SQL Basics and Marketing-Specific Queries

HNG Data Analysis Stage Three- HNG username(ACSP)

RELEVANT LINKS FOR THE TASK

- https://drive.google.com/file/d/11Ap001mY8Sa32pD5e_WoRyXE0jnXibxe/view?usp=sharing

- https://github.com/ACSP-Tech/DATA-ANALYSIS/blob/main/HNG-3.ipynb

- https://hng.tech/hire/data-analysts

CONTENT LAYOUT

INTRODUCTION

DATA CLEANING

CAMPAIGN TOTAL IMPRESSION

HIGHSET ROI CAMPAIGN

Top 3 Locations

Average Engagement **Score**

Overall CTR (

Most Cost-Effective Campaign

CTR Above 5% Threshold

Ranking Channels

CONCLUSION AND OVERVIEW

1. INTRODUCTION

SCHEMA BEFORE CLEANING:

	column_name name	data_type character varying
1	campaign_id	numeric
2	conversion_rate	numeric
3	acquisition_cost	money
4	roi	numeric
5	clicks	numeric
6	impressions	numeric
7	engagement_score	numeric
8	date	character varying
9	customer_segme...	character varying
10	company	character varying
11	campaign_type	character varying
12	target_audience	character varying
13	duration	character varying
14	channel_used	character varying
15	location	character varying

❑ Marketing Campaign Dataset

➢ The above named was given dataset for this task.
➢ This shape of the dataset is **200005 rows and 15 columns.**
➢ The schema as imported is given in the screenshot at the left.
➢ The first 20 rows are been displayed in the screenshot below.

SCREENSHOT OF THE FIRST 20 ROWS OF THE DATASET:

Data Output   Messages   Notifications

Showing rows: 1 to 20   Page No: 1   of 1

	campaign_id [PK] numeric	company character varying	campaign_type character varying	target_audience character varying	duration character varying	channel_used character varying	conversion_rate numeric	acquisition_cost numeric	roi numeric	location character varying	date date	clicks numeric	impressions numeric	engagement_score numeric	customer_segment character varying
1	1	Innovate Industries	Email	Men 18-24	30 days	Google Ads	0.04	16174.00	6.29	Chicago	2021-01-...	506	1922	6	Health & Wellness
2	2	NexGen Systems	Email	Women 35-44	60 days	Google Ads	0.12	11566.00	5.61	New York	2021-02-...	116	7523	7	Fashionistas
3	3	Alpha Innovations	Influencer	Men 25-34	30 days	YouTube	0.07	10200.00	7.18	Los Angeles	2021-03-...	584	7698	1	Outdoor Adventure...
4	4	DataTech Solutions	Display	All Ages	60 days	YouTube	0.11	12724.00	5.55	Miami	2021-04-...	217	1820	7	Health & Wellness
5	5	NexGen Systems	Email	Men 25-34	15 days	YouTube	0.05	16452.00	6.5	Los Angeles	2021-05-...	379	4201	3	Health & Wellness
6	6	DataTech Solutions	Display	All Ages	15 days	Instagram	0.07	9716.00	4.36	New York	2021-06-...	100	1643	1	Foodies
7	7	NexGen Systems	Email	Women 35-44	60 days	Website	0.13	11067.00	2.86	Los Angeles	2021-07-...	817	8749	10	Tech Enthusiasts
8	8	DataTech Solutions	Search	Men 18-24	45 days	Google Ads	0.08	13280.00	5.55	Los Angeles	2021-08-...	624	7854	7	Outdoor Adventure...
9	9	Alpha Innovations	Social Media	Women 35-44	15 days	Facebook	0.09	18066.00	6.73	Chicago	2021-09-...	861	1754	6	Tech Enthusiasts
10	10	TechCorp	Email	Women 35-44	15 days	Instagram	0.09	13766.00	3.78	Los Angeles	2021-10-...	642	3856	3	Tech Enthusiasts
11	11	NexGen Systems	Display	Men 25-34	45 days	Email	0.12	8590.00	3.49	New York	2021-11-...	321	6628	10	Tech Enthusiasts
12	12	Innovate Industries	Influencer	Men 25-34	60 days	Google Ads	0.05	17502.00	3.59	Los Angeles	2021-12-...	659	8948	1	Foodies
13	13	TechCorp	Social Media	Men 25-34	60 days	Facebook	0.09	17189.00	4.91	Chicago	2021-01-...	677	8817	10	Tech Enthusiasts
14	14	TechCorp	Email	Men 25-34	45 days	Instagram	0.14	9975.00	7.06	New York	2021-01-...	994	2201	4	Health & Wellness
15	15	TechCorp	Display	All Ages	45 days	Website	0.04	11346.00	5.28	Chicago	2021-01-...	482	8470	1	Outdoor Adventure...
16	16	Innovate Industries	Social Media	Women 35-44	60 days	YouTube	0.11	9407.00	2.91	New York	2021-01-...	299	1512	5	Health & Wellness
17	17	Innovate Industries	Display	Women 35-44	45 days	Website	0.08	5478.00	4.53	Houston	2021-01-...	931	2488	3	Fashionistas
18	18	Alpha Innovations	Influencer	Men 18-24	15 days	Instagram	0.14	9485.00	4.48	Miami	2021-01-...	218	9264	9	Health & Wellness
19	19	Alpha Innovations	Social Media	Men 25-34	60 days	Google Ads	0.07	19224.00	6.08	New York	2021-01-...	182	5798	1	Foodies
20	20	DataTech Solutions	Influencer	Men 25-34	15 days	Google Ads	0.09	10258.00	3.83	Miami	2021-01-...	193	3677	1	Tech Enthusiasts

