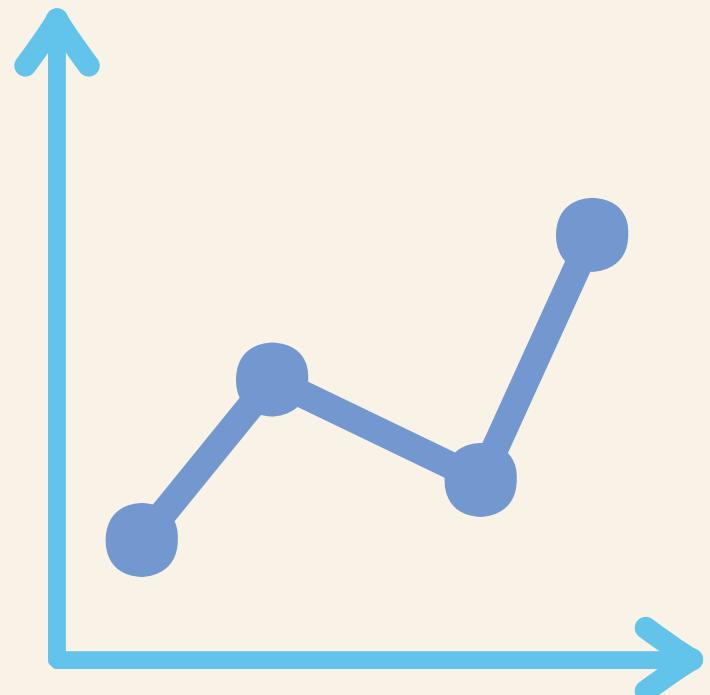


DSA210 PROJECT



Berinay Zümra
Sariel

An Insight into
My Instagram
Habits and Study
Patterns

Outline

01 - Introduction and Null Hypotheses

02 - Data Collections and Evaluations

03 - Data Visualizations

04 - Hypothesis Testing & Conclusions

05 - Further Attempts

03.1 Forest App

03.2 Instagram

03.3 Analogies

01. Introduction and Hypothesis

- In this project, I am exploring my studying habits and my social media usage by first exploring them individually, and then combining them in order to come up with an outcome.
- The social media platform is chosen as Instagram, since it is mostly what I use for procrastination.

Null Hypotheses (H_0):

1. Hourly Correlation: There is no monotonic relationship between hourly study hours and Instagram usage, especially in last 3 semesters.
2. Daily Correlation: There is no monotonic relationship between daily study hours and Instagram usage, in last 3 semesters.

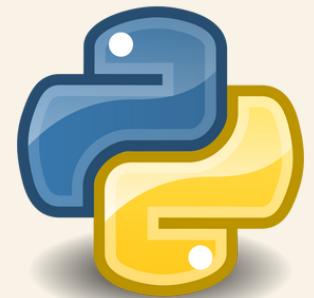


Note:

The reason I am focusing on last three semesters is that:

- I have more accurate recordings in last 3 semesters, since I have become more responsible and busier.
- I did not want to include my intellectual hobbies that I spent my time during prep school or the time I spent on my part time jobs, since I want to focus on my academics more.

02 - Data Collections and Evaluations



where and how?

For my Forest Data, I have requested to export my data from the app, which was pretty easy and fast to access. I have imported in .csv format.

For Instagram Data, I have requested from the app again, in .json format, which took longer time.

I have preprocessed them in pyCharm, made them anonymized and ready-to-use.

liked_posts.json

```

110
111 ]
112 },
113 {
114   "title": "cattos.gattos",
115   "string_list_data": [
116     {
117       "href": "https://www.instagram.com/reel/DCUlrrRIi6l/",
118       "value": "\u00f0\u009f\u0091\u008d",
119       "timestamp": 1732005524
120     }
121   ],
122 },
123 {
124   "title": "sabancikedy",
125   "string_list_data": [
126     {
127       "href": "https://www.instagram.com/p/DCiBij4NAJL/",
128       "value": "\u00f0\u009f\u0091\u008d",
129       "timestamp": 1732005501
130     }
131   ],
132 },
133 {
134   "title": "sabancikedy",
135   "string_list_data": [
136     {
137       "href": "https://www.instagram.com/p/DCiBQJ2NiMC/",
138       "value": "\u00f0\u009f\u0091\u008d",
139       "timestamp": 1732005312
140     }
141   ]

```

extracted_likes_time	
timestamp	
2022-09-01	07:42:01
2022-09-01	12:39:30
2022-09-01	18:35:56
2022-09-01	20:08:36
2022-09-01	20:17:46
2022-09-02	08:12:25
2022-09-02	13:23:20
2022-09-02	15:58:25
2022-09-03	12:01:51
2022-09-03	14:15:07
2022-09-03	14:51:10
2022-09-03	14:58:48
2022-09-03	16:47:30
2022-09-03	16:47:56
2022-09-03	16:48:51
2022-09-04	06:52:26
2022-09-04	11:14:11
2022-09-04	18:44:49
2022-09-05	10:42:19
2022-09-06	04:29:49
2022-09-06	04:53:51
2022-09-06	07:52:13
2022-09-06	08:28:24
2022-09-06	11:17:28
2022-09-06	11:43:36

2024-10-29T15:43:54.339+0300	2024-10-29T16:43:54.339+0300	MATH204	Tangerine Tree	TRUE
2024-10-29T21:37:44.563+0300	2024-10-29T23:07:44.563+0300	CS204	Spooky Tree	TRUE
2024-10-29T23:29:00.034+0300	2024-10-29T23:59:00.034+0300	CS204	Mooncake Tree	TRUE
2024-10-30T00:26:38.607+0300	2024-10-30T00:56:38.607+0300	CS204	Mooncake Tree	TRUE
2024-10-30T11:25:43.388+0300	2024-10-30T12:25:43.388+0300	CS204	Mooncake Tree	TRUE
2024-10-30T15:43:48.365+0300	2024-10-30T17:43:48.365+0300	CS204	Narcissus	TRUE
2024-10-30T19:07:04.553+0300	2024-10-30T19:34:35.877+0300	CS204	Lavender	TRUE
2024-10-30T20:05:11.673+0300	2024-10-30T22:05:11.673+0300	HUM201	Jacaranda	TRUE
2024-10-31T20:40:00.774+0300	2024-10-31T22:40:00.774+0300	ECON202	Osmanthus	TRUE
2024-11-01T11:02:37.308+0300	2024-11-01T12:02:37.308+0300	MATH204	Tangerine Tree	TRUE
2024-11-01T12:11:43.263+0300	2024-11-01T12:41:43.263+0300	MATH204	Tangerine Tree	TRUE
2024-11-01T14:31:08.097+0300	2024-11-01T16:31:08.097+0300	MATH204	Tangerine Tree	TRUE
2024-11-01T17:07:05.778+0300	2024-11-01T17:52:05.778+0300	M	2024-10-26 16:53:25.842000+03:00	2024-10-26 18:53:25.842000+03:00
2024-11-01T20:53:46.407+0300	2024-11-01T22:53:46.407+0300	M	2024-10-28 14:05:30.758000+03:00	2024-10-28 15:05:30.758000+03:00
2024-11-02T09:55:27.184+0300	2024-11-02T11:25:27.184+0300	M	2024-10-28 19:04:20.182000+03:00	2024-10-28 21:04:20.182000+03:00
2024-11-02T12:31:00.065+0300	2024-11-02T14:01:00.065+0300	M	2024-10-29 10:20:44.438000+03:00	2024-10-29 12:00:44.438000+03:00
2024-11-02T12:31:00.065+0300	2024-11-02T14:01:00.065+0300	M	2024-10-29 12:51:01.928000+03:00	2024-10-29 14:21:01.928000+03:00
2024-11-02T14:10:01.000+0300	2024-11-02T14:40:01.000+0300	M	2024-10-29 14:39:51.243000+03:00	2024-10-29 14:59:51.243000+03:00
2024-11-02T15:32:25.373+0300	2024-11-02T17:02:25.373+0300	E	2024-10-29 15:43:54.339000+03:00	2024-10-29 16:43:54.339000+03:00
2024-11-02T19:16:13.148+0300	2024-11-02T20:16:13.148+0300	M	2024-10-29 21:37:44.563000+03:00	2024-10-29 23:07:44.563000+03:00
2024-11-02T20:34:08.721+0300	2024-11-02T21:34:08.721+0300	M	2024-10-29 23:29:00.034000+03:00	2024-10-29 23:59:00.034000+03:00
2024-11-03T19:26:21.656+0300	2024-11-03T21:26:21.656+0300	E	2024-10-30 00:26:38.607000+03:00	2024-10-30 00:56:38.607000+03:00
2024-11-04T13:14:11.998+0300	2024-11-04T14:59:11.998+0300	C	2024-10-30 15:43:48.365000+03:00	2024-10-30 17:43:48.365000+03:00
2024-11-04T16:48:53.944+0300	2024-11-04T18:48:53.944+0300	C	2024-10-30 19:07:04.553000+03:00	2024-10-30 19:34:35.877000+03:00
2024-11-04T18:49:18.565+0300	2024-11-04T20:19:18.565+0300	C	2024-10-30 20:05:11.673000+03:00	2024-10-30 22:05:11.673000+03:00
2024-11-05T18:59:57.195+0300	2024-11-05T19:05:48.347+0300	D	2024-10-31 20:40:00.774000+03:00	2024-10-31 22:40:00.774000+03:00
2024-11-06T10:50:41.813+0300	2024-11-06T12:50:41.813+0300	D	2024-11-01 11:02:37.308000+03:00	2024-11-01 12:02:37.308000+03:00
2024-11-06T18:46:39.182+0300	2024-11-06T20:46:39.182+0300	D	2024-11-01 12:11:43.263000+03:00	2024-11-01 12:41:43.263000+03:00
			2024-11-01 14:31:08.097000+03:00	2024-11-01 16:31:08.097000+03:00
			2024-11-01 17:07:05.778000+03:00	2024-11-01 17:52:05.778000+03:00
			2024-11-01 20:53:46.407000+03:00	2024-11-01 22:53:46.407000+03:00
			2024-11-02 09:55:27.184000+03:00	2024-11-02 11:25:27.184000+03:00
			2024-11-02 12:31:00.065000+03:00	2024-11-02 14:01:00.065000+03:00
			2024-11-02 14:10:01+03:00	2024-11-02 14:40:01+03:00
			2024-11-02 15:32:25.373000+03:00	2024-11-02 17:02:25.373000+03:00
			2024-11-02 19:16:13.148000+03:00	2024-11-02 20:16:13.148000+03:00
			2024-11-02 20:34:08.721000+03:00	2024-11-02 21:34:08.721000+03:00
			2024-11-03 19:26:21.656000+03:00	2024-11-03 21:26:21.656000+03:00
			2024-11-04 13:14:11.998000+03:00	2024-11-04 14:59:11.998000+03:00

