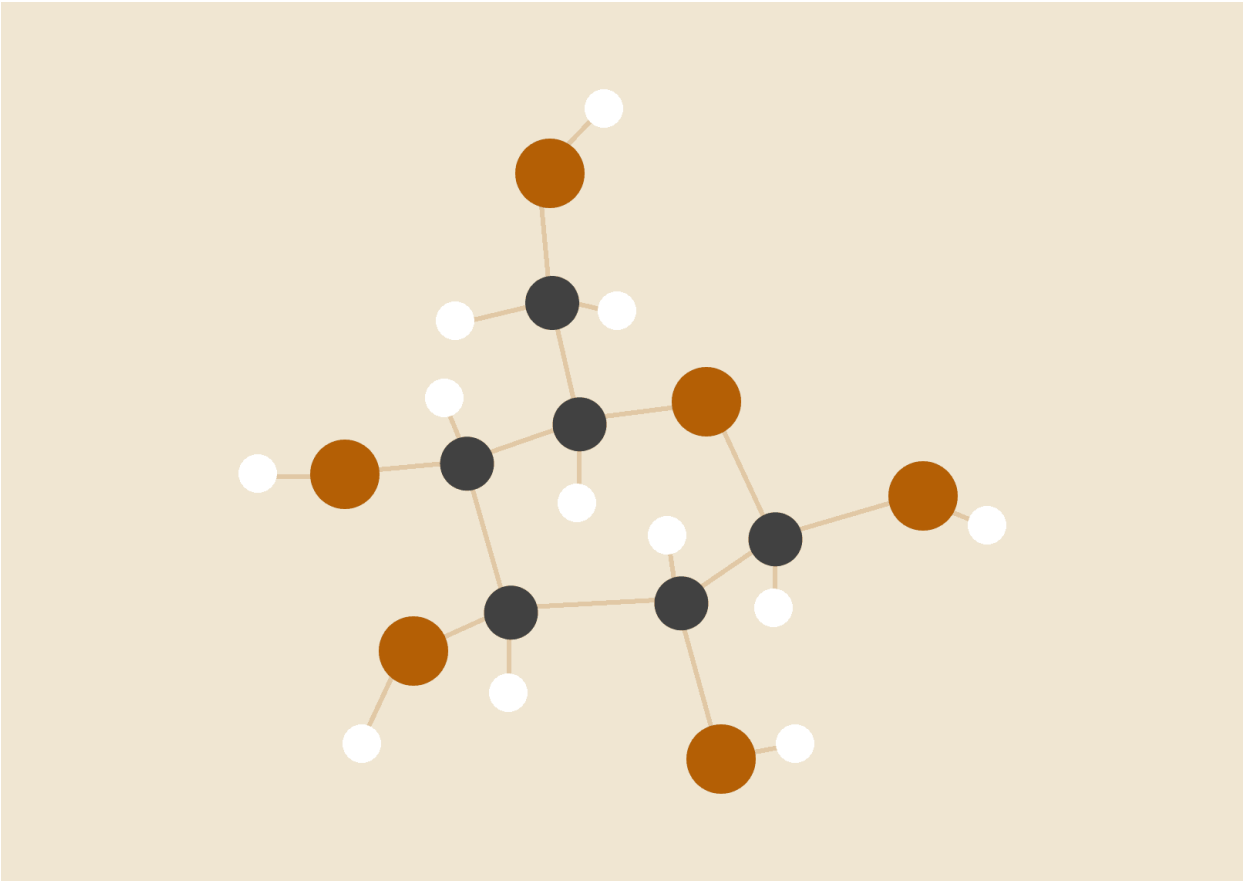
DATA ANALYST TASK

My Thriving Child - Interview Task



Maryam Zarin

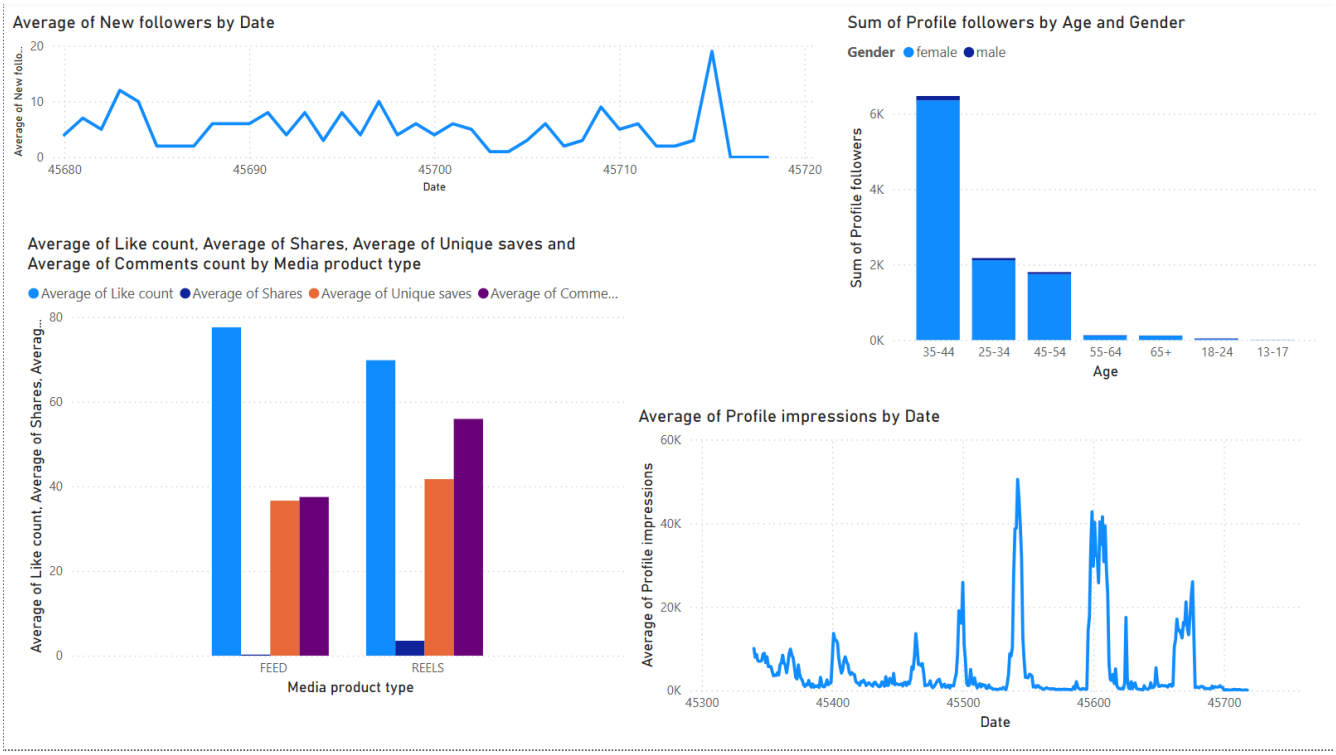
BSc Computer Science

github.com/maryamzarin

13th June 2025

Part 1: Instagram Performance Dashboard

- A dashboard that includes relevant charts/tables



- Insights on post engagement, reach, and effectiveness

Feed Posts seem to get more likes but reels get more comments and unique saves on average. We can also see the majority of the demographic consists of women aged 35-44.

1. What are some key metrics you would track to measure the performance of Facebook and Instagram posts?

Initially the views/ number of unique viewers/ people reached is a good indicator of reach. In terms of performance Likes, Shares and Saves as well as comments and users liking comments/ the most popular comments. Looking at the viewer to like ratio would be a good indicator of how popular a particular post is and to further this we can look at the demographic and see who the post is most popular with.

2. How would you determine if an Instagram Reel is performing well?

If an Instagram reel is getting lots of engagement, such as views, likes, shares, comments,

saves and others reading comments and liking, then the reel is performing well.

3. How would you A/B test different ad creatives on Facebook?

I would create the two ads and then split a demographic (for example Women aged 35-40) in half and show each half one of the ads and see how group A and group B both reacted to the ad. Were there more clicks for one ad in particular? If there are comments enabled what are they saying? If it's an ad shared as a post how many likes are they receiving?

4. What insights can you gain from the provided dataset?

We can plot engagement against time, and see our target demographic which is made up of the follow base.

5. What additional analyses would you recommend?

If we connect the engagement from posts to profile follows so we can see which posts brought more users to the profile and resulted in follows.

PART 2: Python Challenge

1. Data Cleaning

• Read the dataset into a Pandas DataFrame.

