



EDA Internship Task Week 5

Instagram Post Analysis

Find Dataset "Instagram data.csv" and implement EDA techniques.

Import following libraries at start:

- import pandas as pd
- import plotly.express as px
- import plotly.graph_objects as go
- import plotly.io as pio
- from wordcloud import WordCloud
- pio.templates.default = "plotly_white"

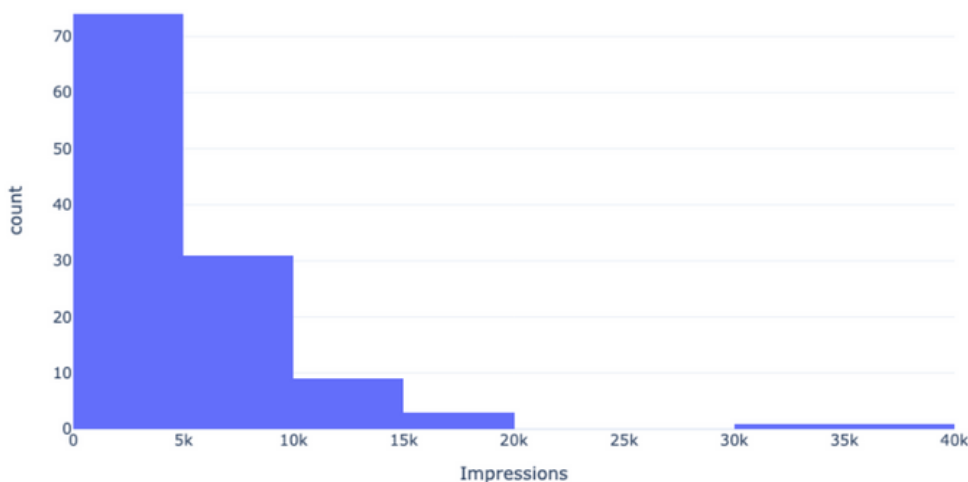
Q.1: Show column names and have a look at their info.

Q.2: Show the descriptive statistics of the data.

Q.3: Check if your data contains any missing values

Q.4: When you start exploring your data, always start by exploring the main feature of your data. For example, as we are working on a dataset based on Instagram Reach, we should start by exploring the feature that contains data about reach. In our data, the Impressions column contains the data about the reach of an Instagram post. So let's have a look at the distribution of the Impressions:

