

Data Visualization Pitch

Data columns (total 15 columns):

#	Column	Non-Null Count	Dtype
0	person_name	5000 non-null	object
1	age	5000 non-null	int64
2	date	5000 non-null	object
3	gender	5000 non-null	object
4	platform	5000 non-null	object
5	daily_screen_time_min	5000 non-null	int64
6	social_media_time_min	5000 non-null	int64
7	negative_interactions_count	5000 non-null	int64
8	positive_interactions_count	5000 non-null	int64
9	sleep_hours	5000 non-null	float64
10	physical_activity_min	5000 non-null	int64
11	anxiety_level	5000 non-null	int64
12	stress_level	5000 non-null	int64
13	mood_level	5000 non-null	int64
14	mental_state	5000 non-null	object

Mental Health Social Media Indicators

Student name: Asmi Mahesh Patole
Student ID: 945252

Student e-mail address: a.patole@campus.unimib.it

MOTIVATION

Social media is an ongoing trend in todays time which means being on social media equivalents to being "cool" specially in youngsters.

But has anyone ever wondered the impact it causes? The numer of deaths social media has caused?.

Well just to make this topic a bit clear... according to Centers of Disease Control and Prediction in the year 2020 around 46,000 people lost their file to suicide that had been caused due to cyber bullying, lack of confidence on social media, less followers and what not.

In this presentation I would like to give a clear image on the side effects it causes and the way it shapes todays mind.

1. Mental Health Social Media Indicators

Dataset: impact of social media

URL: <https://www.kaggle.com/datasets/sonalshinde123/social-media-mental-health-indicators-dataset>

Explanation: This dataset contains the relationship between social media usage, screen time behavior and daily lifestyle such as sleep duration and interaction quality.

Key features:

- Person identity
- Daily screen time
- Social media screen time
- Negative interaction counts
- Positive interaction counts
- Mood
- Anxiety
- Stress level
- Mental state

Research questions

- If social media is a new platform for connecting people why are so many feeling left out?
- How does the constant need to be "perfect" and gets likes in social media apps like instagram shape our mind.
- How much does validation increase anxiety and lead to mood swings

About Data

- The topic selected of this project was completely on todays faced concern. I needed a project that spoke well about todays neglected discussion. Social Media perfectly fit the puzzle but also is a huge topic in todays generation which needs to be studied.
- The dataset i picked is from a opensource kaggle.
- It is a public domain and available for everyone to use for educational purposes.
- The dataset is well structured a does not need any data cleaning
- This dataset consists of 5000 rows and 15 columns, capturing detailed information about individuals, social media useage patterns, life habits and their corresponding impact on health.
- The dataset contains of personal data, behavior, anxiety, mood, interactions made with friends and stranger, sleep and physical activities.

Methodology

- **Data Collection:** The data was collected from an opensource kaggle.
- **Data Processing & Cleaning:** For the data visualization project I used Python as my main course with various libraries like pandas, seaborn, matplotlib, matplotlib venn and numpy.
- **Use of AI Tools:** For this dataset I used claude AI and gemini which helped me suggest shorter codes for my visualization.
- **Analytical Techniques:** In this project I have used Trend Analysis, outlier detection, descriptive statistics, correlation technique and time series.

Insights from the Data

The following analysis shows the relationship between social media usage , screen time behavior, daily life style factors and interaction quality.

From a methodology point of view, the analysis was carried out using PYTHON(programming language).

1. Mixed type of dataset including : numerical data, identifiers and categorical data.
2. The dataset is based on platforms such as instagram, twitter, snapchat, facebook, whatsapp , tiktok and youtube.