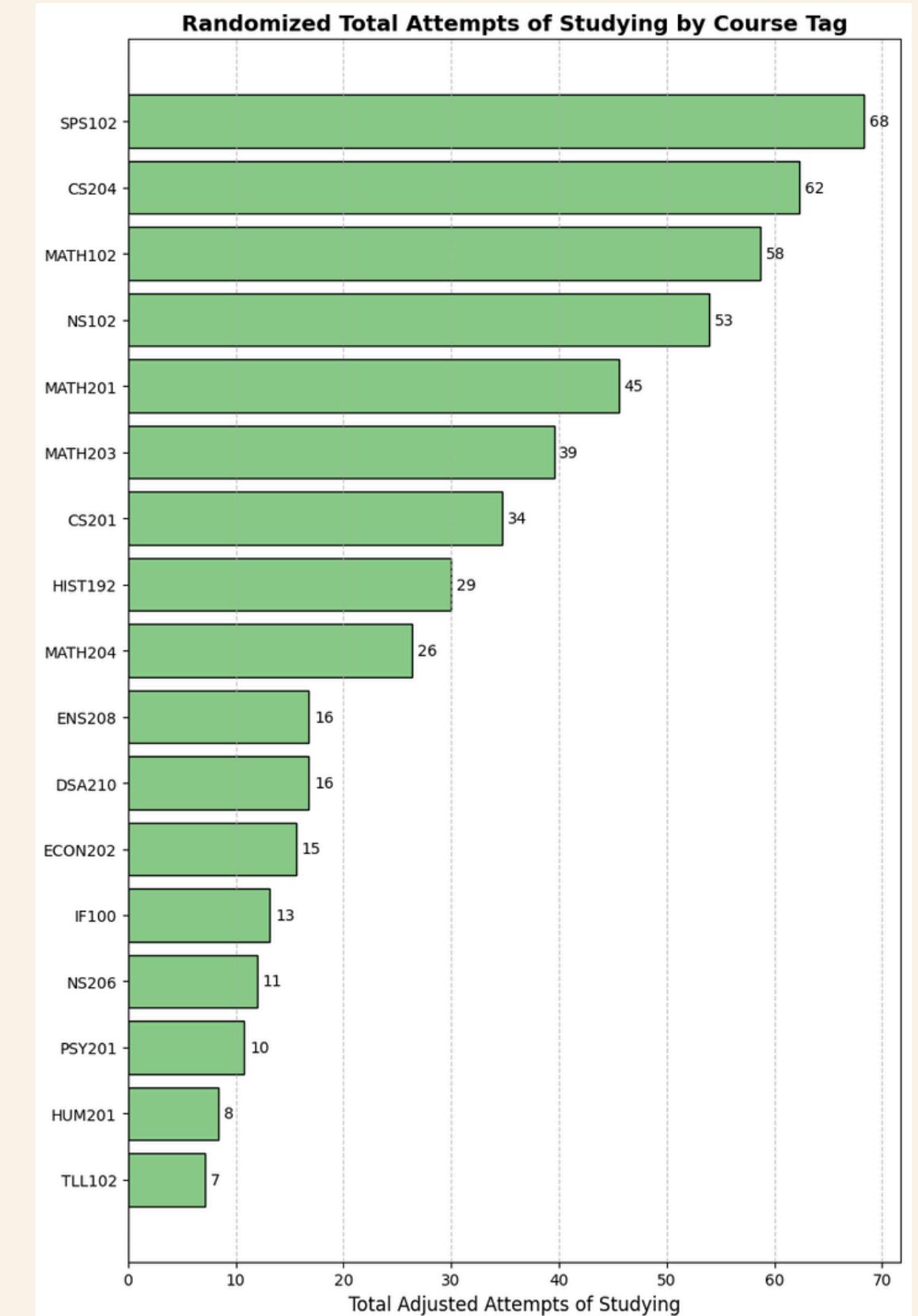
from this -> to that..



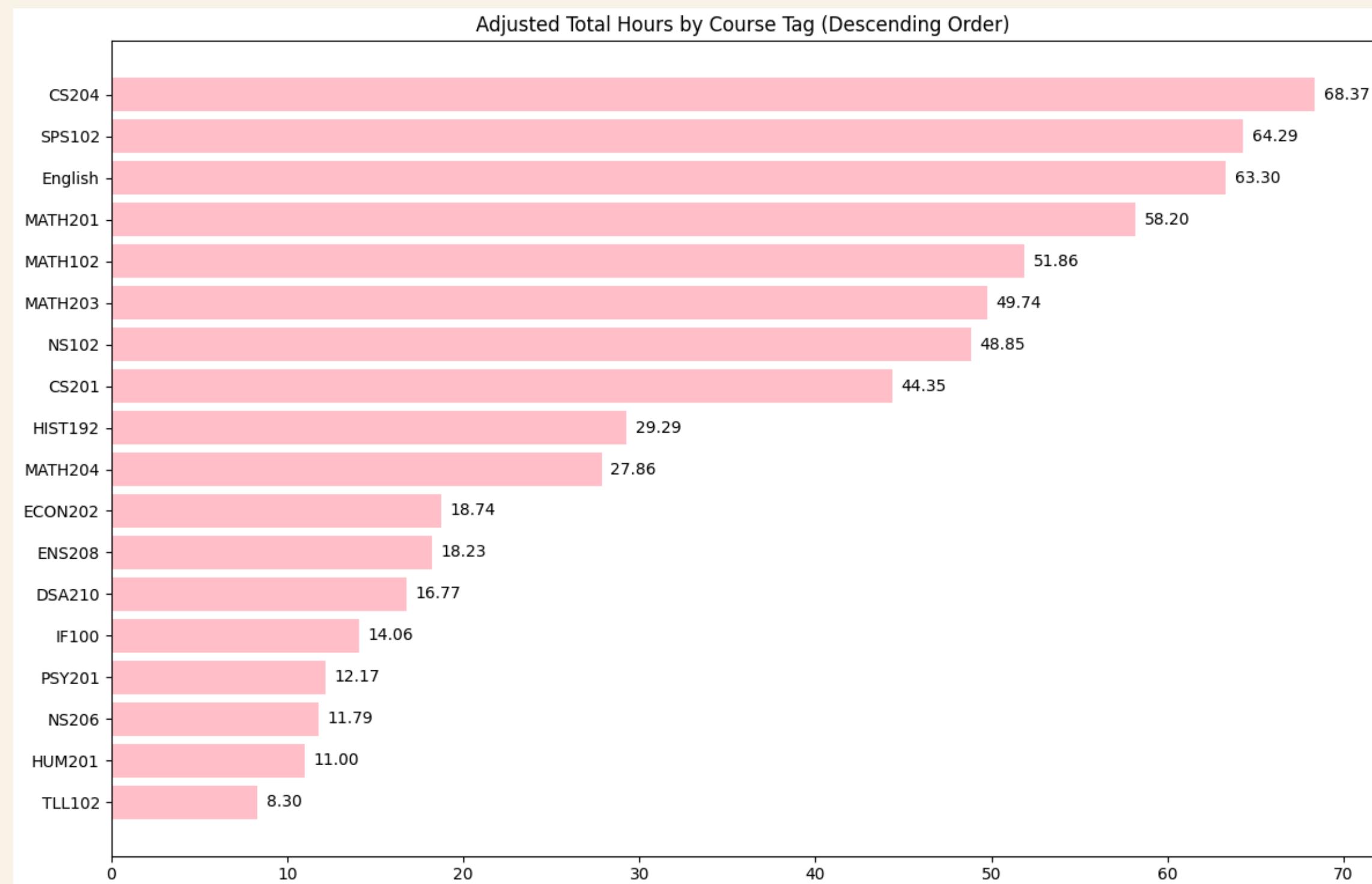
03 - Data Visualizations

03.1 Forest App Data

- I have firstly calculated my “study attempts” with randomization.
- By “study attempts” I mean the times where I have sat on the table just to study for that particular course.
- The reason why I used “randomization” was because of some badly denoted data in my set.



03.1 Forest App Data, cont.

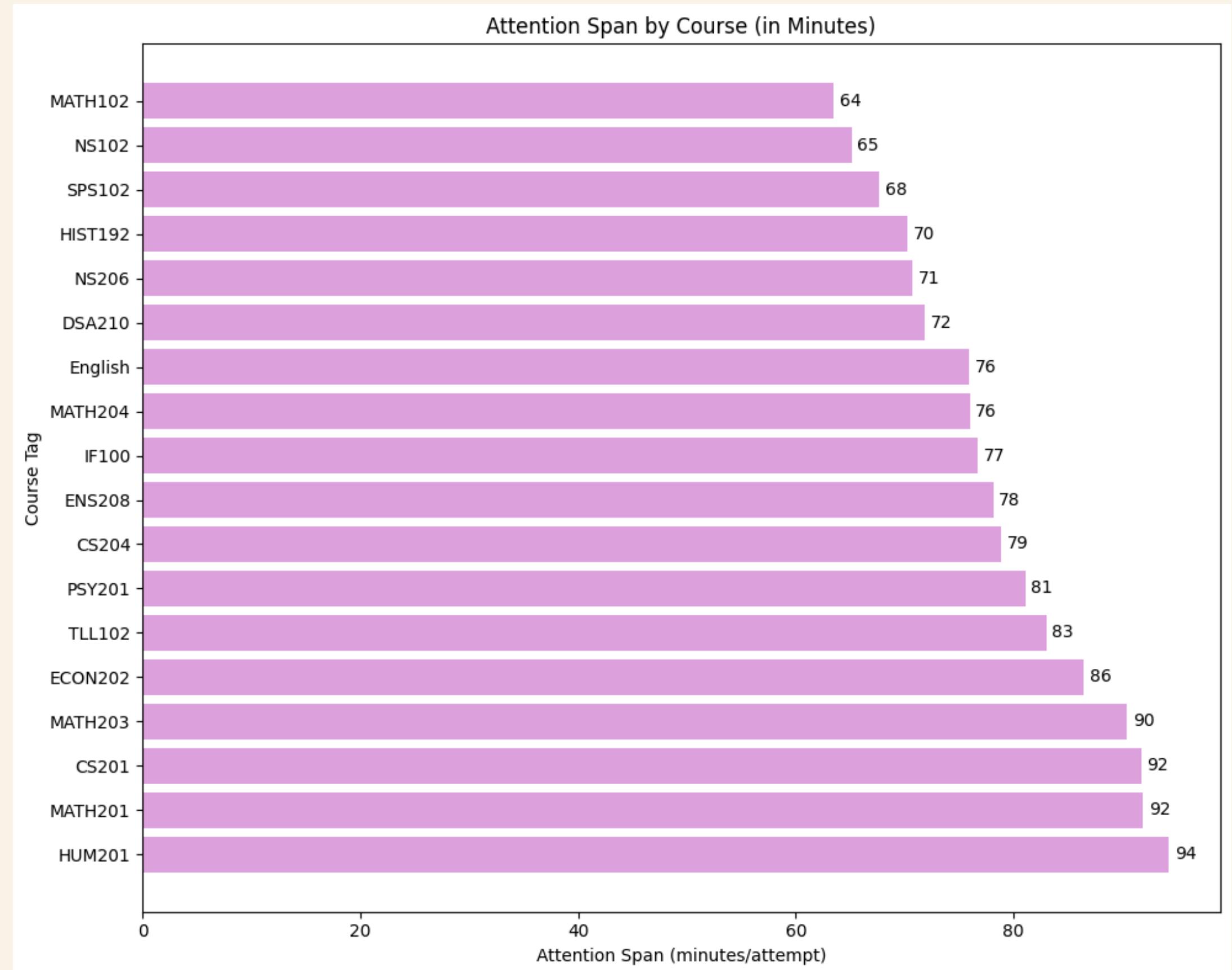


After that, I have calculated the total amount of time spent on each course, respectively.