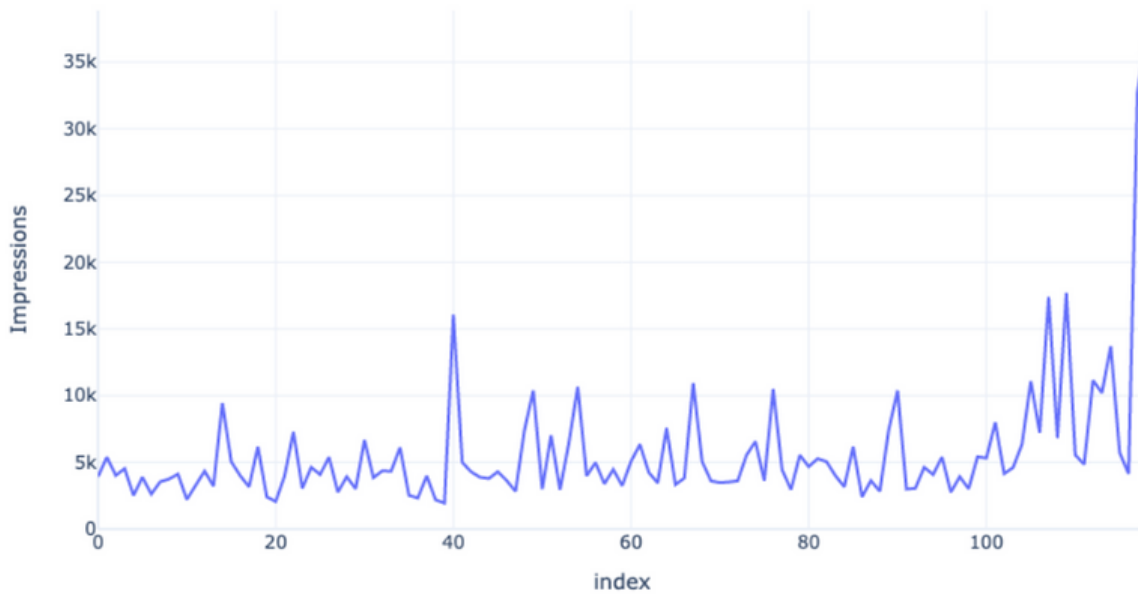
Distribution of Impressions





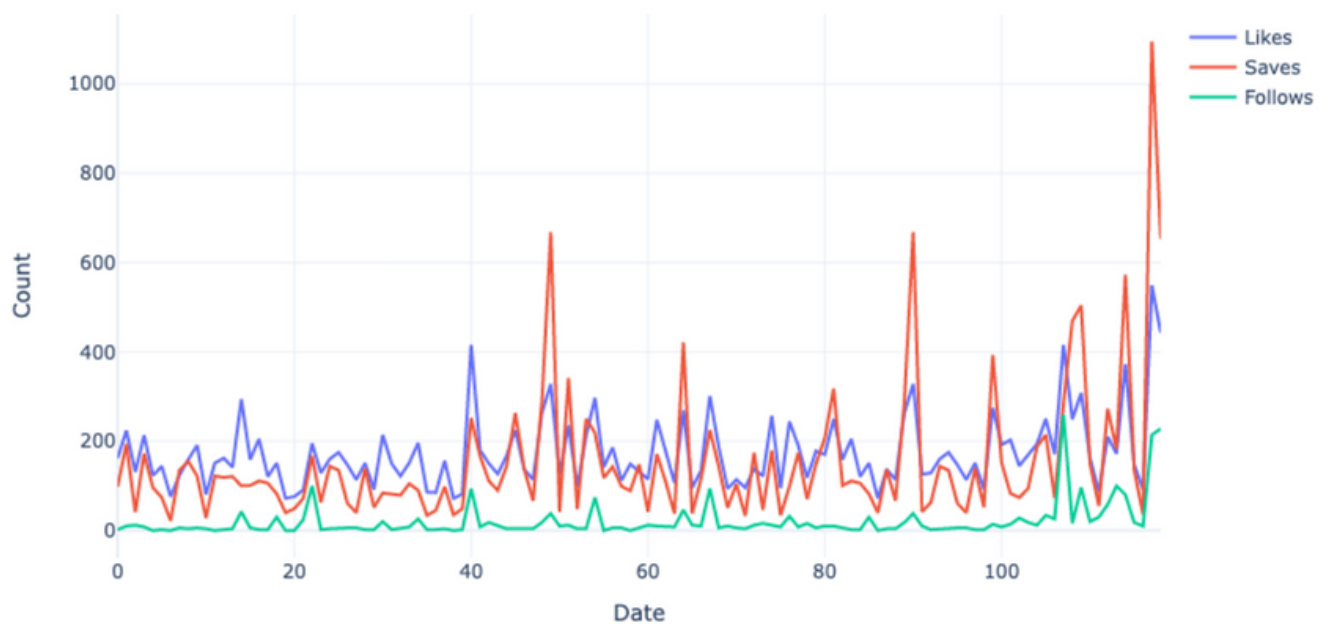
Q.5: Have a look at the number of impressions on each post over time as shown below

Impressions Over Time



Q.6: Have a look at all the metrics like Likes, Saves, and Follows from each post over time as shown below

