



Social media marketing

Supervisor

Mohamed Abdelmawla

Track

Microsoft Data Engineer

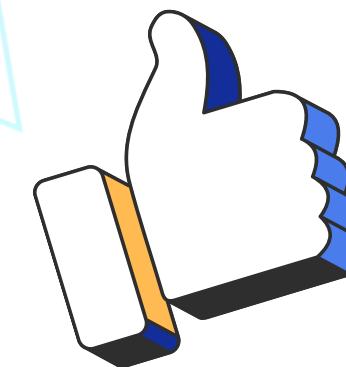




رواد مصر الرقمية



وزارة الاتصالات
وتكنولوجيا المعلومات



Overview

- Our Team
- Look back
- Project Overview
- Problem Statement

- Data Collection & Analysis
- Exploratory Data Analysis
- Dashboard
- Insights



OUR TEAM



AMR KHALED



IBRAHIM HEGAZI



YOUSEF ALSEBAEY



AMMAR YASSER



ASHRAF MORGAN



MAHMOUD OSAMA





The Weather Dataset

Weather Data for +1200 Cities Worldwide. Updated every Sunday!

Data Card Code (15) Discussion (1) Suggestions (0)

About Dataset

Feel free to [FORK THIS NOTEBOOK](#) in order to correctly load the data for your project!

Overview:

In the beginning

We were going to work on weather data To analyze the impact of population density on weather patterns and compare regions with similar latitudes and longitudes to understand how varying population densities influence local weather conditions such as temperature, humidity, and air quality.

But we faced some problems

PreProcessing problems

While working on Collecting the data, several issues arose, including:

- **Incomplete Data**

Problem: Where the population, humidity levels, and air temperature columns were missing

Solution 1: Since there was no actual up to date data available, therefore we collected the approximates of these values for each 1235 city manually (Worked for the time being)

- **Lots of missing values**

Problem: Lots of missing cells in some of the most detrimental columns like min, max, avg temperatures, and city names.

solution 1: To handle the missing cities, we merged the cities table with weather table to get rid of the unavailable data as much as possible.

solution 2: To handle the temperature values, we filled the missing values with the mean.



Database Steps and problems

While working on creating the database, several issues arose, including:

- **Database Design:** No Problem
- **Database Normalization:** No Problem

- **Loading the Data into SQL Server:**

Problem: Data was big and lots of data went missing because of the SQL Server loading Capabilities

Solution 1: Dividing the large dataset into smaller batches (did not work)

Solution 2: Loading the dataset into Power BI to have all the Value (worked)



Database Steps and problems 2



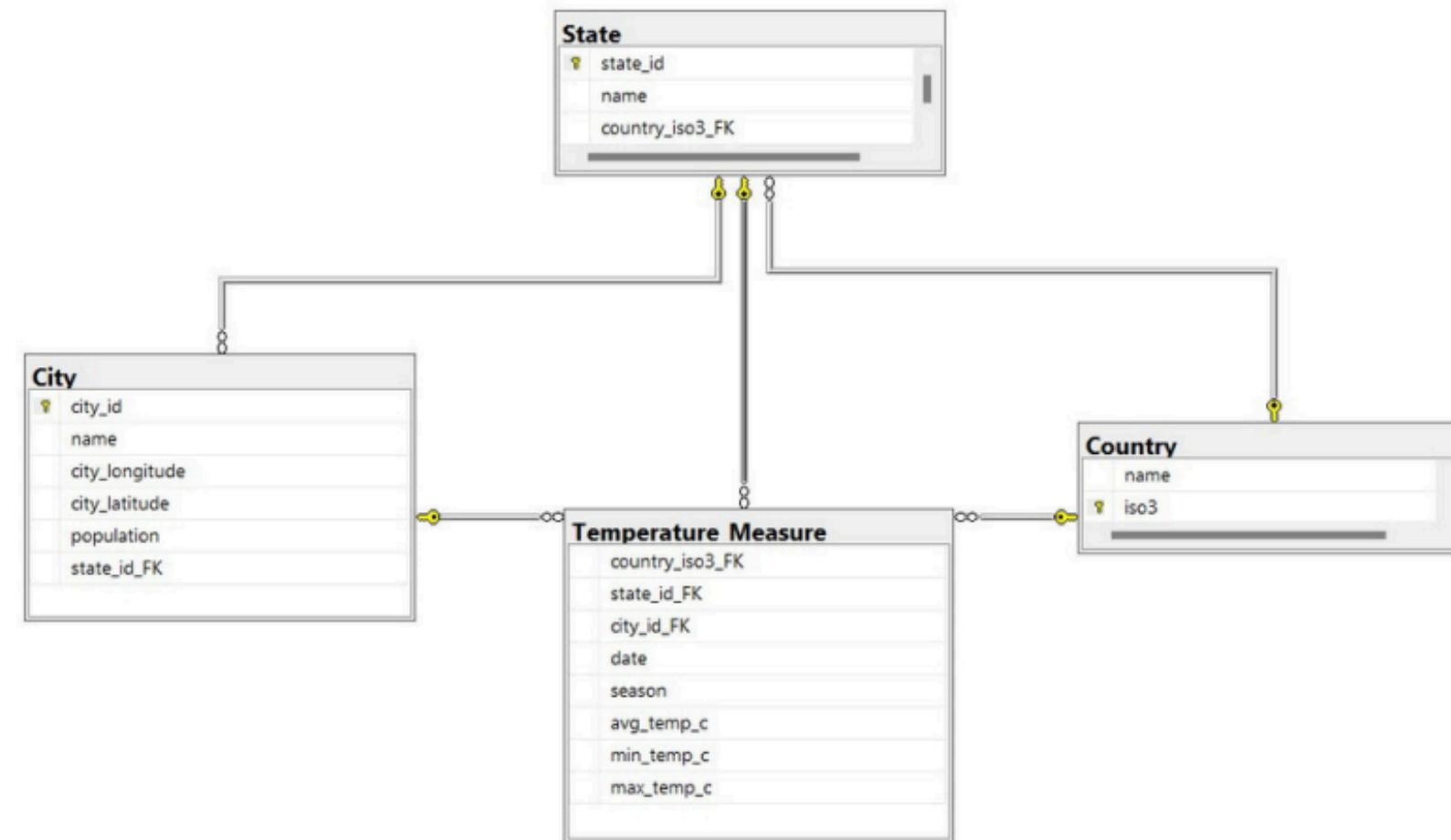
- SQL Code to enter the wrong dataset

into the right tables (no problem)

- Data inconsistency:

Problem: we had two tables that include the longitude and the latitude sadly these values were not the same

Solution 1: Dropping one of the tables, country table. (worked)

