



# **DATA ANALYST INTERNSHIP CHALLENGE 2024**

**BY: KRISTINA CHUANG  
VIENNA, JULY 2024**

# CONTENT

- 01** PROJECT OVERVIEW
- 02** GOALS
- 03** TASK 1: PROGRAMMATIC ADVERTISING ANALYSIS
- 04** TASK 2: WORLD CUP 2022 ANALYSIS
- 05** ABOUT ME

# OVERVIEW

- This project is an analysis of programmatic advertising data and the impact of the 2022 World Cup on advertising metrics.
- The work is divided into three main tasks:
  - analyzing ad impressions and conversions,
  - generating insights from the World Cup 2022 data
  - preparing the current presentation to summarize the findings.
- Code files can be found in the project folder.



# GOALS AND OBJECTIVES

## Objective n° 1

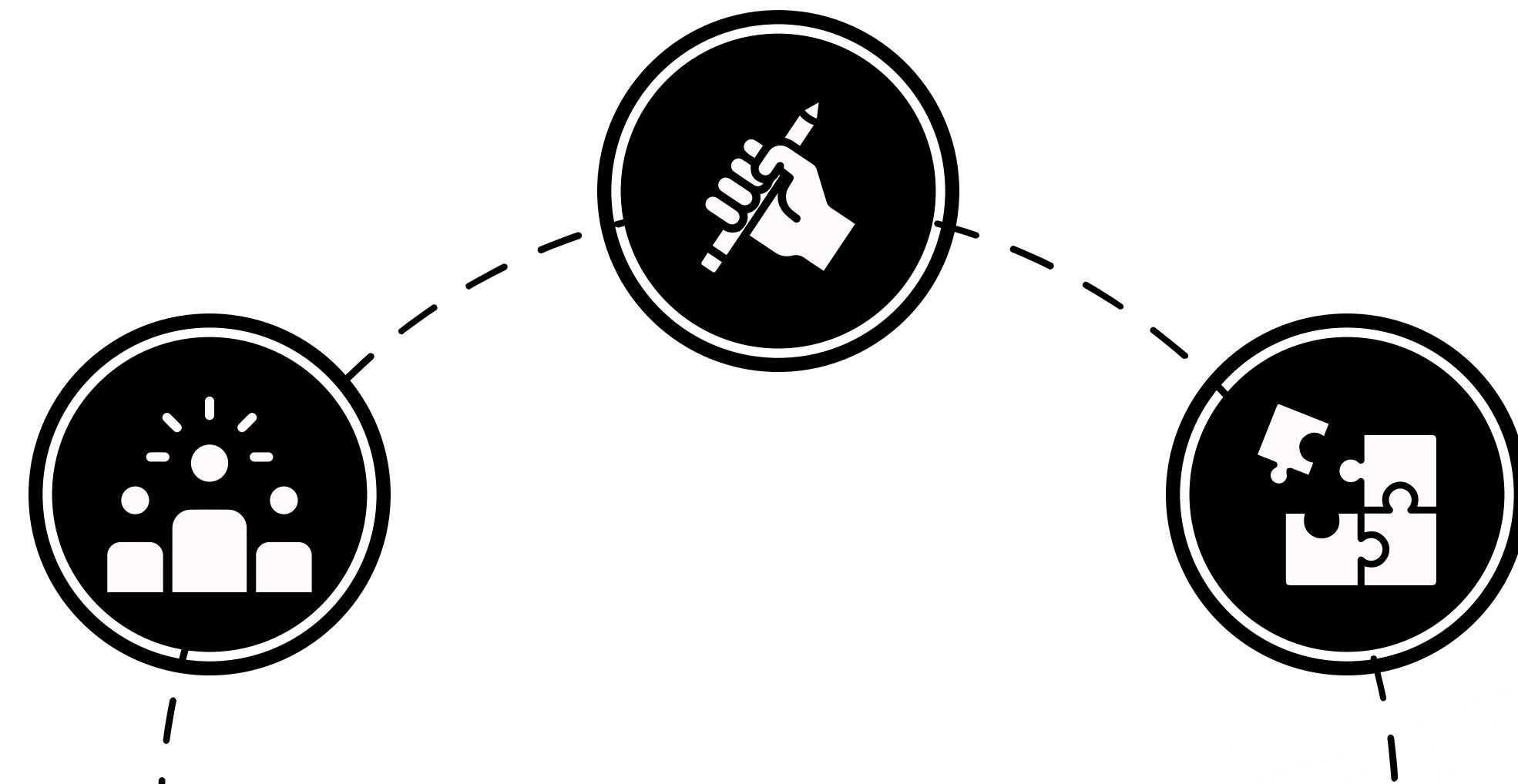
leverage data analysis  
to provide insights into  
programmatic  
advertising

## Objective n° 2

understand the impact of major sporting  
events, specifically the World Cup 2022, on  
advertising metrics

## Objective n° 3

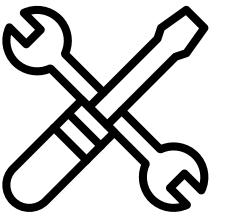
help stakeholders make  
informed decisions to  
optimize ad spend and  
improve conversion  
rates



# TASK 1: PROGRAMMATIC ADVERTISING ANALYSIS

- Given the schema of 3 database tables:
  - impressions
  - clicks
  - conversions

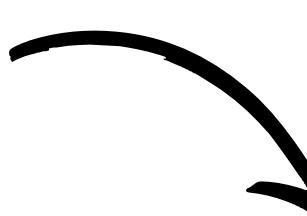
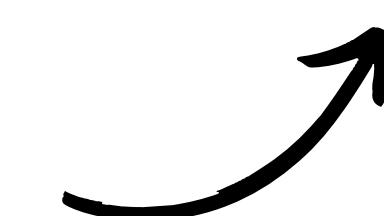
- Answer the questions:
  - What is the Click-Through Rate (CTR) in Austria?
  - For each converted user, how many impressions were served and what were the timestamps of the first and last impressions?
  - What further insights might be interesting for stakeholders?