• Handle any missing or inconsistent data

```
In [116]: # 1. Data Cleaning, Read data into data frame and handle missing data

insta_xls = pd.ExcelFile("Copy of Instagram_Analytics - DO NOT DELETE (for interview purposes).xlsx")
profileOverviewDF = insta_xls.parse(0)
postEngageDF = insta_xls.parse(1)
ageGenderDF = insta_xls.parse(2)
topCitiesDF = insta_xls.parse(3)
postEngageDF
```

Out[116]:

	Date	Media ID	Media caption	Media product type	Media impressions	Media reach	Like count	Comments count	Shares	Unique saves	Video views	
0	2025-01-19	17902204887015249	Comment 'CLASS' to sign up to my free mastercl...	FEED	868	814	10	8	1	2	0	673eb16a83d01ab25bd18d6511109e6f720
1	2025-01-18	18071205646706820	Comment 'CLASS' to sign up to my free mastercl...	FEED	741	695	8	2	0	1	0	55102c6de8e1a68e015c44e3c8a600d0608
2	2025-01-17	18027615518313508	Comment 'CLASS' to sign up to my free mastercl...	FEED	1821	1666	24	14	0	13	0	68123274789468c3beb60c484c03735b62f
3	2025-01-16	18034591967437473	Comment 'CLASS' to sign up to my free mastercl...	REELS	0	1280	13	27	1	5	0	905354b956819e1cf3738a4541695cc1659
4	2025-01-15	18012211340694193	Comment 'CLASS' to sign up to my free mastercl...	REELS	0	753	6	8	1	2	0	48a803f148ba92cef3d43ac7944ce9bb44f
...	...	...	...	...	...	...	...	...	...	...	...	
56	2024-08-09	18347817988138734	Tears ran down my face as I dropped my son off...	FEED	2040	1910	32	104	0	13	0	1e619a202f708d84dd4a2a150b89b4600d4
57	2024-08-08	18050919739826668	If you snap at your child because you're stres...	FEED	2989	2831	157	2	0	78	0	04bc7ef29acf5ea0e1f8e35d4cfa6130df5

2. Basic Analysis

• Calculate the average engagement rate for Instagram posts.

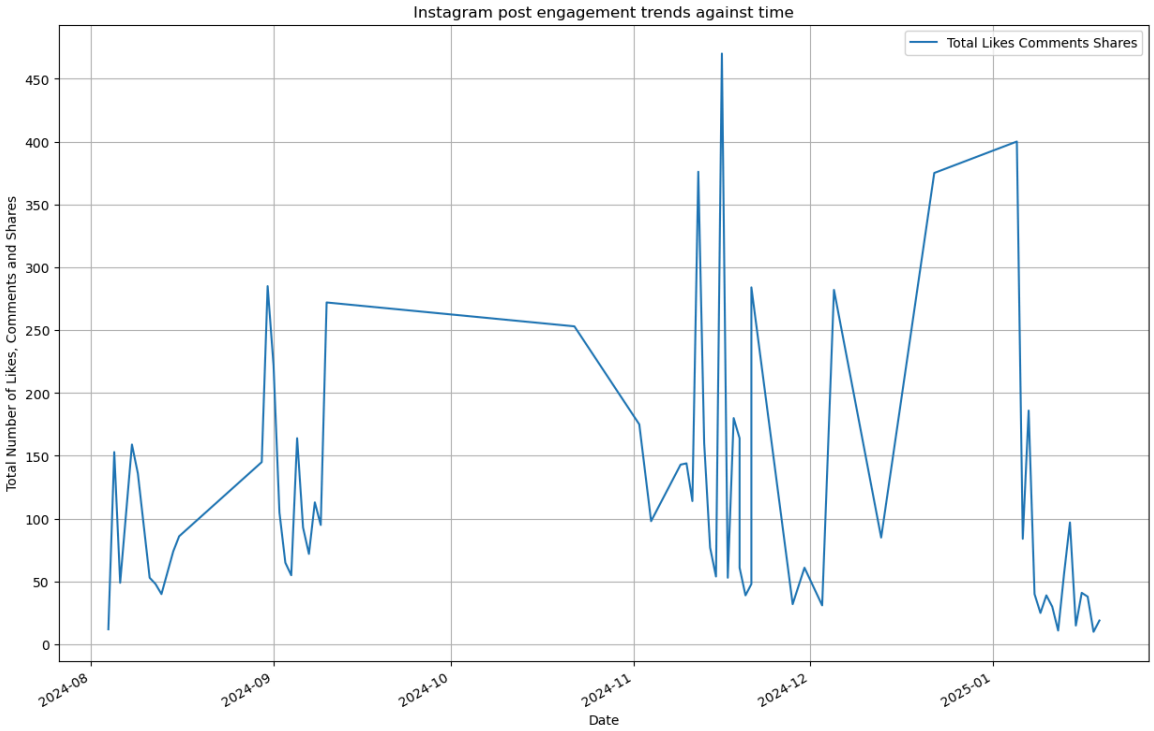
```
In [26]: # 2. Basic Analysis
         # Instagram Post
         # average engagement rate for Instagram posts.

         postEngageDF["Total Engagement"] = postEngageDF["Media impressions"] +postEngageDF["Media reach"]+ postEngageDF["Like count"]+pos
         postEngageDF["Total Likes Comments Shares"] = postEngageDF["Like count"]+postEngageDF["Comments count"]+postEngageDF["Shares"]
         avLike = round(postEngageDF.loc[:, "Like count"].mean(), 1)
         avComment = round(postEngageDF.loc[:, "Comments count"].mean(), 1)
         avShares = round(postEngageDF.loc[:, "Shares"].mean(), 1)

         print("Average number of likes per post is",avLike,". Average number of comments per post is",avComment,". Average number of shar

         Average number of likes per post is 74.6 . Average number of comments per post is 44.4 . Average number of shares per post is
         1.4 .
```