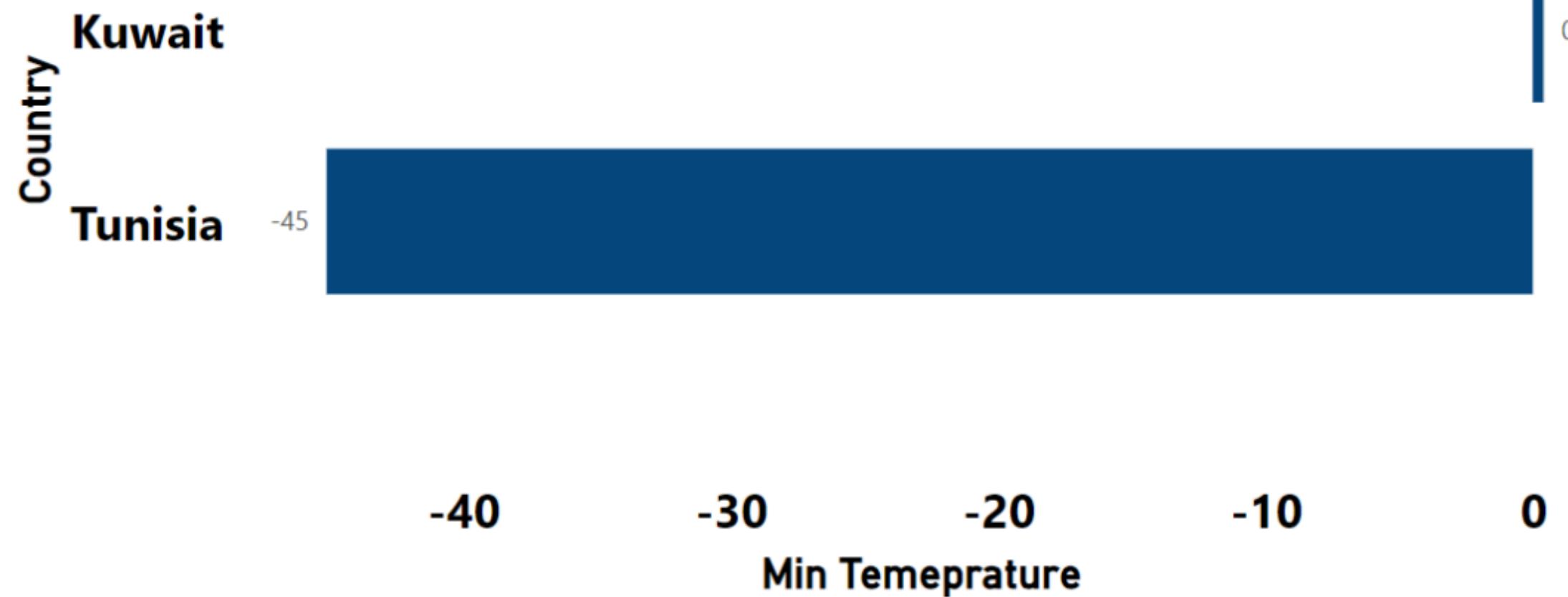




Problems During Data Discovery

We faced several problems at the Exploratory Data Analysis stage Through which we aimed to analyze data sets to understand their main characteristic

The recorded data is illogical and contains incorrect recordings !!



Social Media Analysis



وزارة الاتصالات
وتكنولوجيا المعلومات



"Not all wealth is money, it can be Data."

**"While the digital and AI is the dominant on our Earth,
social media is the key of communication and content
where hearts believe"**



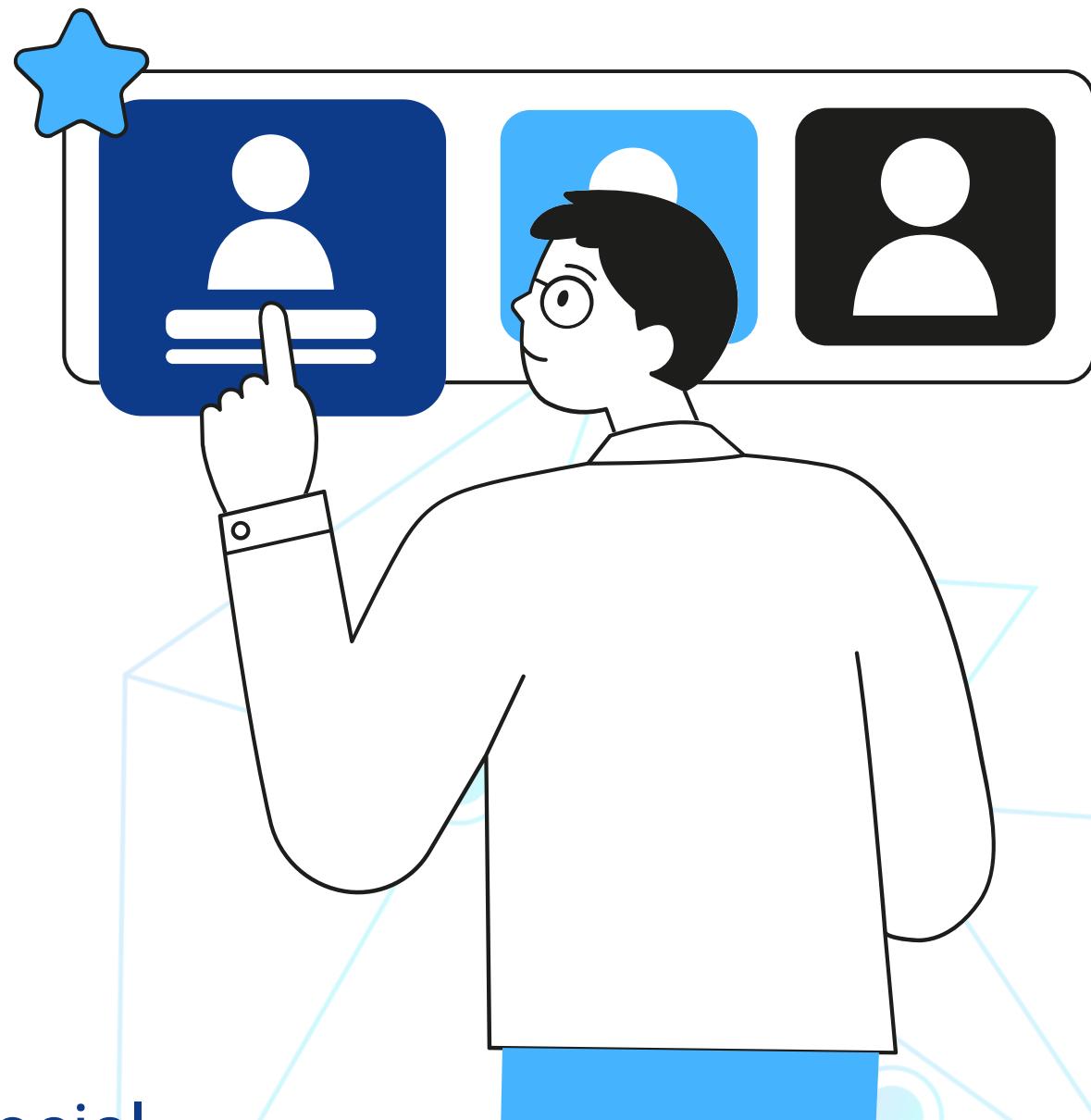
Problem Statement

Businesses and marketers generate vast amounts of data across social media platforms, but struggle to extract meaningful insights.

The complexity and volume of data, including user engagement, audience behavior, and content performance, make it challenging to optimize strategies and measure success.