## Tool Box

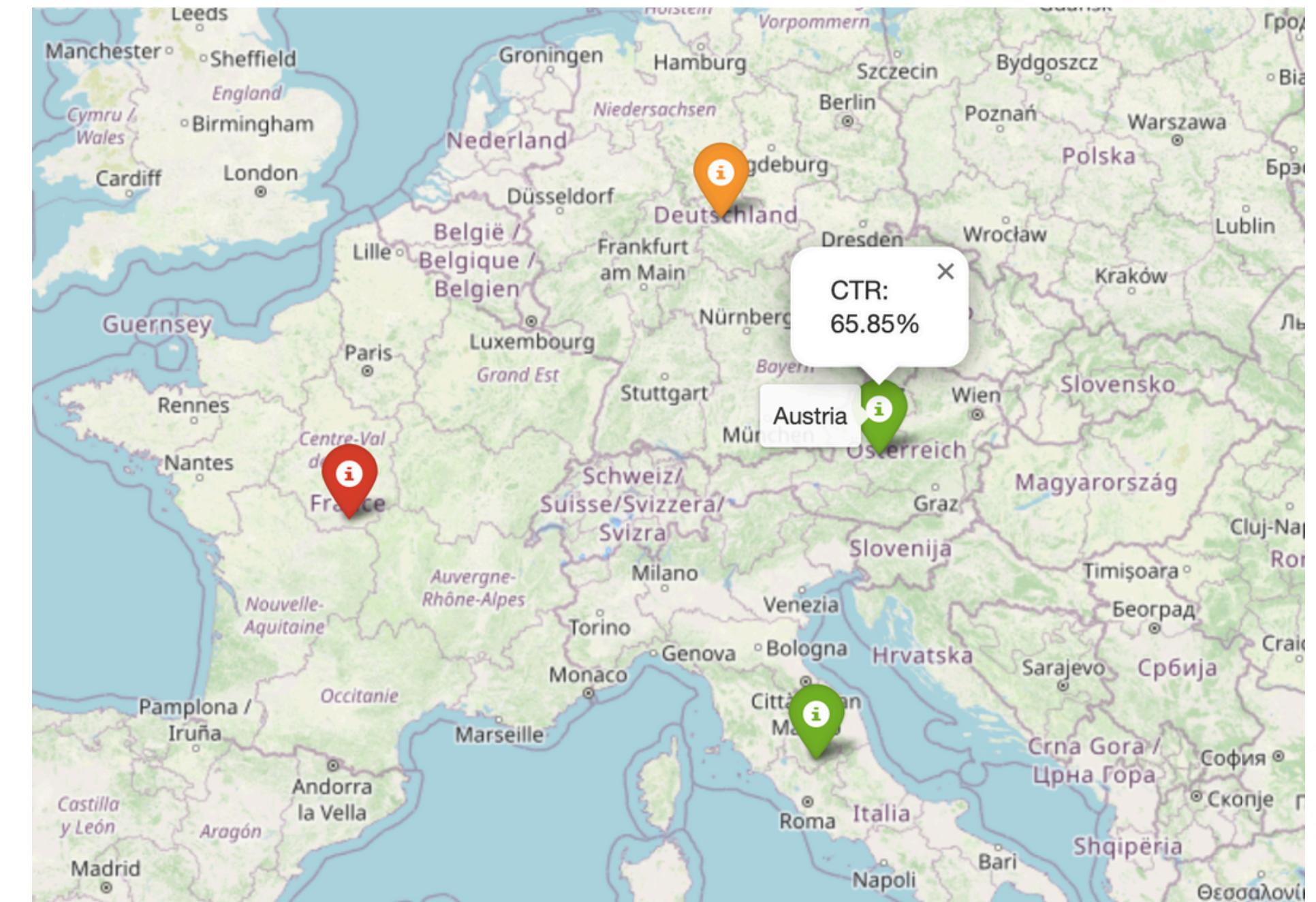
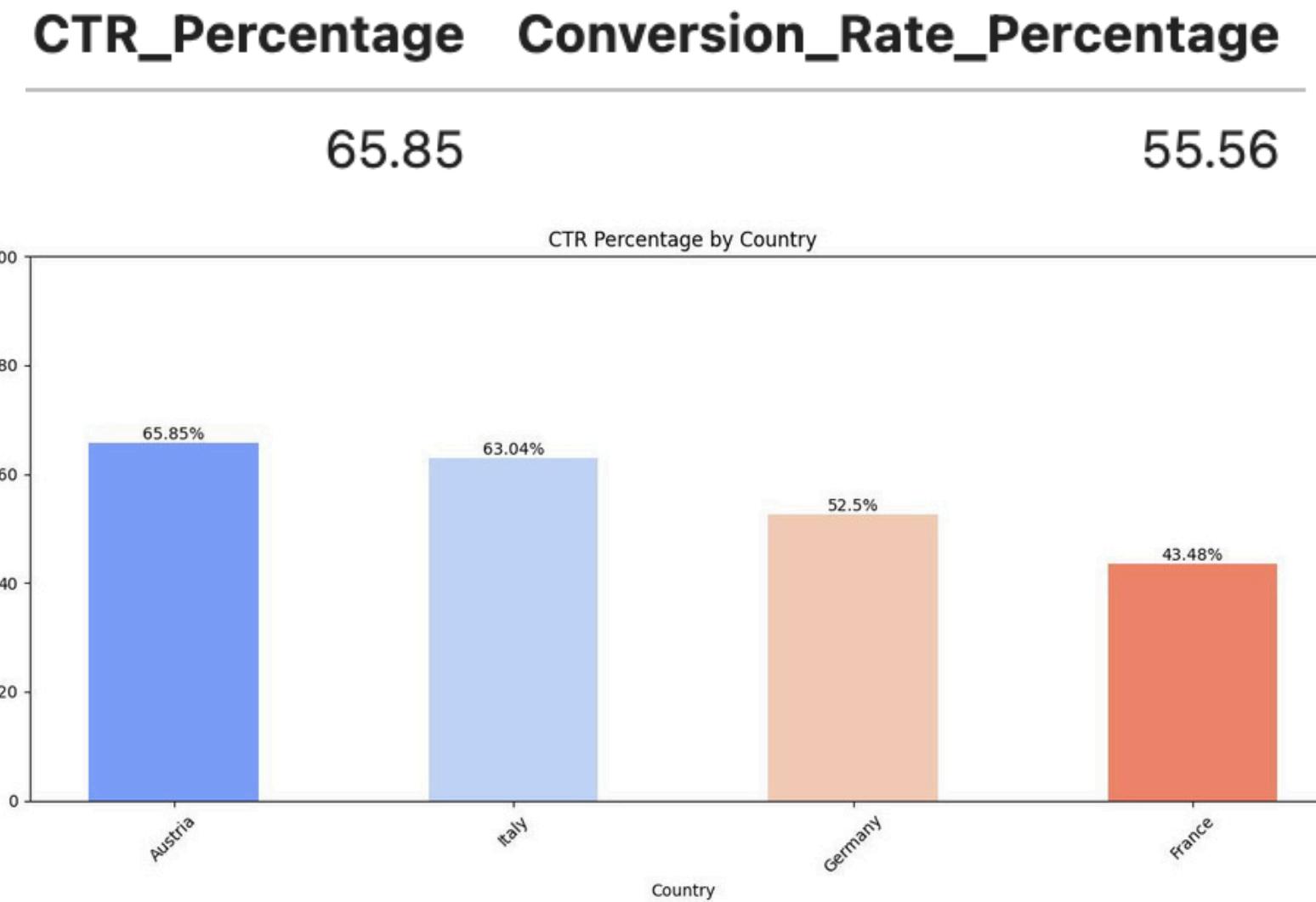
- SQL
- Python
- Jupyter Lab
- SQLite
- Libraries:
  - Sqlite3
  - Pandas
  - Matplotlib
  - Seaborn
  - Folium