Metrics Over Time



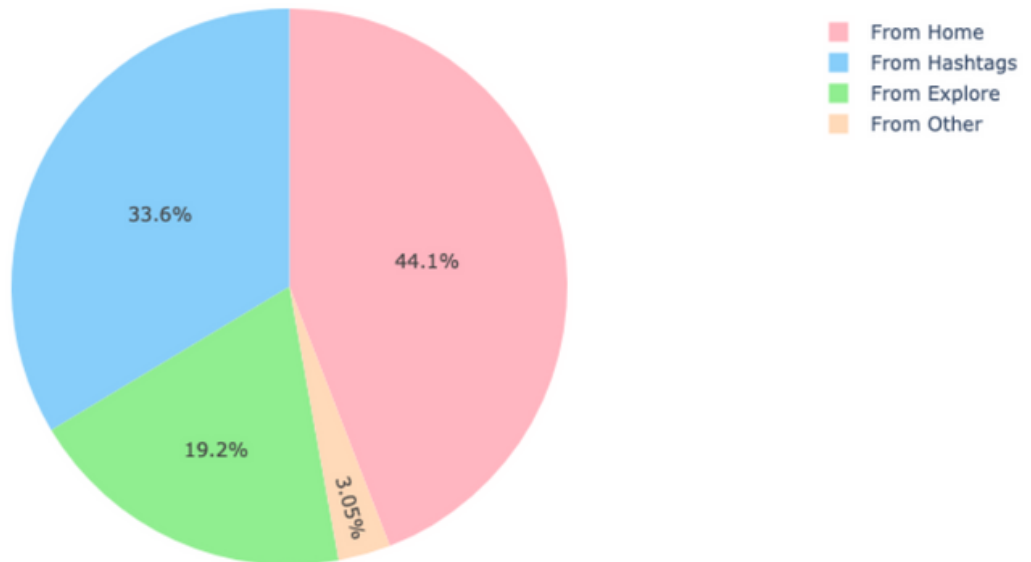


Artificial Intelligence Community of Pakistan

--Evolution to Revolution--

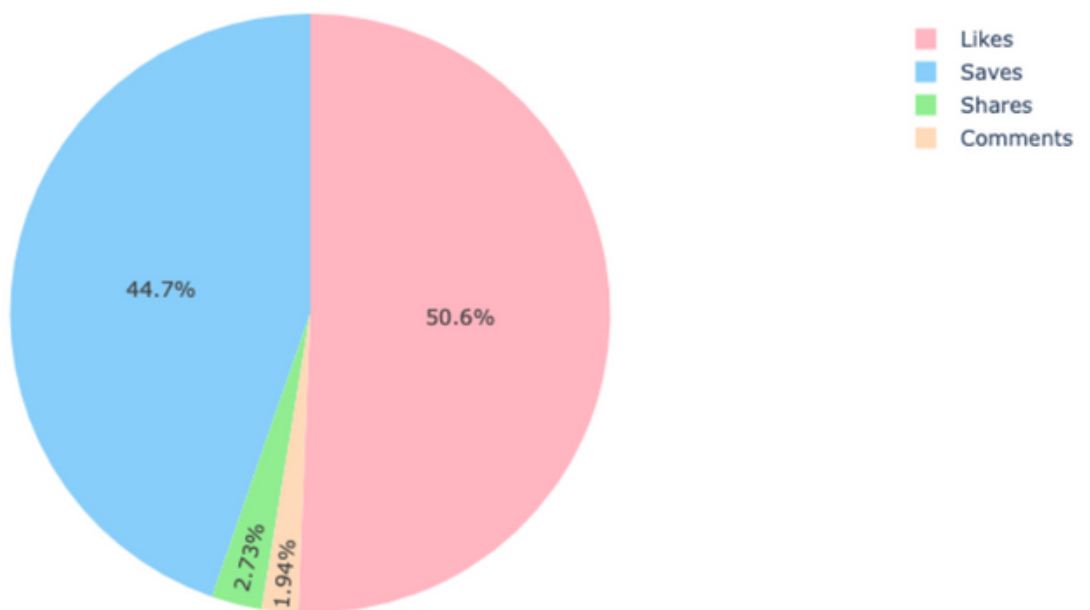
Q.7: Have a look at the distribution of reach from different sources as shown below

