

## **Social Media Reach by Reihaneh**

This project analyzes the reach of Instagram posts to understand how the platform's algorithm influences post visibility. By manually collecting data from each post over a week, the project explores metrics such as impressions from users' home feeds, hashtags, profile views, and more. Using Python, the analysis visualizes patterns and applies machine learning techniques to predict future reach, offering insights for those interested in maximizing engagement on Instagram. It's ideal for data science students who want hands-on experience with social media analytics.

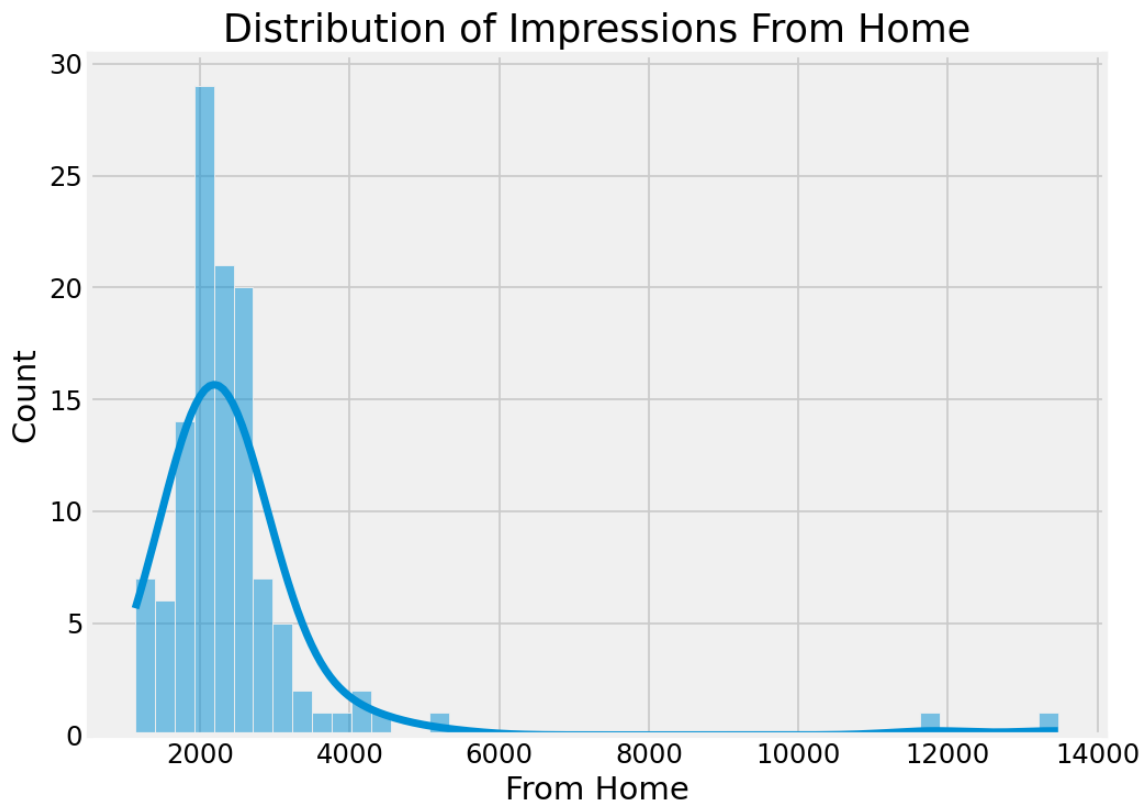
In this project, several Python libraries are imported to facilitate data manipulation, visualization, and modeling. Key libraries include:

- Pandas for data handling and manipulation.
- NumPy for numerical operations.
- Matplotlib and Seaborn for data visualization.
- Plotly for interactive plots.
- WordCloud for generating word clouds.
- Scikit-learn for machine learning tasks, particularly regression modeling.

The dataset, stored in a CSV file ("Instagram.csv"), is loaded into a Pandas DataFrame with the `read_csv` function, and its first few rows are displayed to give an overview of the data's structure and contents.