2. DATA CLEANING – DATA TYPES
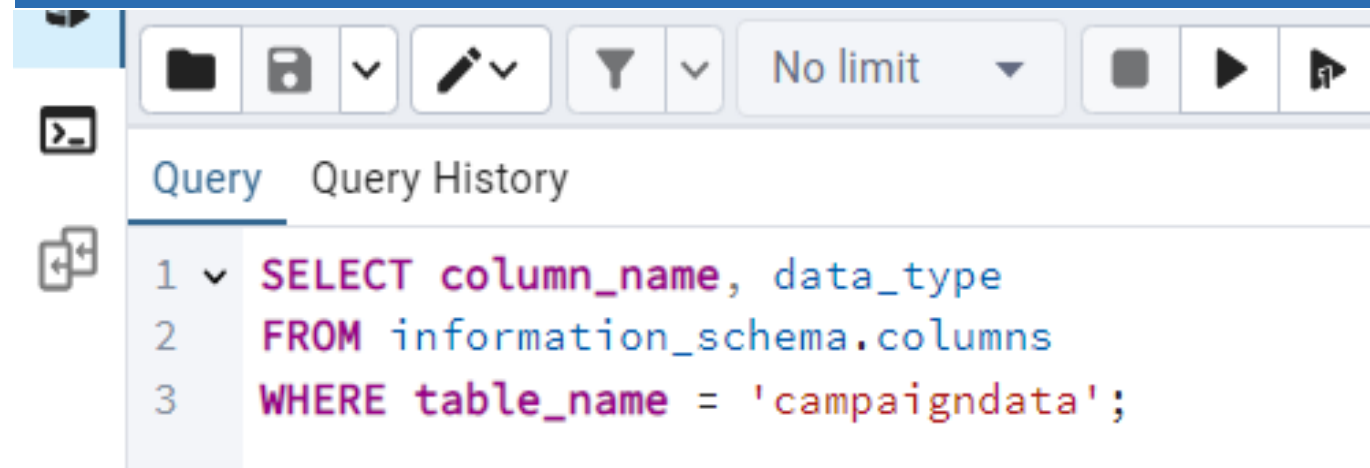
SCHEMA AFTER CLEANING:

	column_name 🔒 name	data_type 🔒 character varying
1	campaign_id	numeric
2	conversion_rate	numeric
3	acquisition_cost	numeric
4	roi	numeric
5	date	date
6	clicks	numeric
7	impressions	numeric
8	engagement_score	numeric
9	company	character varying
10	campaign_type	character varying
11	target_audience	character varying
12	duration	character varying
13	channel_used	character varying
14	customer_segme...	character varying
15	location	character varying

❑ **Marketing Campaign Dataset**
  ➢ The dataset was cleaned to make acquisition cost and date to the data type needed for the analysis.
  ➢ The schema after cleaning the data types is given in the screenshot at the left.
  ➢ The schema as imported is given in the screenshot at the left.
  ➢ The SQL query code used to obtain the schema information is displayed in the screenshot below.

SQL query code used to obtain the schema information:

```
1  SELECT column_name, data_type
2  FROM information_schema.columns
3  WHERE table_name = 'campaigndata';
```

3. TOTAL IMPRESSIONS FOR EACH CAMPAIGN

❑ Marketing Campaign Dataset

➢ The **Campaign ID is an auto-incrementing unique identifier** ranging from 1 to 200,005, representing different marketing campaigns.

➢ Each campaign has a varying number of total impressions, as shown by the orange area in the chart.

➢ The light blue area represents the campaign ID count, which **increases linearly** as more campaigns are recorded.

➢ The **spikes and dips** in impressions indicate that some campaigns performed significantly **better or worse** than others.

STACKED LINE CHART OF CAMPAIGN ID AND TOTAL IMPRESSION FOR EACH CAMPAIGN
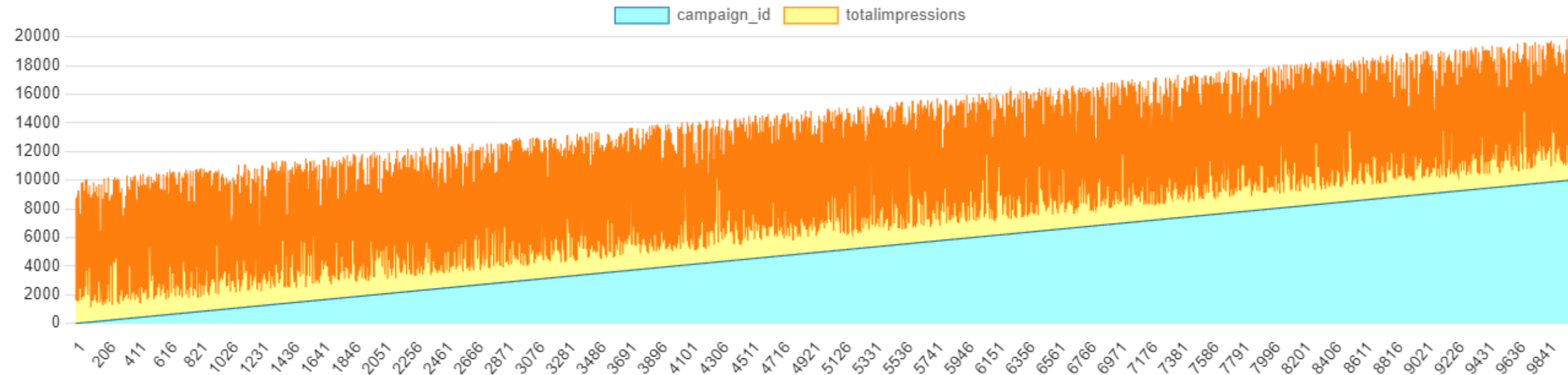
4. CAMPAIGN WITH THE HIGHEST ROI

❑ Campaign ID - 160 Campaigns with the Highest Return on Investment.