Reach from Different Sources



Q.8: Have a look at the distribution of engagement sources as shown below

Engagement Sources



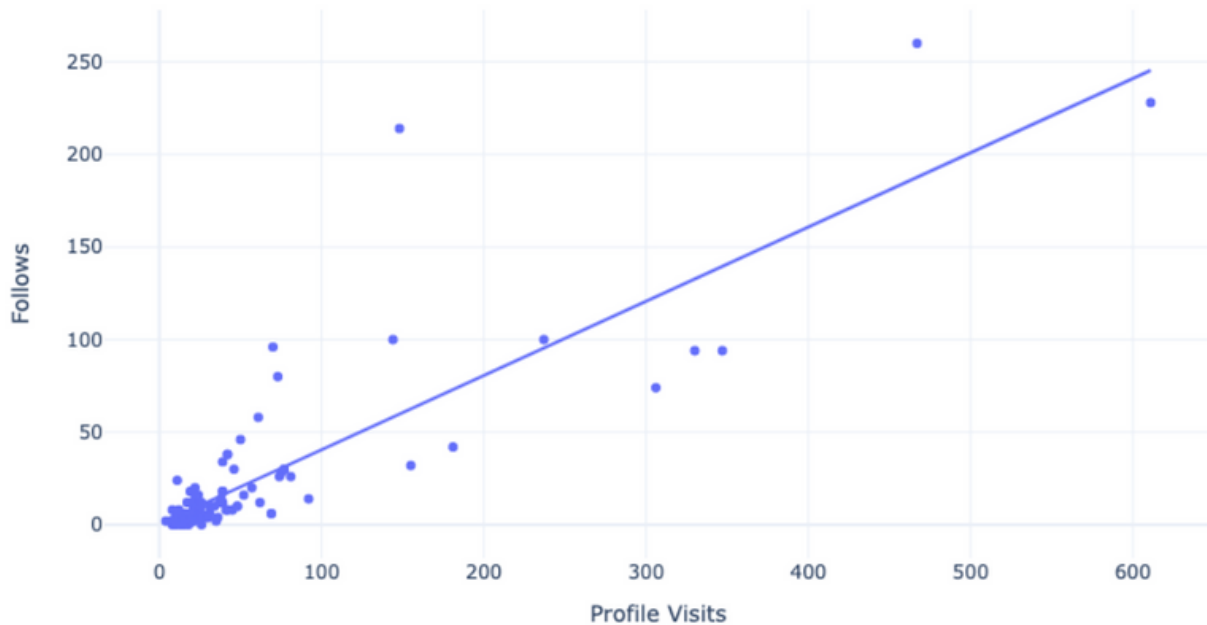


Artificial Intelligence Community of Pakistan

--Evolution to Revolution--

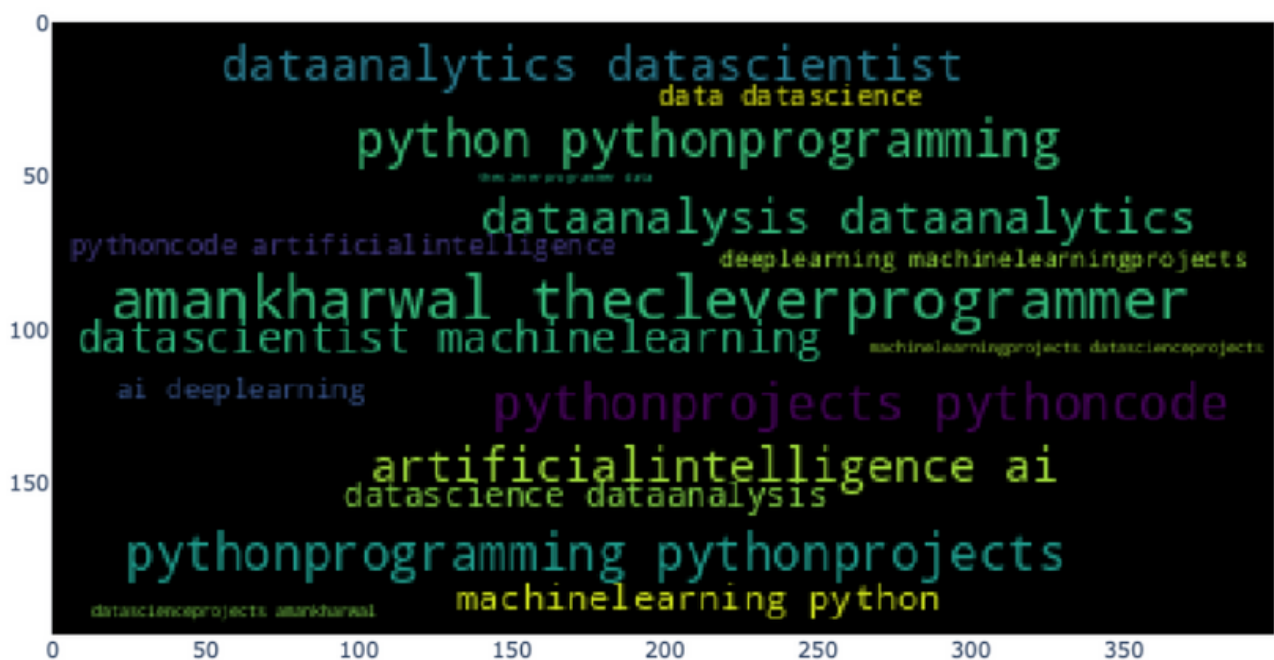
Q.9: Have a look at the relationship between the number of profile visits and follows as shown below

