audience Level has the highest count, followed by several other levels with progressively lower counts. This indicates that a specific target audience segment is significantly more prevalent in the dataset compared to others.

❖ Identify which Customer_Segment has the highest Conversion_Rate for each Language using a bar chart.

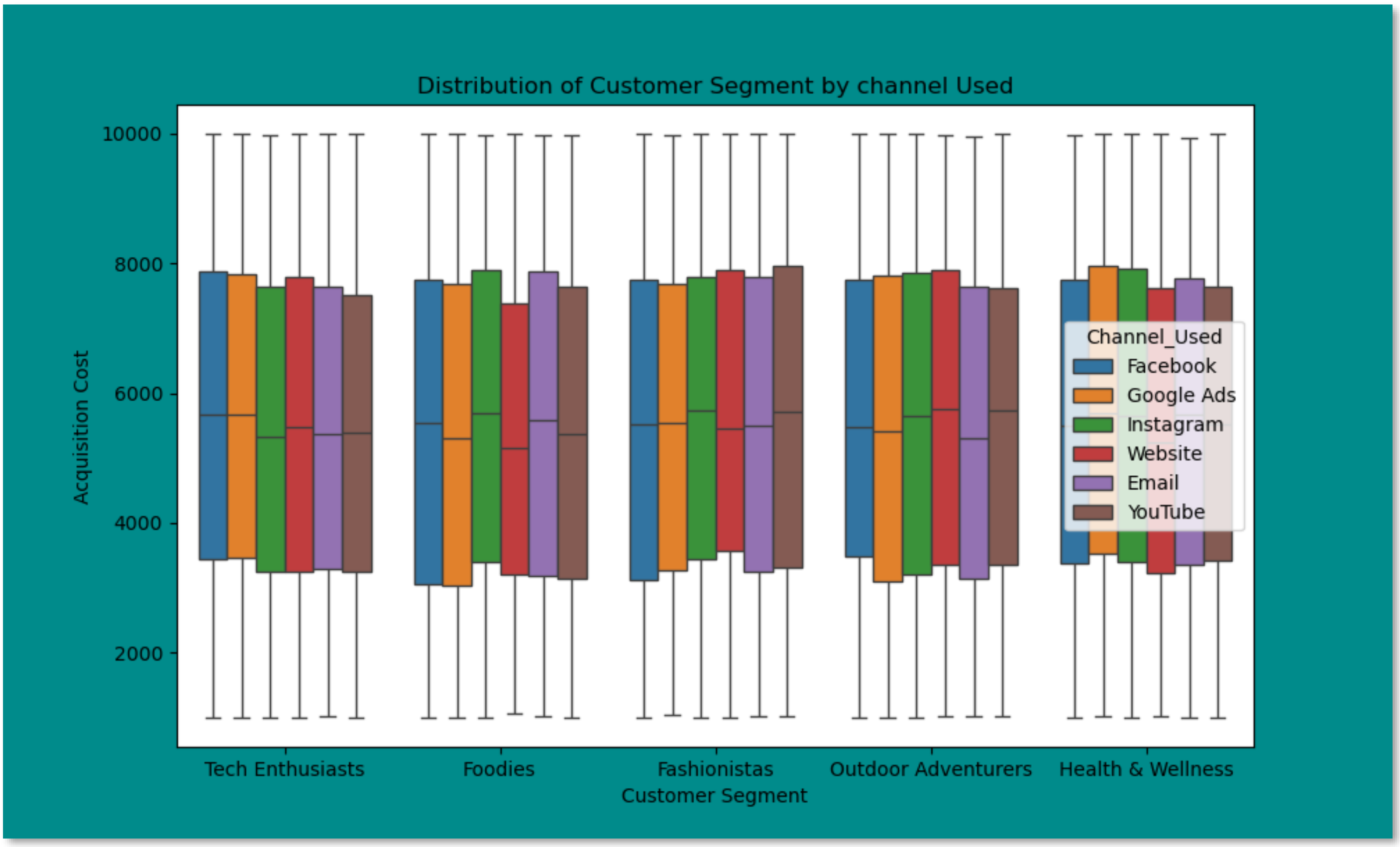
```
# Identify which Customer_Segment has the highest Conversion_Rate for each Language using a bar chart.
plt.figure(figsize=(12, 6),facecolor='m')
sns.barplot(data= df,x='Language',y='Conversion_Rate',hue='Customer_Segment',dodge=True ,estimator='mean',legend='brief',edgecolor='black')
plt.title('Highest Conversion Rate by Customer Segment for Each Language')
plt.xlabel('Language')
plt.ylabel('Conversion Rate')
plt.savefig('Highest Conversion Rate by Customer Segment for Each Language')
plt.show()
```



The bar chart presents the highest conversion rate for each customer segment across different languages. For each language, the bars represent the conversion rate for different customer segments: Tech Enthusiasts, Foodies, Fashionistas, Outdoor Adventurers, and Health & Wellness. The chart reveals that conversion rates vary across different customer segments and languages. Some customer segments consistently show higher conversion rates across multiple languages, while others demonstrate lower performance. This visualization helps identify which customer segments are most responsive within each language, enabling marketers to tailor their strategies for specific language groups and customer segments.

❖ Visualize the distribution of Acquisition_Cost across each Customer_Segment, categorized by Channel_Used, using a box plot.

