

# Social Media Addiction Analysis

## Tanay Dangaich

### Question:

Can we predict whether a user is addicted to social media or not based on the number of hours they spend on social media?

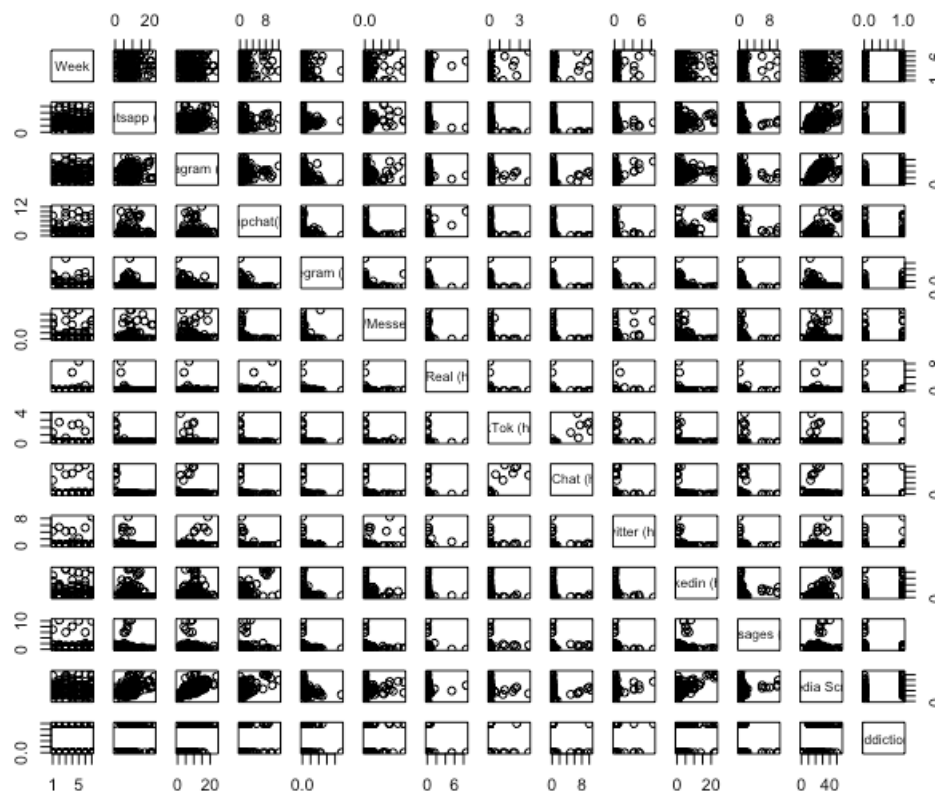
### Hypothesis:

Null Hypothesis (H0): There is no significant relationship between the number of hours spent on social media and social media addiction.

Alternative Hypothesis (HA): There is a significant relationship between the number of hours spent on social media and social media addiction.