Profile Visits vs. Follows



Q.10: Have a look at the type of hashtags used in the posts using a wordcloud as shown below

Hashtags Word Cloud



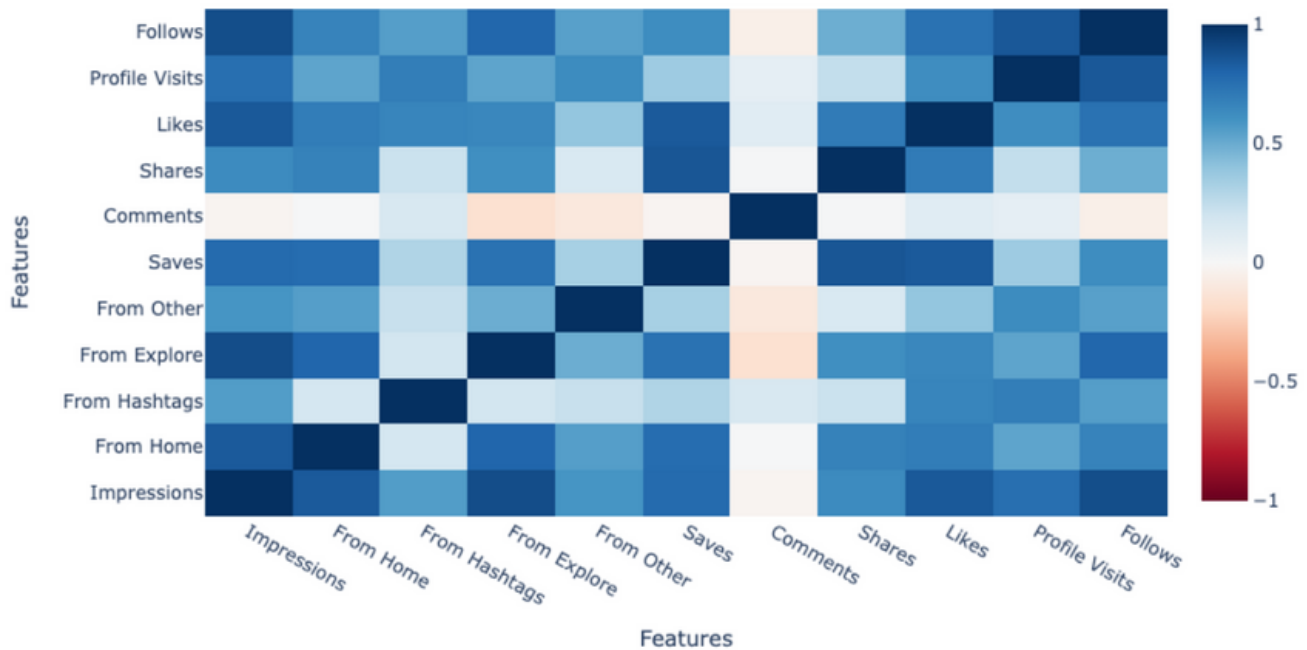


Artificial Intelligence Community of Pakistan

--Evolution to Revolution--