# WORKFLOW AND MOCK DATA SET

- Characteristics:
    - randomly generate 150 impressions
      - 4 countries
      - 4 months
      - 70 users
      - 60% chance of click
      - 35% chance of conversion
  - Record workflow in Jupyter.
  - Create SQLite local database with given datatypes
  - Generate mock data
- DATA PREP** 
- Data type inconsistency handling
  - SQL queries to extract information on CTR and Conversions
- DATA QUERY** 
- Bar charts
  - map
  - graph interpretation
  - insights
  - generate further questions
- VISUALIZATION,  
INSIGHTS AND  
FURTHER  
QUESTIONS**

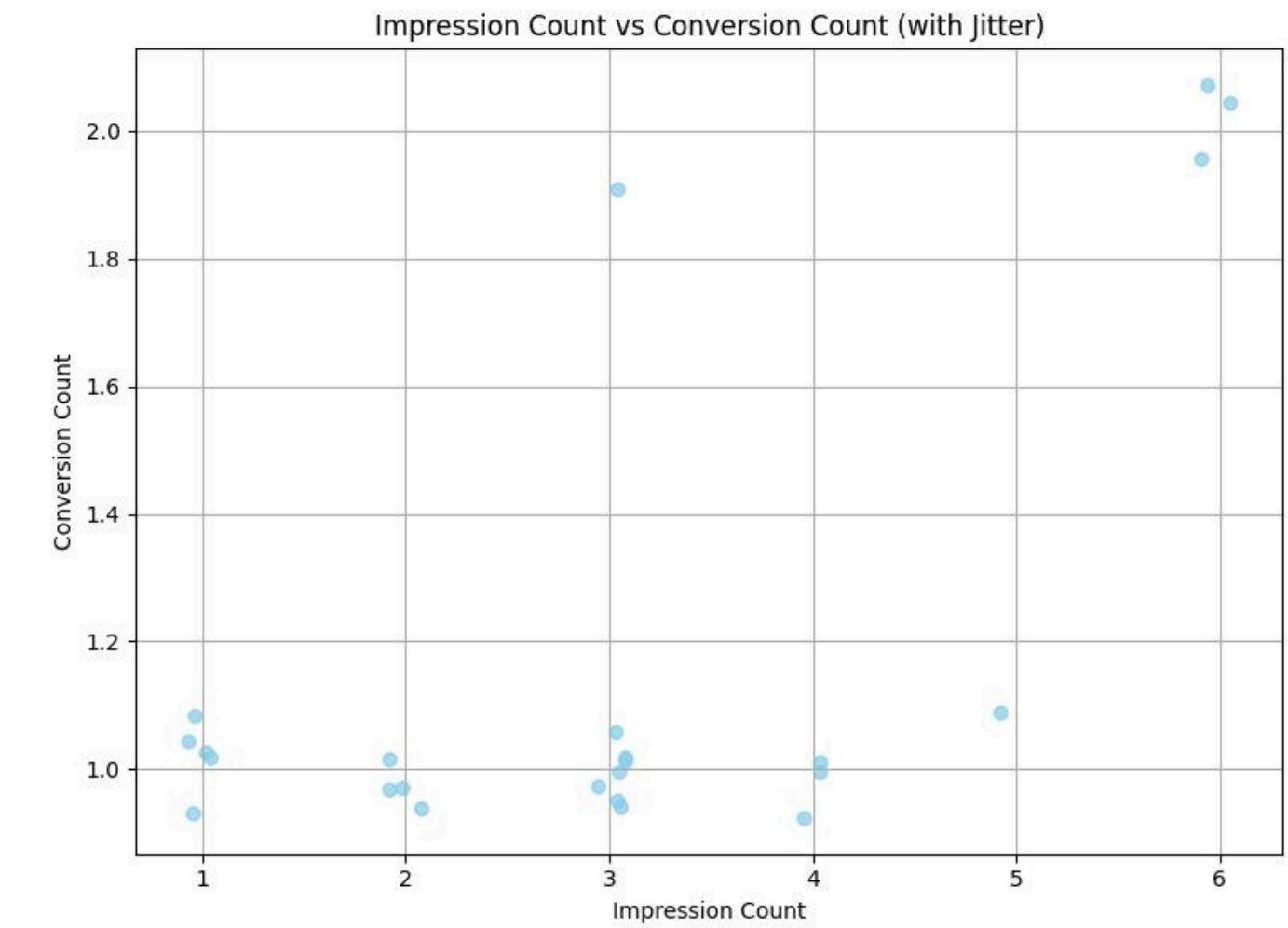
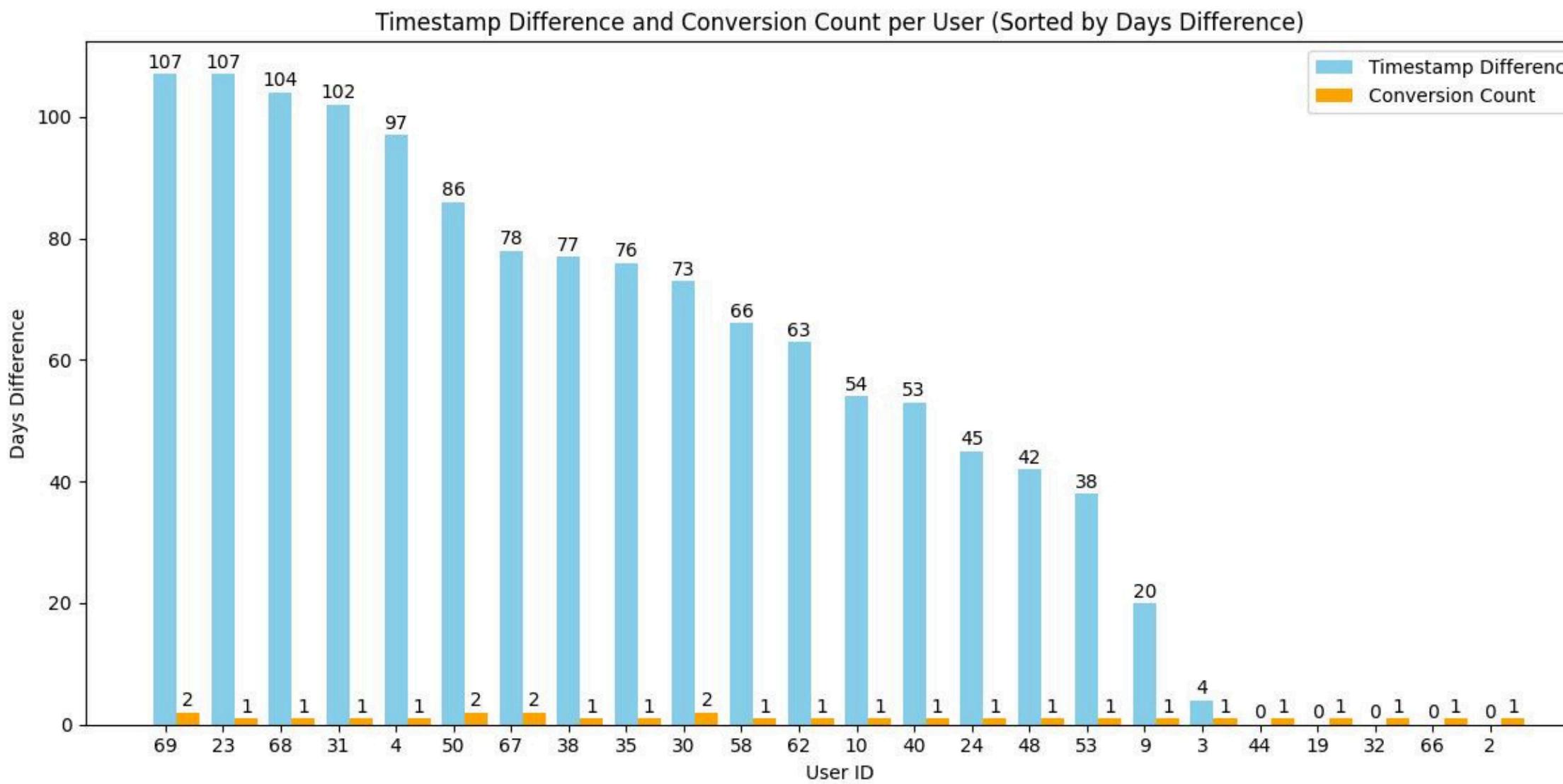
# WHAT IS THE CTR (%) FOR IMPRESSIONS SERVED IN "AUSTRIA"?