This approach allowed us to highlight:

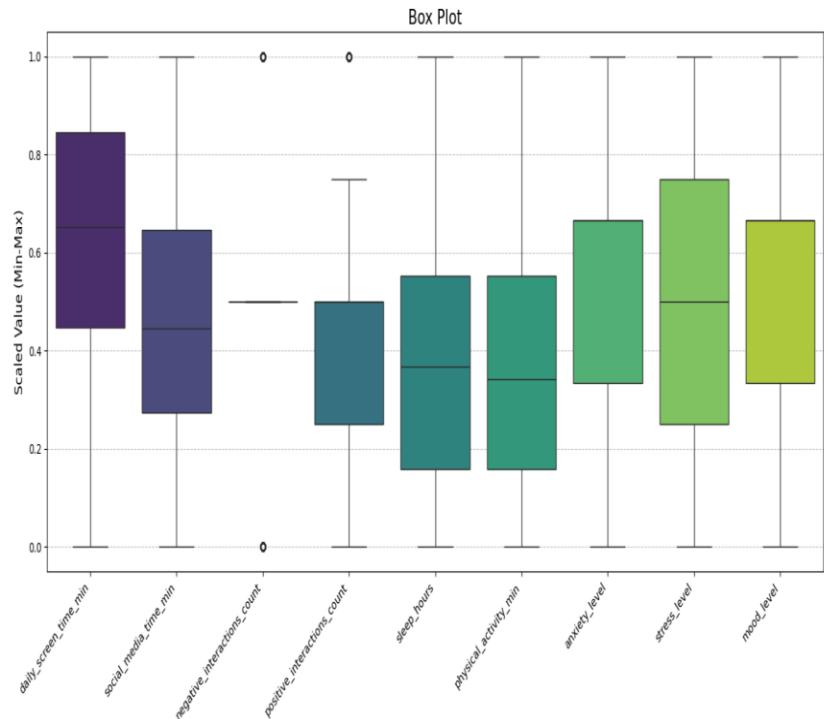
1. The unknown factors of using social media
2. How does screen time affect our mental health

INSIGHT 1:

- The visualization represents the interquartile range(IQR) containing the middle 50% of the data.
- The line side the box is called as the median.
- The whiskers extend t show the range of data points.
- Thw circles represent outliers.

OBSERVATIONS:

- Daily screen time min has the highest number of observations with relatively high values overall.
- Negative interactions count and positive interaction counts show compressed distributions at the bottom with very low medians and several outliers at the top.
- The data spread for stress level shows nearly the entire 0-1 range.
- Negative outliers also have an outlier near 0.



INSIGHT 2: THE INTERACTIONS

- In this data, people experienced twice as many good interactions as bad ones on social media.

Atypical person had around 2 positive experiences and less than 1 negative interaction:

TOTAL NUMBERS; the dataset consists of a total of 9177 positive interactions and 4321 negative interactions.

OBSERVATIONS

- Each person got about 2 positive interactions per day.
- Each person got less than 1 negative interaction per day.

Total Positive Interactions: 9177

Total Negative Interactions: 4321

Average Positive Interactions: 1.8354

Average Negative Interactions: 0.8642

«GRAPHICIAL REPRESENTATION»