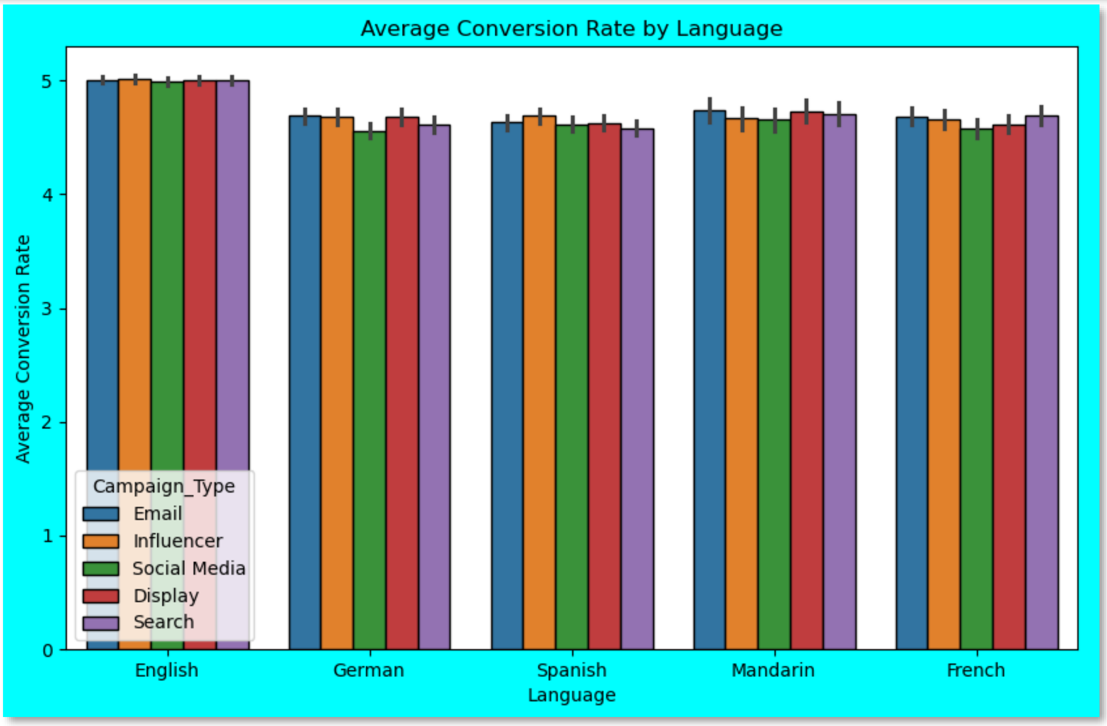
```
#Visualize the distribution of Acquisition_Cost across each Customer_Segment, categorized by Channel_Used, using a box plot.
plt.figure(figsize=(10,6),facecolor='darkcyan')
sns.boxplot(data= df,x='Customer_Segment',y='Acquisition_Cost',hue='Channel_Used')
plt.title('Distribution of Customer Segment by channel Used')
plt.xlabel('Customer Segment')
plt.ylabel('Acquisition Cost')
plt.savefig('Distribution of Customer Segment by channel Used')
plt.show()
```



The box plot illustrates the distribution of Acquisition Cost across different Customer Segments, categorized by the Channel Used. It shows how the cost of acquiring customers varies depending on the customer segment and the marketing channel employed. The plot reveals that there are variations in acquisition costs across different customer segments. Some segments might be more expensive to reach than others. Additionally, the choice of channel also influences acquisition costs. Certain channels might be more cost-effective for specific customer segments. The presence of outliers suggests that there might be some extreme cases where acquisition costs were significantly higher or lower than the typical range. This visualization can help identify the most cost-effective channels for acquiring different customer segments, inform budget allocation decisions, and guide further analysis into the factors contributing to the variations in acquisition costs.

❖ Analyze average Conversion_Rate by Language using a bar chart to compare the effectiveness of campaigns conducted in different languages.

```
#Analyze average Conversion_Rate by Language using a bar chart to compare the effectiveness of campaigns conducted in different languages.
plt.figure(figsize=(10, 6),facecolor='cyan')
sns.barplot(data=df, x ='Language', y='Conversion_Rate',hue='Campaign_Type',estimator='mean',edgecolor='black')
plt.title('Average Conversion Rate by Language')
plt.xlabel('Language')
plt.ylabel('Average Conversion Rate')
plt.savefig('Average Conversion Rate by Language')
plt.show()
```



The bar chart presents the average conversion rates for different campaign types across various languages. Each group of bars represents a language (English, German, Spanish, Mandarin, and French), and within each group, the bars show the average conversion rate for different campaign types: Email, Influencer, Social Media, Display, and Search.

Key Observations:

- **Language Variation:** The overall conversion rates seem to fluctuate slightly across languages, but the differences are not substantial.
- **Campaign Type Performance:** There are noticeable differences in conversion rates between campaign types within each language. Email campaigns consistently show higher conversion rates across most languages, suggesting they are generally more effective.
- **Campaign Type Consistency:** Some campaign types, like Email, tend to have consistently higher conversion rates across multiple languages.

Insights:

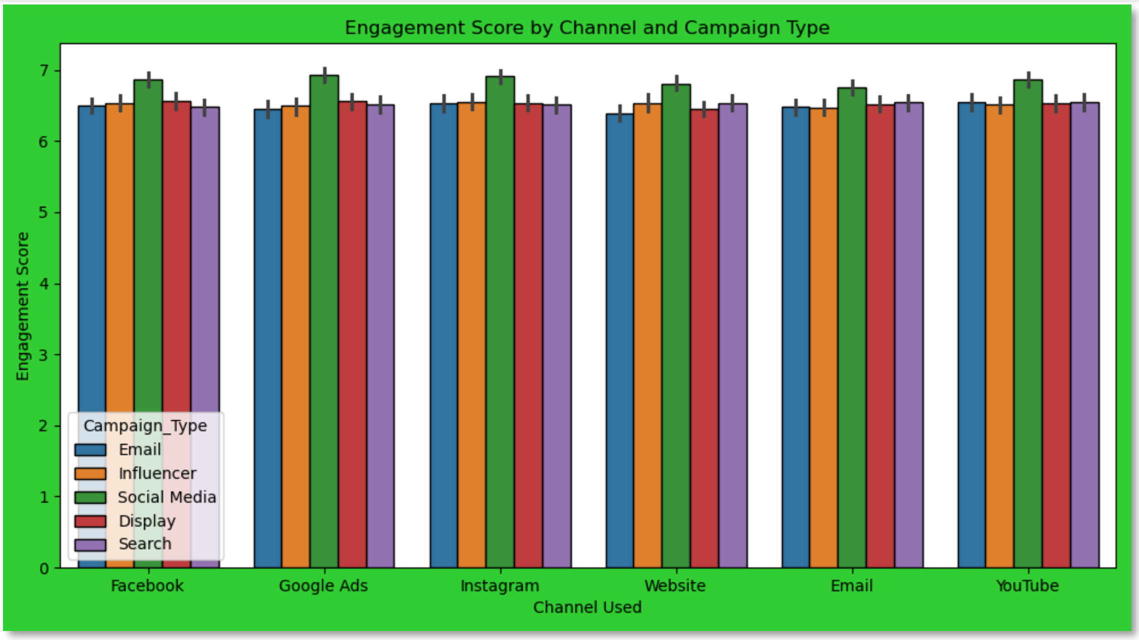
- **Channel Optimization:** This chart suggests that optimizing the use of Email campaigns could be beneficial across different language markets.
- **Content Localization:** The variations across languages highlight the importance of localized content and messaging that resonates with different cultural and linguistic preferences of each target audience.
- **Further Analysis:** Investigating the factors contributing to the variations in conversion rates (e.g., target audience, channel used, creative elements) could provide deeper insights.

✓ Channel Effectiveness:

❖ Compare the Engagement_Score for different Channels_Used, segmented by Campaign_Type, using a bar chart.

```python
# Compare the Engagement_Score for different Channels_Used, segmented by Campaign_Type, using a bar chart

plt.figure(figsize=(12, 6))
sns.barplot(data=df, x='Channel_Used', y='Engagement_Score', hue='Campaign_Type',edgecolor='black')
plt.title('Engagement Score by Channel and Campaign Type')
plt.xlabel('Channel Used')
plt.ylabel('Engagement Score')
plt.show()
```



The bar chart presents the average Engagement Score for different campaign types across various channels. Each group of bars represents a channel (Facebook, Google Ads, Instagram, Website, Email, YouTube), and within each group, the bars show the average Engagement Score for different campaign types: Email, Influencer, Social Media, Display, and Search.