- With a Click-Through Rate of 65.85% the Impressions served in Austria were the most effective in capturing user engagement compared with other countries in Europe.
- Moreover, The conversion rate for Austria is also high at 55.56%, indicating that the impressions met the expectations of more than half of the users who clicked.



# FOR EACH CONVERTED USER, HOW MANY IMPRESSIONS WERE SERVED AND WHAT WERE THE TIMESTAMPS OF THE FIRST AND LAST IMPRESSIONS?

- 24 unique user ids resulted in conversions
- They were served between 1 and 6 impressions in a period of 4 months from 21 November 2022, with a range of 120 days.
- 4 users clicks resulted in 2 separate conversions.
- No correlation between the number of impressions served and conversion
- bar plot below shows the difference in days between the first and last impressions served and the counts of conversions per user.



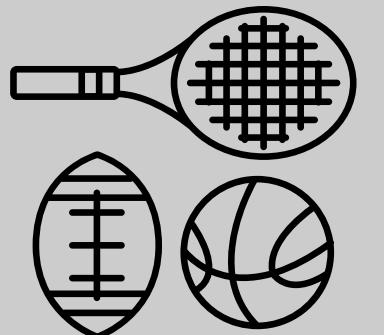
\*each point corresponds to one user

# WHAT FURTHER INSIGHTS MIGHT BE INTERESTING FOR STAKEHOLDERS?



## User Behavior Insights:

- Click Patterns & Conversion Journey: Analyze the time differences between impressions, clicks, and conversions.
- User Engagement: Measure the duration from impression to user attention capture.
- Decision Time: Assess how long users stay on the landing page before making a deposit.



## Content Performance Insights:

- Sport Popularity: Identify which sports content (e.g., soccer, baseball) resonates most with the audience by country.
- Conversion Rates by Sport: which sports content leads to higher deposits?
- How effective is the dynamic display by sport type?



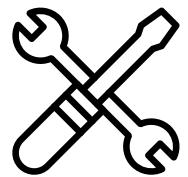
## Financial Insights:

- Currency impact: Analyze conversion rates and deposit values across different currencies to understand financial impact.
  - What is the users spending behavior in various markets based on the currency of the transaction?
- Revenue Analysis:
  - Calculate total revenue generated from conversions in different regions.
  - Identify profitability across different ad types to identify the most lucrative segments.

## TASK 2: WORLD CUP 2022 ANALYSIS

### What was the impact of a major sporting event on advertising metrics?

- Challenges:
  - Optimising repetitive generation of insights encouraging clients to invest more in advertising during sporting events.
  - Identify and compute KPIs for advertising metrics.
  - Highlight potential increases in conversions and cost reductions for advertisers during a major sporting event.



#### Tool Box

- Python
- Jupyter Lab
- Libraries:
  - Pandas
  - Matplotlib

# WORKFLOW FOR WORLD CUP DATA

- Variables identification
- KPIs identification
- datetime type casting

**DATA INSPECTION,  
CLEANING,  
TRANSFORMATION  
AND EDA.**

- identify world cup 2022 period.
- add new columns for every KPI computed
- generate KPIs comparison tables