➢ **The highest Return on Investment (ROI) is 8,** achieved by a select number of campaigns.

➢ Only **160 campaigns (0.08%)** out of the total dataset attained this maximum ROI value.

➢ A query filtering the highest ROI and limiting results to 1 returns **Campaign ID 168,** belonging to **NexGen Systems**.

➢ The blue circle in the chart scatter plot represent the ROI values achieved by 160 different campaigns, with each campaign ID displayed on the x-axis.

➢ The uniform distribution of the filtered results indicates that multiple campaigns share similar ROI values of 8.

Scatter plot of campaign with the highest return on Investment

Scatter plot of Campaign with the Highest ROI
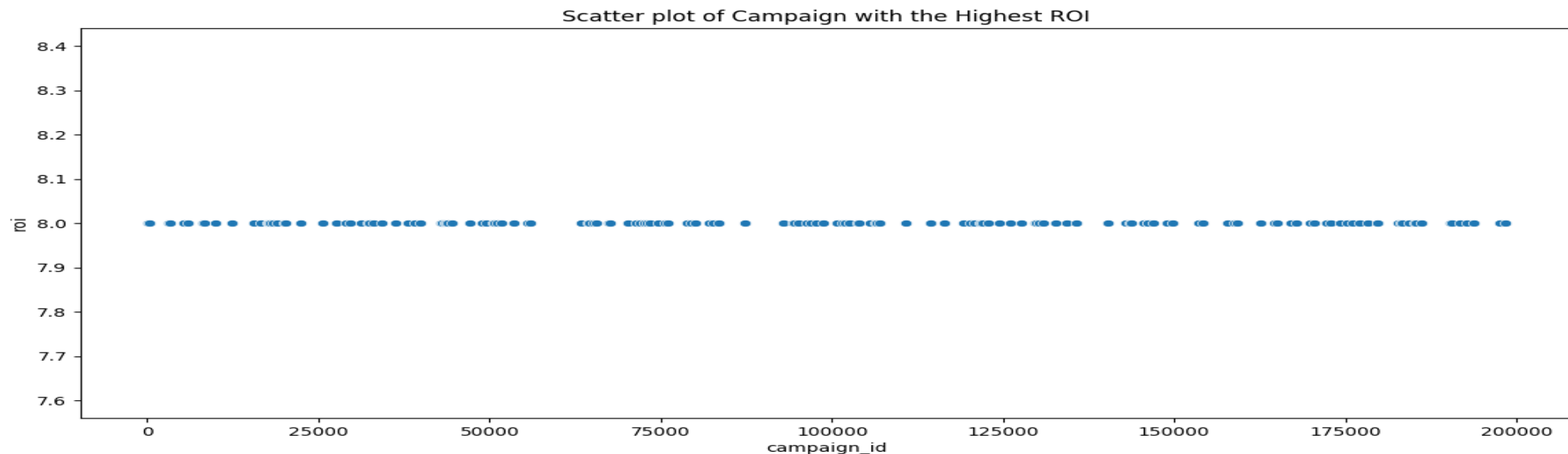
5. TOP THREE LOCATION WITH THE MOST IMPRESSIONS

❑ Top three -New York, Miami and Chicago

➢ The blue pie slice represents New York, which has the highest total impressions of 221,359,758.

➢ The **orange** slice corresponds to **Miami**, with **221,347,730** impressions, ranking second.

➢ The **green** slice represents **Chicago**, with **219,999,364** impressions, slightly lower than Miami.

➢ The **three locations** account for the highest share of total impressions, indicating strong audience engagement in these regions.