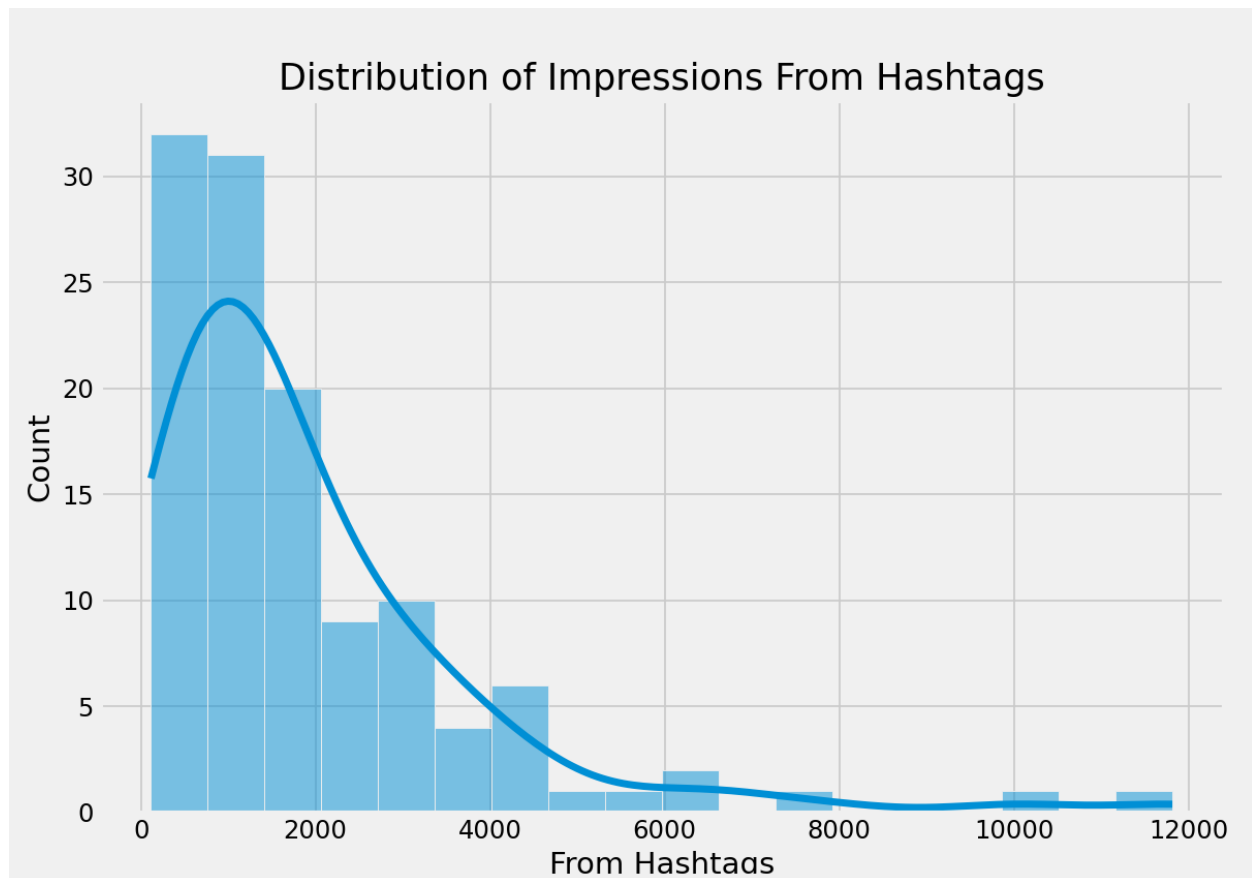
### Distribution of Impressions from Home:

This plot visualizes the distribution of impressions that Instagram posts receive from users' home feeds. By analyzing this distribution, we can identify patterns in how often posts appear in followers' home feeds, giving insights into the reach and effectiveness of posts in a key area of the Instagram algorithm.



