

Facebook users

waheeb Algabri

Project Proposal

Exploring Facebook Users by Country Dataset and also do Comparative Analysis of Facebook Usage in Saudi Arabia and Yemen

Introduction

Social media platforms have become an essential part of our daily lives, and Facebook is one of the most popular ones. This project aims to explore the distribution of Facebook users across different countries and regions, using the Facebook Users by Country dataset. We also aim to conduct a comparative analysis of Facebook usage in Saudi Arabia and Yemen. The objective is to understand the differences in Facebook adoption, user numbers, and usage patterns between these two countries. By examining these variations, we can gain insights into the factors influencing Facebook usage and the digital landscape in these regions.

```
library(tidyverse)
library(ggplot2)
```

Reading the data

```
data <- read.csv("data.csv")
```

```
glimpse(data)
```

```
## Rows: 226
## Columns: 5
## $ Name      <chr> "India", "United States", "Indonesia", "Brazil", "Phil~
## $ Users     <chr> "416.6M", "240M", "176.5M", "139M", "91M", "78M", "75.~
## $ Facebook_Users. <chr> "29.16%", "70.59%", "63.6%", "64.23%", "77.55%", "60.7~
## $ Date_of_Data <chr> "2021-06", "2020-12", "2021-06", "2020-12", "2021-06",~
## $ Population <chr> "1,428,627,663", "339,996,563", "277,534,122", "216,42~
```

```
str(data)
```

```
## 'data.frame':   226 obs. of  5 variables:
## $ Name          : chr  "India" "United States" "Indonesia" "Brazil" ...
## $ Users          : chr  "416.6M" "240M" "176.5M" "139M" ...
## $ Facebook_Users.: chr  "29.16%" "70.59%" "63.6%" "64.23%" ...
## $ Date_of_Data   : chr  "2021-06" "2020-12" "2021-06" "2020-12" ...
## $ Population     : chr  "1,428,627,663" "339,996,563" "277,534,122" "216,422,446" ...
```

```
summary(data)
```

```
##      Name           Users      Facebook_Users.    Date_of_Data
## Length:226      Length:226      Length:226      Length:226
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
## Population
## Length:226
## Class :character
## Mode  :character
```

```
knitr::kable(head(data, 7))
```

Name	Users	Facebook_Users.	Date_of_Data	Population
India	416.6M	29.16%	2021-06	1,428,627,663
United States	240M	70.59%	2020-12	339,996,563
Indonesia	176.5M	63.6%	2021-06	277,534,122
Brazil	139M	64.23%	2020-12	216,422,446
Philippines	91M	77.55%	2021-06	117,337,368
Mexico	78M	60.72%	2020-12	128,455,567
Vietnam	75.9M	76.82%	2021-06	98,858,950

```
# Identify countries with the highest and lowest Facebook users
```

```
highest_users <- data %>% arrange(desc(Users)) %>% head(7)
```

```
lowest_users <- data %>% arrange(Users) %>% head(7)
```

```
knitr::kable(head(highest_users, 7))
```

Name	Users	Facebook_Users.	Date_of_Data	Population
Anguilla	9K	56.61%	2020-12	15,899
Cyprus	989K	78.48%	2020-12	1,260,138
Mayotte	95.5K	28.42%	2020-12	335,995
Saint Lucia	92K	51.04%	2020-12	180,251
Mauritania	927.3K	19.07%	2020-12	4,862,989
Philippines	91M	77.55%	2021-06	117,337,368
Aruba	91K	85.63%	2020-12	106,277

```
knitr::kable(head(lowest_users, 7))
```

Name	Users	Facebook_Users.	Date_of_Data	Population
Botswana	1.1M	42.57%	2020-12	2,675,352
Tajikistan	1.1M	11.2%	2021-06	10,143,543
Jamaica	1.1M	38.93%	2020-12	2,825,544
Slovenia	1.1M	50.34%	2020-12	2,119,675
Zimbabwe	1.3M	7.82%	2020-12	16,665,409
Sudan	1.3M	2.7%	2020-12	48,109,006
Lithuania	1.4M	51.5%	2020-12	2,718,352

Name	Users	Facebook_Users.	Date_of_Data	Population
------	-------	-----------------	--------------	------------

Let's take a quick look at yemeni users

```
# Filter the dataset for Yemen
yemen_data <- data %>%
  filter(Name == "Yemen")

# Print the filtered data

knitr::kable(yemen_data)
```