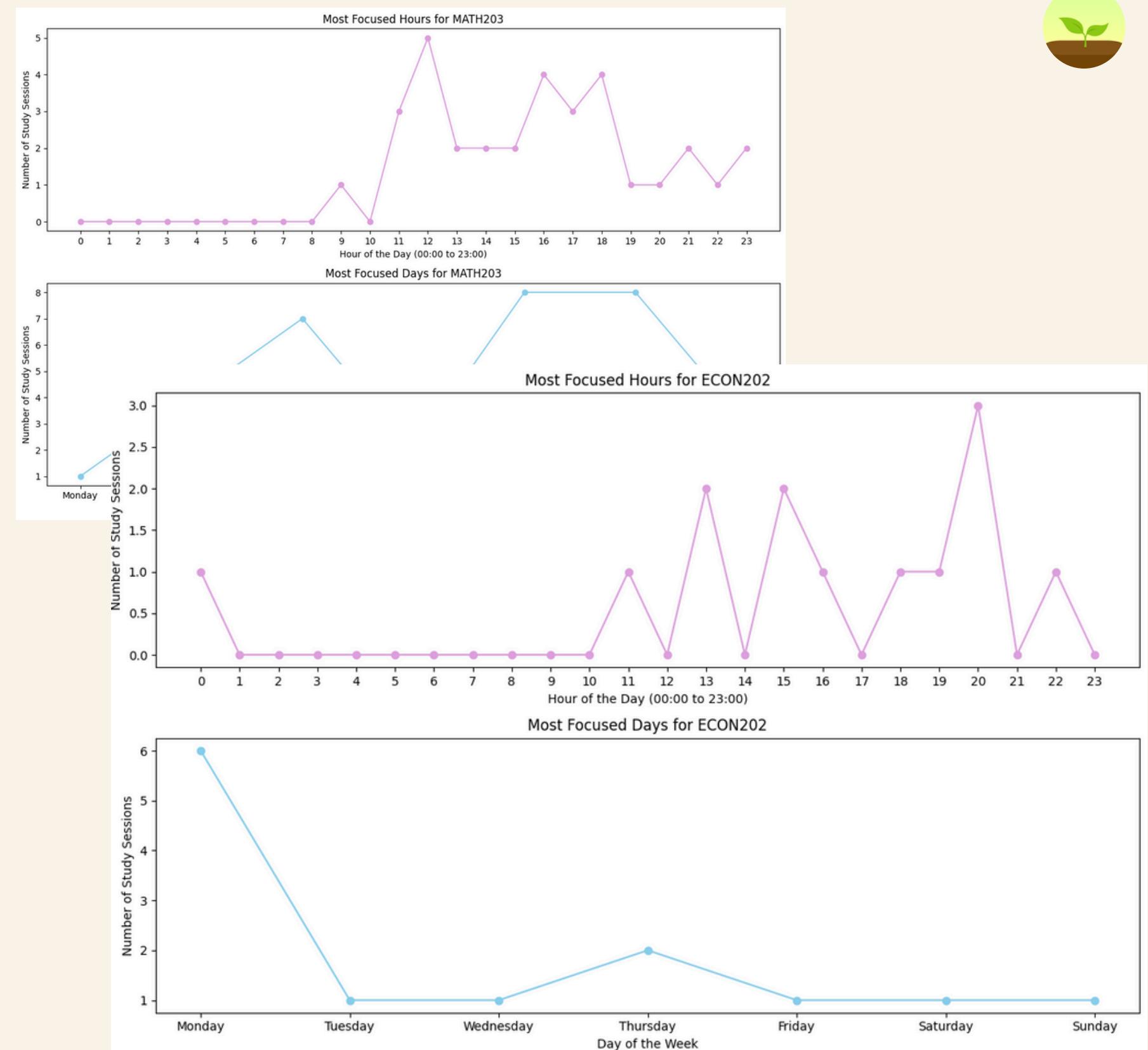
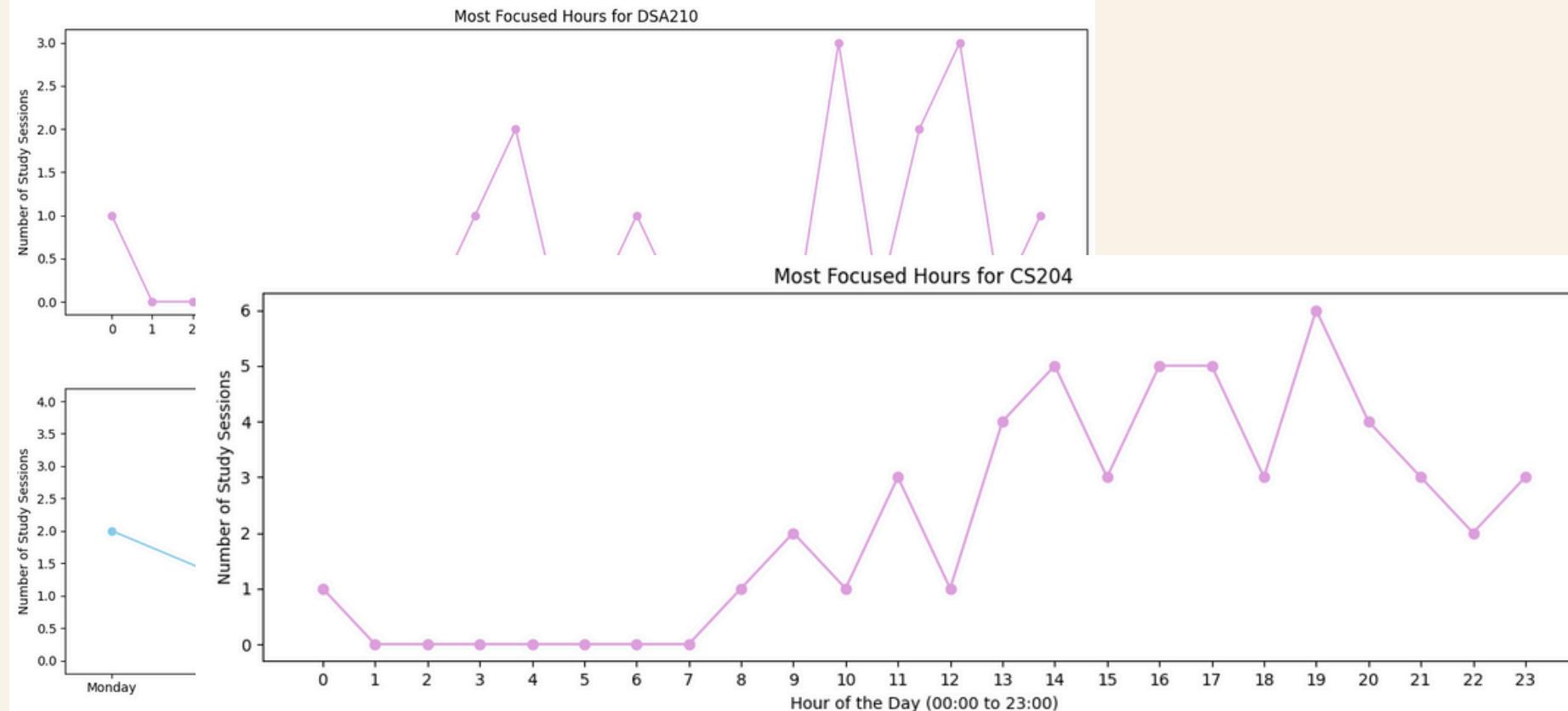
03.1 Forest App Data, cont.



After finding the number of times I have attempted to study and the aggregate time spent on each course, I thought it would be fun to explore my attention span.

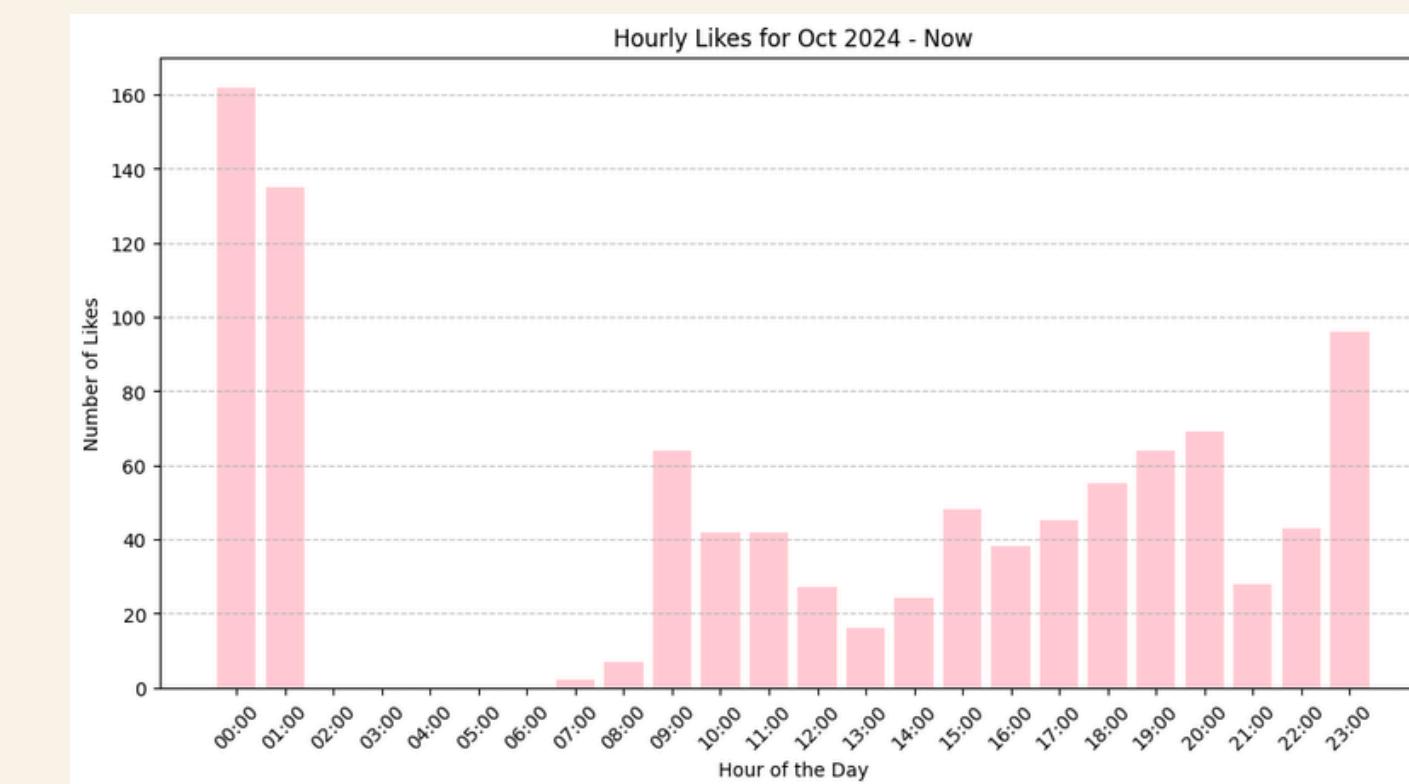
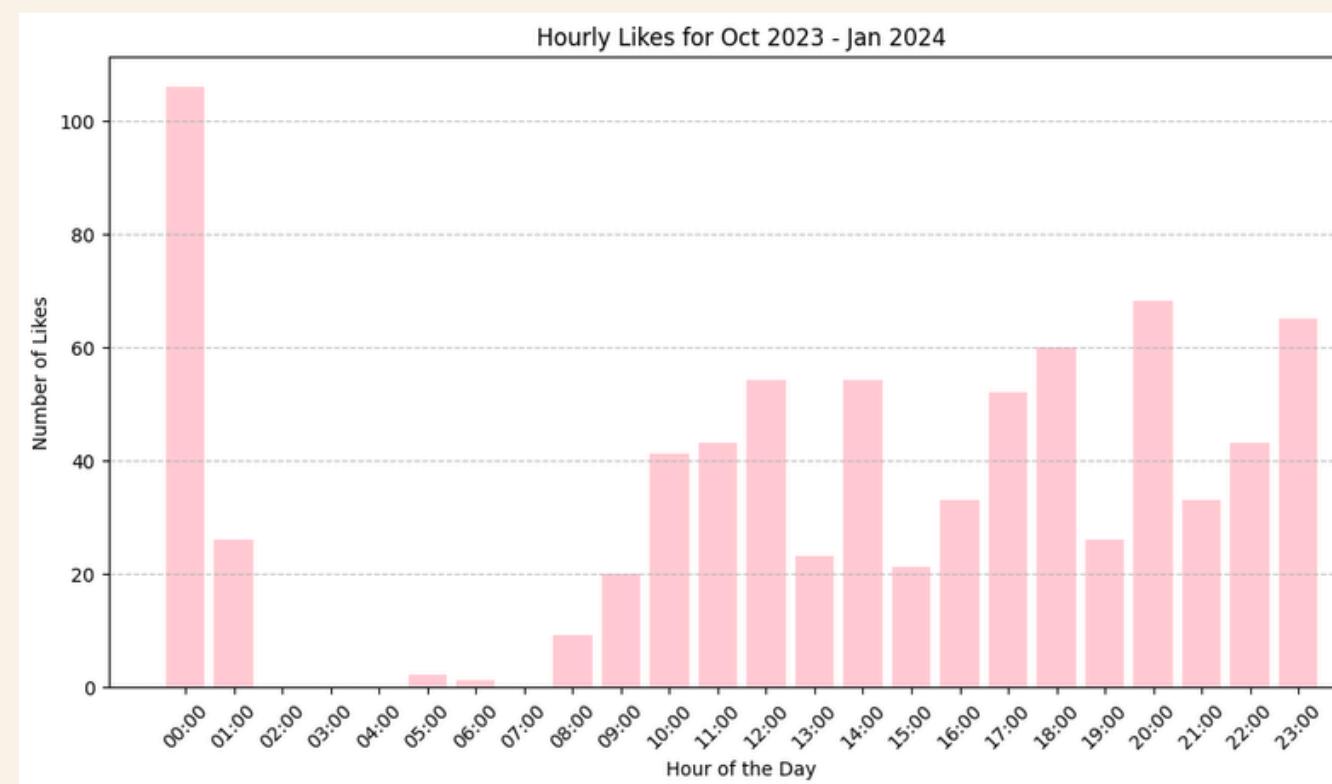
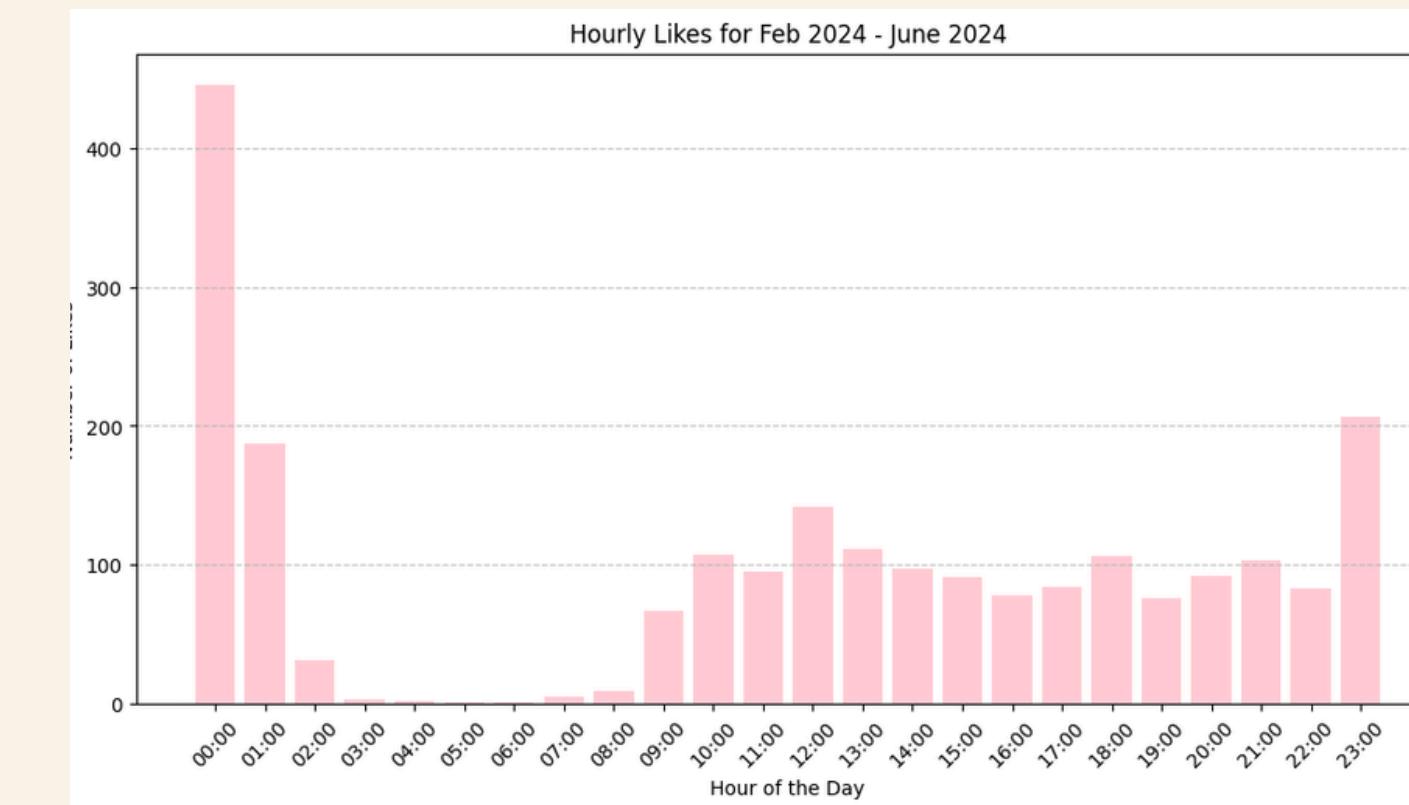
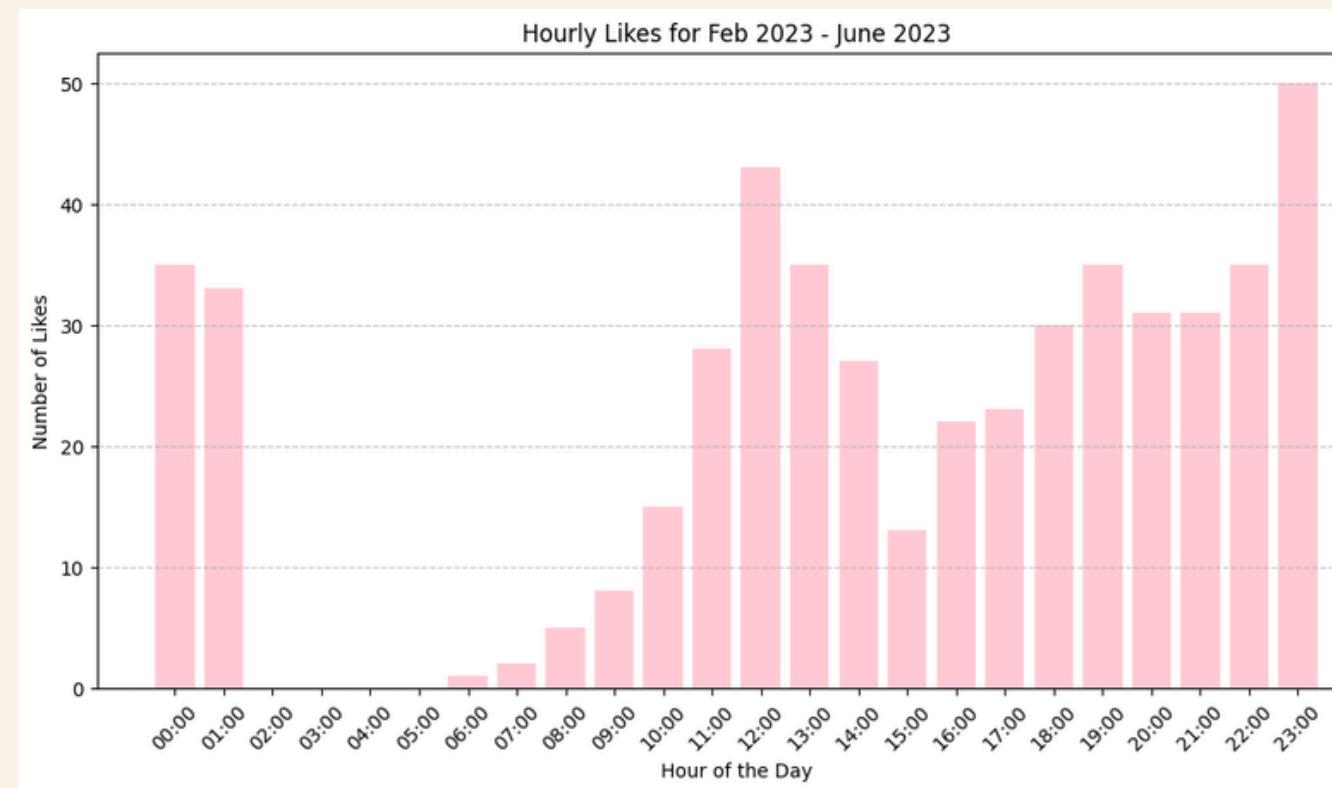