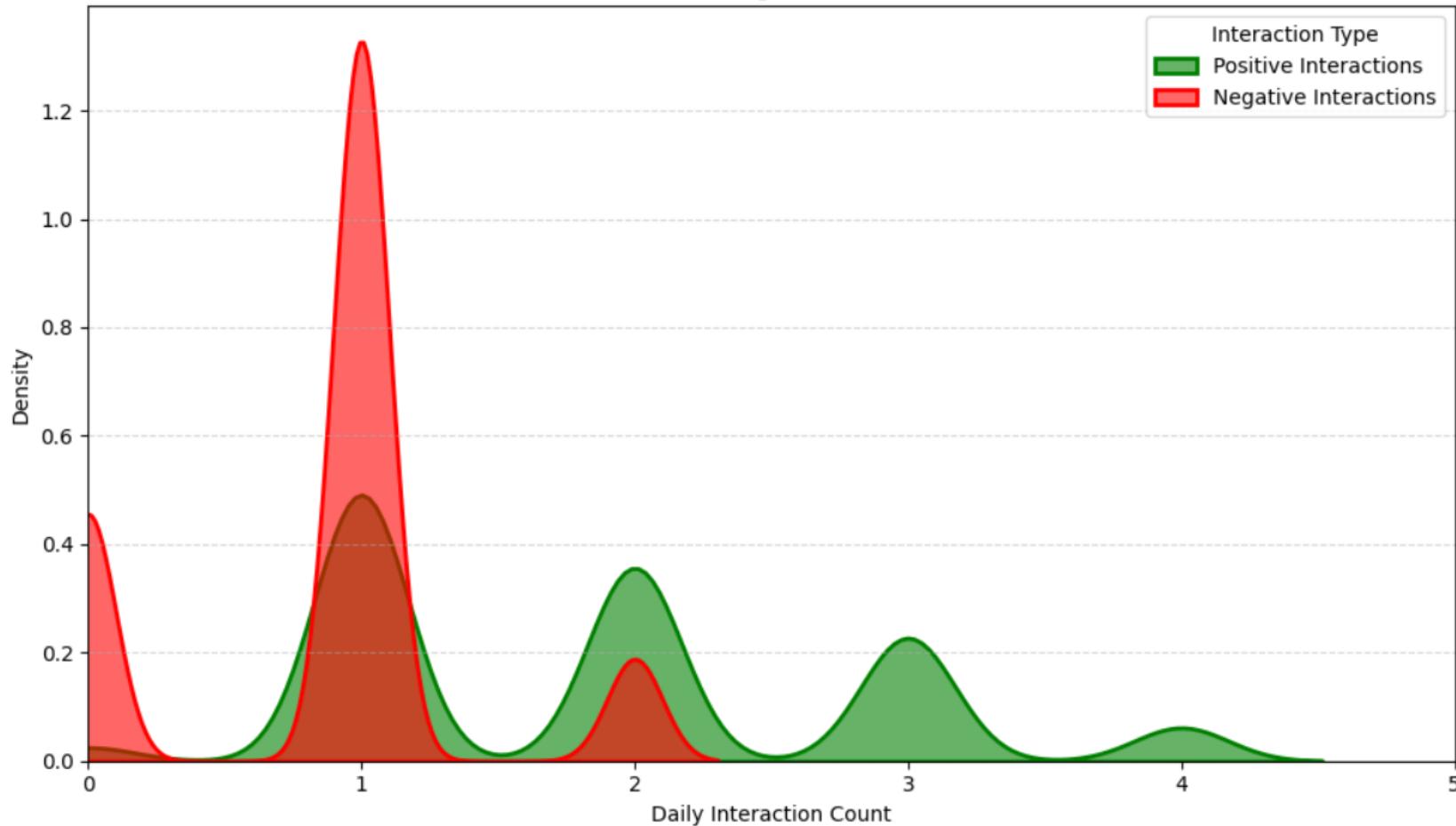
The density plot visualizes how positive and negative social media interactions are distributed.

The red curve represents negative interactions showing a sharp concentrated peak at approximately 1 interaction per day. This represents that negative interactions are constant across participants with most encountering around one negative interaction per day.

In contrast, the green curve representing positive interactions displays a much broader, multi-modal distribution. The peak occurs around 2 interactions per day with additional peaks at 3 and 4 interactions. The wide spread clearly shows more positive engagement experienced than negative.

The overlapping region represents instances where participants experienced similar frequencies of both interaction types.

Distribution of Positive vs. Negative Social Media Interactions



INSIGHT 3: ANXIETY, STRESS AND MENTAL HEALTH

This 3D dimensional scatter plot represents the relationship between three mental health variables in the dataset.