Top 3 Locations with the Most Impressions
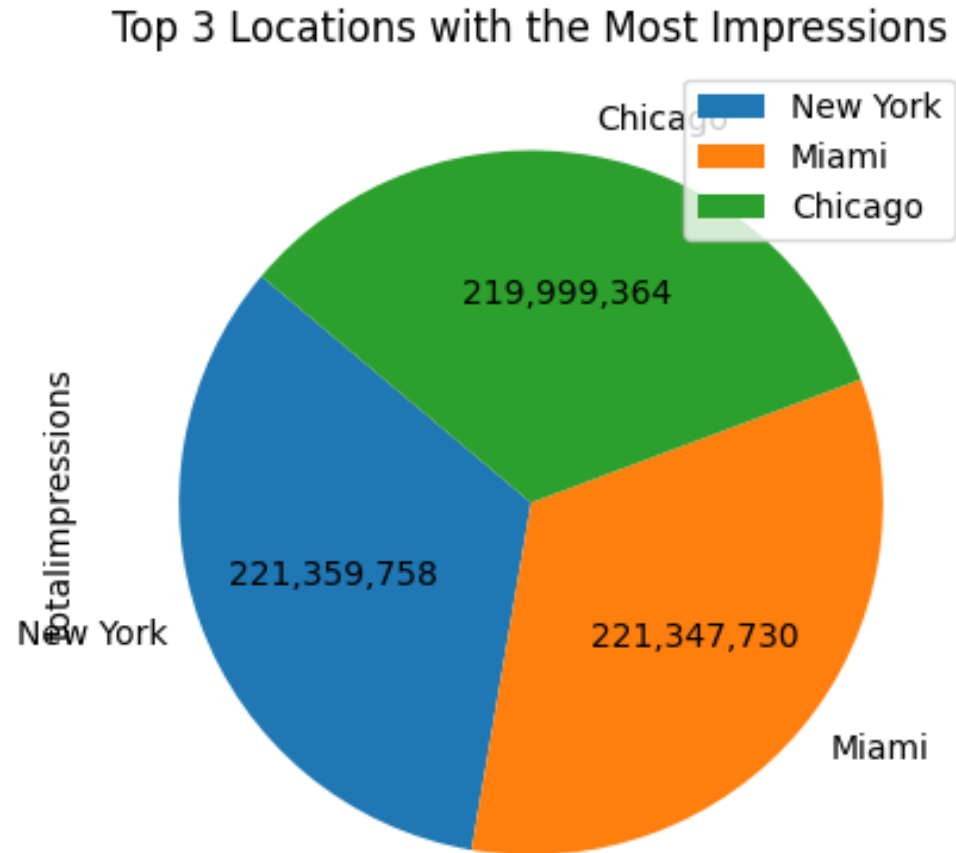
6. AVERAGE ENGAGEMENT SCORE BY TARGET AUDIENCE

❑ The highest **average engagement score** is observed among **Men aged 18-24** , indicating strong interaction from this demographic.

➢ **Women aged 25-34** and **Men aged 25-34** show moderate engagement levels, suggesting a balanced interest from both groups..

➢ **Women aged 35-44** and the **All Ages category** have the lowest engagement scores, implying a relatively lower interaction rate from these segments.
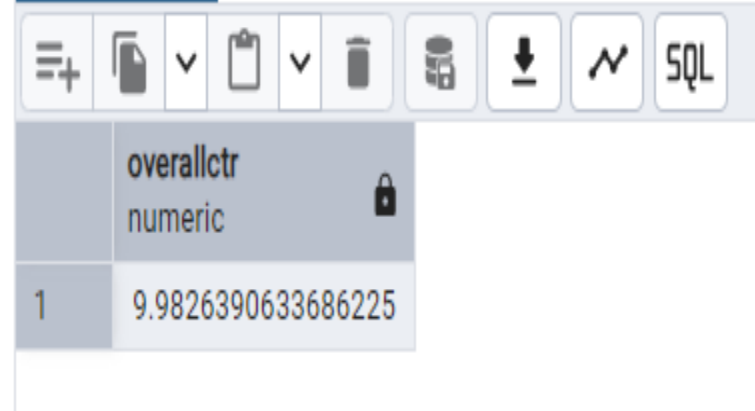
Data Output | Messages | Notifications

	target_audience character varying 🔒	avgengagementscore numeric 🔒
1	All Ages	5.4868693935683766
2	Men 18-24	5.5150152760873345
3	Men 25-34	5.4919798121127324
4	Women 25-34	5.4927398595456477
5	Women 35-44	5.4865702479338843