Without a centralized system to track and analyze key metrics, decision-makers face difficulties in improving engagement and refining social media campaigns.

There is a need for an automated solution to analyze and visualize social media data, providing actionable insights through a comprehensive dashboard for better decision-making.



Project Overview

- The primary goal of this project is to analyze social media data to extract valuable insights and create an interactive dashboard that provides key metrics and trends
- By leveraging data from various social media platforms, this project aims to assist businesses, influencers, and marketers in making data-driven decisions to optimize their social media strategies and understand audience behavior





Objectives

01

Analyze social media data
to extract valuable insights
on content performance
and audience behavior

02

Develop an interactive dashboard
that visualizes key performance
indicators (KPIs) for easy and
real-time data monitoring

03

Enable data-driven
decision-making to optimize
social media marketing
strategies and enhance
audience engagement



Strengths

- Comprehensive data collection from multiple social media platforms
- User-friendly interface for easy navigation and access to insights
- Customizable dashboards tailored to specific metrics

Opportunities

- Growing demand for social media insights in marketing strategies
- Potential to expand features (e.g., AI-driven insights, predictive analytics)

S

W

O

T



Weaknesses

- Potential for data overload, making actionable insights difficult to extract
- High subscription costs for advanced features
- Limitations on accessing historical data

Threats

- Intense competition in the social media analysis tool market
- Increasing privacy regulations limiting data collection capabilities
- User fatigue from the abundance of available tools



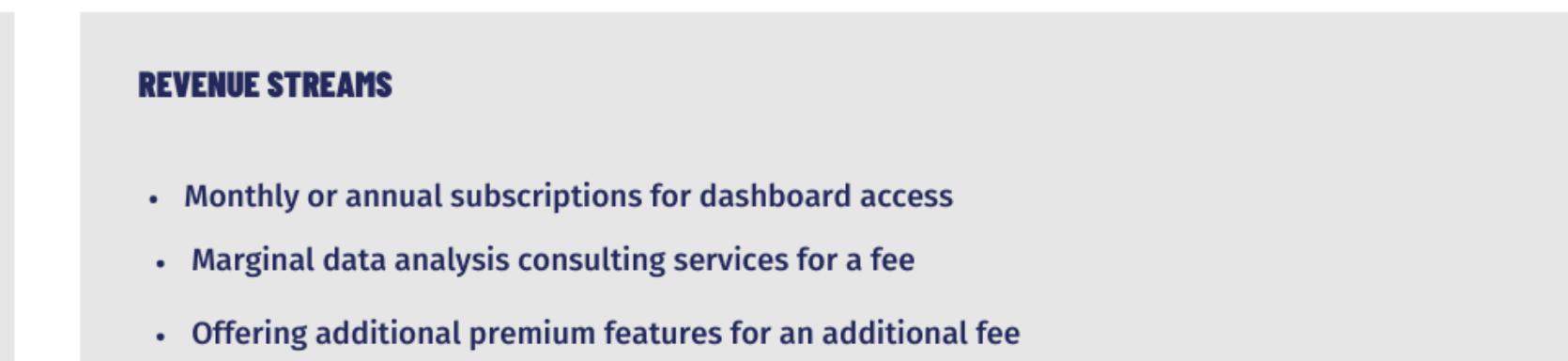
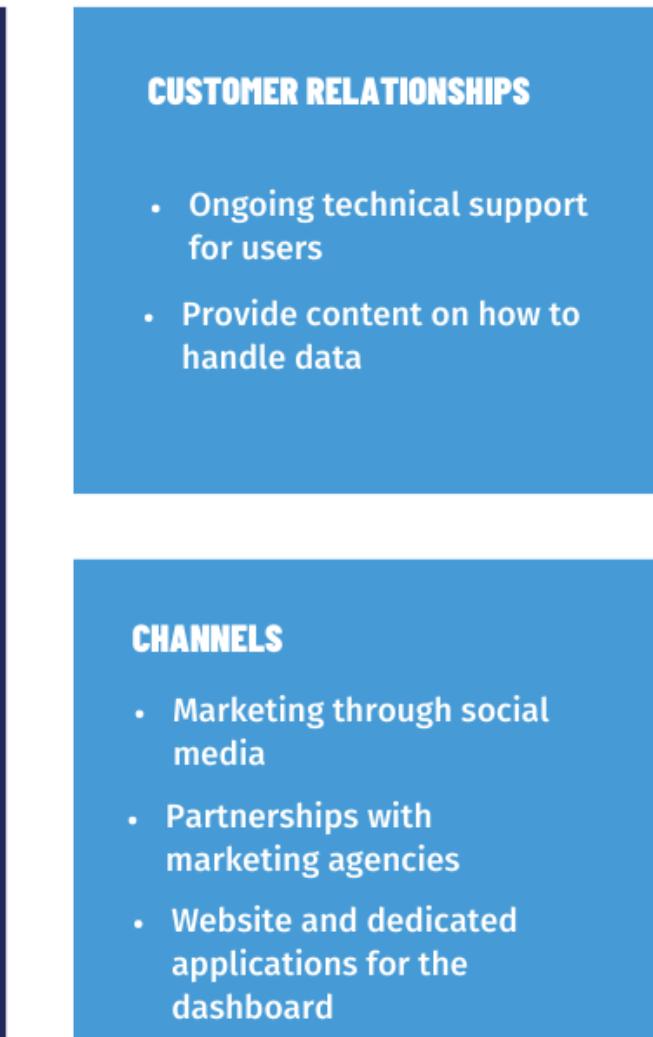


Social Media Marketing Analysis

وزارة الاتصالات
وتقنيه لجها المعلومات



Business Model Canvas





Data Collection

- kaggle

Social Media Sentiments Analysis Dataset 

Exploring Emotions, Trends, and Interactions in the Digital Tapestry 

Data Card Code (38) Discussion (1) Suggestions (0)



Usability ⓘ
10.00

License
CC0: Public Domain

Expected update frequency
Never

- World bank data

 WORLD BANK GROUP

Home About Data Research Learning News Projects & Operations Publications Countries Topics English ▾

