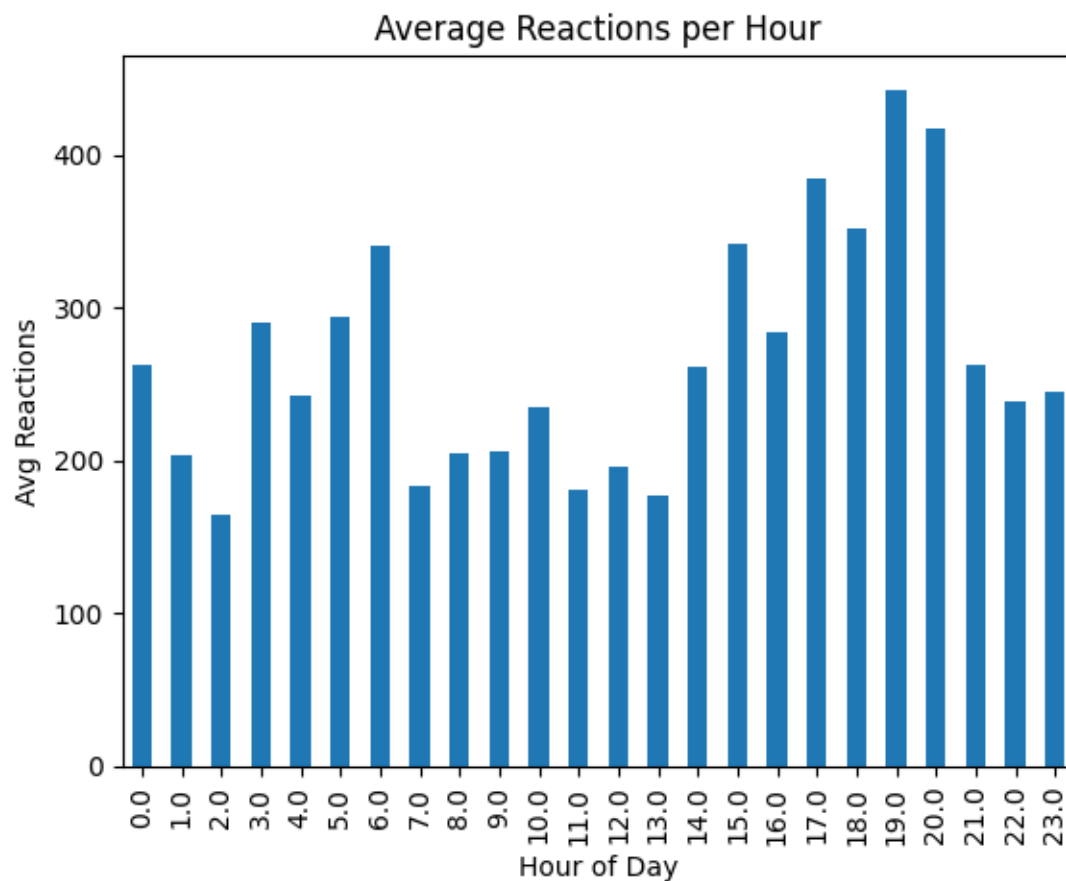Facebook Marketplace Dataset

Questions

1. How does the time of upload (status_published) affects the num_reaction?

=>

Average Reactions per Hour

2. Is there a correlation between the number of reactions (num_reactions) and other engagement metrics such as comments (num_comments) and shares (num_shares)? If so, what is the strength and direction of this correlation?