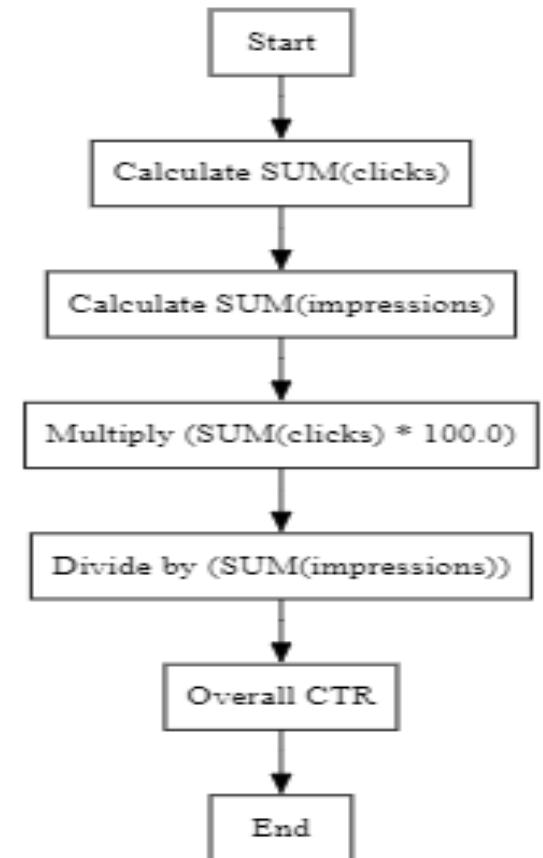Average Engagement Score by Target Audience
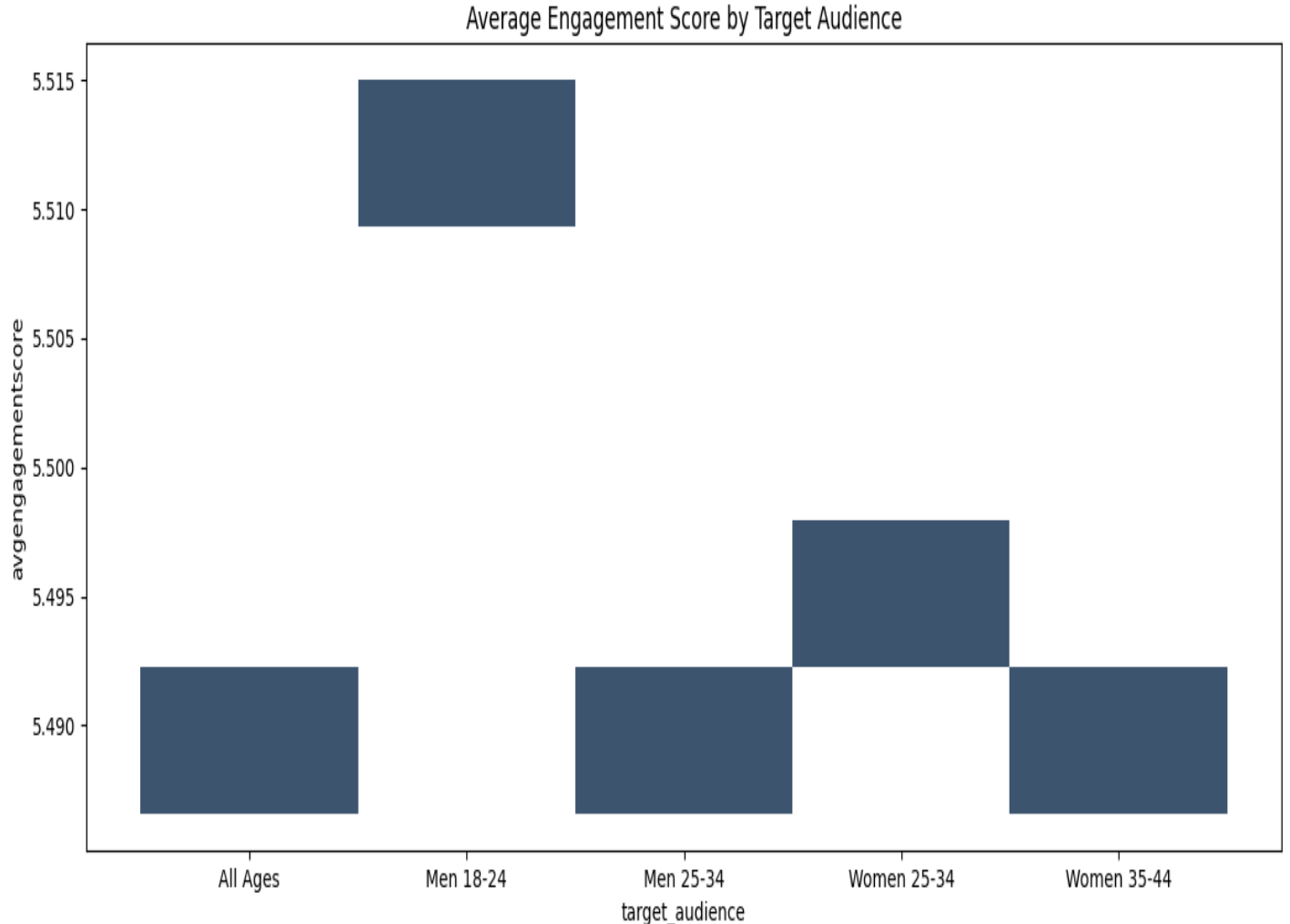
7. OVERALL CTR (Click-Through Rate)

❑ The SQL Query execution resulted in an Overall CTR of 9.98% (from the screenshot).

➢ Click-Through Rate (CTR) measures how effectively our ads are resonating with our target audience

➢ This 9.98% CTR signifies that nearly 1 in 10 users who were shown our ad found it compelling enough to click and learn more.".

➢ The flowchart shows the step by step process taken to calculate for the overall ctr.

Data Output   Messages   Notifications

	overallctr numeric 🔒
1	9.98263906333686225

Start

↓

Calculate SUM(clicks)

↓

Calculate SUM(impressions)

↓

Multiply (SUM(clicks) * 100.0)

↓

Divide by (SUM(impressions))

↓

Overall CTR

↓

End

# 8. MOST COST-EFFECTIVE CAMPAIGN

❑ The query result shows that campaign ID: 42484 (Alpha Innovations) has the most effective cost per conversion, which is $3.3428 approximately

➢ Cost Per Conversion (CPC) measures the efficiency of our advertising spend by determining how much we spend to acquire a single conversion.

➢ The lower the CPC, the better, as every organization aims to minimize cost.