This page is in English Español Français عربي 中文

 Log in Now  TWEETS  LIKE  SHARE +

DataBank

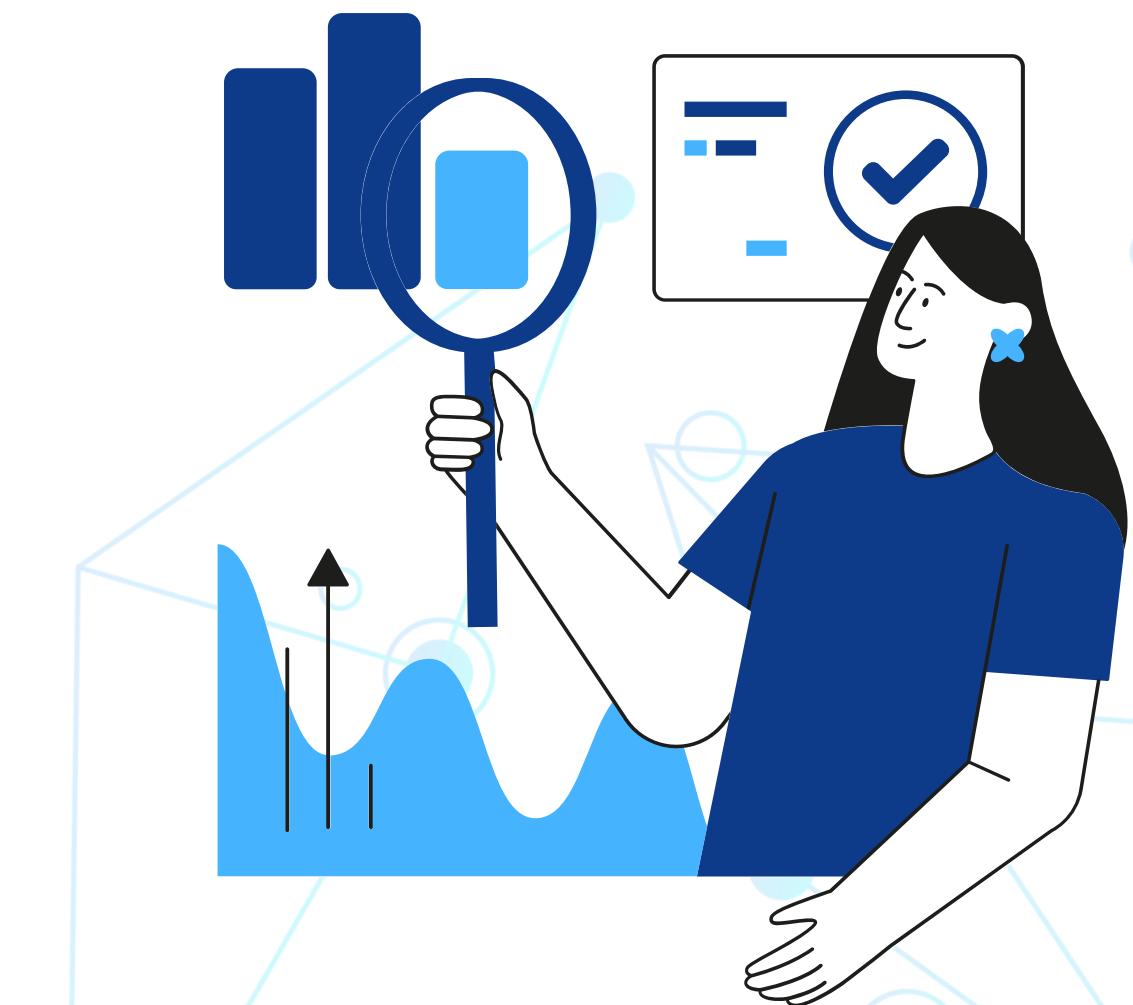
DataBank Home

- Databases
- Create Report
- Saved Reports
- Saved Datasets

Explore. Create. Share: Development Data

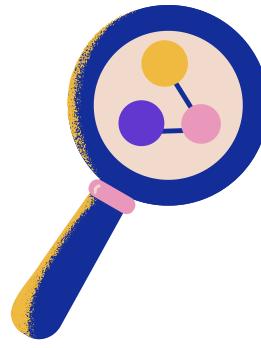
DataBank is an analysis and visualisation tool that contains collections of time series data on a variety of topics. You can create your own queries; generate tables, charts, and maps; and easily save, embed, and share them. Enjoy using DataBank and let us know what you think!

FAQs Feedback





Exploratory Data Analysis



What are the platforms within the dataset?

```
unique_platform = sentiment_fact['Platform'].nunique()  
unique_platform
```

5

```
platform_counts = sentiment_fact['Platform'].value_counts()  
platform_counts
```

Platform	
Facebook	2014
TikTok	2013
Twitter	2012
Instagram	1986
Snapchat	1975
Name:	count, dtype: int64

- There are five platforms in the data that we will analyze through, which are TikTok, Facebook, Instagram, Twitter, and Snapchat.



Exploratory Data Analysis



Are there missing values in the data

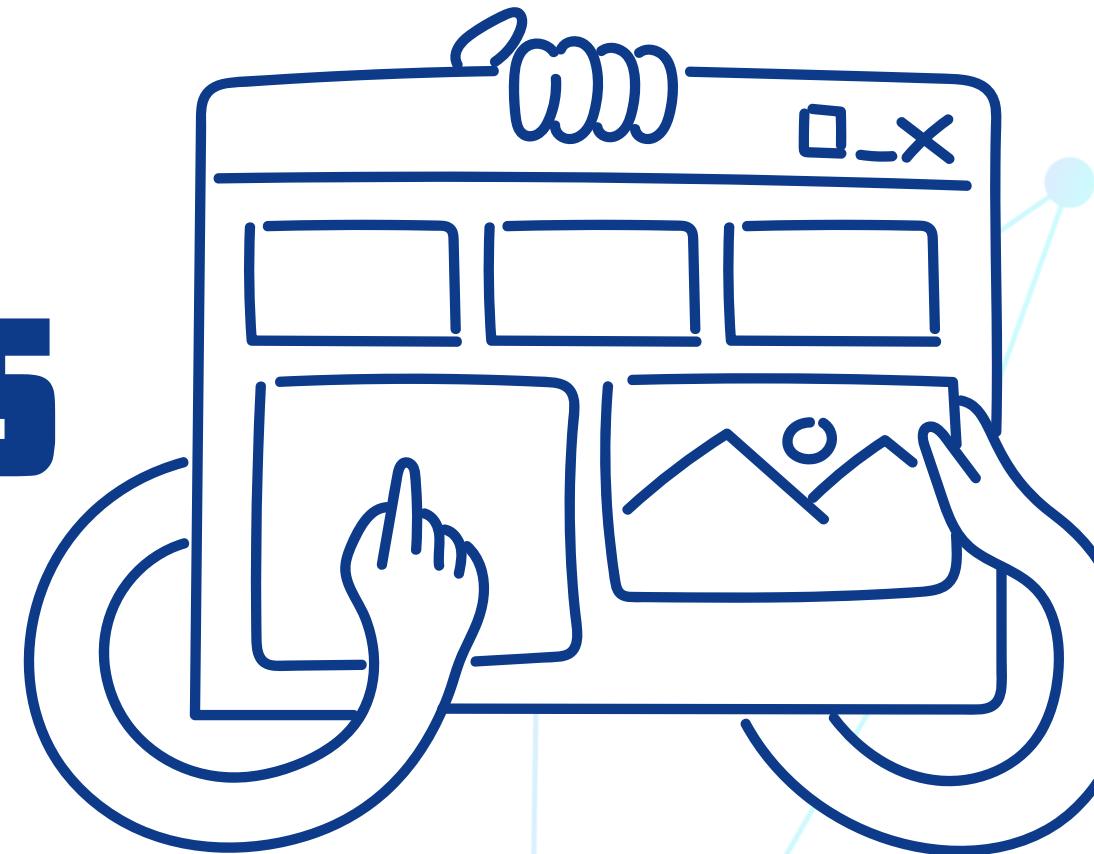
```
sentiment_fact.isnull().sum()
```

SentimentID	0
PostID	0
UserID	0
DateID	0
SentimentScore	0
SentimentCategory	0
Date	0
Month	0
Quarter	0
Year	0