Key Observations:

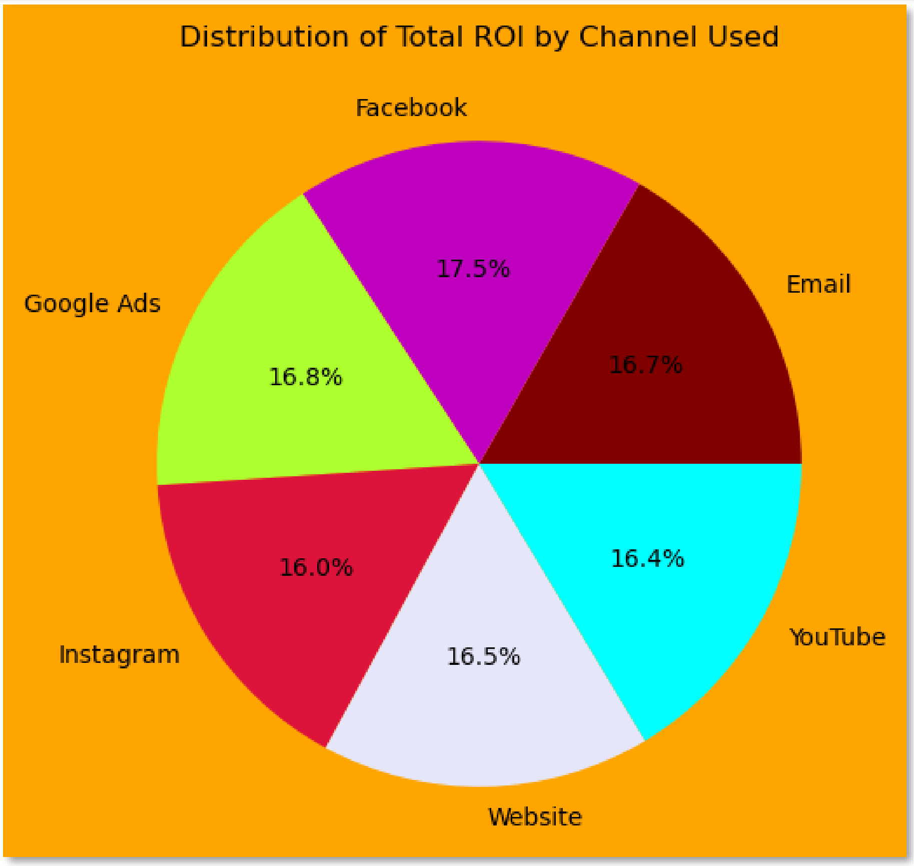
- **Channel Variation:** The overall Engagement Scores seem to fluctuate slightly across channels.
- **Campaign Type Performance:** There are noticeable differences in Engagement Scores between campaign types within each channel. Some campaign types consistently show higher Engagement Scores across most channels.
- **Campaign Type Consistency:** Some campaign types, like E mail, tend to have consistently higher Engagement Scores across multiple channels.

Insights:

- **Channel Optimization:** This chart suggests that optimizing the use of E mail campaigns could be beneficial across different channels.
- **Content Strategy:** The variations across channels highlight the importance of creating content that resonates with the specific audience and platform characteristics of each channel.
- **Further Analysis:** Investigating the factors contributing to the variations in Engagement Scores (e.g., target audience, creative elements, call to action) could provide deeper insights.

❖ Show the distribution of total ROI across different Channels_Used using a pie chart.

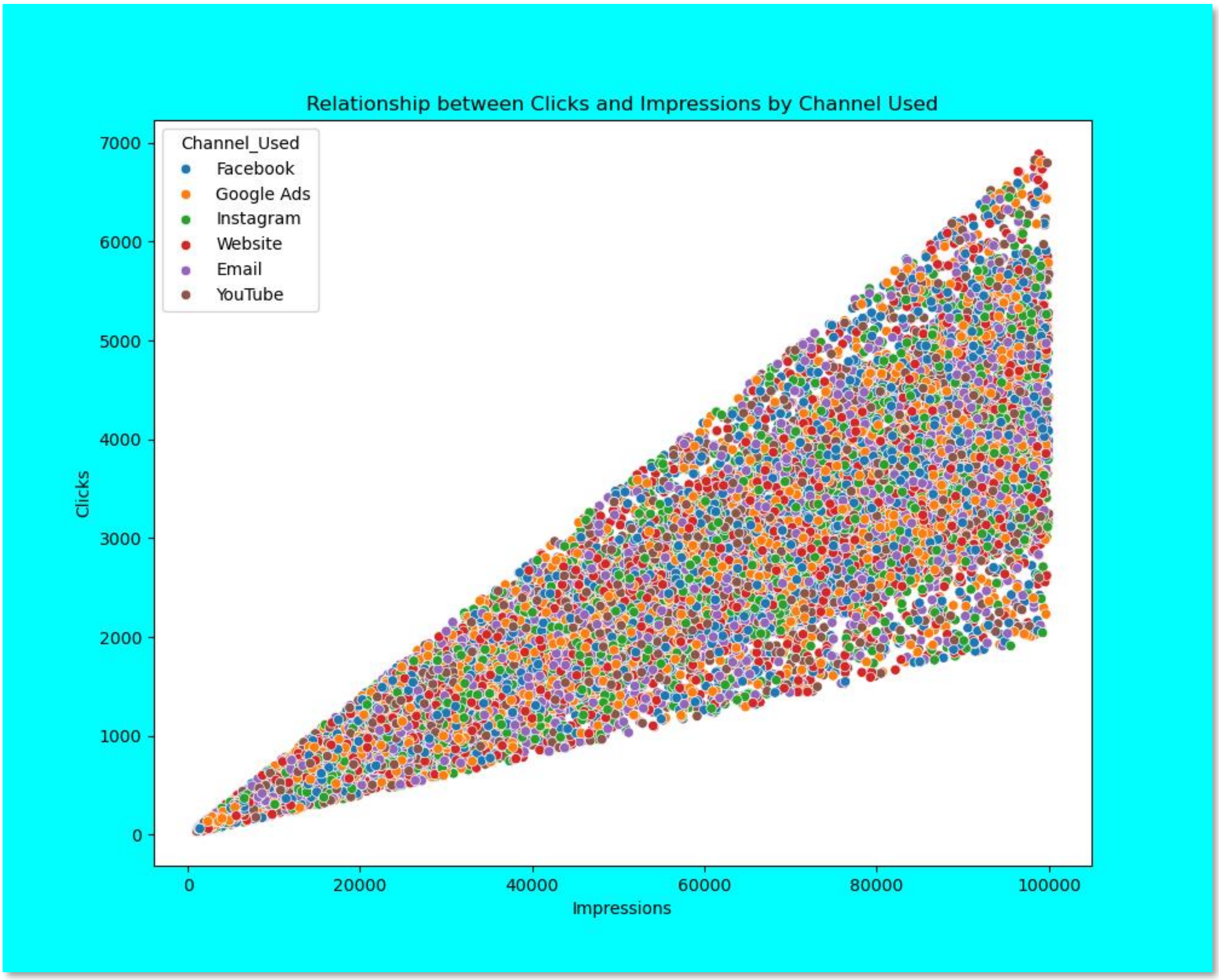
```
#Show the distribution of total ROI across different Channels_Used using a piechart.
df_grouped = df.groupby('Channel_Used').sum().reset_index()
plt.figure(figsize=(12,6 ), facecolor='orange')
plt.pie(df_grouped['ROI'], labels=df_grouped['Channel_Used'], autopct='%1.1f%%', colors=['maroon', 'm', 'greenyellow','crimson','lavender','aqua'],)
plt.title('Distribution of Total ROI by Channel Used',loc='center')
plt.savefig('Distribution of Total ROI by Channel Used')
plt.show()
```