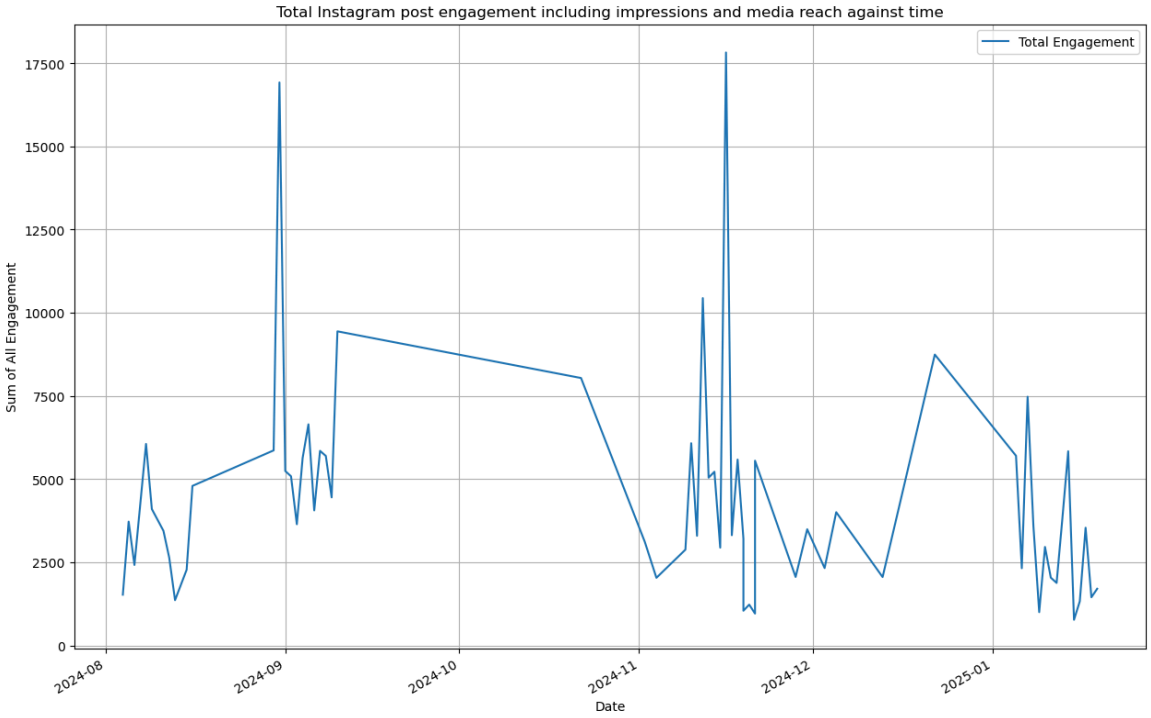
• Identify the top-performing post based on engagement (likes, comments, shares).

```
In [30]: # Top Performing post based on engagement(likes, comments, shares).

         avShares = postEngageDF.loc[:, "Total Likes Comments Shares"].max()

         postEngageDF.loc[postEngageDF["Total Likes Comments Shares"].idxmax()]

Out[30]: Date                                                    2024-11-16 00:00:00
         Media ID                                                  18077447254583656
         Media caption                 8 reminders you need to hear today as a divorc...
         Media product type                                                     FEED
         Media impressions                                                      9096
         Media reach                                                            8063
         Like count                                                              430
         Comments count                                                           40
         Shares                                                                    0
         Unique saves                                                            197
         Video views                                                               0
         RowHash                       49d7c44fe51fcba432d743f096c8b1e73494959e48e49d...
         Total Engagement                                                      17826
         Total Likes Comments Shares                                             470
         Name: 28, dtype: object
```