03.1 Forest App Data, cont.



and many more...

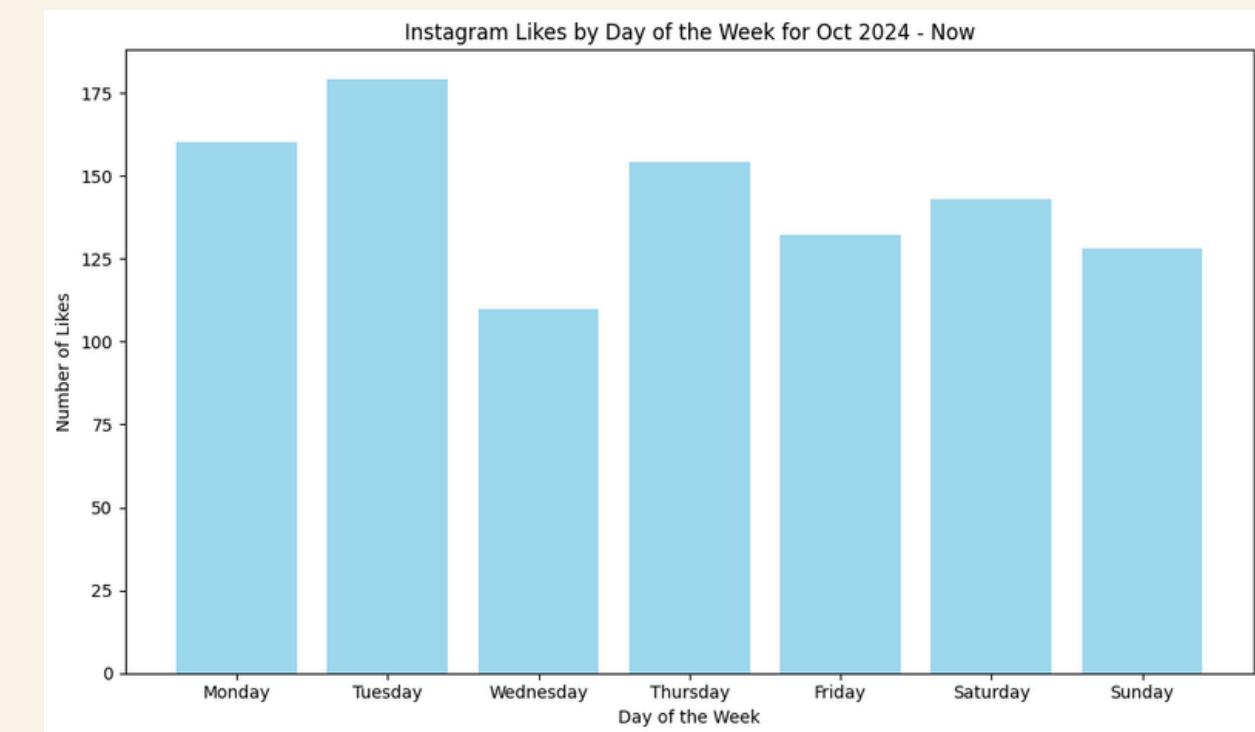
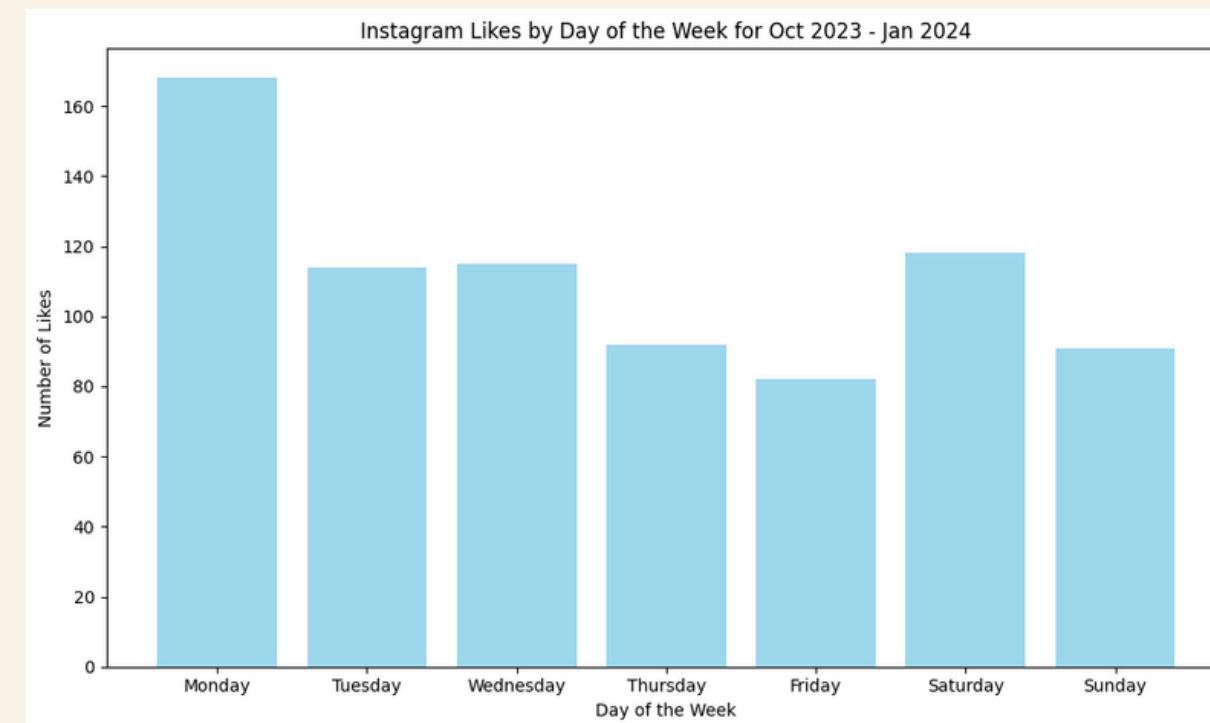
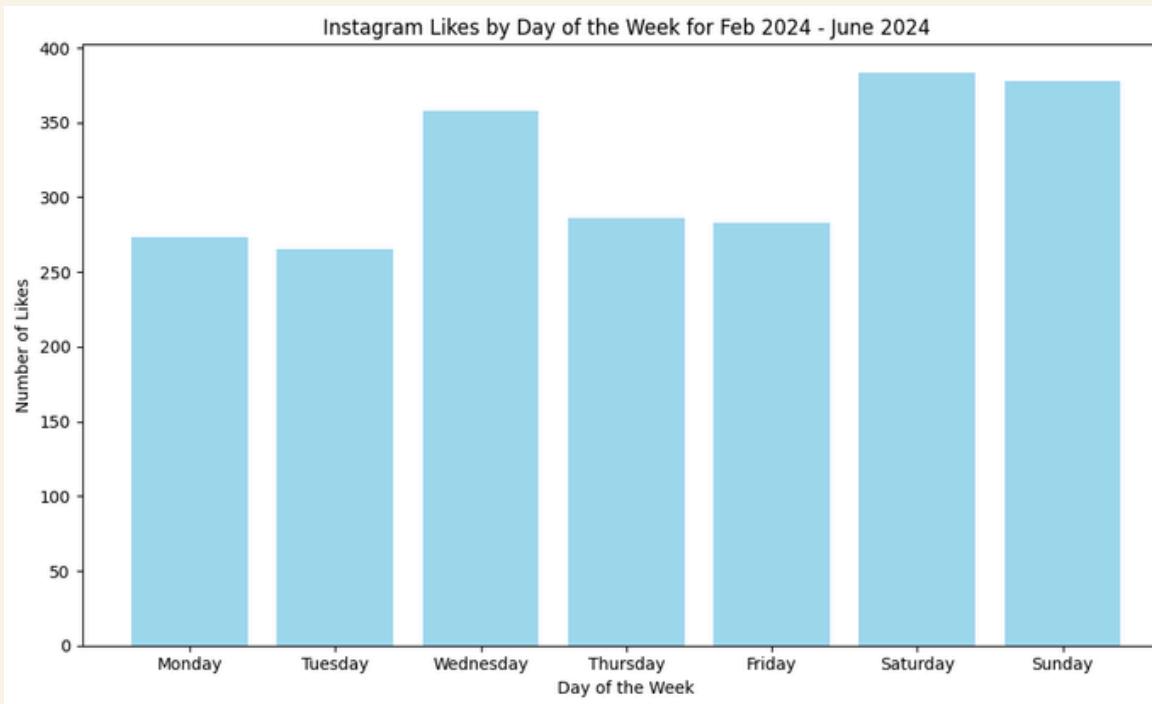
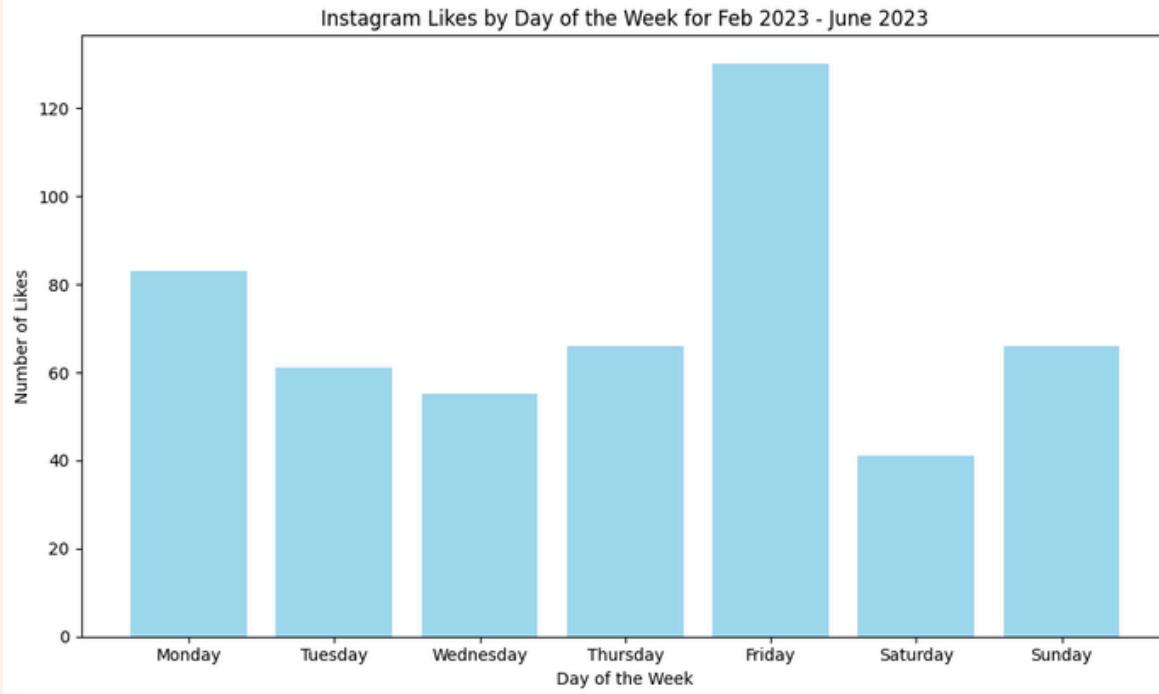
03.2 Instagram App Data