The pie chart illustrates the distribution of Total ROI across different channels used in marketing campaigns. Facebook holds the largest share with 17.5%, followed closely by E mail at 16.7%. Google Ads and YouTube each account for 16.8% and 16.4% of the total ROI, respectively. Instagram and Website contribute 16.0% and 16.5% to the overall ROI. This visualization provides a clear picture of how the total ROI is distributed among the different channels used in the marketing campaigns

❖ Plot a scatter plot to show the relationship between Clicks and Impressions for each Channel_Used.

```
#Plot a scatter plot to show the relationship between Clicks and Impressions for each Channel_Used.
plt.figure(figsize=(10,8), facecolor='aqua')
sns.scatterplot(data= df,x='Impressions',y='Clicks',hue='Channel_Used')
plt.title('Relationship between Clicks and Impressions by Channel Used')
plt.xlabel('Impressions')
plt.ylabel('Clicks')
plt.savefig('Relationship between Clicks and Impressions by Channel Used')
plt.show()
```



The scatter plot visualizes the relationship between Clicks and Impressions for different channels used in marketing campaigns. Each point on the plot represents a data point with its corresponding values for Clicks and Impressions. The different colors represent the various channels used, such as Facebook, Google Ads, Instagram, Website, Email, and YouTube.

Key Observations:

- **Positive Correlation:** There is a clear positive correlation between Clicks and Impressions. As the number of Impressions increases, the number of Clicks also tends to increase. This suggests that a higher number of impressions generally leads to a higher number of clicks, which is an expected outcome in digital marketing campaigns.
- **Channel-Specific Patterns:** While the overall trend shows a positive correlation, there might be subtle variations in the relationship between Clicks and Impressions for different channels. Some channels might exhibit a stronger correlation than others, indicating differences in their click-through rates.
- **Outliers:** There might be some outliers (data points that deviate significantly from the general trend) for certain channels. These outliers could represent campaigns with unusually high or low click-through rates.

Insights:

- **Campaign Performance:** The plot can be used to assess the performance of different channels in terms of generating clicks. Channels with a higher number of clicks for a given number of impressions are generally considered more effective.
- **Budget Allocation:** The information can be used to optimize budget allocation across different channels, focusing on those with higher click-through rates.
- **Further Analysis:** Investigating the factors contributing to the variations in click-through rates for different channels (e.g., creative elements, targeting strategies, audience engagement) could provide deeper insights.

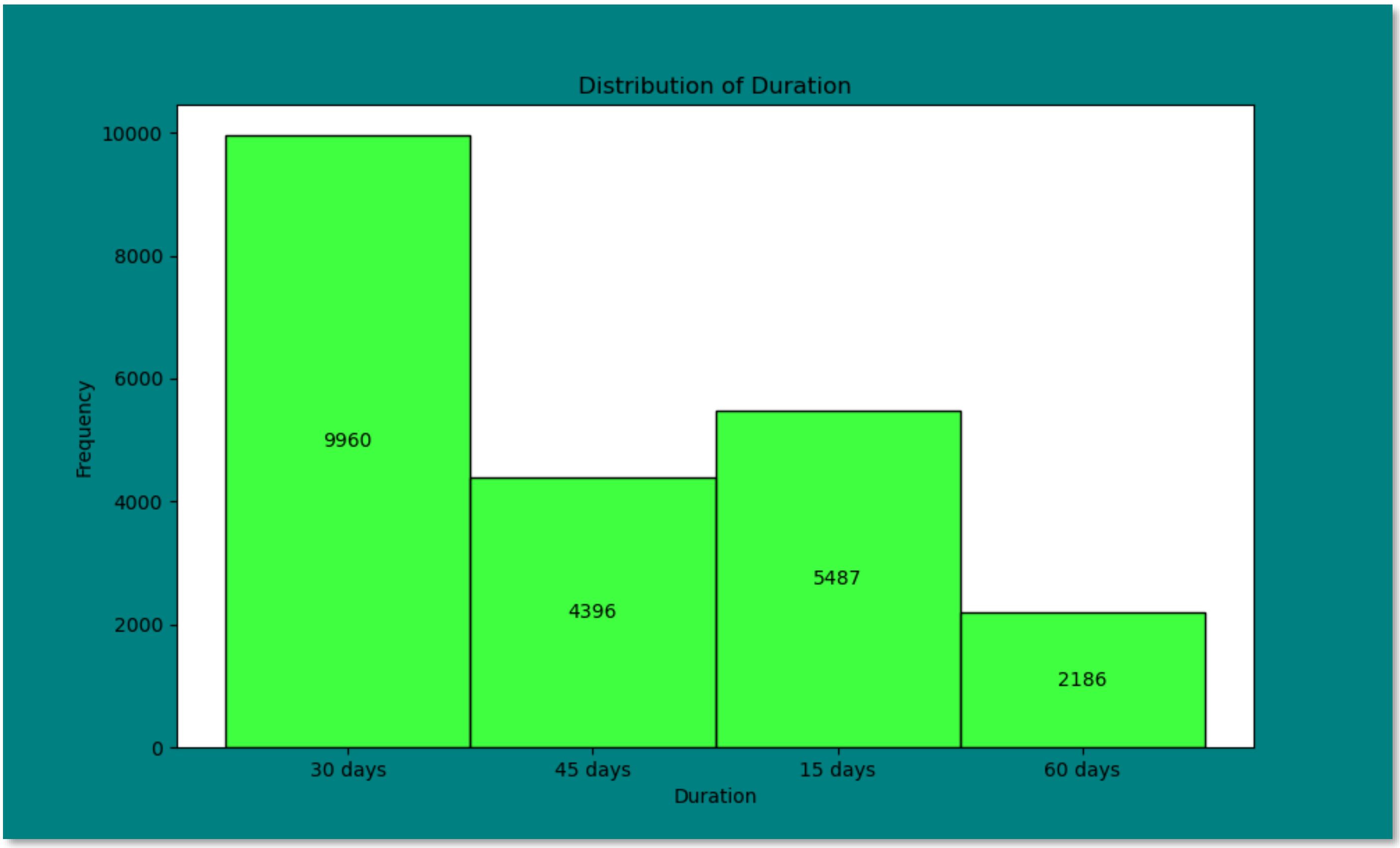
Note:

- The density of points in different regions of the plot can also provide insights into the distribution of click-through rates across different levels of impressions.