- There is no data loss and the data is intact and clean



Dashboards



Dashboard 1



[Clear all slicers](#) →

Year

All

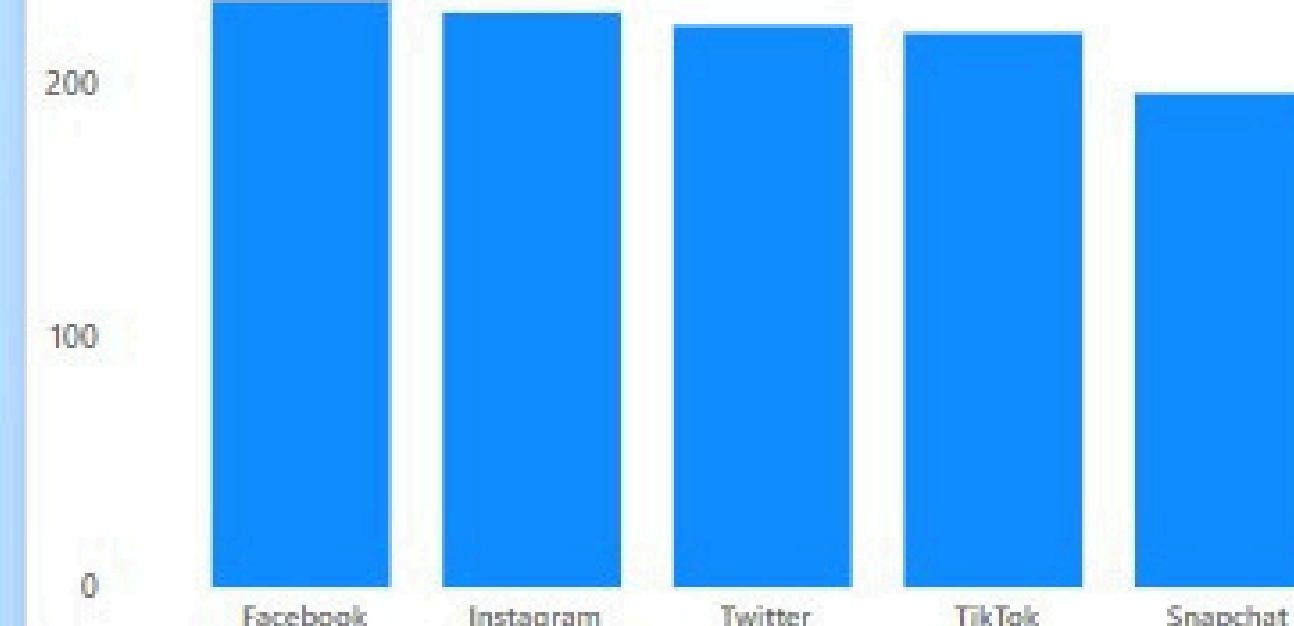
Post Content

Games

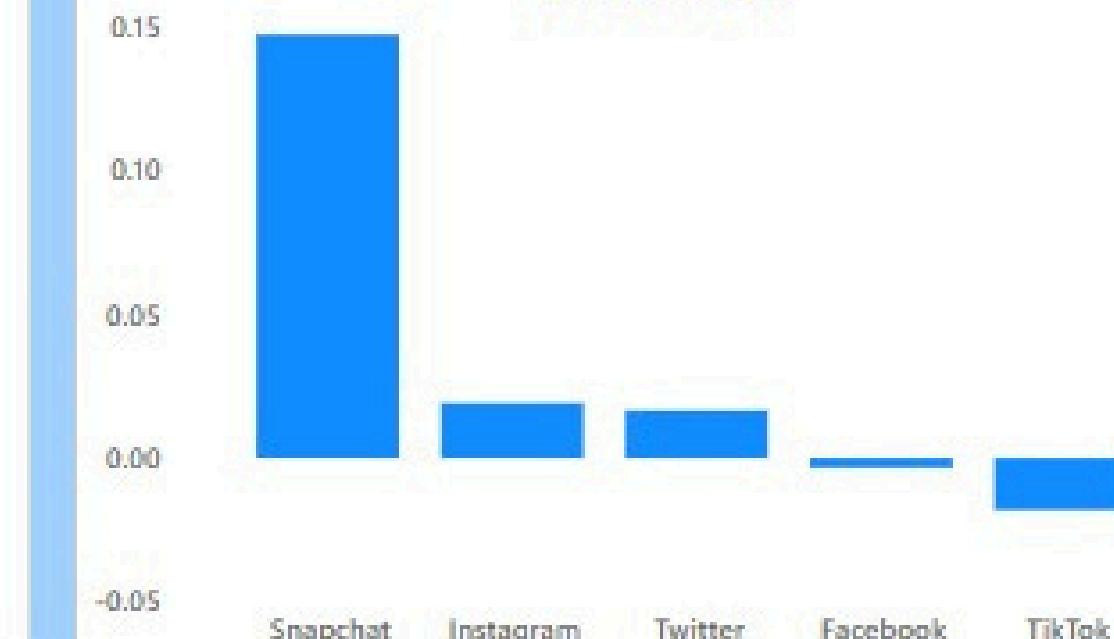


Evaluating Platform Success Based on Post Volume and Content Type

Post Volume and Content Type



Sentiment Score Distribution by Year Across Platforms



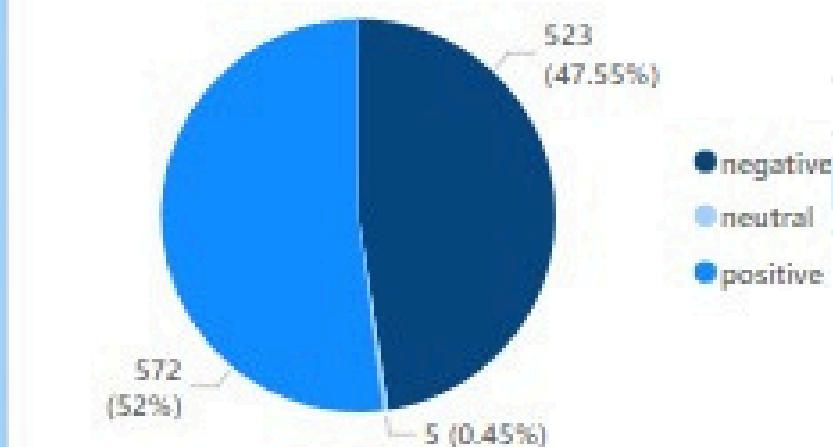
Top Global Content Creators Ranked by Follower Count

1106	19996	Tokyo
1191	9996	Paris
1854	9993	Tokyo
3174	9979	New York
414	9991	New York
4422	9997	New York

Top Global Content Creators Ranked by Sentiment Score

1002	0.78	Tokyo
1004	0.03	New York
101	0.89	Cairo
1020	-0.96	Tokyo
1021	0.51	Paris
1023	-0.22	Tokyo