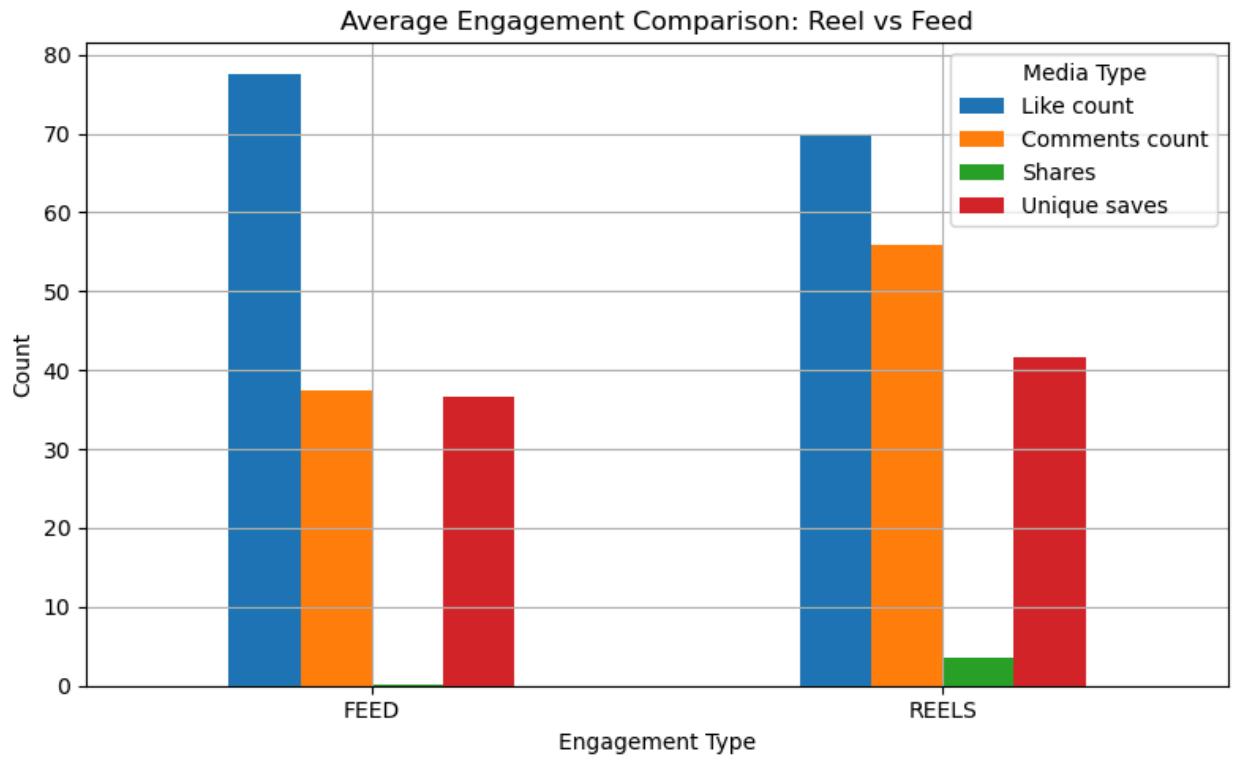
3. Visualization

• Create a simple line chart showing post engagement trends over time.





• Plot a bar chart comparing different post types (Reels vs. Static Posts).



4. Bonus Task (Optional)

• Write a function that predicts whether a post will perform well based on previous engagement data (e.g., using a simple threshold model).

Please submit your Python code in a Jupyter Notebook (.ipynb) or Python script (.py).