✓ Time-Based Analysis:

❖ Plot the distribution of Duration using a histogram.

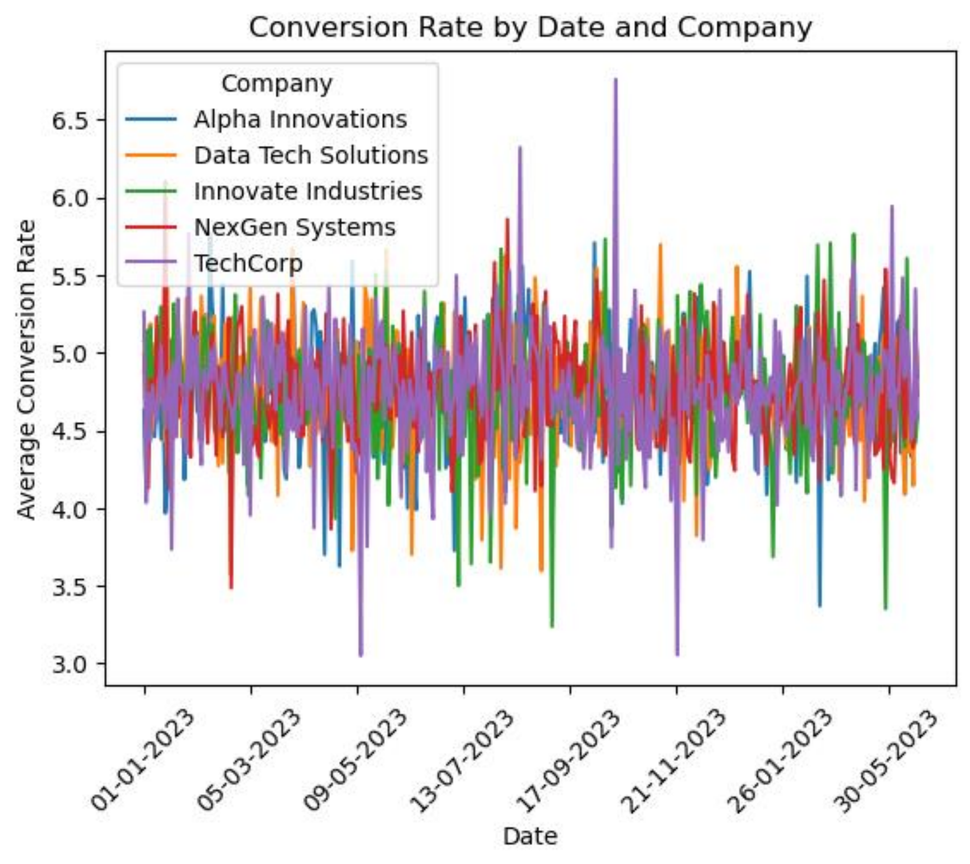
```
# Plot the distribution of Duration using a histogram
plt.figure(figsize=(10,6),facecolor='teal')
ax=sns.histplot(df['Duration'],color='lime',bins=10,edgecolor='black')
ax.bar_label(ax.containers[0],label_type='center')
plt.title("Distribution of Duration")
plt.xlabel('Duration')
plt.ylabel('Frequency')
plt.savefig('Distribution of Duration')
plt.show()
```



The bar chart presents the distribution of campaign durations. The x-axis represents the different durations (30 days, 45 days, 15 days, 60 days) and the y-axis represents the frequency or count of campaigns with that duration. The chart shows that the most frequent campaign duration is 30 days, followed by 45 days, 15 days, and 60 days. This visualization provides insights into the typical duration of campaigns and can be used to inform planning and decision-making regarding campaign timelines.

❖ Analyze how the overall Conversion_Rate has changed over Date for each Company using a line chart.

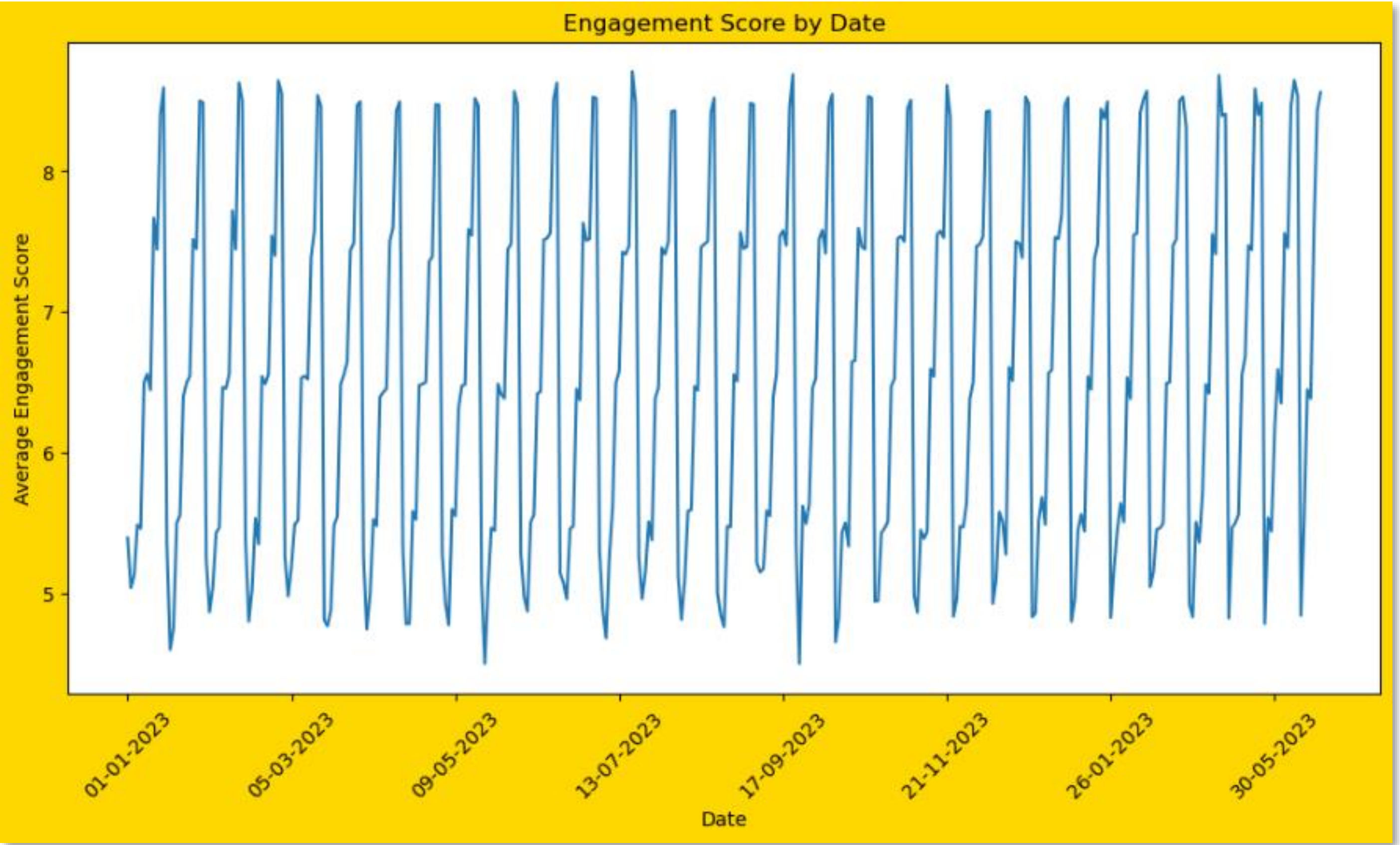
```
conversion_rate_by_date_company = df.groupby(['Date', 'Company'])['Conversion_Rate'].mean().unstack()
conversion_rate_by_date_company.plot(kind='line')
plt.xlabel('Date')
plt.ylabel('Average Conversion Rate')
plt.title('Conversion Rate by Date and Company')
plt.xticks(rotation=45)
plt.legend(title='Company')
plt.show()
```