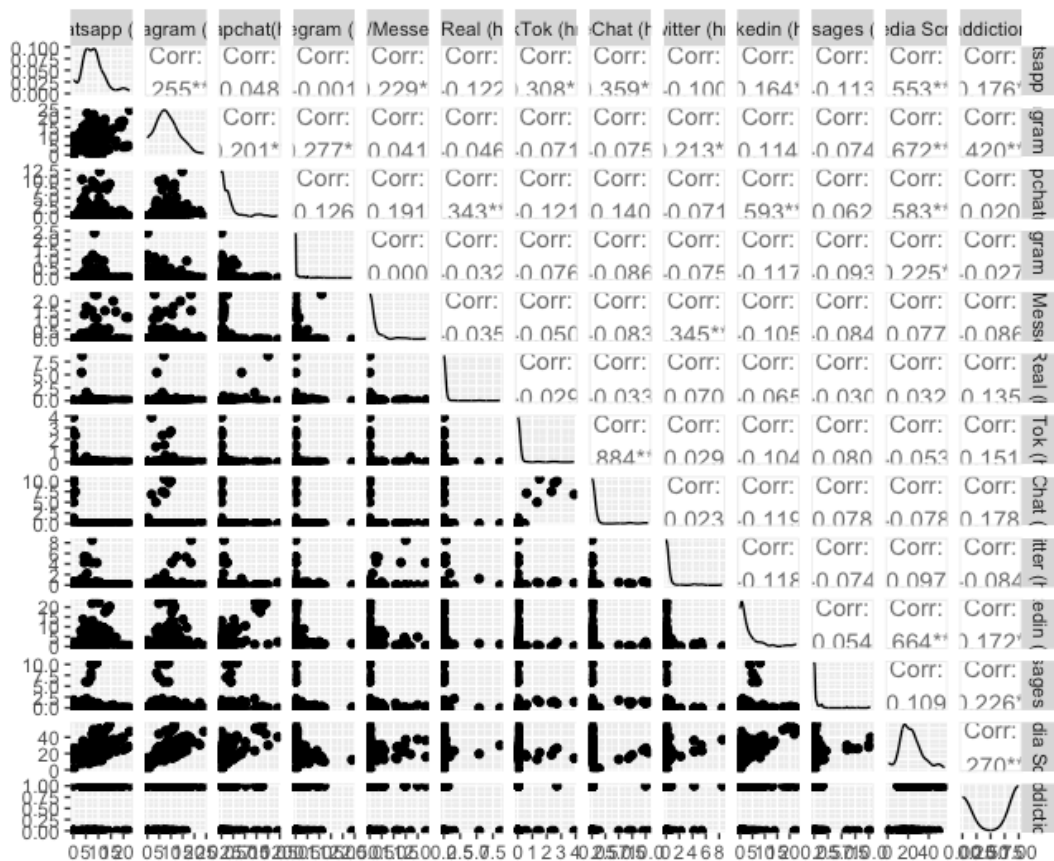
### Methodology:

1. Missing data – Took care of missing data by taking the mean of the columns and replaying 'NA' with the mean value of that column.
2. Exploratory Data Analysis
  - a. Scatter plot – We can see that the data does not have linearity between the variables.



- b. Scatter and correlation plot – We can see that there is correlation between a few variables –