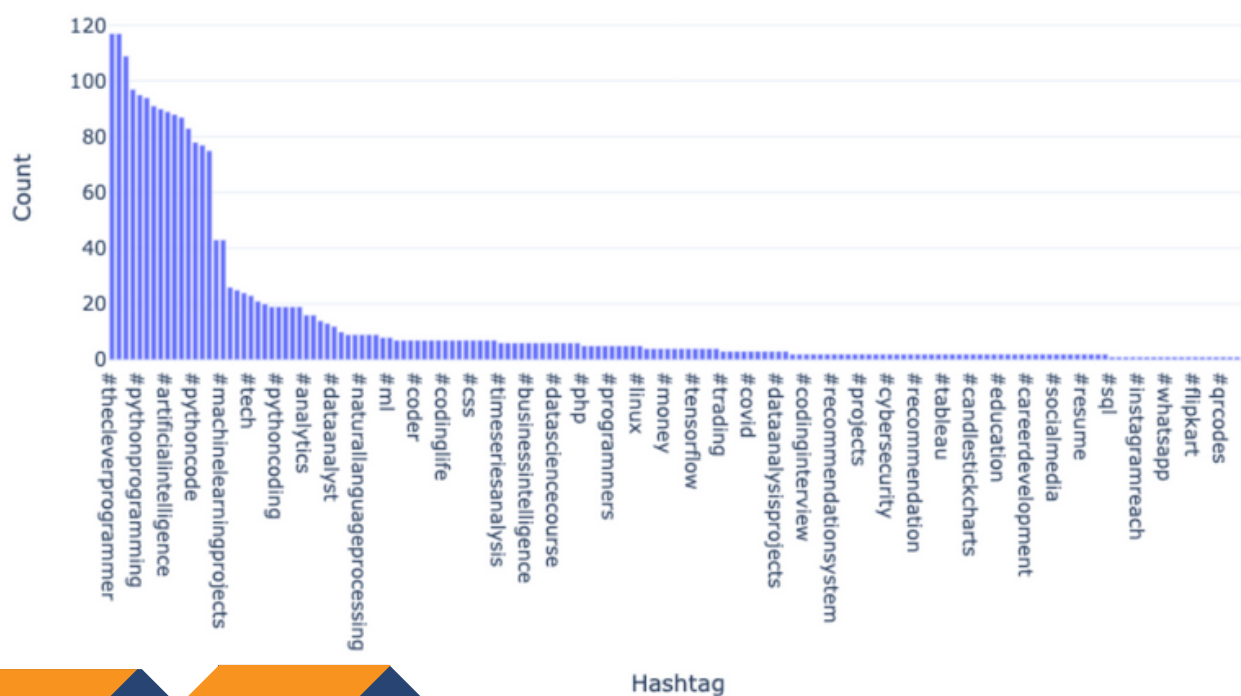
Q.11: Have a look at the correlation between all the features as shown below

Correlation Matrix



Q.12: Have a look at the distribution of hashtags to see which hashtag is used the most in all the posts as shown below