The line chart visualizes the average conversion rate over time for different companies. Each line represents a company (Alpha Innovations, Data Tech Solutions, Innovate Industries, NexGen Systems, and TechCorp), and the x-axis represents the date. The chart shows how the average conversion rate fluctuates for each company over the given time period. Some companies exhibit higher conversion rates during certain periods, while others show more consistent performance. This visualization allows for the comparison of conversion rate trends across different companies and can be useful for identifying periods of high and low performance for each company.

❖ Examine the trend of Engagement_Score over Date with a line chart.

```
#Examine the trend of Engagement_Score over Date with a linechart
plt.figure(figsize=(12,6),facecolor='gold')
engagement_score_by_date = df.groupby('Date')['Engagement_Score'].mean()
engagement_score_by_date.plot(kind='line')
plt.xlabel('Date')
plt.ylabel('Average Engagement Score')
plt.title('Engagement Score by Date')
plt.xticks(rotation=45)
plt.show()
```

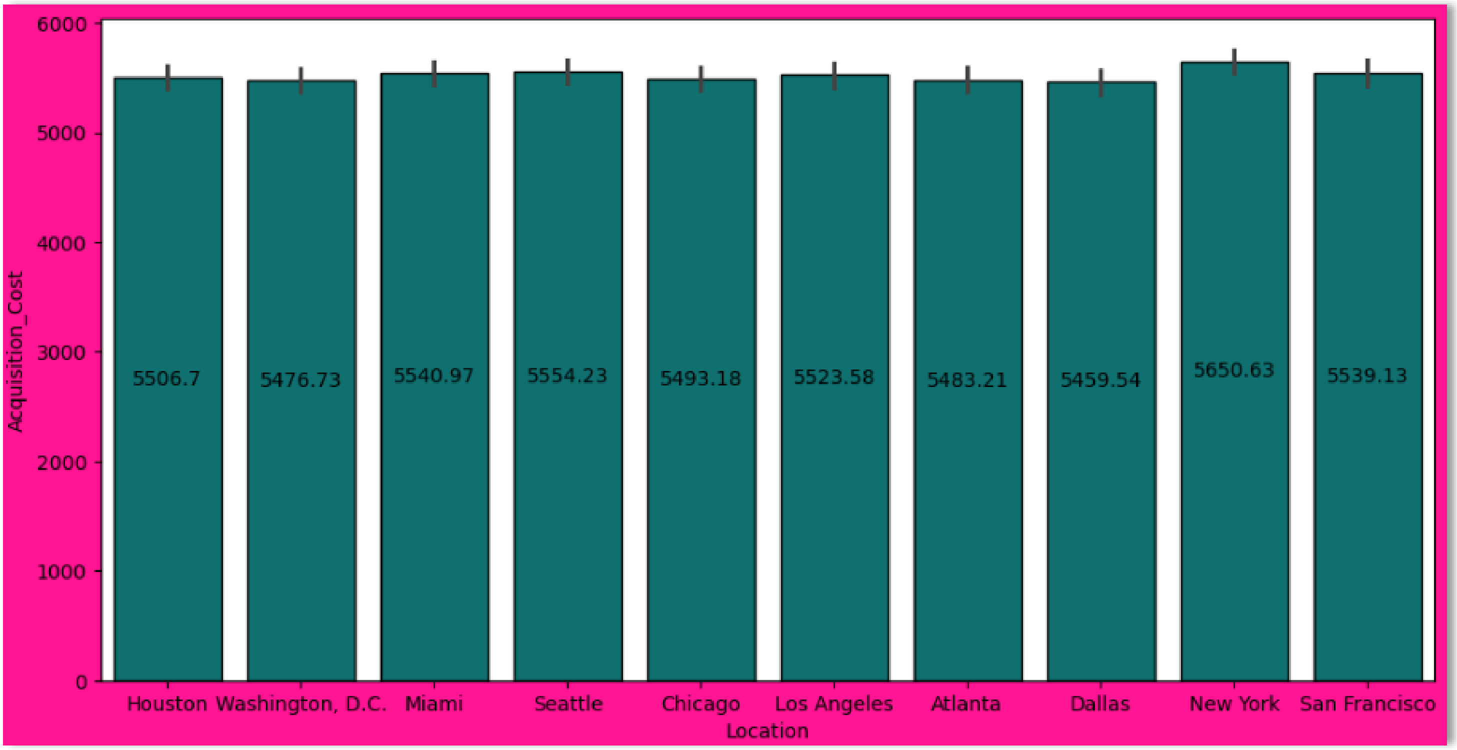


The line chart illustrates the trend of average Engagement Score over time. The x-axis represents the date, and the y-axis represents the average Engagement Score. The line shows how the average Engagement Score fluctuates over the given time period. The Engagement Score exhibits significant fluctuations over time, with periods of high Engagement Score followed by periods of low Engagement Score. There doesn't appear to be a consistent upward or downward trend in Engagement Score over the entire period. The fluctuations seem to be somewhat random, although there might be some seasonality in the data with potential recurring patterns of high and low Engagement Score at certain times of the year. However, more data and analysis would be needed to confirm this. Understanding the factors that influence Engagement Score can help marketers optimize their campaigns and improve audience engagement. Additionally, the data could be used to build predictive models that forecast future Engagement Score trends, allowing for proactive adjustments to marketing strategies.