➢ Cost per Conversion is equal to Acquisition Cost divided by conversation

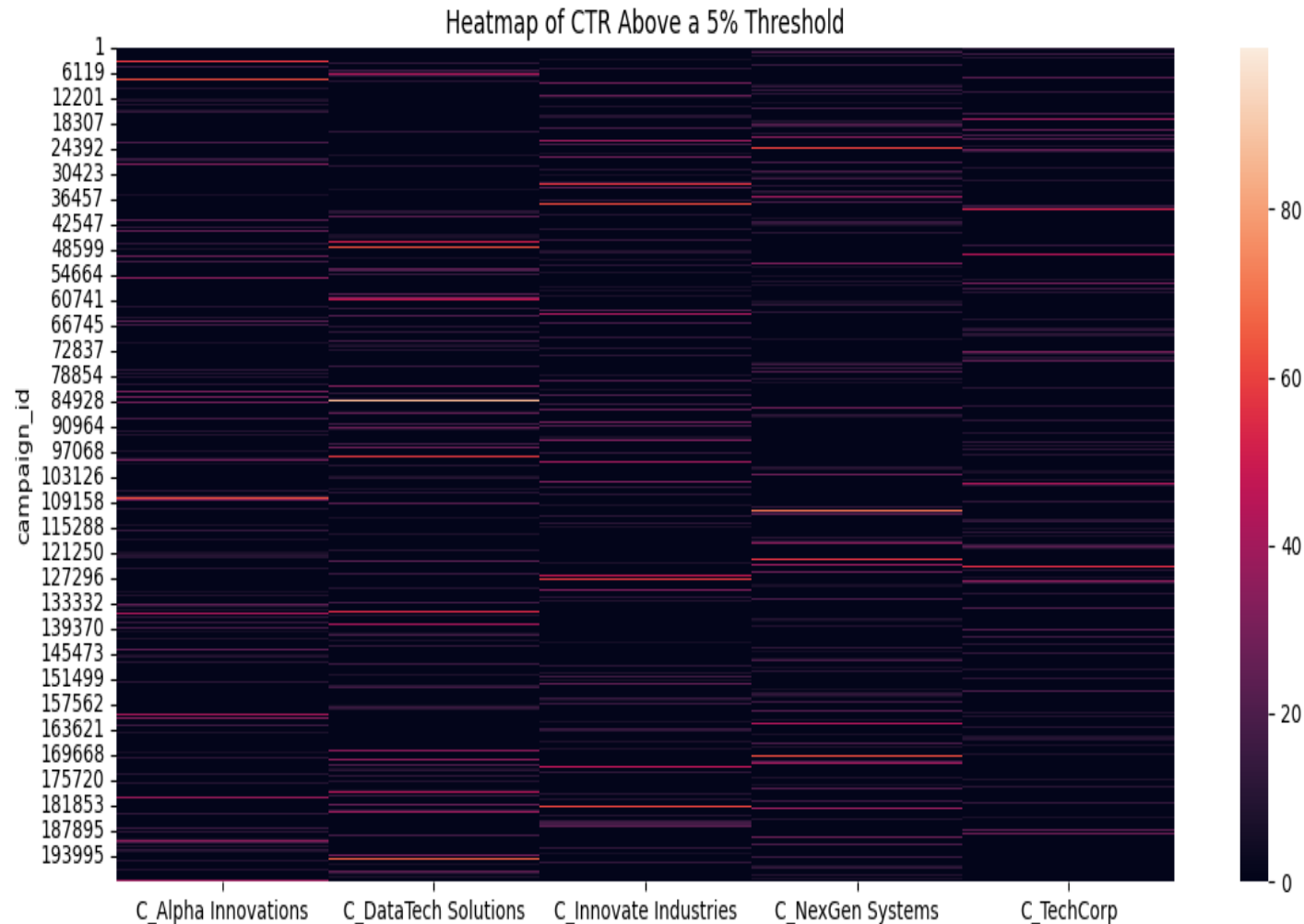➢ The conversation was calculated as shown in the flow chart.

SCREENSHOT OF THE OUTPUT

Data Output   Messages   Notifications

	campaign_id [PK] numeric	company character varying	costperconversion numeric
1	42484	Alpha Innovations	3.3426760093426760