Overall Sentiment Distribution for Global Content





Dashboard 2

وزارة الاتصالات
وتقنيه لجها المعلومات



Clear all slicers

Platform

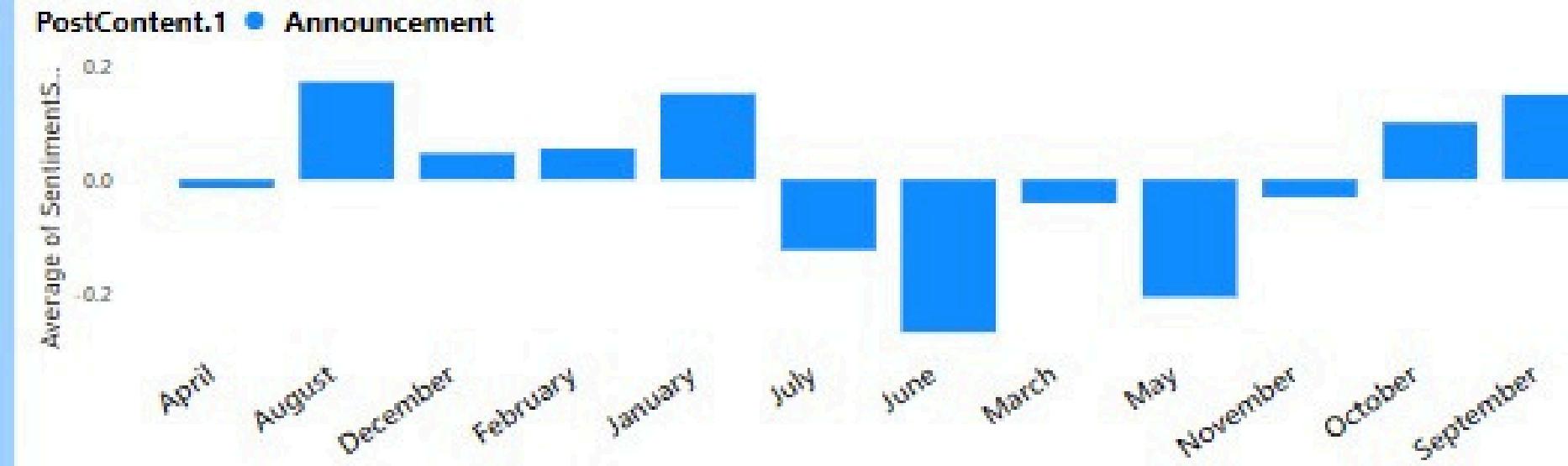
Instagram

PostContent.
Announce...



Analysis Platform And Content

Content Performance: Positive vs. Negative Sentiment Over Months by Platform

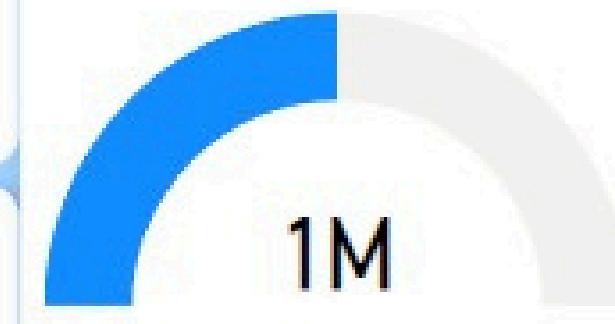


Followers By Platform And Content

PostContent.1 ● Announcement



Followers



Dashboard 3



Global Comparative Analysis of Content Quality and Sentiment Across Major Cities

Year

All

Post Content

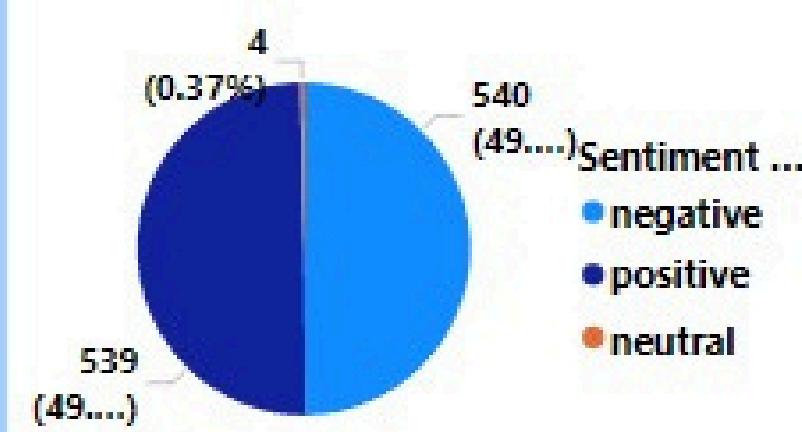
Course



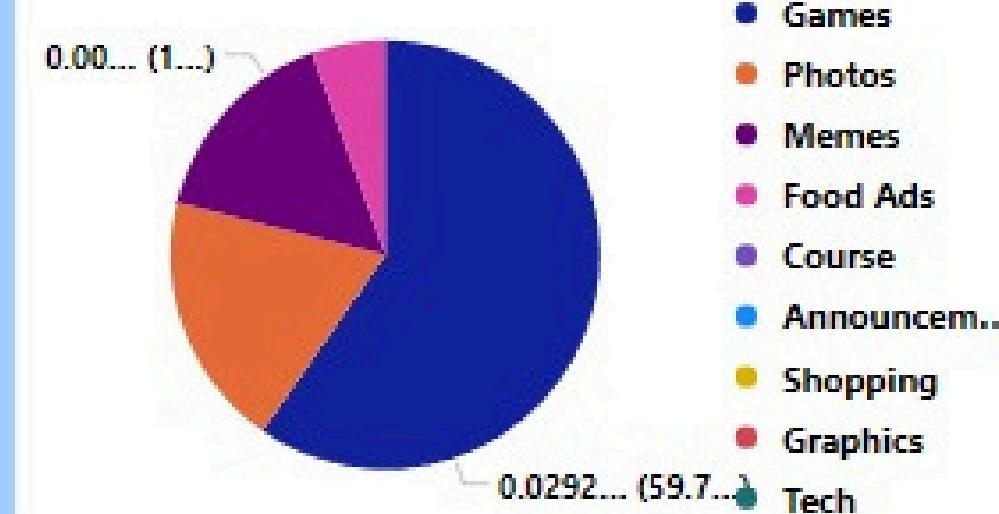
Content Quality Distribution Across Global Cities