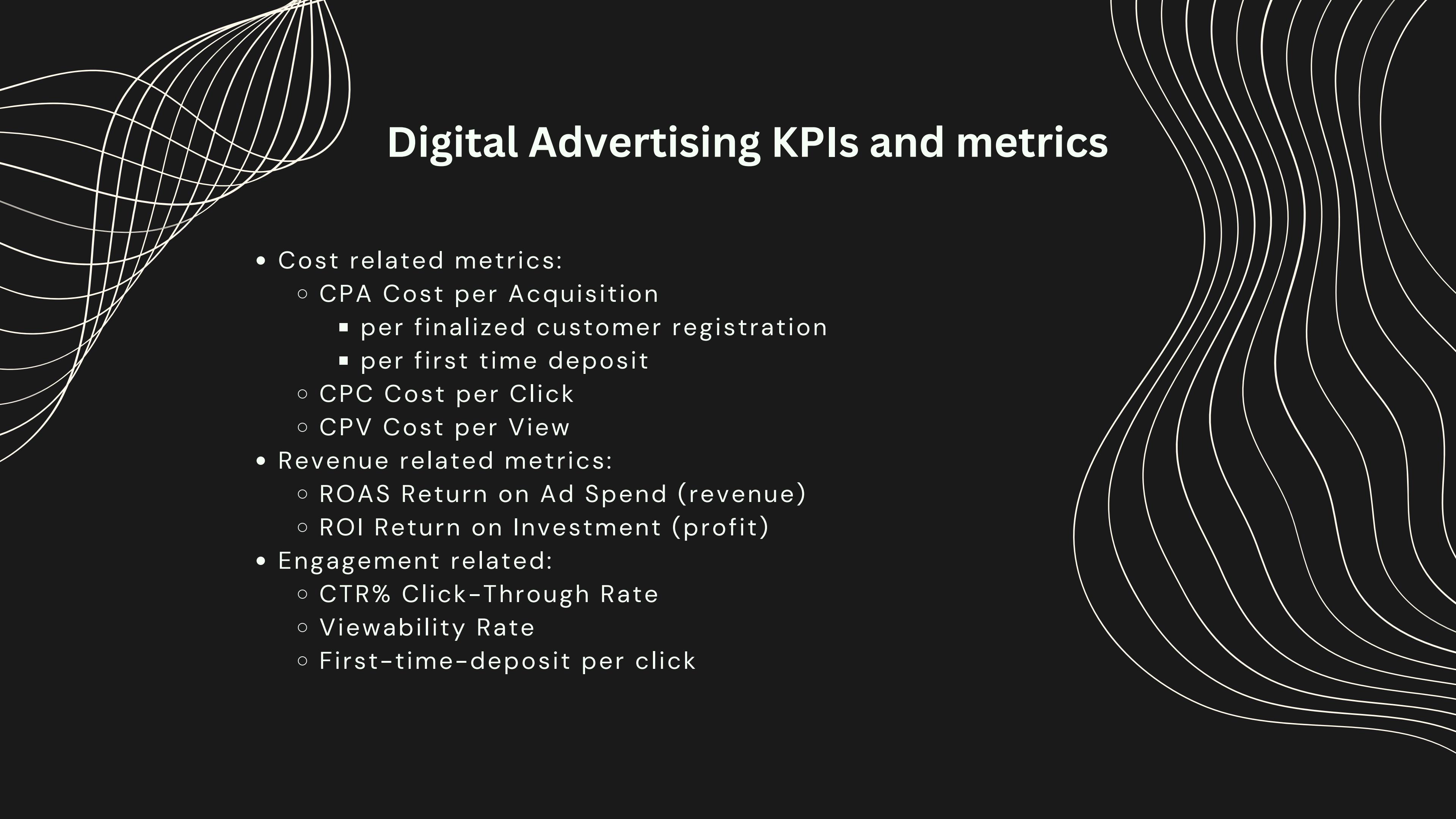
**KPI  
COMPUTATION**

- Aggregate and sort computed KPIs by country and world cup period.
- Automate for every KPI comparison

**AGGREGATION  
AND  
FUNCTION  
DEFINITION**

**VISUALIZATION,  
ANALYSIS AND  
INSIGHTS**