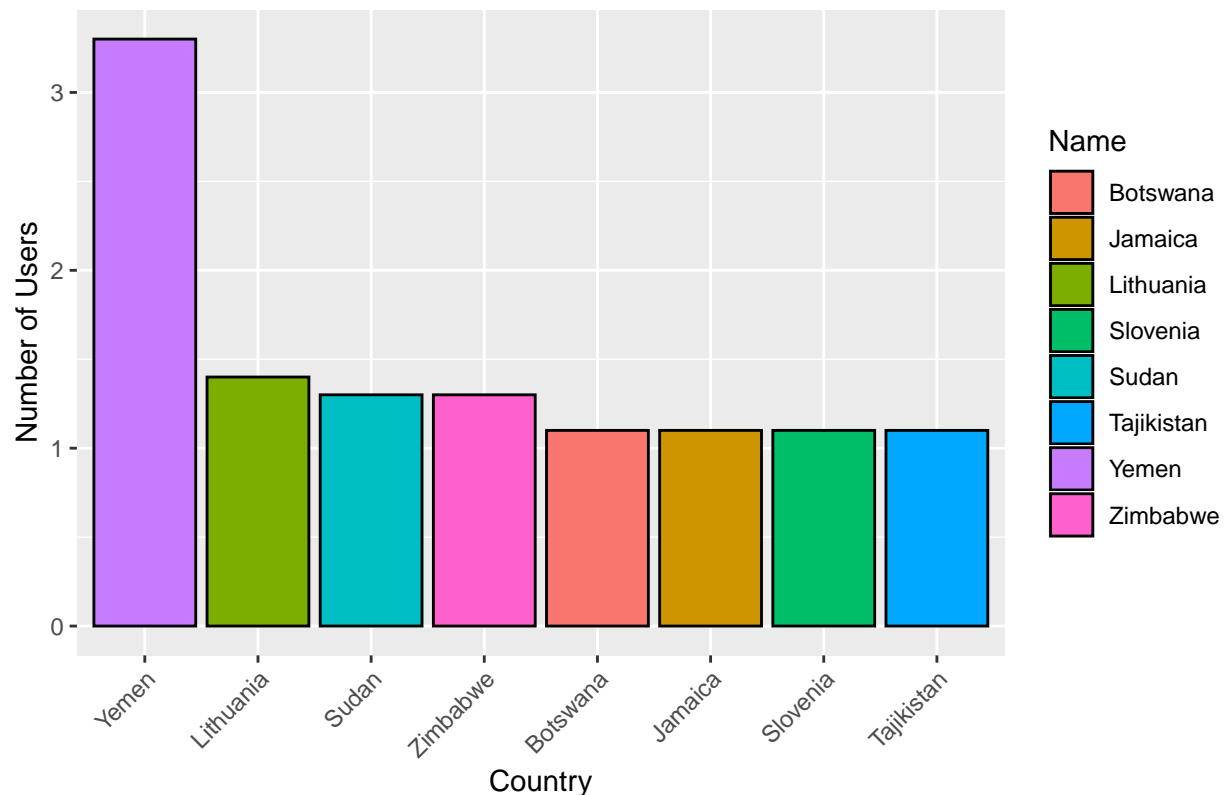
Name	Users	Facebook_Users.	Date_of_Data	Population
Yemen	3.3M	9.7%	2021-03	34,449,825

```
# Combine the filtered datasets for Yemen and the top 7 lowest countries
combined_data <- rbind(yemen_data,lowest_users )

# Convert the Users column to numeric
combined_data$Users <- as.numeric(gsub("M", "", combined_data$Users))

# Plot the comparison
ggplot(combined_data, aes(x = reorder(Name, -Users), y = Users, fill = Name)) +
  geom_bar(stat = "identity", color = "black") +
  labs(title = "Comparison of Facebook Users - Yemen vs. Top 7 lowest Countries",
       x = "Country", y = "Number of Users") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

Comparison of Facebook Users – Yemen vs. Top 7 lowest Countries



So let's take a quick look at saudi data which makes it easier for us to compare between users

```
# Filter the dataset for Yemen and Saudi Arabia
```

```
saudi_data <- data %>%
  filter(Name %in% c("Saudi Arabia"))
knitr::kable(saudi_data)
```

Name	Users	Facebook_Users.	Date_of_Data	Population
Saudi Arabia	26.3M	71.13%	2021-03	36,947,025