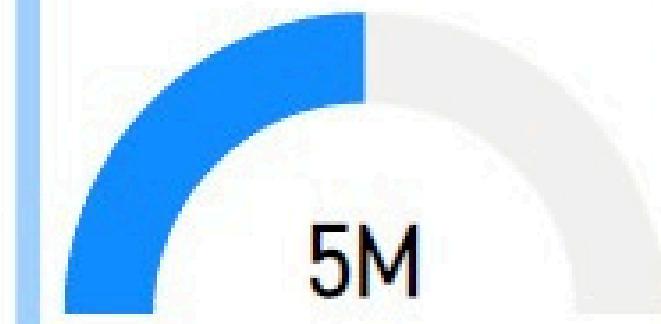
Number Of Feedback Based On Category



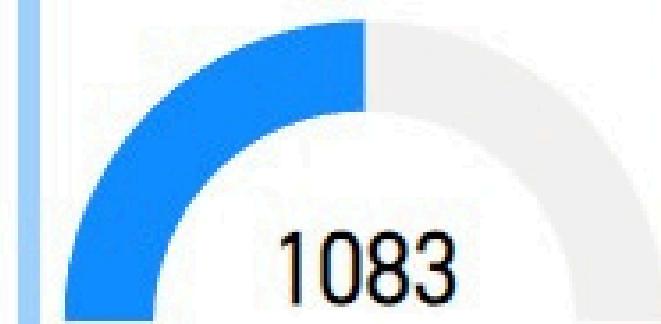
SentimentScore For Content all year



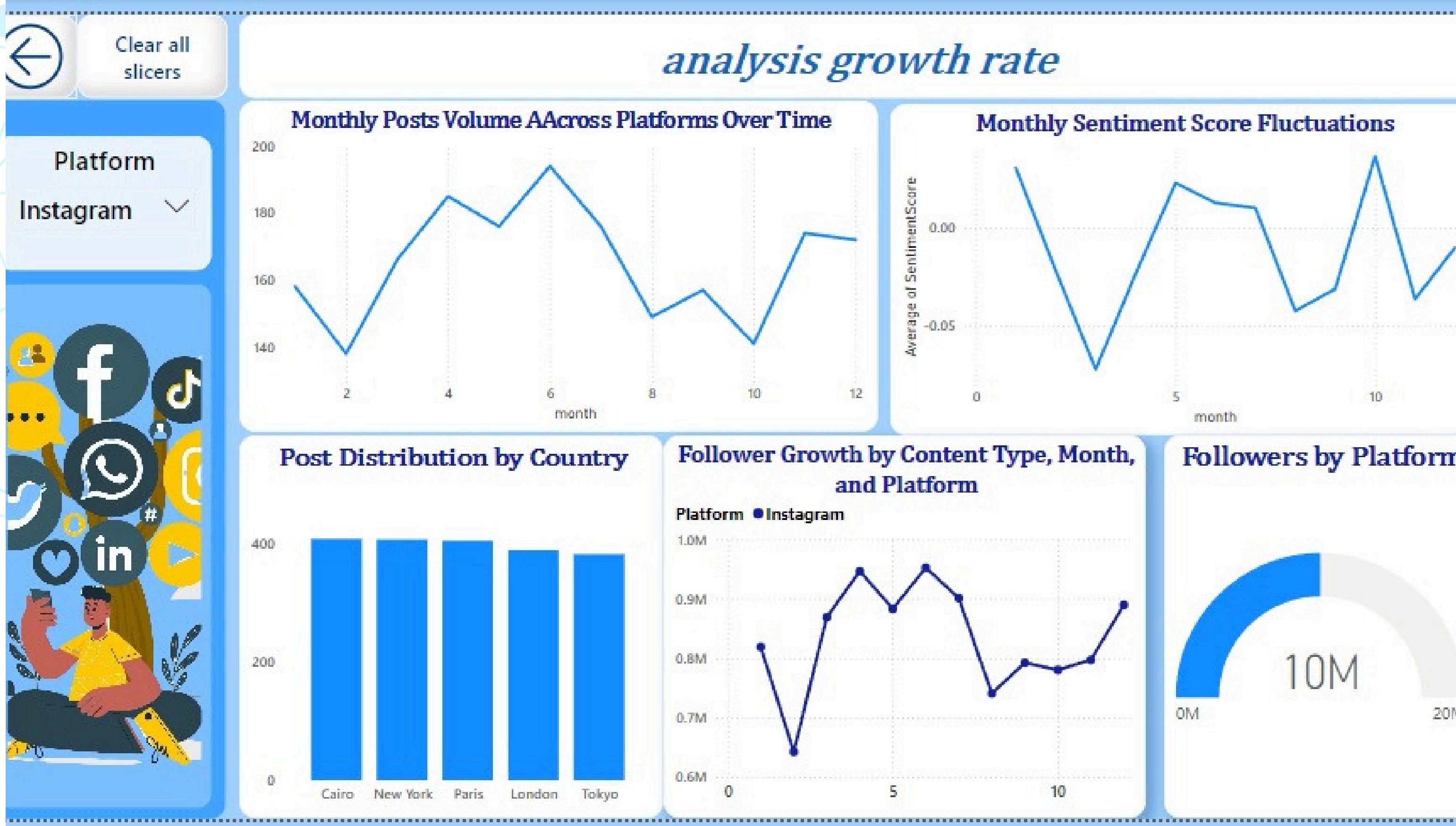
Followers



Total number Of Feedback



Dashboard 4





Dashboard 5

وزارة الاتصالات
وتقنيه لجيا المعلومات



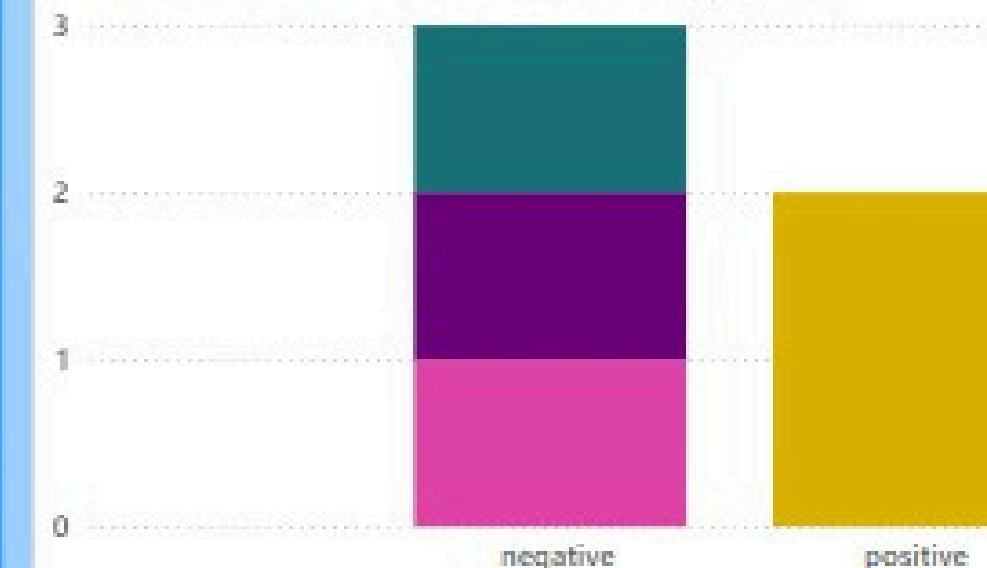
User Sentiment and Content Insights Across Global Platform