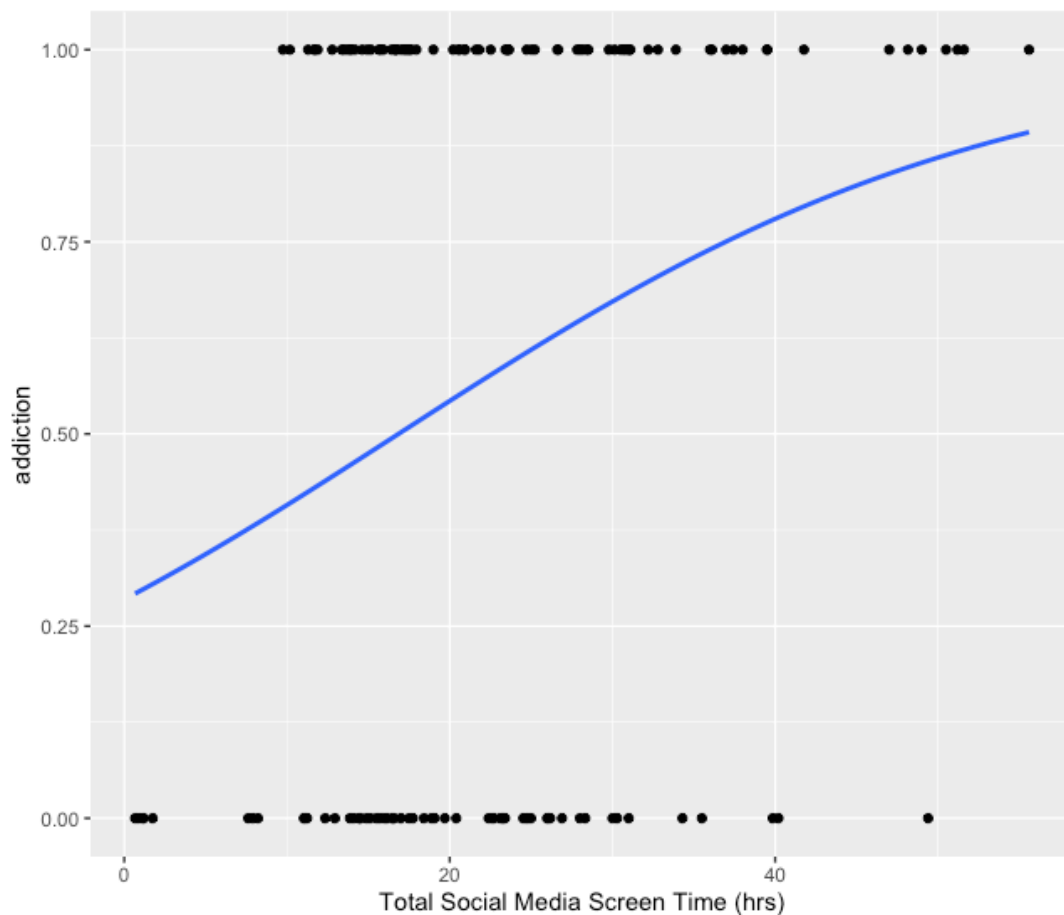
#### Social Media



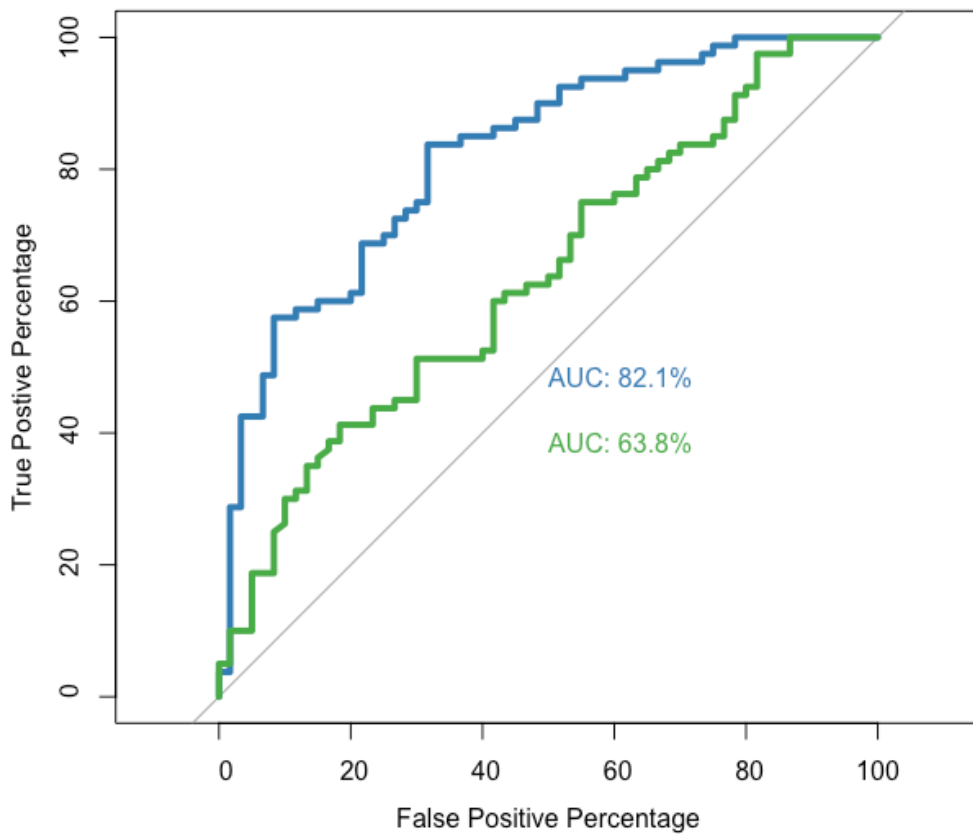
- i. Positive correlation
  1. LinkedIn and Social Media screen time
  2. WeChat and TikTok
  3. Twitter and Facebook/ Messenger
  4. LinkedIn and Snapchat
  5. Snapchat and Social media screen time
  6. BeReal and Snapchat
  7. Instagram and number of times opened.
  8. Social Media screen time and Instagram
  9. Instagram and WhatsApp
  10. Social Media screen time and WhatsApp
- ii. Negative correlation
  1. Telegram and Instagram
  2. TikTok and WhatsApp
  3. WeChat and WhatsApp
- iii. We can see that Instagram is correlated with number of times screen opened, which means it has a significant contribution towards users having social media addiction. (Why is Instagram the only app having significant relation with the numbers of times the users have opened their phone? Is this related