I have explored my hourly Instagram Usage as the first step.



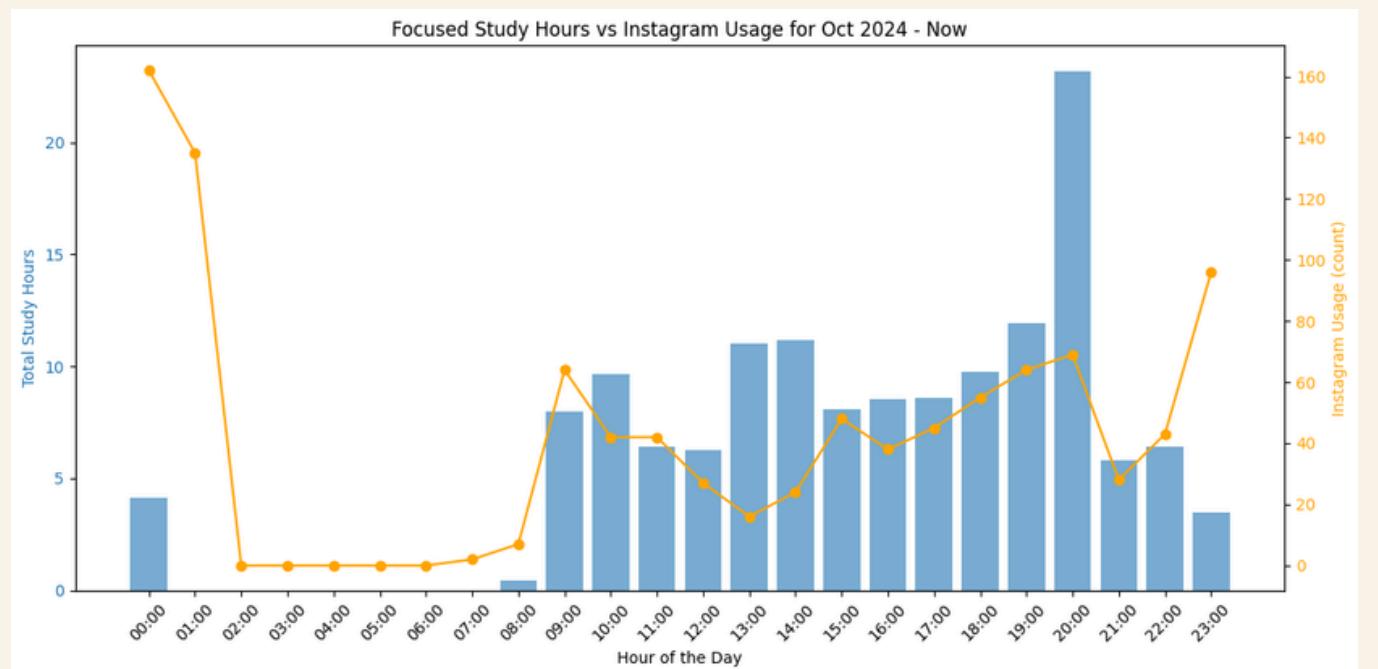
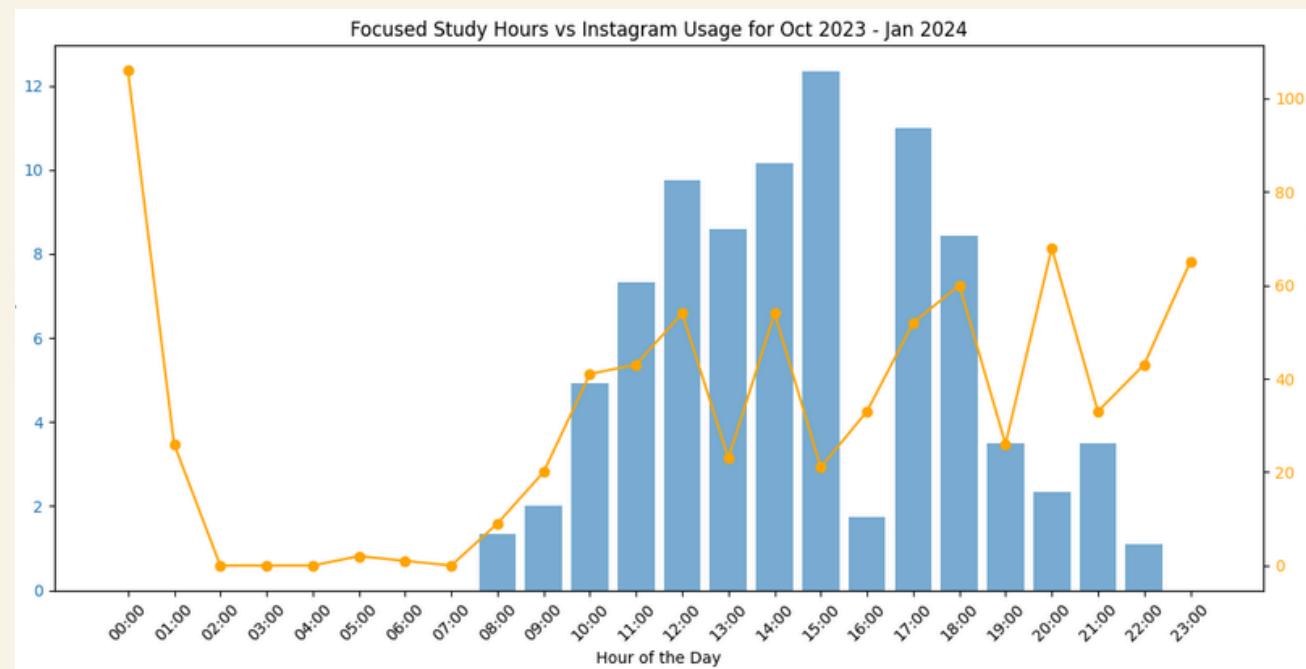
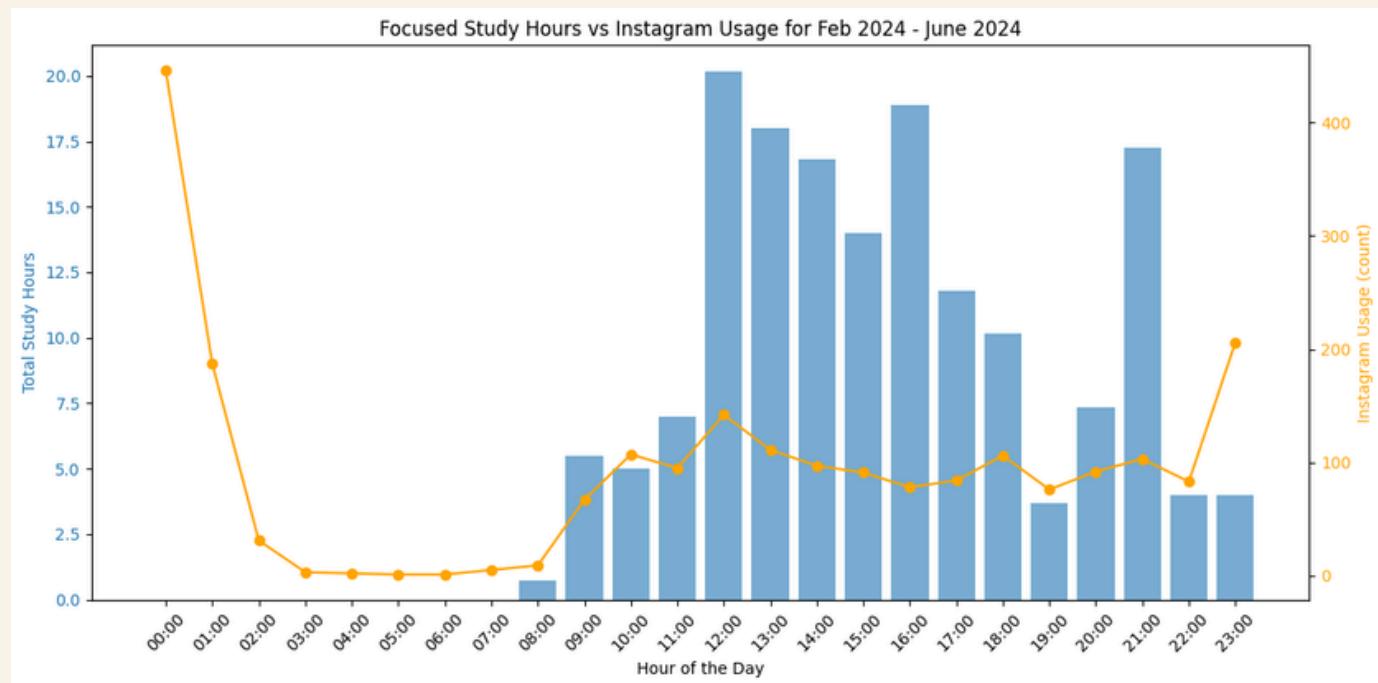
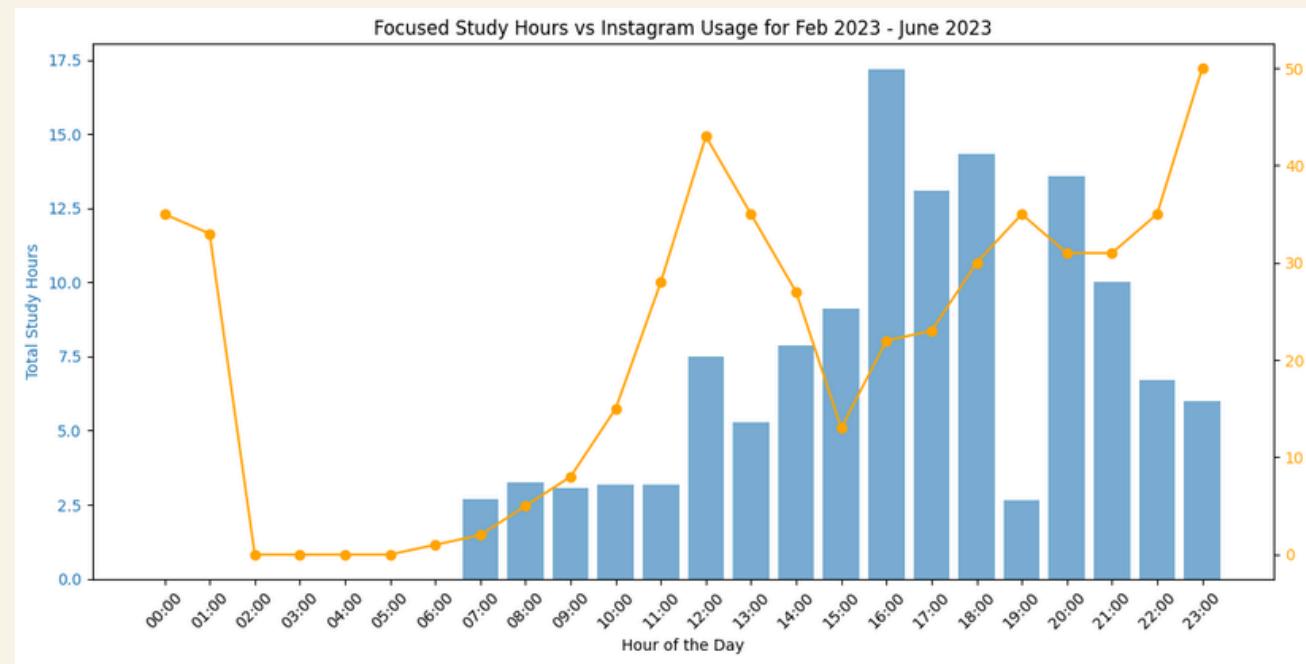
03.2 Instagram App Data, cont.