- grouped bar charts to compare absolute and relative KPIs during and outside world cup period.



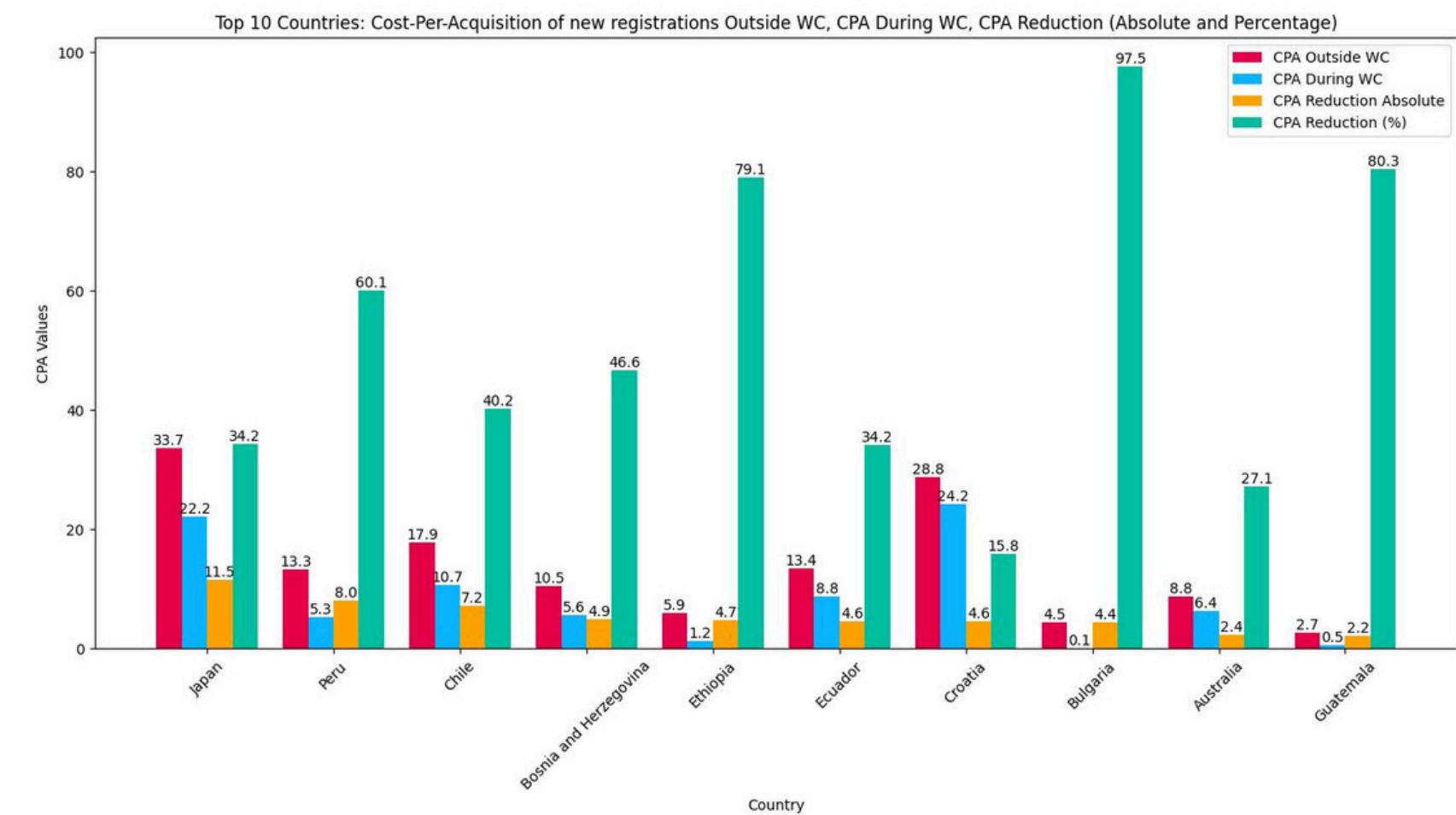
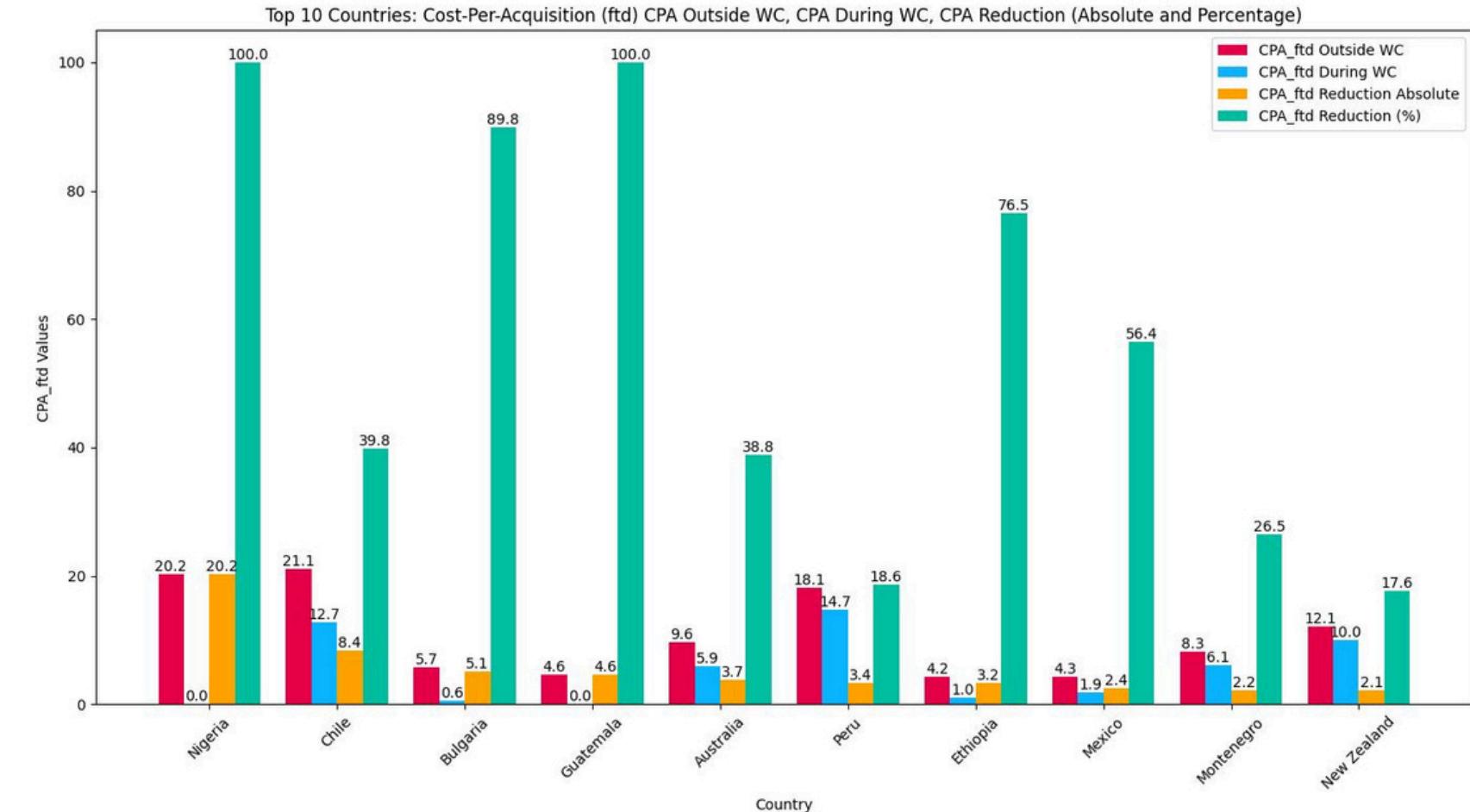
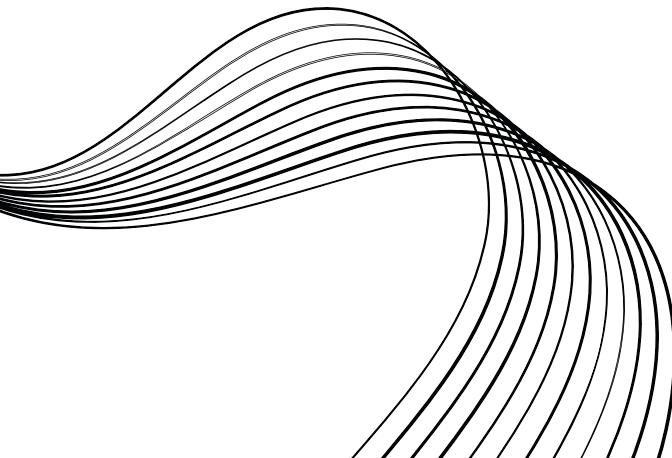
# Digital Advertising KPIs and metrics