=>

```
                num_reactions   num_comments   num_shares
num_reactions       1.000000       0.150843     0.250723
num_comments        0.150843       1.000000     0.640637
num_shares          0.250723       0.640637     1.000000
```
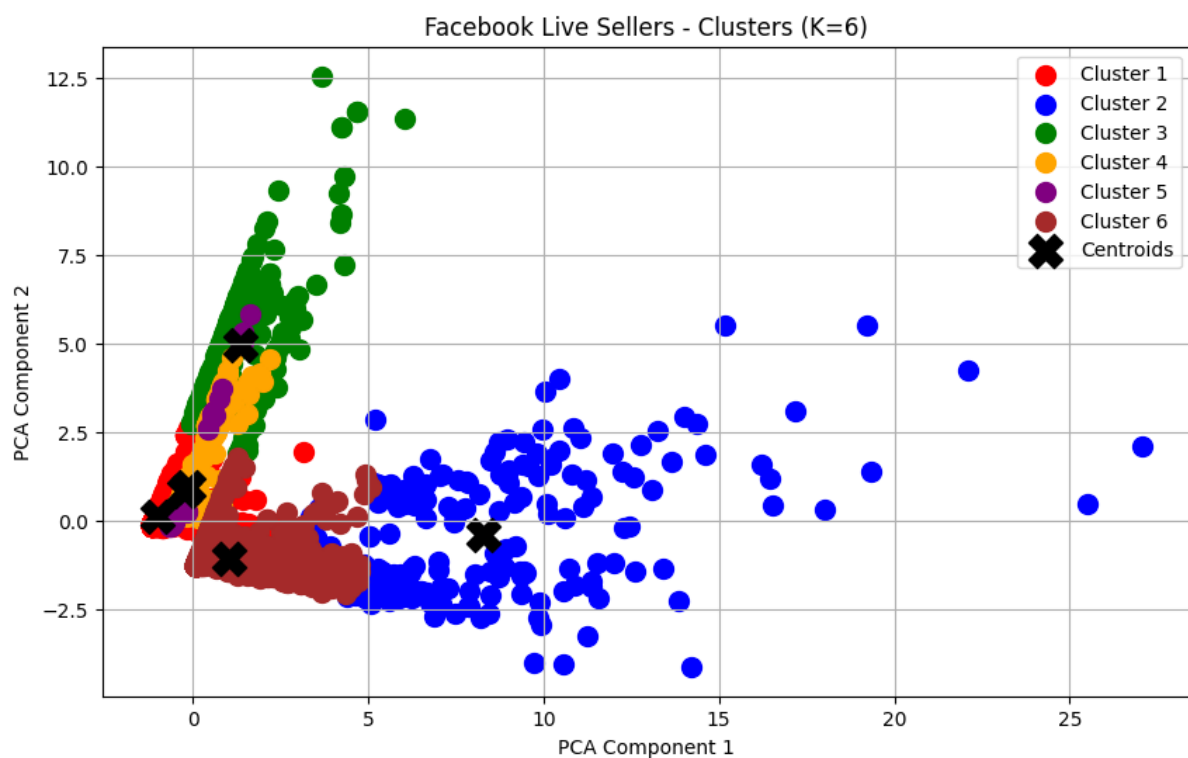
Yes, there is a correlation between num_reactions and other engagement metrics like num_comments and num_shares.
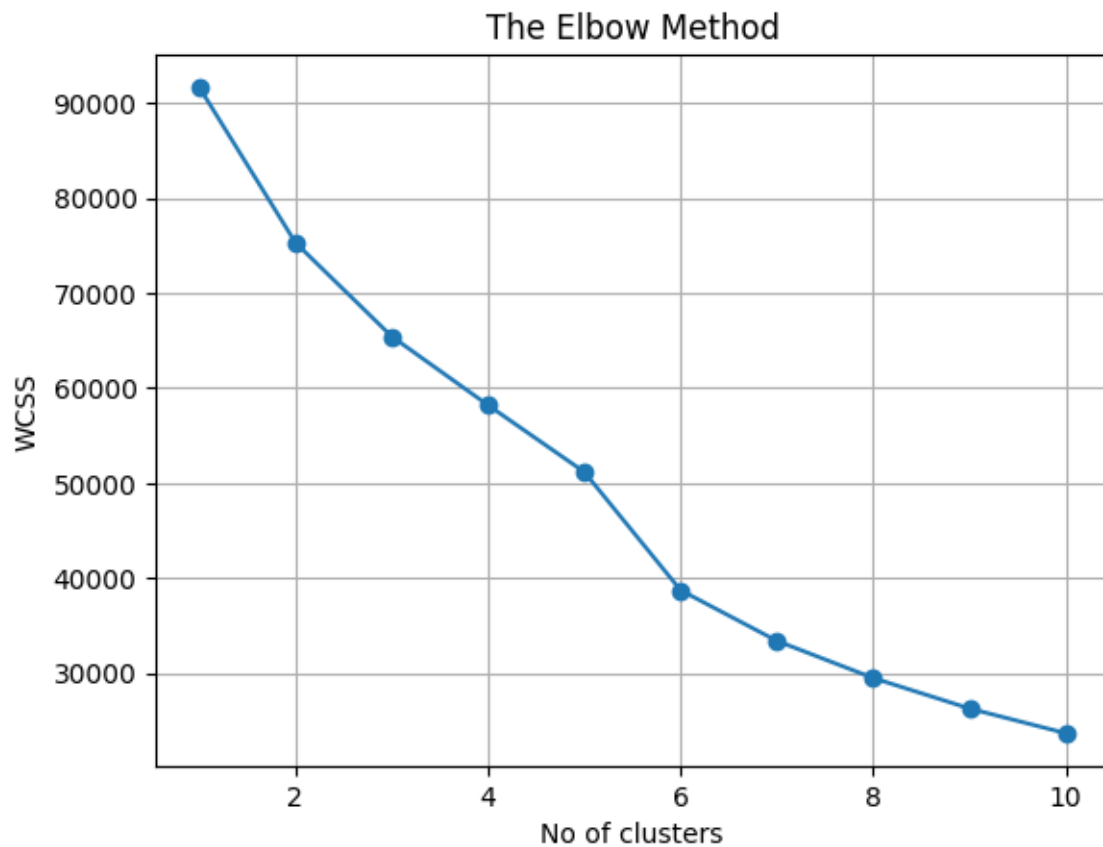
- The correlation between **reactions and comments** is weak (**r ≈ 0.15**)

- The correlation between **reactions and shares** is weak-to-moderate (**r ≈ 0.25**)

- The **strongest correlation** is between **comments and shares** (**r ≈ 0.64**), suggesting posts that generate discussions are more likely to be shared.