✓ Geographic Analysis:

❖ Determine which location has the highest Acquisition_Cost using a bar chart.

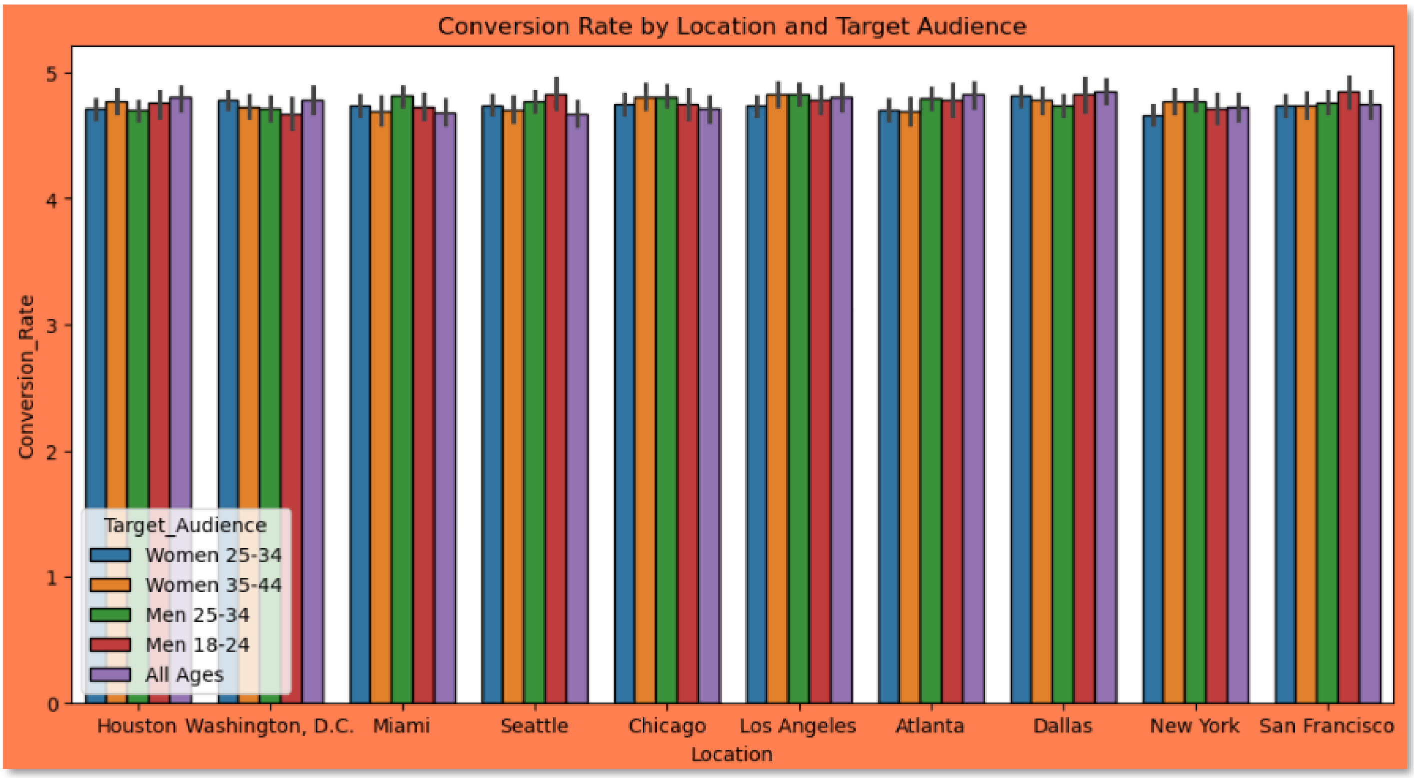
```
# Determine which location has the highest Acquisition_Cost using a bar chart.
plt.figure(figsize=(12,6),facecolor='deeppink')
ax=sns.barplot(data=df, x='Location',y='Acquisition_Cost',color='teal',edgecolor='black')
ax.bar_label(ax.containers[0],label_type='center')
plt.show()
```



The bar chart presents the average Acquisition Cost for different locations. The x-axis represents the different locations (Houston, Washington, D.C., Miami, Seattle, Chicago, Los Angeles, Atlanta, Dallas, New York, San Francisco), and the y-axis represents the average Acquisition Cost. The chart shows that the average Acquisition Cost is relatively similar across all locations, with minor variations. This suggests that the cost of acquiring customers is not significantly influenced by the location in this dataset.

❖ Visualize the Conversion_Rate by different Location, categorized by Target_Audience, using a bar chart.