- Cost related metrics:
  - CPA Cost per Acquisition
    - per finalized customer registration
    - per first time deposit
  - CPC Cost per Click
  - CPV Cost per View
- Revenue related metrics:
  - ROAS Return on Ad Spend (revenue)
  - ROI Return on Investment (profit)
- Engagement related:
  - CTR% Click-Through Rate
  - Viewability Rate
  - First-time-deposit per click

# CPA: Cost per Acquisition

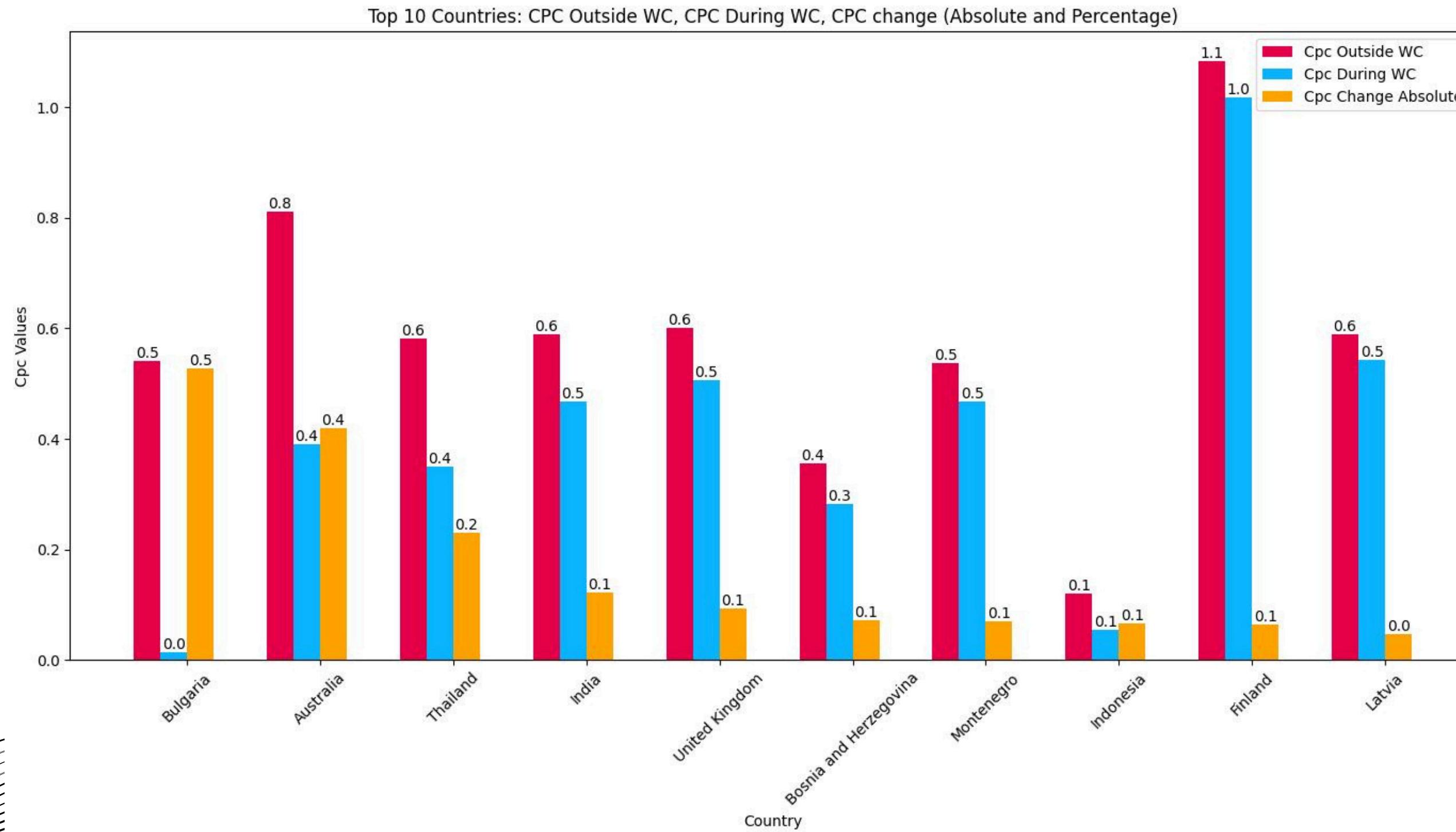
## Top 10 countries with cost reduction

- Analysis depending on the definition of new customer:
  - placed a first deposit (top graph)
  - completed the registration process (bottom graph)
- Chile shows a considerable reduction in cost for both new registrations and new conversions during the world cup period. Suggesting the country has an interest in football related events.
- Japan shows the largest decrease in absolute cost per registrations, however this is not reflected in actual conversions.
- green bar shows the relative change in cost reduction compared to outside world cup 2022 period



# CPC: Cost per Click

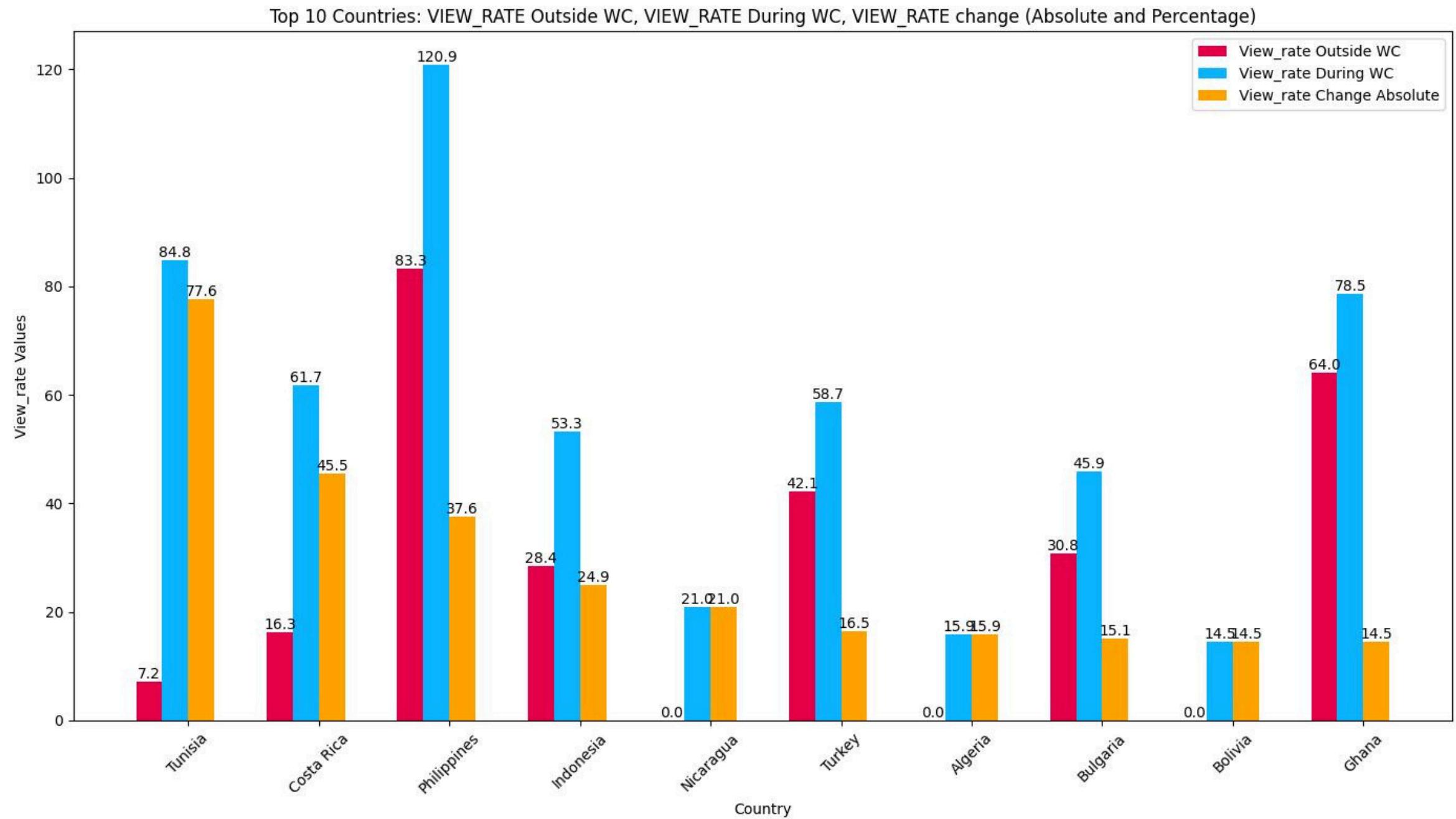
## Top 10 countries with cost reduction in USD per click