FLOWCHART OF CPC CALCULATION
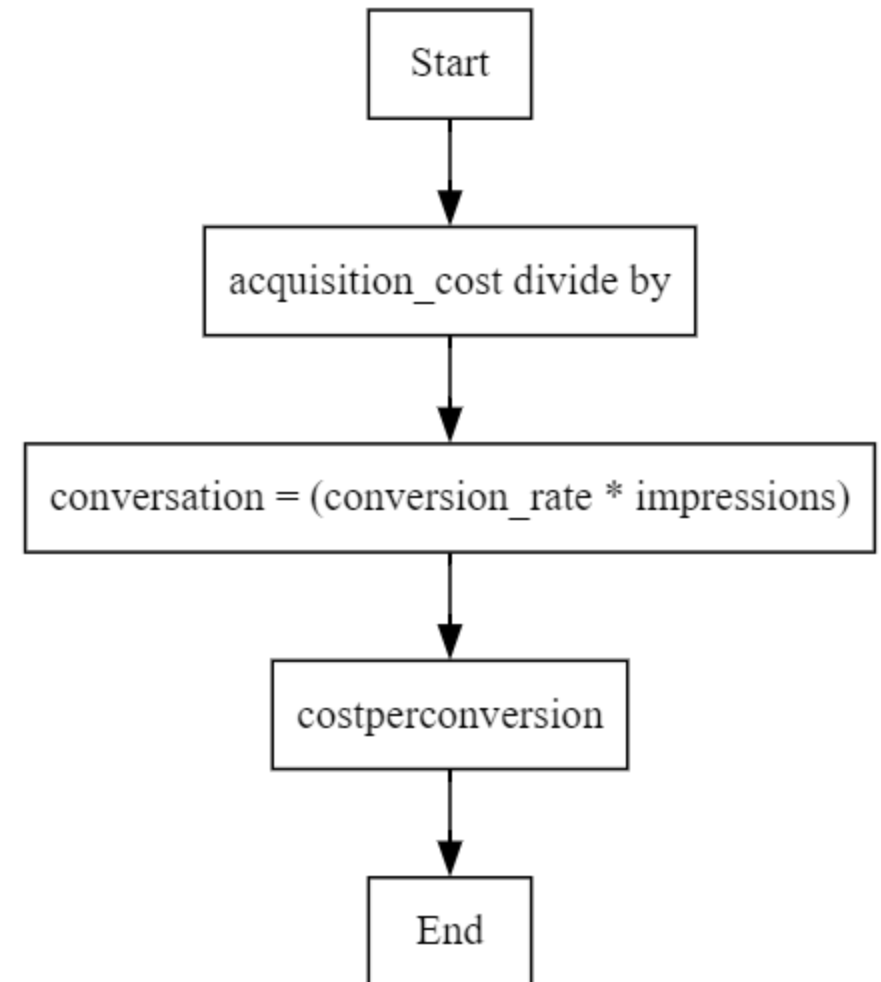
9. CAMPAIGNS WITH CTR ABOVE 5% THRESHOLD

Heatmap of CTR Above a 5% Threshold

- ❑ Certain companies (e.g., C_TechCorp and C_NexGen Systems) seem to have more campaigns with high CTR values, as indicated by the brighter color patches.
  - ➢ Each column represents a different company on the x-axis of the heatmap.
  - ➢ Y-axis: Each row represents a campaign ID that had a CTR above 5%.
  - ➢ The values in the columns correspond to CTR values for campaigns belonging to these companies.
  - ➢ Darker areas (black or deep purple) indicate campaigns with relatively lower CTR values (but still above 5%).
  - ➢ Brighter areas (red to white) indicate campaigns with significantly higher CTR values.

10. RANK CHANNELS BY TOTAL CONVERSIONS

Rank Channels by Total Conversions
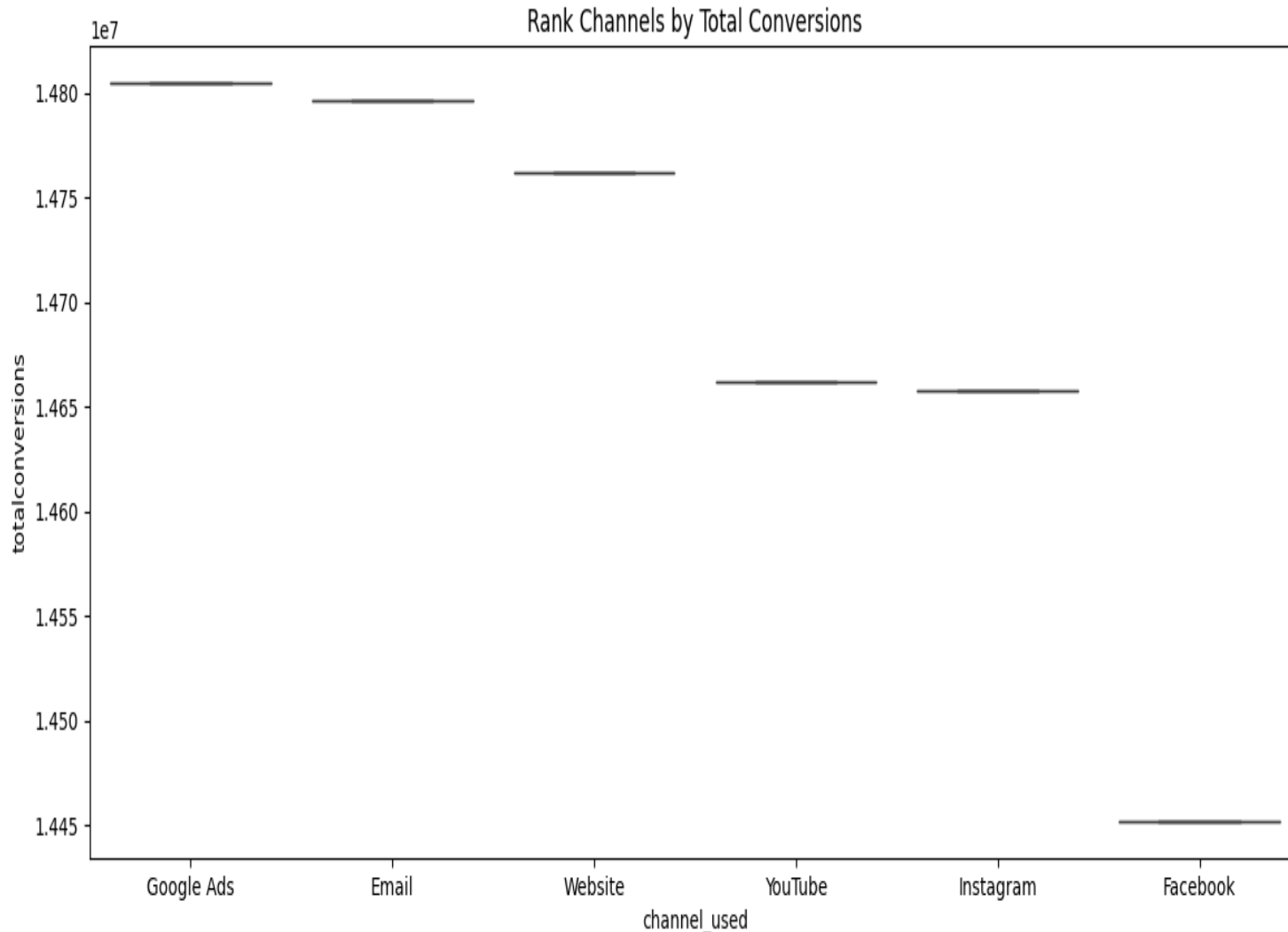
❑ Google Ads, Email, and Website have the highest total conversions, all close to 14.8 million.