...and my daily Instagram Usage as the second step.



03.3 Data Analogy



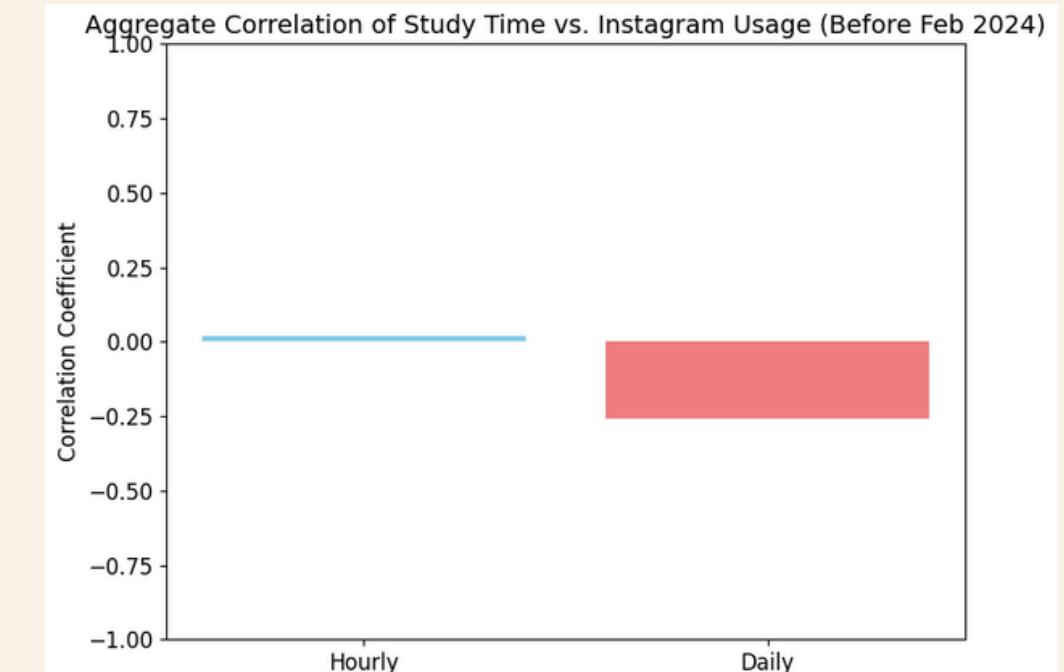
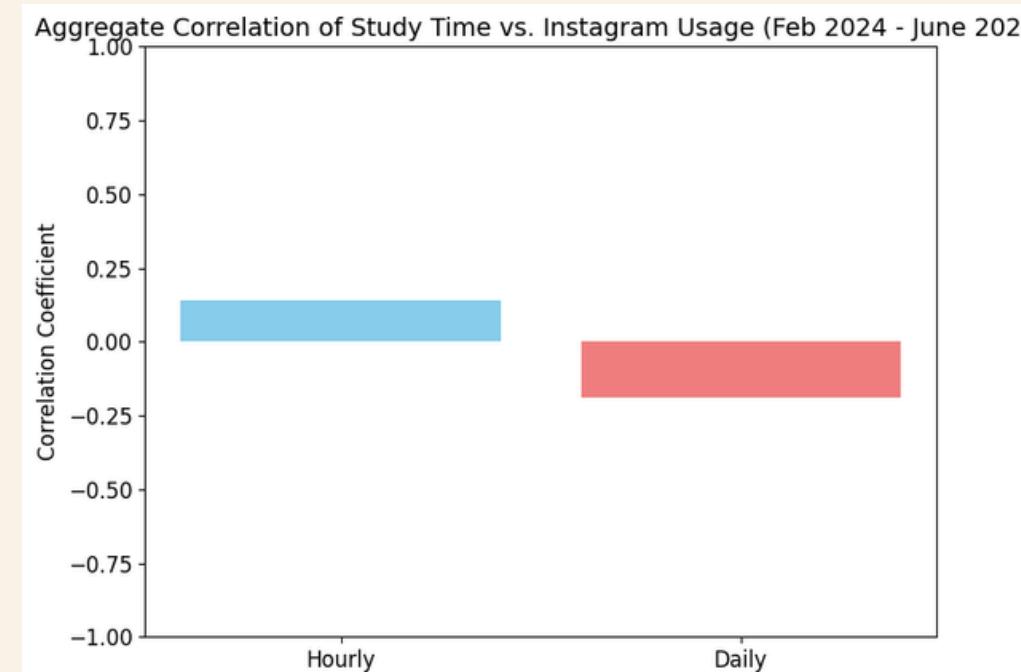
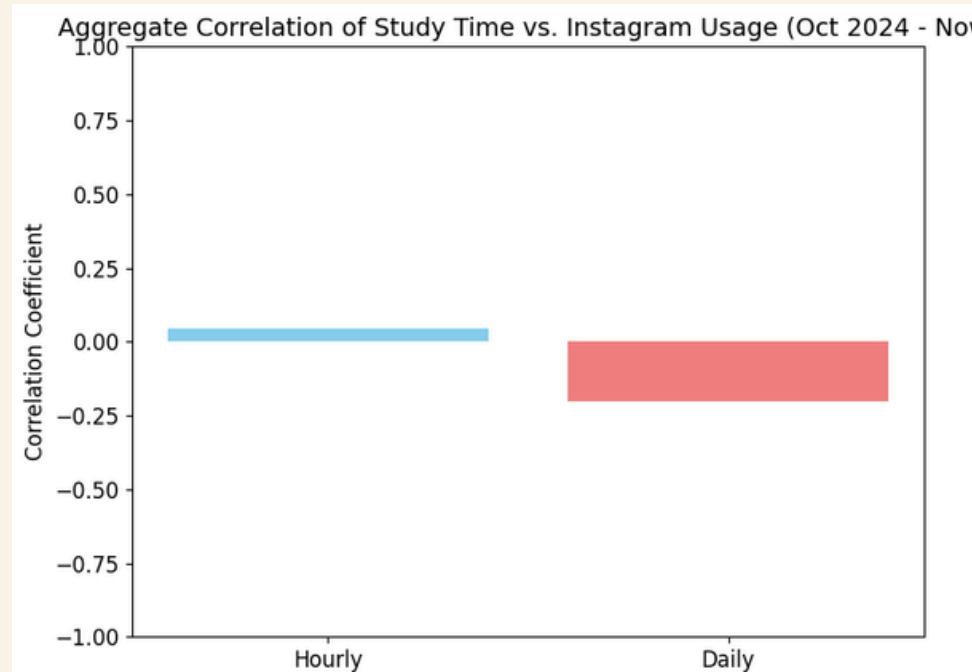
when explored separately, datasets might not give bright conclusions at the first glance.

so, I plotted them together to catch any relationship.