User ID
1441
1442
1443



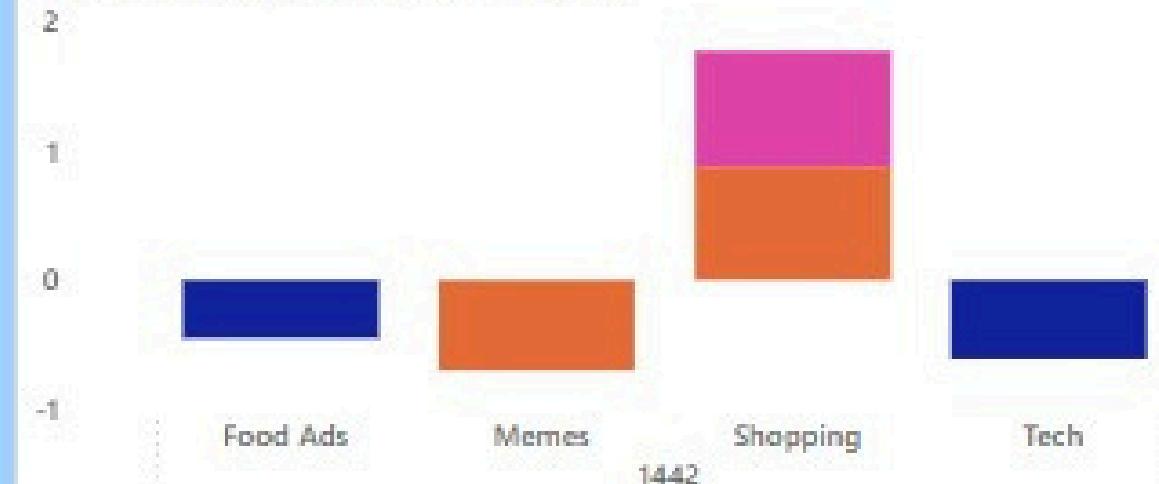
Sentiment Analysis by Content Type

PostContent.1 ● Food Ads ● Memes ● Shopping ● Tech



Platform Insights for Specific Content Types and Users

Platform ● Instagram ● Snapchat ● Twitter



Location

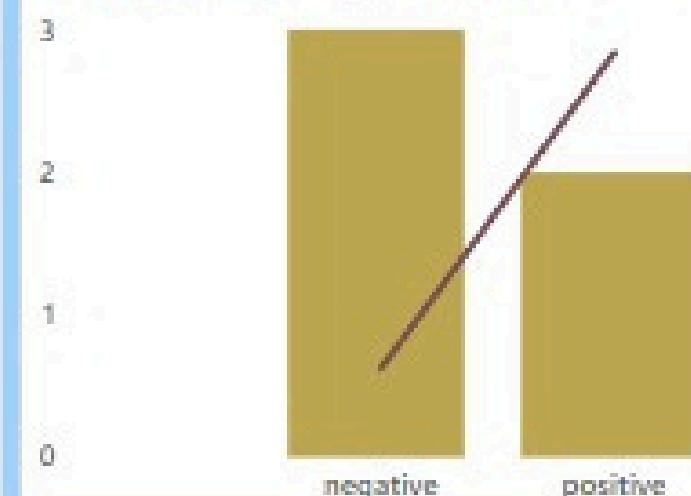


Followers

13K

Correlation Between Sentiment and The Number Of feedback Among Creators

UserID ● 1442 ● Average of SentimentScore



Dashboard 6



Top and Bottom 10 Content Creators Ranked By Performance

- Post Content**
- Games
 - Announce...
 - Course
 - Food Ads
 - Games
 - Graphics



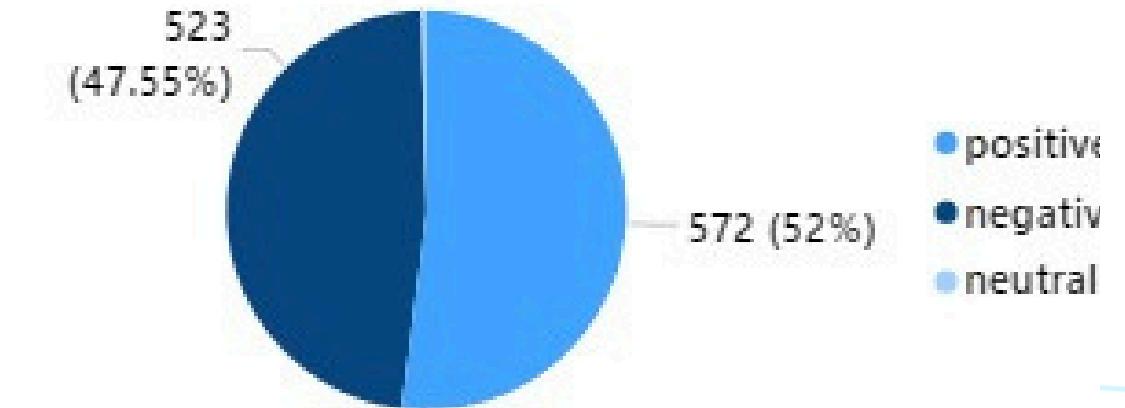
Geographical Distribution of Top-Ranked Content Creators



Top 10 Content Creators by Quality and Content Type

185	0.99	Games
2352	0.99	Games
2388	0.99	Games
245	1.00	Games
2559	0.99	Games
2710	0.99	Games
28	1.00	Games

Sentiment Analysis of Top and Bottom Content Creators



Bottom 10 Content Creators Ranked by Performance Quality

1763	-0.99	Games
1794	-0.99	Games
2410	-0.97	Games
2965	-0.98	Games
3103	-0.98	Games
3203	-1.00	Games



insights



- **Best city to target based on your content type:**

- 1 - Announcement content: Tokyo
- 3 - Food content: London
- 5 - Graphics content: Tokyo
- 7 - Shopping content: NewYork
- 9- Tech content: Paris

- 2 - Course content: NewYork
- 4 - Games content: London
- 6 - Memes content: NewYork
- 8 - Photos content: London



insights



وزارة الاتصالات
وتكنولوجيا المعلومات



- **The platforms with the most number of positive feedback (Sentiment Score)**

- 1 - Announcement content: Facebook
- 3 - Food content: Instagram
- 5 - Graphics content: Twitter
- 7 - Shopping content: Facebook
- 9- Tech content: TikTok

- 2 - Course content: Facebook
- 4 - Games content: Snapchat
- 6 - Memes content: Facebook
- 8 - Photos content: Facebook



insights



- **The platforms with the most number Users(Followers)**

- 1 - Announcement content: Facebook
- 3 - Food content: Snapchat
- 5 - Graphics content: TikTok
- 7 - Shopping content: TikTok
- 9- Tech content: Instagram

- 2 - Course content: Twitter
- 4 - Games content: Instagram
- 6 - Memes content: Snapchat
- 8 - Photos content: TikTok





Recommendations

Recommendations for Campaign Success

- Focus on High-Engagement Platforms : Prioritize **Snapchat** and **TikTok** for campaigns due to their higher positive sentiment scores.
- Focus on High-Engagement Influencers Collaborate with influencers from **Tokyo**, **New York**, and **Cairo**, where follower counts and sentiment are strong, to amplify campaign reach and ensure positive audience interaction.
- Content creators would benefit from focusing their efforts on posting during the months of **May**, **June**, and **July**. This period aligns with the start of summer vacation, which tends to have a positive impact on public sentiment.





Future Work



- 1- Add a date for each time the follower who follow or unfollow a user**
- 2- give us more information about the engagement time for each post to know the beat time to post**
- 3-The learning model helps content creators publish at the right time according to the type of content that achieves the highest reach**





Any Questions?





Thank You