# Viewability Rate %

## Top 10 countries with increase engament

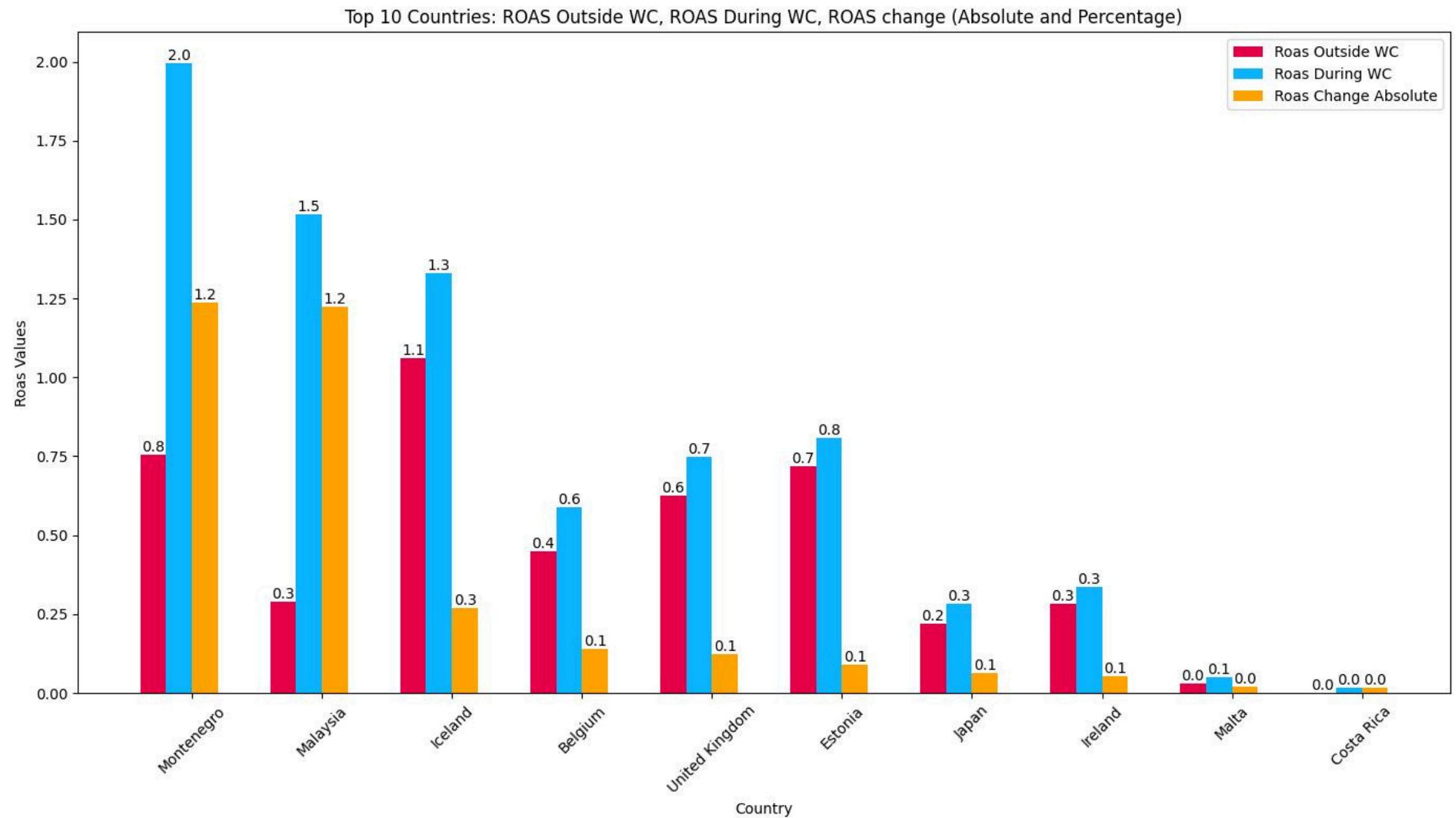
- Viewability rate: percentage of impressions actually viewed by user.



# ROAS Return on As Spend

## Top 10 countries with increase Revenue per dollar spent

- measure the revenue generated for every dollar spent on advertising.
- Use to evaluate the effectiveness of advertising campaigns in driving sales and revenue.



# **Adapting world cup analysis to Euro 2024 and other sporting events**

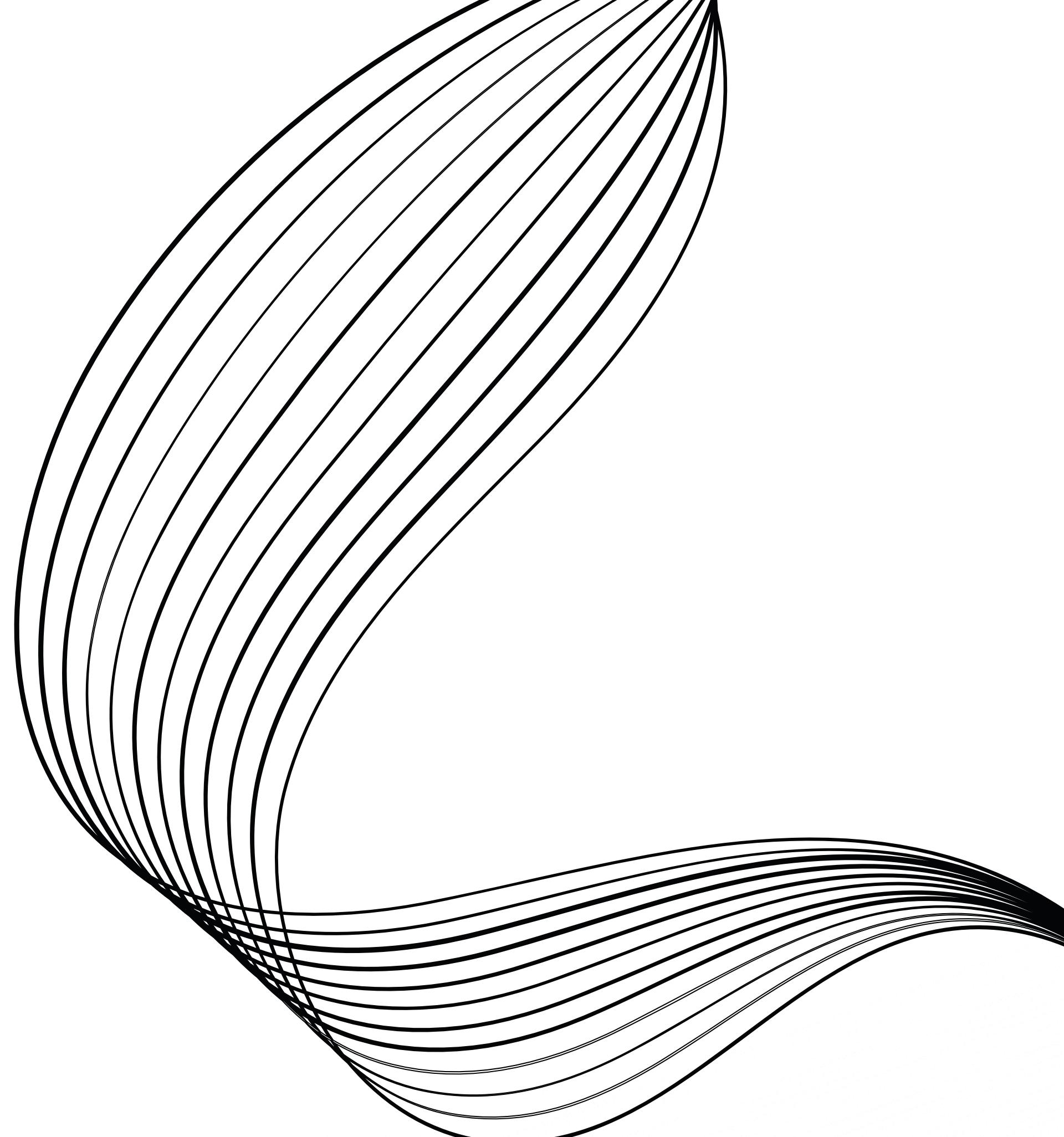
- Collect Ads data with the same metrics in and around Euro 2024 dates
- Focus analysis in broader Europe region compared to non-European regions
- Particular attention to metrics in participating countries
- Carry out a comparative analysis of metrics change of World Cup data 2022 v/s Euro 2024 data, to support findings and insights

## **Further questions to investigate**

- Perform Time series analysis to identify peak engagement times for both events.
- Are there specific times or types of content that yield better engagement and conversion rates?
- Do different operating systems or platforms affect engament and conversions?

# **THANK YOU**

for considering me for  
the Data Analyst  
Internship position.  
I truly appreciate the  
opportunity.



# ABOUT ME

Kristina Chuang

Former globe trotter.

Current business informatics graduate.

Enthusiastic about solving problems with data.

In search for a job