  ➤ X-Axis (channel_used): Represents different marketing channels.

  ➤ Y-Axis (totalconversions): Represents the total number of conversions achieved through each channel.

  ➤ YouTube and Instagram are still effective but slightly lagging behind with a total conversions, around 14.65 million

  ➤ Facebook has the lowest total conversions, around 14.45 million.

11. CONCLUSION

❑ Campaign Performance Insights
  ➢ Google Ads, Email, and Website were the top-performing channels in terms of conversions.
  ➢ Facebook had the lowest total conversions, suggesting room for optimization.
❑ Engagement and ROI Trends:
  ➢ The highest engagement was observed among Men aged 18-24, indicating a key demographic for future targeting.
  ➢ New York, Miami, and Chicago led in total impressions, highlighting strong audience engagement in these regions.
  ➢ Only 160 campaigns (0.08%) achieved the highest ROI of 8, showing that highly successful campaigns are rare.
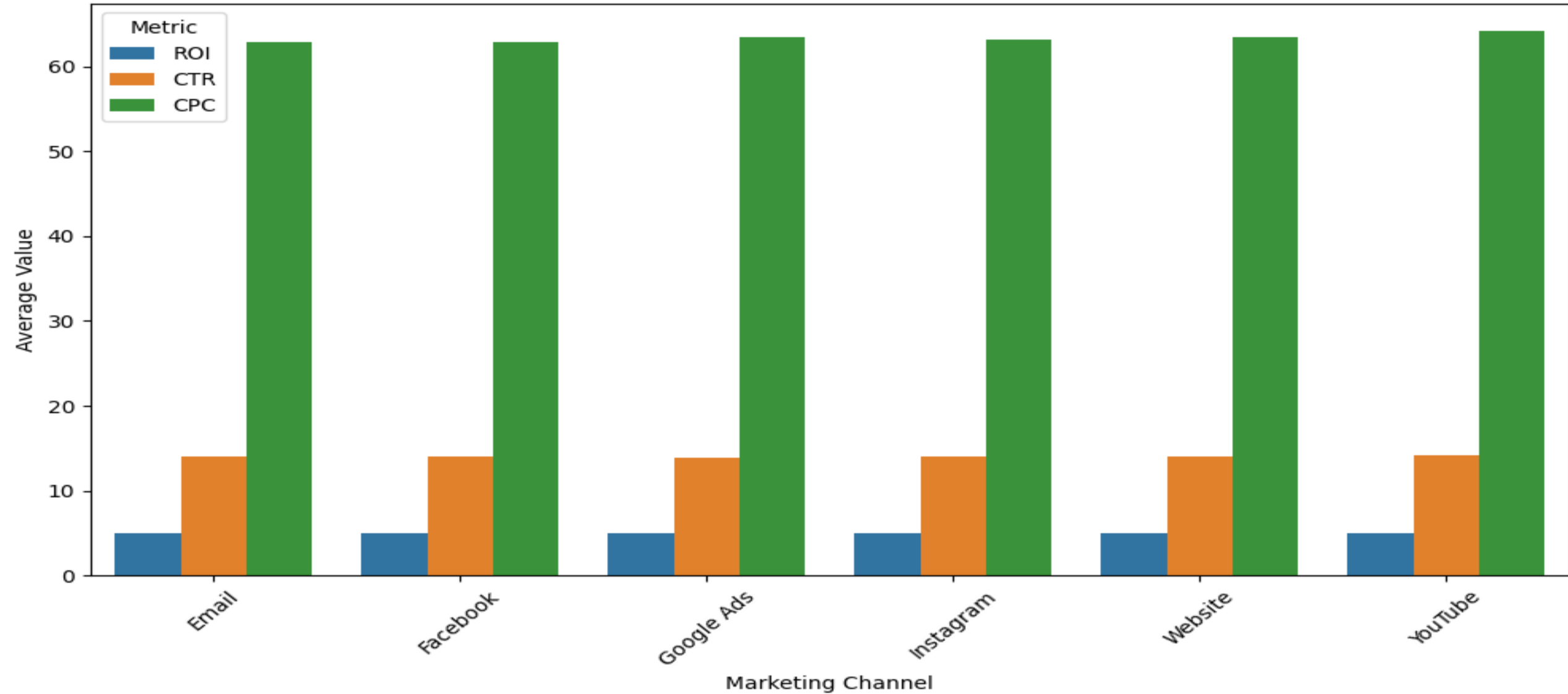❑ Cost Efficiency
  ➢ Alpha Innovations (Campaign ID: 42484) had the most effective Cost Per Conversion ($3.34), emphasizing efficient ad spending.
  ➢ Campaigns from C_TechCorp and C_NexGen Systems had higher CTR values, making them standout performers.
❑ Click-Through Rate (CTR) Effectiveness
  ➢ An overall CTR of 9.98% indicates that nearly 1 in 10 users found the ads compelling enough to click.
  ➢ Heatmap analysis revealed certain companies consistently achieved higher CTRs, signaling effective strategies worth replicating.

11. CONCLUSION – CONTD

Comparison of Average ROI, CTR, and CPC Across Marketing Channels