to how the app is built? Does the number of notifications sent by the app affect number of times phone is opened?)

3. Logistic Regression - Classification of users as social media addicts or not by performing logistic regression on the data. I have converted the addicted or not addicted column into binary dummy variables and replaced 'Social Media Addiction' with 'addiction' column consisting of 0 and 1.



- a. Model 1 - I have used all the features except names of the 'Student', 'Week' and 'Number of times opened (hourly intervals)'. I got the AUC as 82.1% which signifies that the model is a fairly accurate model.
- b. Model 2 – I have used only 'Total Social Media Screen Time (hrs)' to predict whether a user is addicted to social media or not. I got the AUC as 63.8% which signifies that it is barely better than a random coin toss.



We can see that Model 1 performs better at predicting whether a user is addicted to social media or not as compared to Model 2. There is a lot of scope to improve the model further.

### **Conclusion:**

Based on the logistic regression analysis performed on the dataset of social media addiction, we can conclude that the total number of hours spent on social media can be a predictor of whether a user is addicted to social media or not.

The logistic regression model had a good AUC of 82.1% in predicting the addiction status of users based on their social media usage.

The results suggest that social media addiction is a growing concern and can have negative effects on users' lives. Therefore, it is important to educate users about responsible social media usage and to develop interventions to reduce social media addiction.

### Future Work:

1. How does Instagram app affect social media addiction?
2. Is there a relationship between social media applications and geographic location of the user and demographic representation of the users?
3. Are the users using social media for personal use or business?
4. Is mood affected by social media usage?
5. Is there a relationship between notifications sent by the apps with addiction? Does the sound of the notification play a significant part in the user's decision to open the app?

### Appendix:

Type of variable	Column/ Variable name	Description
Categorical	Student	Name of the student
Categorical	Week	Week of entry
Quantitative	WhatsApp (hrs)	Number of hours spent on WhatsApp during the week.
Quantitative	Instagram (hrs)	Number of hours spent on Instagram during the week.
Quantitative	Snapchat (hrs)	Number of hours spent on Snapchat during the week.
Quantitative	Telegram (hrs)	Number of hours spent on Telegram during the week.
Quantitative	Facebook/ Messenger (hrs)	Number of hours spent on Facebook/ Messenger during the week.
Quantitative	BeReal (hrs)	Number of hours spent on BeReal during the week.
Quantitative	TikTok (hrs)	Number of hours spent on TikTok during the week.
Quantitative	WeChat (hrs)	Number of hours spent on WeChat during the week.
Quantitative	Twitter (hrs)	Number of hours spent on Twitter during the week.
Quantitative	LinkedIn (hrs)	Number of hours spent on LinkedIn during the week.
Quantitative	Messages (hrs)	Number of hours spent on Messages during the week.
Quantitative	Total Social Media Screen time (hrs)	Cumulative number of hours spent on all the apps listed throughout the week.
Quantitative	Number of times opened (hourly intervals)	Number of times a user has unlocked their phone through the week.
Categorical	Social Media Addiction	Whether a user is addicted to social media or not.