3. Use the columns status_type, num_reactions, num_comments, num_shares, num_likes, num_loves, num_wows, num_hahas, num_sads, and num_angrys to train a K-Means clustering model on the Facebook Live Sellers dataset.

=>

4. Use the elbow method to find the optimum number of clusters.

The Elbow Method

1 clusters: WCSS = 91649.99999999943

2 clusters: WCSS = 75295.01493442431

3 clusters: WCSS = 65423.1164000412

4 clusters: WCSS = 58198.4684828752

5 clusters: WCSS = 51167.45012130091

6 clusters: WCSS = 38734.129567133015

7 clusters: WCSS = 33401.968119760735

8 clusters: WCSS = 29499.397811254927

9 clusters: WCSS = 26290.919366923255

10 clusters: WCSS = 23689.685854319596

From **5 → 6**, there's a **huge drop of 12,433**, which is **significant**.
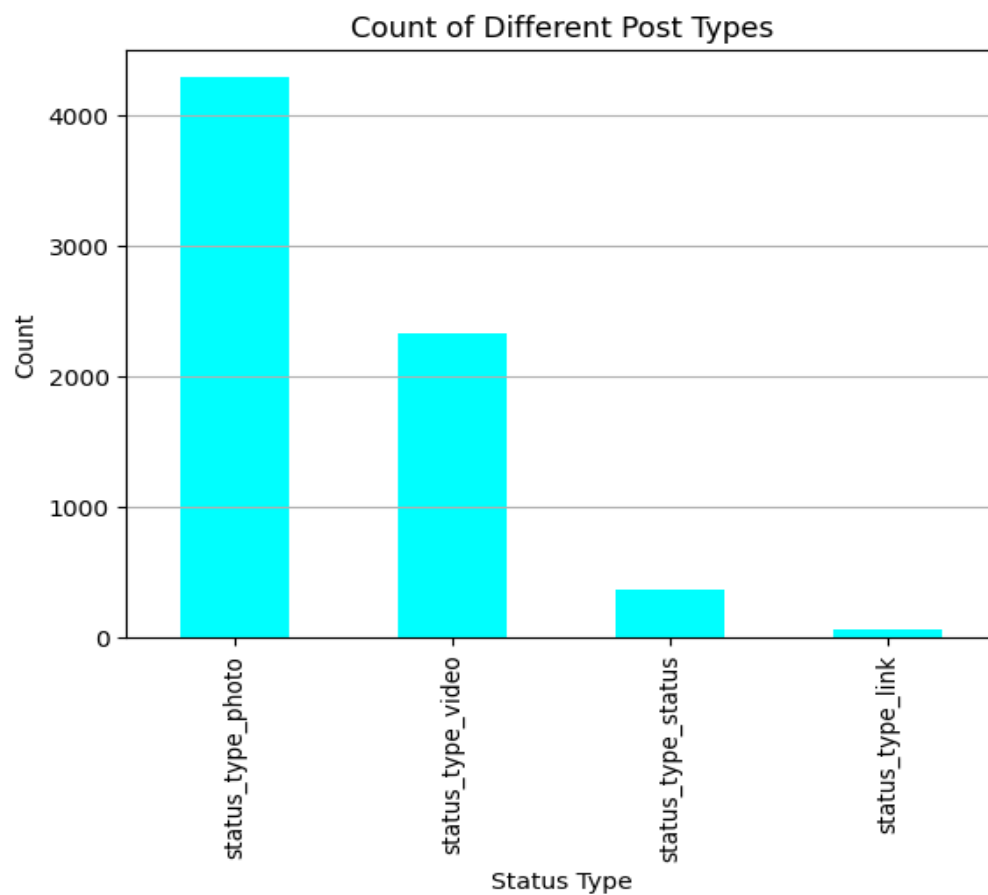
So taking 6 clusters

5. What is the count of different types of posts in the dataset?

status_type_photo   4288.0

status_type_video   2334.0

status_type_status   365.0

status_type_link   63.0

Count of Different Post Types

6. What is the average value of num_reaction, num_comments, num_shares for each post type?

Photo:

 num_reactions   181.290345

num_comments   15.993470

num_shares   2.553871

dtype: float64

Video:

 num_reactions   283.409597

num_comments   642.478149

num_shares   115.679949

dtype: float64

Status:

num_reactions   438.783562

num_comments   36.238356

num_shares   2.558904

dtype: float64

Link:

num_reactions   370.142857

num_comments   5.698413

num_shares   4.396825

dtype: float64

Average Reactions, Comments, and Shares per Post Type