03.3 Data Analogy, cont.

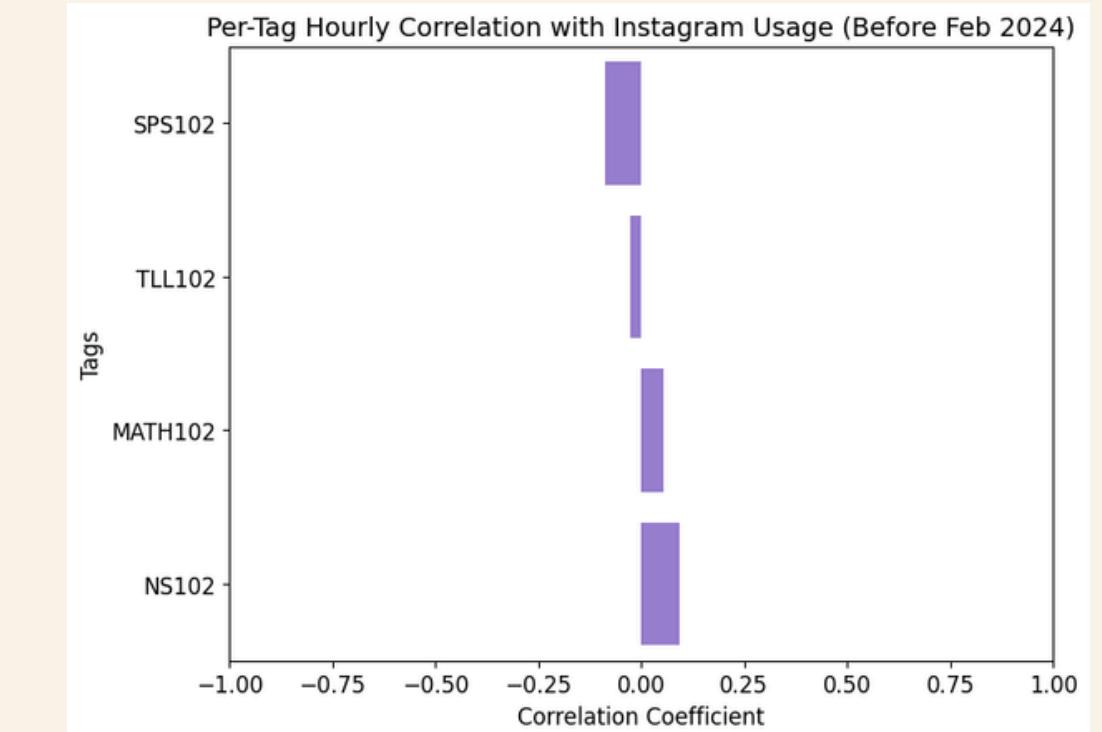
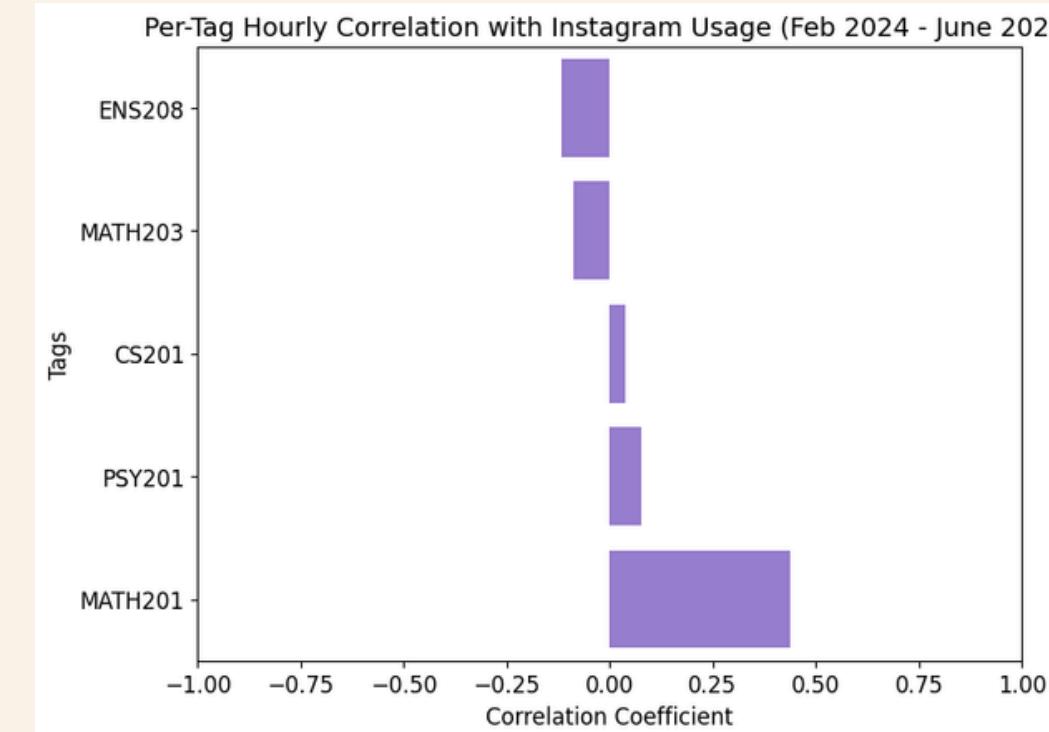
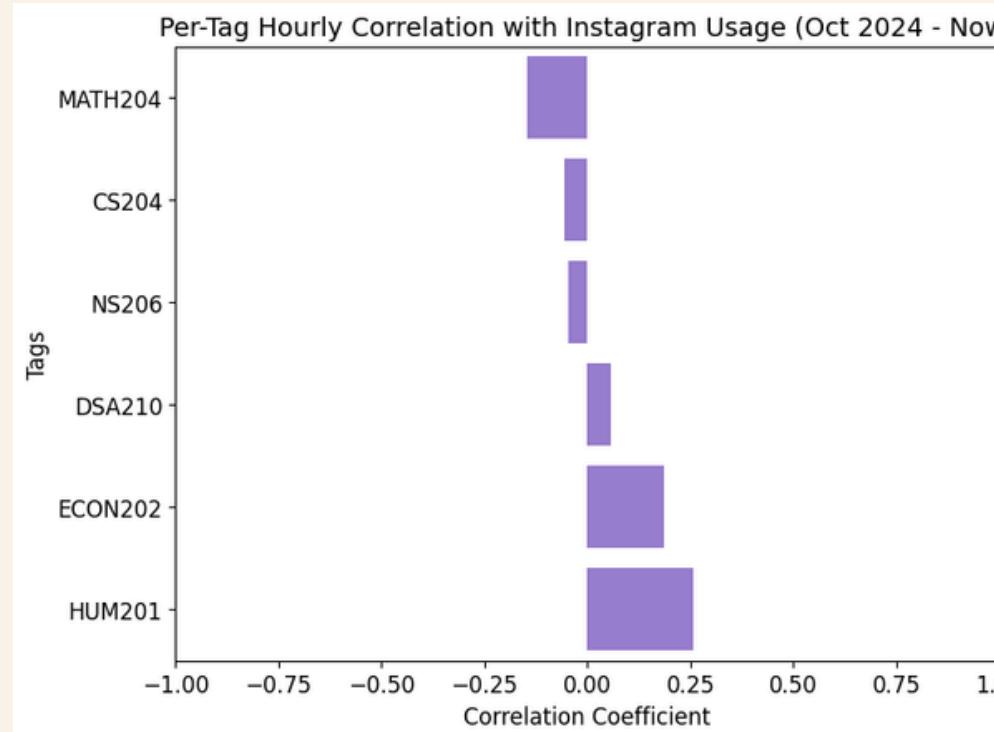
(For aforementioned reasons, I will be focusing on my last three academic semesters for preparation for my hypotheses testings.)



- In all three periods, the bars for hourly correlation are close to zero (or slightly positive).
- And in all three periods, correlation bars for daily Instagram usage are consistently negative (below zero) -> This implies that on days when Instagram usage is higher overall, total daily study hours tend to be lower.

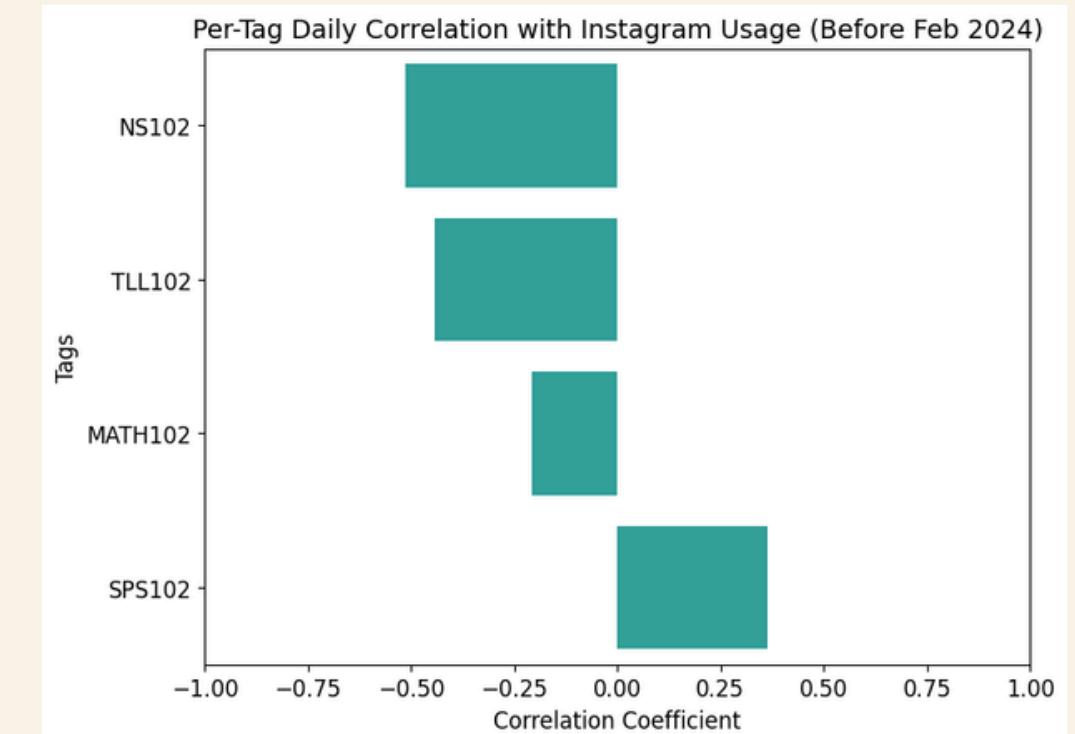
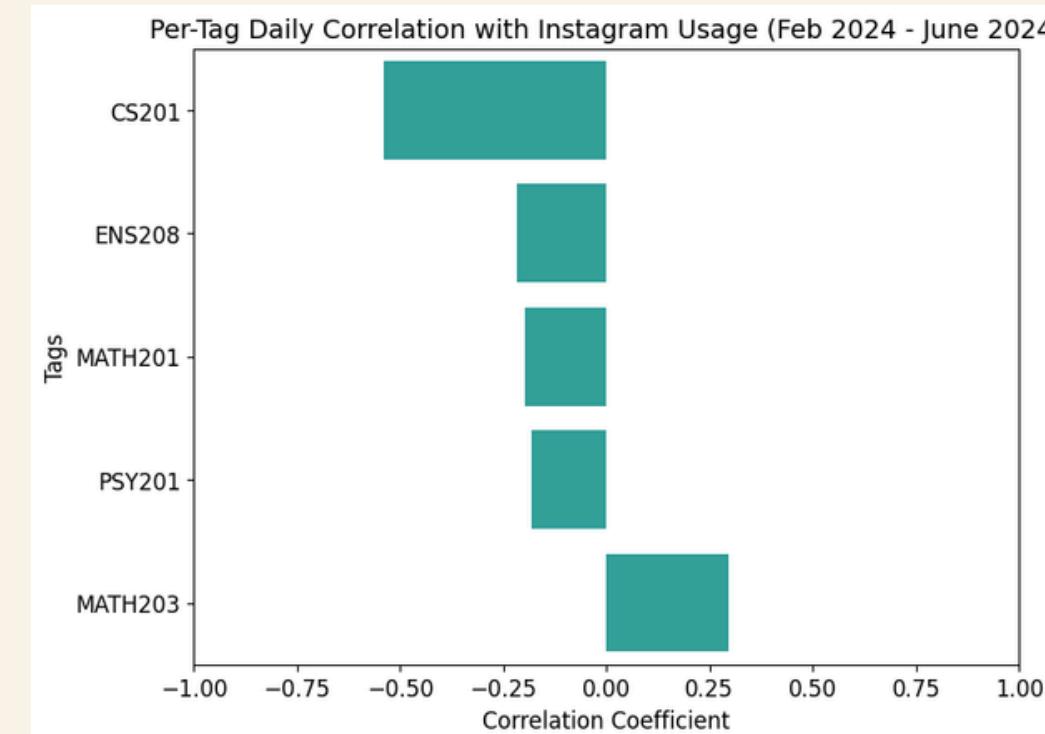
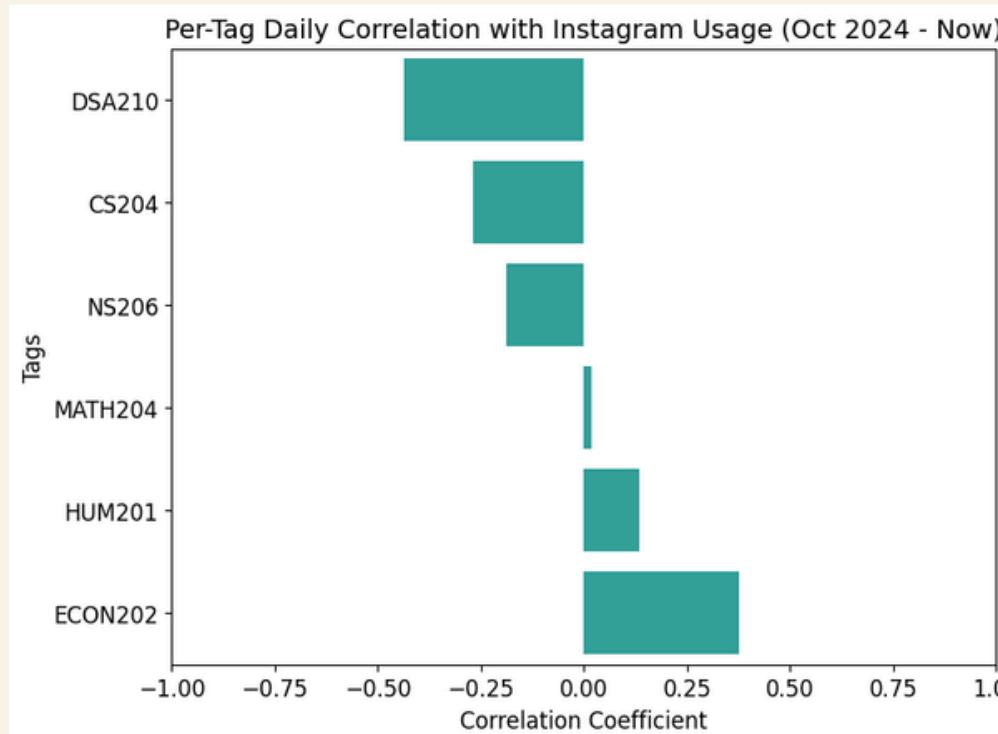