```
# Create a new dataset combining Yemen and Saudi Arabia data
```

```
comparison_data <- rbind(yemen_data, saudi_data)
```

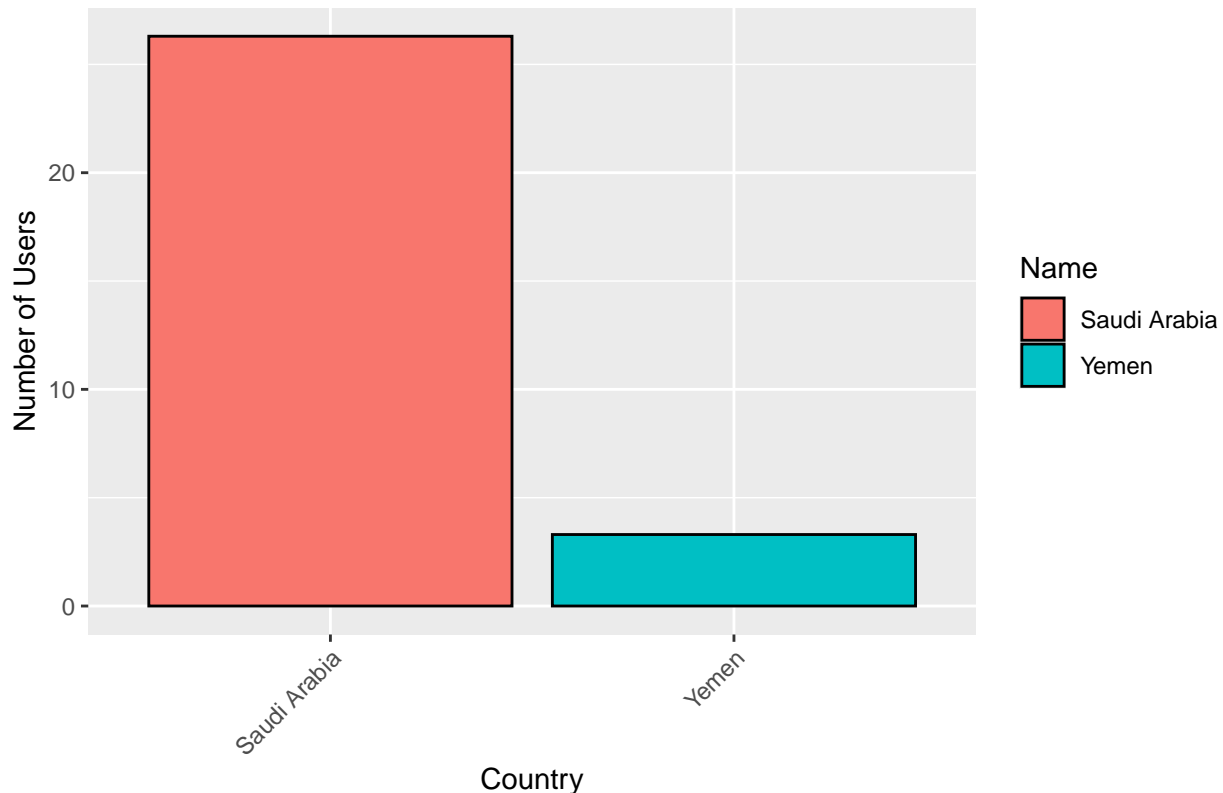
```
# Convert the Users column to numeric
```

```
comparison_data$Users <- as.numeric(gsub("M", "", comparison_data$Users))
```

```
# Plot the comparison
```

```
ggplot(comparison_data, aes(x = Name, y = Users, fill = Name)) +
  geom_bar(stat = "identity", color = "black") +
  labs(title = "Comparison of Facebook Users - Yemen vs. Saudi Arabia",
       x = "Country", y = "Number of Users") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

Comparison of Facebook Users – Yemen vs. Saudi Arabia



From the data, we can observe the following:

- Saudi Arabia has a significantly higher number of Facebook users compared to Yemen. Saudi Arabia has approximately 26.3 million users, while Yemen has around 3.3 million users.
- In terms of Facebook user percentage, Saudi Arabia has a higher percentage (71.13%) compared to Yemen (9.7%). This indicates a larger proportion of the population in Saudi Arabia actively using Facebook compared to Yemen.
- The data is from the same date, which allows for a fairer comparison between the two countries.

Conclusion

Based on the factors mentioned, the following conclusion can be drawn regarding the difference in Facebook usage between Saudi Arabia and Yemen: The variation in Facebook usage between Saudi Arabia and Yemen can be attributed to multiple factors. Firstly, Saudi Arabia has a higher rate of internet penetration compared to Yemen, enabling a larger population to access and utilize online platforms like Facebook. This can be attributed to the investments made in internet infrastructure in Saudi Arabia, resulting in better connectivity and accessibility. Secondly, the disparity in Facebook usage can be influenced by socioeconomic factors. Saudi Arabia's higher GDP per capita and stronger economy contribute to greater internet access, affordability of devices, and overall digital literacy. These factors create an environment that fosters higher Facebook usage among the population. Additionally, government policies and regulations play a role in shaping Facebook usage. Saudi Arabia has actively promoted digital transformation and social media usage through initiatives such as the Vision 2030 plan. This has led to improvements in internet connectivity and the establishment of incentives for businesses to establish an online presence, thereby contributing to higher Facebook usage.

rates. Lastly, cultural and social factors also come into play. Saudi Arabia has experienced rapid social and cultural changes, including increased openness to social media and digital platforms. Social media platforms like Facebook have gained popularity as a means of socializing, content sharing, and staying connected. In contrast, Yemen may have different cultural dynamics and societal norms that impact the adoption of social media.