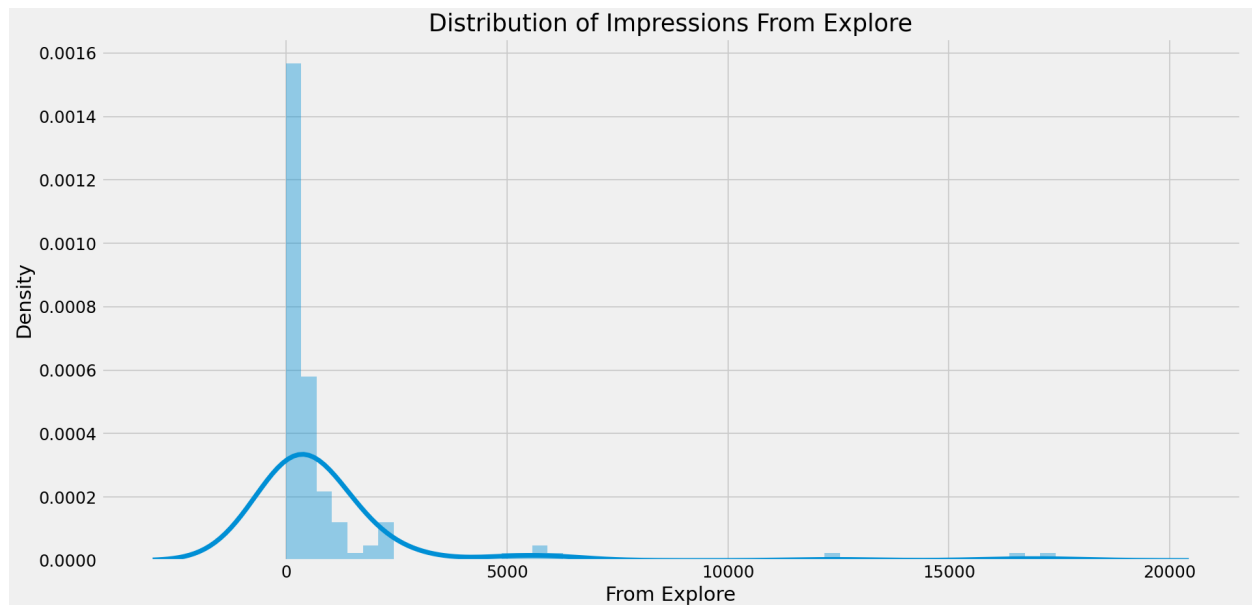
### Distribution of Impressions from Hashtags:

This plot shows the distribution of impressions generated from hashtags used in Instagram posts. By examining this distribution, we gain insights into how effectively hashtags contribute to a post's reach, helping to understand how specific hashtags influence post visibility within the broader Instagram community.



### Distribution of Impressions from Explore:

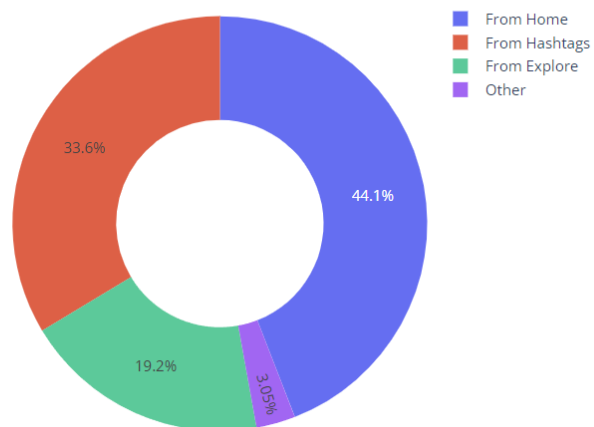
This plot displays the distribution of impressions that posts receive from Instagram's Explore page. Analyzing this distribution helps reveal how frequently posts appear on Explore, providing insights into the platform's recommendation algorithm and how it may impact reach for broader, non-follower audiences.



