Countries with the highest Instagram audience reach and reasons behind it

(COMP3125 Individual Project)

Davud,Azizov  
*School of Computing and Data Science*

***Abstract*—** **This research investigates what drives Instagram adoption across nations, analyzing 2024 data through clustering and correlation methods. Results reveal Middle Eastern countries leading global usage (95.6%) despite lower GDP, while traditional economic indicators show surprisingly weak correlation with adoption rates. The findings suggest cultural factors and regional characteristics, rather than economic development, primarily influence social media penetration patterns worldwide.**

***Keywords—Instagram usage, social media penetration, regional analysis, cultural adoption***

# **I.Introduction (*Heading 1*)**

Social media platforms have literally transformed how people connect globally, with Instagram being one of the leading platforms. This individual project aims to analyze the countries with the highest Instagram audience reach and the reasons behind it, such as is it related to countries’ GDP level, their territory, their religion, etc.

Considering the relationship, we have with social media in modern society makes this topic important and interesting. Considering also how attached people are to their phones, scrolling through social media for hours, it makes sense that the richest countries, with the highest smartphone penetration rates, should dominate Instagram use. There is, however, a surprising pattern that appears in the data: only one country (Turkey) appears on both lists of top Instagram users and top GDPs. Contrary to what we might expect, higher GDP usually indicates busier workers with less time to use social media - a discovery that confronts our assumptions about digital behavior.

# **II.Datasets**

## ***A.Source of dataset (Heading 2)***

I got all datasets used in this project from Statista – very credible and widely-used source. These are the datasets were used:

1. Instagram penetration rate by country and region – as of April 2024
2. Global GDP(Gross Domestic Product) rankings by country – 2024 forecast based on the first 3 months of 2024
3. Social media advertising revenue – as of December 2023
4. Social network penetration rates by region – as of 2024

## ***B. Character of the datasets***

The all datasets were downloaded in Excel format. Before analysis, they all needed some preprocessing. Here are the characteristics of each dataset:

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Variables** | **Units** | **Processing steps** |
| Instagram Usage | Country name, Usage rate | Percentage (%) | Removed first few rows/columns, kept only two columns |
| GDP Rankings | Country name, 2023 GDP | Billion in USD | Only top 20 countries, removed 2024 column, used 2023 |
| Ad Revenue | Country name, Ad spend | Million in USD | Made country column to match with other datasets |
| Regional Penetration | Region name, Penetration rate | Percentage (%) | Removed first empty few rows/columns |

To integrate data – such as merging 2 different ones, I used pandas on country name with the help of seaborn(“hue”). No unit conversions were done since I used all as it is. Additionally, the data cleaning process involved removing headers/footers, handling missing values, etc.

# **III.Methodology**

## ***A.K-means clustering Analysis***

Used to find a distinct patterns in Instagram usage rates across countries.

**Implementation:**

km = KMeans(n\_clusters=3,random\_state=42)

y\_predicted = km.fit\_predict(values)

df1['cluster'] = y\_predicted

The advantages of this method are it is very simple to implement, works well with the data I have, and make it easier to see the patterns. On the other hand, requiring pre-specifying number of clusters is a disadvantage – I hope it’d create clusters on its own.

## ***B.Correlation Analysis***

Examines the relationships between Instagram usage, ad spend and GDP

**Implementation:**

merged\_df = pd.merge(df4, df2[['Country', '2023']], on='Country', how='inner')

merged\_df.columns = ['Country', 'Ad\_Spend', 'GDP\_2023']

The advantages of this method is also very easy to implement and easily visualized. Disadvantages are it only detects linear relationships and very sensitive to outliers.

## ***C. Data Visualization***

1. **Scatter Plots**
   * Shows relationships between variables
   * Implemented using matplotlib and seaborn
2. **Bar Charts**
   * Regional comparisons
   * Usage rate distributions
3. **Heatmaps**
   * Correlation visualization
   * Pattern identification

# **IV.Results**

**A. Global Instagram Usage Patterns**

K-means clustering analysis revealed three distinct groups of countries based on Instagram usage rates:

**Cluster Results:**

* High Usage (85-96%): Middle Eastern countries (Bahrain, UAE, Kazakhstan)
* Medium Usage (78-85%): Mixed profile countries (Chile, Brazil, Argentina)
* Lower Usage (68-76%): Primarily Western countries (Portugal, Australia)

*# Cluster details with statistics*

print("Cluster 0:", df11[['Country', 'in %']]) *# Lower usage*

print("Cluster 1:", df22[['Country', 'in %']]) *# Medium usage*

print("Cluster 2:", df33[['Country', 'in %']]) *# High usage*

**B. Economic Correlations**

Correlation analysis between Instagram usage, GDP, and ad spending revealed unexpected patterns:

1. Instagram Usage vs GDP:

* Weak negative correlation (-0.25)
* Higher GDP countries often showed lower usage rates
* Notable outliers: UAE (high GDP, high usage)

1. Usage vs Ad Spend:

* Moderate positive correlation (0.45)
* Higher ad spending generally associated with increased usage
* Regional variations significant

**C. Regional Analysis**

Regional penetration rates showed clear geographic patterns:

1. Northern Europe leads (80.2%)
2. Western Europe follows (78.2%)
3. Eastern Asia third (74.9%)

Unexpected findings:

* Middle East's high adoption despite lower regional averages
* North America's relatively lower position
* Significant urban-rural divide within regions

# **V.Discussion**

The deeper we dug into Instagram's global usage patterns, the more we realized both the power and shortcomings of our research. Yes, we spotted some fascinating trends in how different regions and economies use Instagram, but we need to be upfront about what we might have missed. We only had solid data from 20 countries - barely scratching the surface, really. Think about all the emerging markets we couldn't include - there's probably a whole other story there.

Our methods worked well enough - K-means clustering showed us some interesting groupings, and the correlation analysis pointed to some clear patterns. But let's be real: social media adoption isn't as simple as plotting points on a graph. We might have oversimplified things. You can't just draw a straight line between a country's GDP and how many people are posting Stories and Reels. It's messier than that. And while lumping countries into regions helped us spot big patterns, it probably glossed over some important local quirks.

There's so much more we could do with this research. I'd love to see what happens when we factor in things like age breakdowns and whether users are city folks or rural dwellers. Getting data from more developing countries would be huge - that's where some of the most interesting trends might be hiding. And imagine tracking these patterns over time - we might catch sight of where social media is headed next. Maybe ditch the simple correlation charts for something that can handle the messy reality of how culture, money, and tech all mix together.

The real gold here could be for the people developing Instagram and plotting market strategies. Understanding why Instagram looks different in Tokyo versus Toronto versus Tehran could help shape everything from new features to marketing campaigns. And if we could somehow measure cultural factors and how people actually use the app - not just how many people use it - now that would be something. Maybe I will work for those companies in the future to make things different – who knows?!

Sure, this research had its blind spots. But that's what makes this field exciting - there's always more to discover, more angles to explore. What we found is just the beginning of understanding how social media weaves its way into different societies around the globe.

# **VI.Conclusion**

Looking at Instagram's global footprint in 2024 threw up some fascinating surprises. We kept expecting the usual story - richer countries leading the pack - but what I found turned that idea on its head. Who would've thought that Middle Eastern countries would be the Instagram powerhouses? The numbers are striking: Bahrain's hitting 95.6%, Kazakhstan's not far behind at 90.8%, and the UAE's right there at 90.7%.

Here's what really got me thinking: everyone assumes wealthy countries would dominate social media, but our data tells a different story. The link between GDP and Instagram use is surprisingly weak - negative at -0.25. It's almost like money isn't the driving force we all thought it was. Instead, it looks like where you live, and your cultural background have a lot more to do with whether you're posting Stories and Reels.

Take Northern Europe - they're crushing it with 80.2% of people on social media. Western Europe's right behind at 78.2%, and Eastern Asia's holding strong at 74.9%. But it's not about their economies - it's more about their digital setup and how their societies view social media.

This stuff matters big time for anyone trying to make it in the global social media game. If you're running a platform like Instagram or trying to market through it, you can't just look at a country's GDP and call it a day. You've got to dig deeper, understand the local vibe, figure out what makes each region tick.

The big takeaway? Social media's way more complex than we thought. You can't just follow the money - you've got to get why people in different parts of the world connect the way they do. It's a wake-up call for marketers and tech companies: success means really getting to know your audience, wherever they are.

##### **References**

1. Statista Research Department, "Instagram penetration rate among internet users in selected global regions as of January 2023," Statista, Jan. 2023
2. Statista Research Department, "Distribution of Instagram users worldwide as of January 2023, by age group," Statista, Jan. 2023.
3. Statista Research Department, "Social network penetration rate worldwide from 2017 to 2025, by region," Statista Digital Market Insights, 2023.
4. Statista Research Department, "Social media advertising revenue in selected countries worldwide from 2022 to 2027," Statista Digital Market Outlook, 2023.
5.  International Monetary Fund, "Countries with the largest gross domestic product (GDP) in 2023," Statista, Oct. 2023.