Distribution of Hashtags



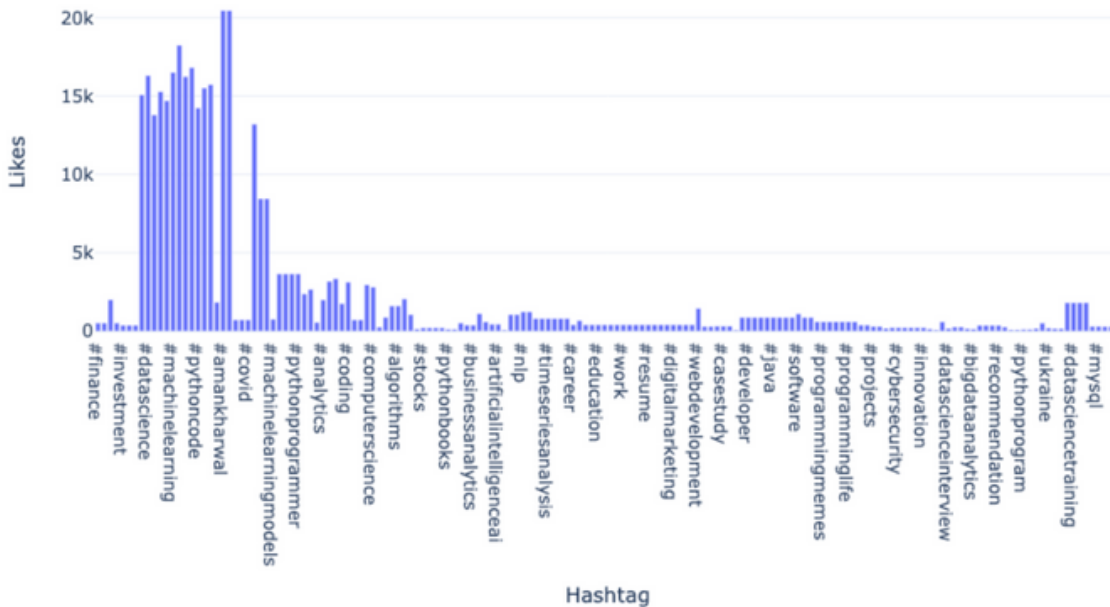


Artificial Intelligence Community of Pakistan

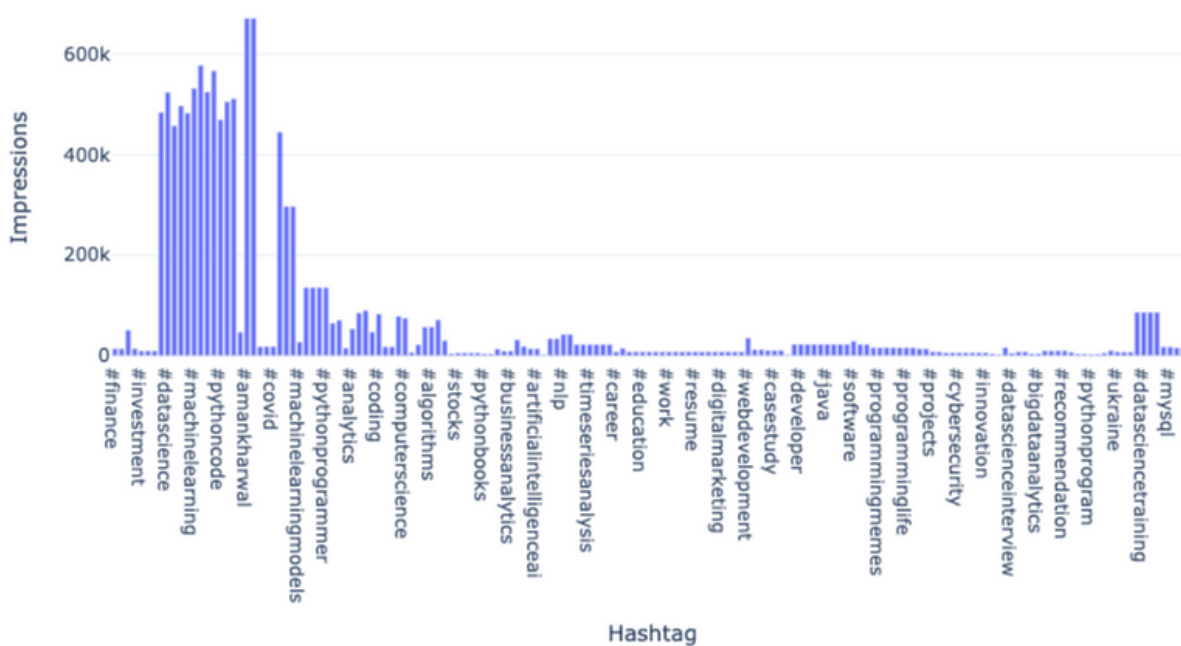
--Evolution to Revolution--

Q.13: Have a look at the distribution of likes and impressions received from the presence of each hashtag on the post as shown below

Likes Distribution for Each Hashtag



Impressions Distribution for Each Hashtag



Q.14: Write summary as per your observation