03.3 Data Analogy, cont.



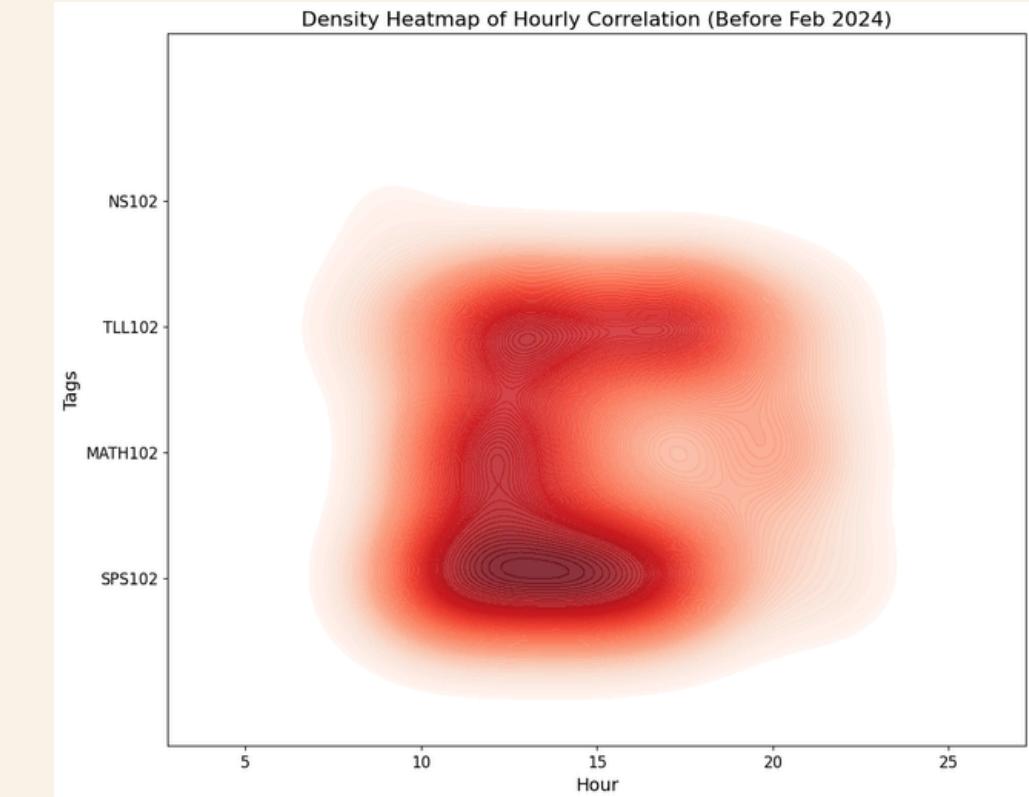
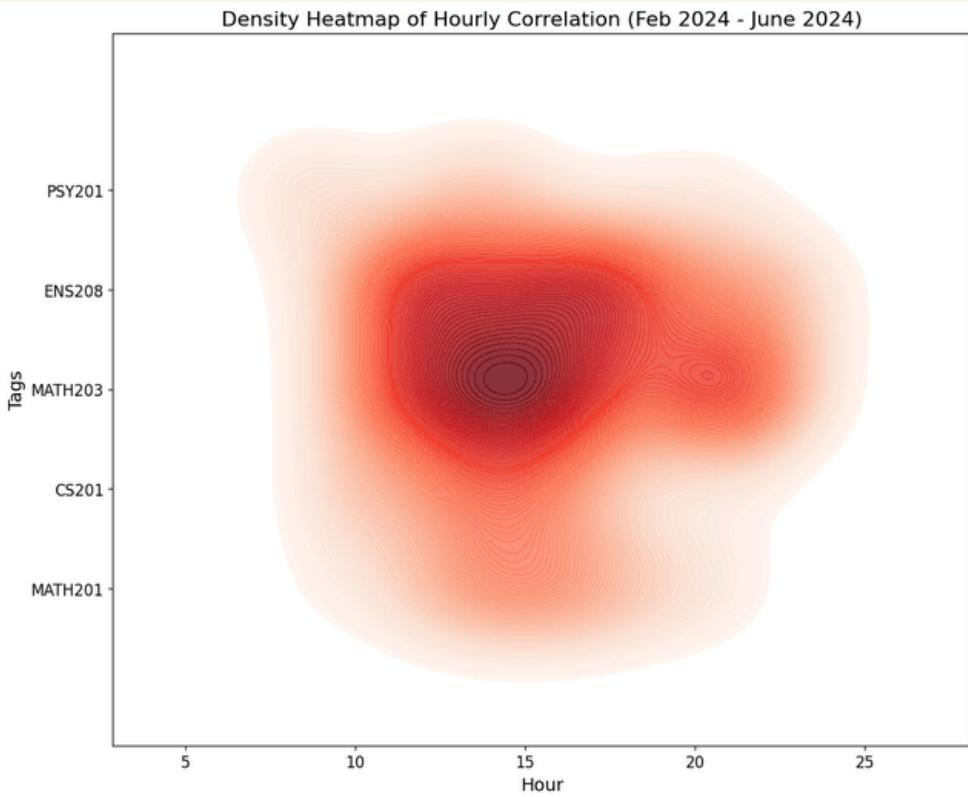
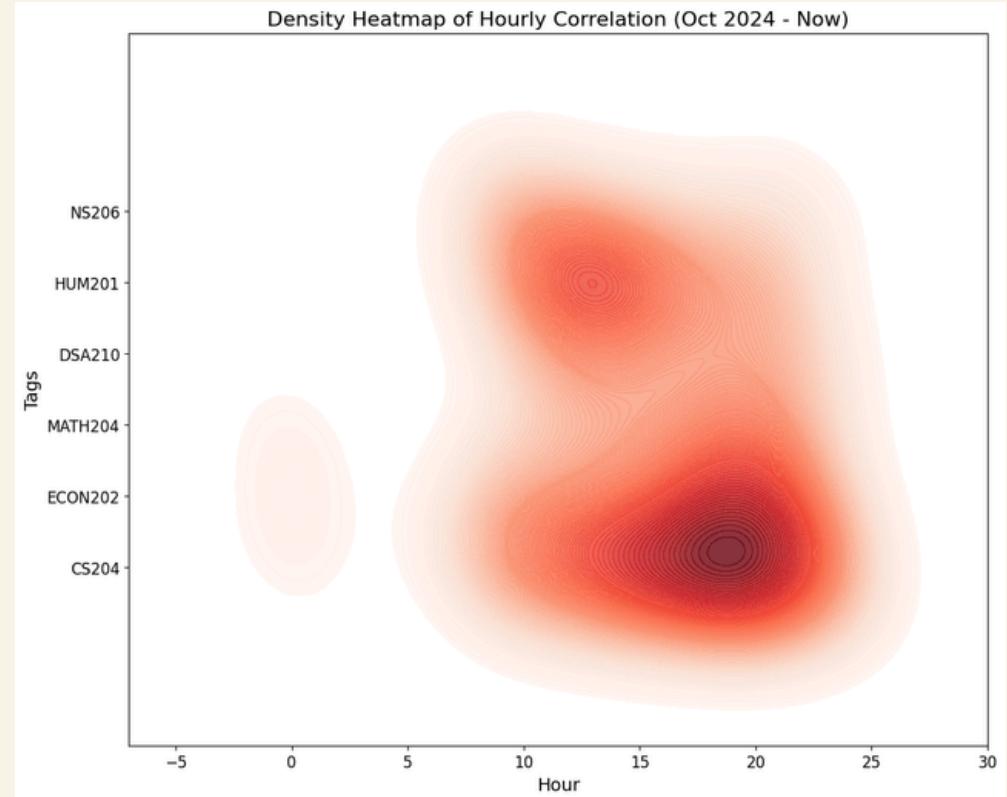
- Seems like I have been taking quick social media breaks while studying for HUM201, ECON202, MATH201 mostly.
- The other courses have almost zero or negative correlation, so my Instagram usage stays unrelated for these.

03.3 Data Analogy, cont.



- *Positive correlations that on days when I use Instagram more, I also study ECON202, MATH203 and SPS102 more.*
- *For the negatively correlated courses like DSA210, CS201 and NS102, Instagram usage takes away my study times - I study less.*

03.3 Data Analogy, cont.

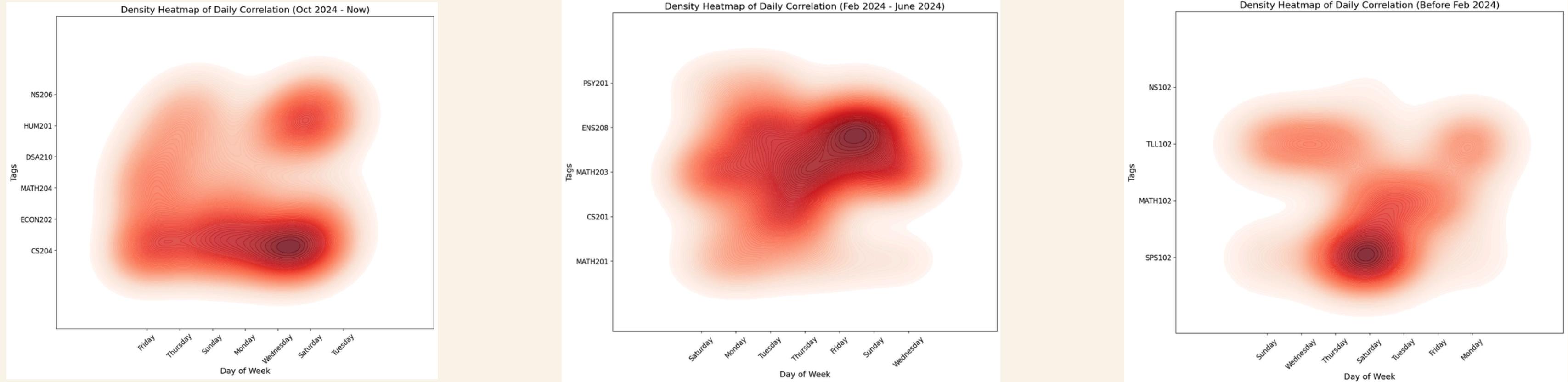


- *in the evenings, Instagram and study hour correlations peak.*

- *in the afternoon, Instagram and study hour correlations peak.*

- *around noon time Instagram and study hour correlations peak.*

03.3 Data Analogy, cont.



- Saturdays consistently show strong alignment in every semester so this is my most productive day or the day I multitask the most with Instagram and study -> we cannot distinct for now.
- Midweek days like Wednesday and Thursday often show strong correlations in more recent periods, indicating a shift in my study routines.

04 - Hypothesis Testing & Conclusions

	Period	Spearman Correlation	Correlation P-value	Chi-Square	Chi-Square P-value	
0	Oct 2024 - Now	0.180603	0.209441	0.017760	0.893981	
1	Feb 2024 - June 2024	-0.095770	0.243693	0.077381	0.780879	
2	Before Feb 2024	-0.088934	0.059993	0.176952	0.674006	

```
Run hypothesis testing × Run : ... /usr/local/bin/python3.11 /Users/berinayzumrasariel/PycharmProjects/DSA210--TERM-PROJECT/GitHubFolder/scripts/hypothesis testing.py --- Hypothesis Testing Results --- +-----+-----+-----+-----+-----+-----+ | Period | Hourly Correlation | Hourly P-value | Hourly Interpretation | Daily Correlation | Daily P-value | Daily Interpretation | +-----+-----+-----+-----+-----+-----+ | 0 | Oct 2024 - Now | 0.286926 | 0.207285 | Hypothesis fails to be rejected at p-value 0.207 | 0.180603 | 0.209441 | Hypothesis fails to be rejected at p-value 0.209 | +-----+-----+-----+-----+-----+-----+ | 1 | Feb 2024 - June 2024 | 0.507398 | 0.011378 | Hypothesis is rejected at p-value 0.011 | -0.0957699 | 0.243693 | Hypothesis fails to be rejected at p-value 0.244 | +-----+-----+-----+-----+-----+-----+ | 2 | Before Feb 2024 | 0.693015 | 0.000349675 | Hypothesis is rejected at p-value 0.000 | -0.0889338 | 0.0599931 | Hypothesis fails to be rejected at p-value 0.060 | +-----+-----+-----+-----+-----+-----+ Process finished with exit code 0
```

It can be concluded that, I reject my Hourly Correlation hypothesis, but fail to reject my Daily Correlation.



05 - Further Attempts



Thanks

-Berinay