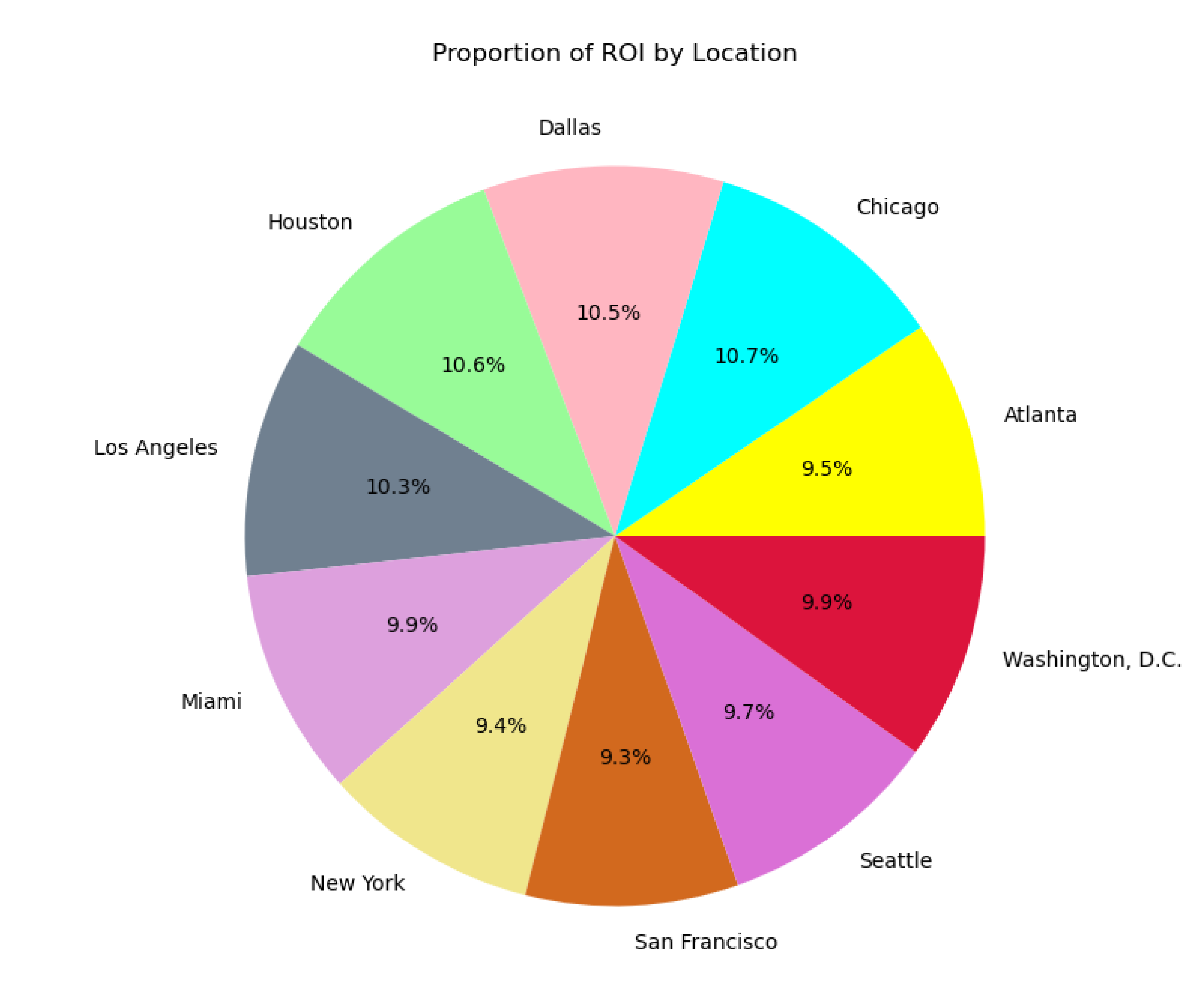
```
# Visualize the Conversion_Rate by different Location, categorized by Target_Audience, using a bar chart.
plt.figure(figsize=(12,6),facecolor='coral')
sns.barplot(data=df, x='Location',y='Conversion_Rate',hue='Target_Audience',edgecolor='black')
plt.title('Conversion Rate by Location and Target Audience')
plt.show()
```



The bar chart presents the conversion rate by location and target audience. Each group of bars represents a location (Houston, Washington, D.C., Miami, etc.), and within each group, the bars show the conversion rate for different target audiences: Women 25-34, Women 35-44, Men 25-34, Men 18-24, and All Ages. The chart reveals that conversion rates can vary across locations and target audiences. Some locations consistently show higher conversion rates across different target audiences, while others demonstrate lower performance. Similarly, some target audiences may have higher conversion rates in certain locations compared to others. This visualization provides valuable insights into the relationship between location, target audience, and conversion rates, which can inform marketing strategies and resource allocation decisions.

❖ Illustrate the proportion of ROI by Location using a pie chart.

```python
# Illustrate the proportion of ROI by Location using a pie chart.
df_grouped = df.groupby('Location').sum().reset_index()
plt.figure(figsize=(8, 8))
plt.pie(df_grouped['ROI'], labels=df_grouped['Location'], autopct='%1.1f%%', colors=['yellow', 'aqua','lightpink',
                                                                                   'palegreen','slategray','plum','khaki',
                                                                                   'chocolate','orchid','crimson'])
plt.title('Proportion of ROI by Location')
plt.savefig('Proportion of ROI by Location')
plt.show()
```



The pie chart illustrates the proportion of ROI (Return on Investment) generated by each location.

Key Observations:

- **Dallas** has the highest proportion of ROI at **10.7%**.
- **Houston** comes in second with **10.6%** of the total ROI.
- **Chicago** and **Dallas** also have a significant share, both contributing around **10.5%** to the total ROI.
- **Los Angeles** and **Washington, D.C.** contribute around **10.3%** and **9.9%**, respectively.
- The remaining locations (Miami, New York, San Francisco, Seattle, and Atlanta) each contribute between **9.3%** and **9.9%** to the total ROI.

Insights:

- **Performance Variation:** The chart highlights that while all locations contribute to the overall ROI, there are variations in their performance. Dallas and Houston appear to be the top performers in terms of ROI contribution.
- **Resource Allocation:** This information can be used to guide resource allocation decisions. More resources could be allocated to locations with higher ROI contributions to potentially maximize overall returns.
- **Further Analysis:** Investigating the factors contributing to the variations in ROI across locations (e.g., market size, competition, marketing strategies) could provide deeper insights.

Note:

- The pie chart provides a visual representation of the relative contribution of each location to the total ROI. However, it's important to consider the overall size of the market in each location when interpreting the results. A higher ROI percentage from a smaller market might not necessarily be more significant than a lower ROI percentage from a larger market.

Conclusion

This analysis explored various aspects of marketing campaign performance, leveraging data visualization techniques to uncover key insights. The findings reveal significant variations in campaign performance across different channels, campaign types, customer segments, and locations. Notably, Email campaigns consistently demonstrated high conversion rates across multiple channels and languages, highlighting the importance of email marketing in the overall marketing strategy. Furthermore, the analysis revealed that while some locations and customer segments showed higher average ROIs and conversion rates, the differences were not always substantial. This suggests that a multi-pronged approach, encompassing a diverse range of channels, targeting various customer segments, and adapting to local nuances, is crucial for maximizing campaign success. The analysis also highlighted the importance of monitoring key metrics like Engagement Score and Acquisition Cost over time to identify trends, optimize campaigns, and ensure continuous improvement. These insights can be used to inform future marketing decisions, optimize resource allocation, and ultimately enhance the overall return on investment for marketing campaigns.

Key Recommendations:

- **Prioritize Email Marketing:** Given the consistent high performance of Email campaigns across channels and languages, investing in and optimizing email marketing strategies should be a key focus.
- **Targeted Segmentation:** Tailoring marketing efforts to specific customer segments and their preferences is crucial for maximizing campaign effectiveness.
- **Channel Optimization:** Analyze channel performance in detail to identify the most cost-effective channels for each customer segment and campaign type.
- **Content Localization:** Adapt content and messaging to resonate with the cultural and linguistic nuances of different target audiences and markets.
- **Continuous Monitoring:** Continuously monitor key metrics like Engagement Score, Conversion Rate, and ROI to identify trends, track campaign performance, and make necessary adjustments to optimize campaign strategies.

This analysis provides a foundation for ongoing monitoring and optimization of marketing campaigns. By continuously analyzing data and refining strategies based on the insights gained, businesses can improve their marketing ROI and achieve sustainable growth.