### Impressions on Instagram Posts from Various Sources:

This pie chart visualizes the proportion of impressions each source contributes to overall post reach on Instagram. By summing impressions from "Home," "Hashtags," "Explore," and "Other" sources, the chart shows the relative impact of each source on total impressions. The chart includes a hole in the center, creating a donut-style view that highlights the contribution of each source, giving a clear snapshot of where most impressions come from.

Impressions on Instagram Posts From Various Sources



## Analyzing Content

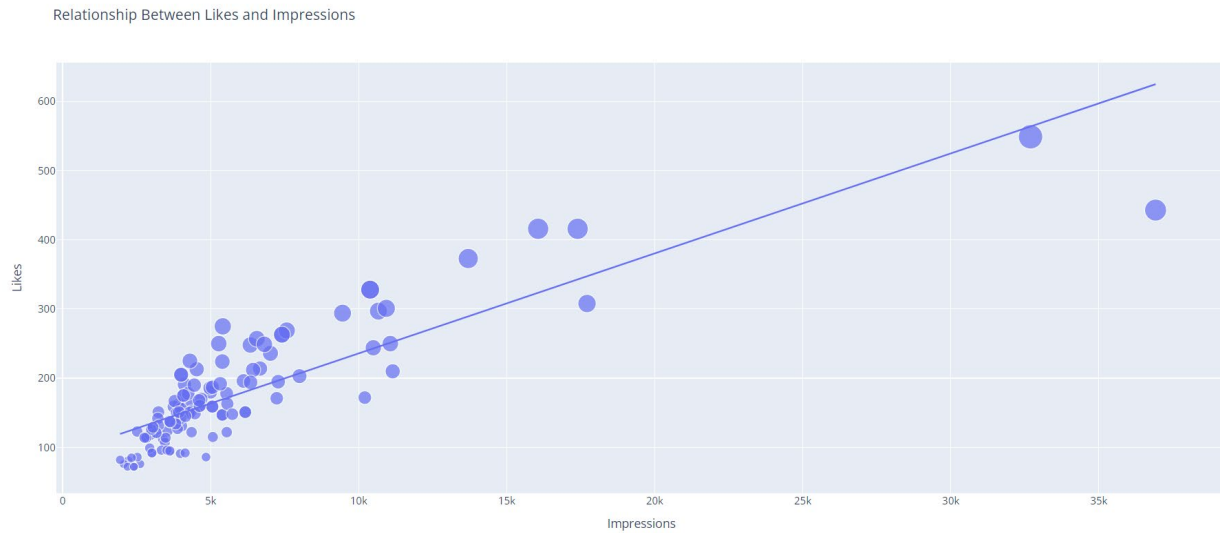
### Word Cloud of Instagram Post Captions:

This word cloud visualizes the most frequently used words in the captions of Instagram posts. By processing the captions and excluding common stop words (like "the" or "and"), the word cloud highlights key terms that appear most often. This provides a visual summary of the language used in posts and can help identify popular themes or topics that may influence engagement on Instagram. The "classic" style adds a clean, minimalistic aesthetic to the visualization.



### Relationship Between Likes and Impressions:

This scatter plot explores the relationship between the number of impressions and likes on Instagram posts. Each point represents a post, with the x-axis showing the number of impressions, the y-axis showing the number of likes, and the size of the points indicating the number of likes. A trendline is added to indicate the general correlation between impressions and likes. This visualization helps to understand how impressions impact engagement, revealing whether posts with higher reach tend to receive more likes.



This project provides a small but insightful example of how Instagram reach data can be analyzed and visualized. Through various charts and visualizations, we've demonstrated how to explore the impact of factors like impressions from the home feed, hashtags, and the Explore page. While this is just an introductory analysis, it highlights how data can be used to understand social media performance. Future work could include more complex models and deeper analysis to predict reach with greater accuracy.