In my project I decided to examine my Instagram data. My initial project idea was to learn how frequently I interact with posts about the same topic. I was going to use the tags of posts for deciding the topic. Thent requested my data from the instagram account. I specifically asked for my liked and saved posts, user data, connections etc. However when I received the data, I realized it didn't contain any information about tags. Then I started to look for another research question. That's when I came up with my current project. Actually that was something I thought about some time ago, but it was something I don't know how to check/ look at. Some time ago I realized that when I start to follow an Instagram account, I frequently see and interact with their posts very frequently, but after some time passes I see that user's post less often. Since now I know how to use data processing tools, and able to analyze it I decided to work on this question of mine.

Let me explain the process in more detail. As I said before I requested the data from Instagram which didn't require much work. I specifically ask the data to be in JSON format because that was something I was familiar with. For my project I used two data sets. The first one was the posts that I like and the second was the users that I followed. I first read my JSON file into a DataFrame, after some processing I acquired a df that looked like this.

	user	timestamp	date
0	makeuprabbit	1606772058	2020-11-30
1	the_female_lead	1606771889	2020-11-30
2	the_frohaven	1606771879	2020-11-30
3	psychicmedium.trish	1606771744	2020-11-30
4	farquaadhater	1606771450	2020-11-30

And from this df I add some other columns that I will use later. Year-month column contains the date without the day information, and post_count contains the total number of liked posts from each user.

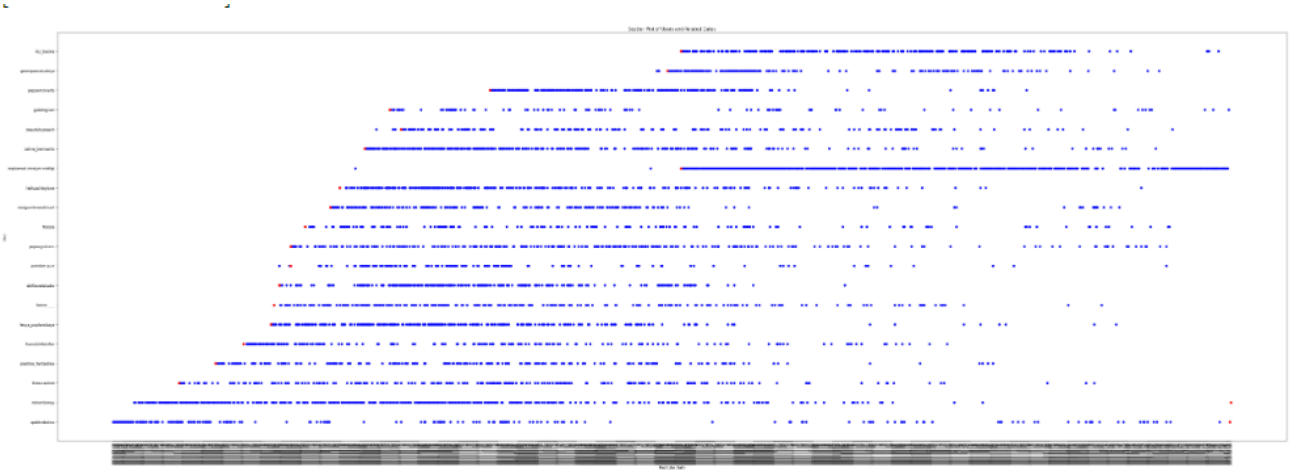
	user	timestamp	date	year-month	post_count
0	makeuprabbit	1606772058	2020-11-30	2020-11	1
1	the_female_lead	1606771889	2020-11-30	2020-11	21
2	the_frohaven	1606771879	2020-11-30	2020-11	1
3	psychicmedium.trish	1606771744	2020-11-30	2020-11	3
4	farquaadhater	1606771450	2020-11-30	2020-11	2

Let's continue with my following data set. Again I read it into a df. This time the date columns contains the date I started the follow that user.

	user	timestamp	date
0	selenay.buse	1705672981	2024-01-19
1	pierrot_lou	1705490392	2024-01-17
2	mastertingus	1704622298	2024-01-07
3	arozear	1704489669	2024-01-05
4	ahsenbucek	1704465984	2024-01-05

Now with the two df I can test my hypothesis. For this I gather the users which have the top 20 most liked posts by me. The reason why I chose to use the top then was to have enough data for my hypothesis. My initial instagram data contained information from the past six years, to understand whether my interaction changed over time I needed enough interaction that roughly matches the timespan.

Then I create a scatter plot of it.



Here the scatter plot looks kinda small, so I will put a bigger version of it at the and pages. Still the science is small, it is easier to see where that creates lines and where it scatters. In this plot the x axis denotes the date of likes and the y axis denotes the top then liked users by me. Blue points are when each post is liked and the red points are when I started to follow that account. As it can be seen after I started following the account the dots became very dense. They even create continuous lines, and after some time they become really scattered which proves my point.

What can be added to this project is that maybe other variables can add to analysis that are related to interactions with the followers. Such as saved posts. Or maybe I can turn the visualized data into a numerical one, that will further agree with my point.