- Anxiety level ranging from 1.0 to 4.0
- Stress level ranging from 6.0 to 9.0
- Mental state ranging from 0.9 to 1.0

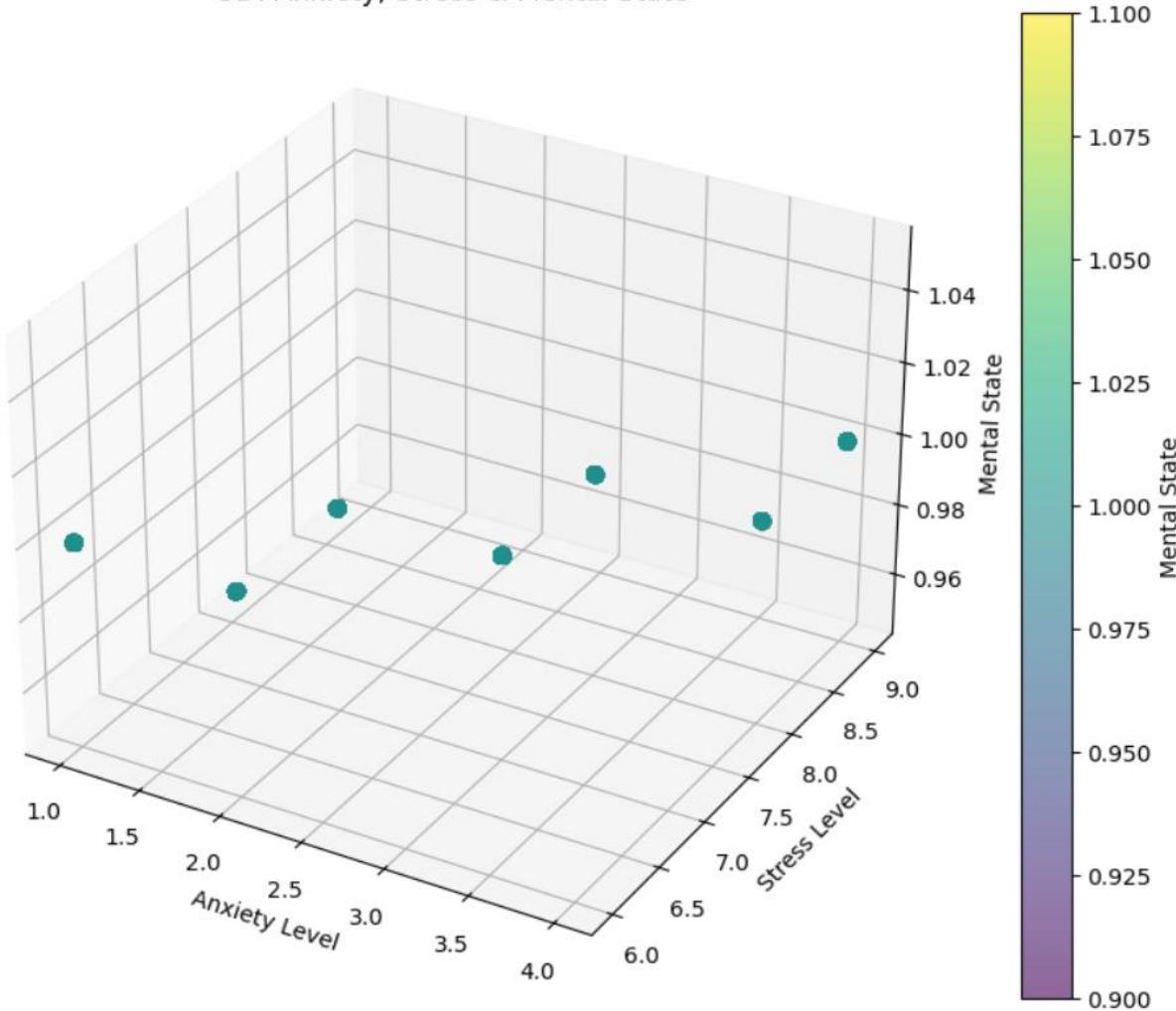
Each dot in the diagram represents a data point showing how each of the three variables are related to each other.

The dots are well spread across different positions but still stay relatively close to the same height.

This states that even though people have different anxiety and stress levels, their overall mental state scores don't vary(they all stay the same within a narrow range)

The uniform coloring shows the all the data points are likely all falling within a narrow range of 0.9 to 1.0 in the color scale.

3D: Anxiety, Stress & Mental State



INSIGHT 3: CIRCULAR AREA CHART

The three axes:

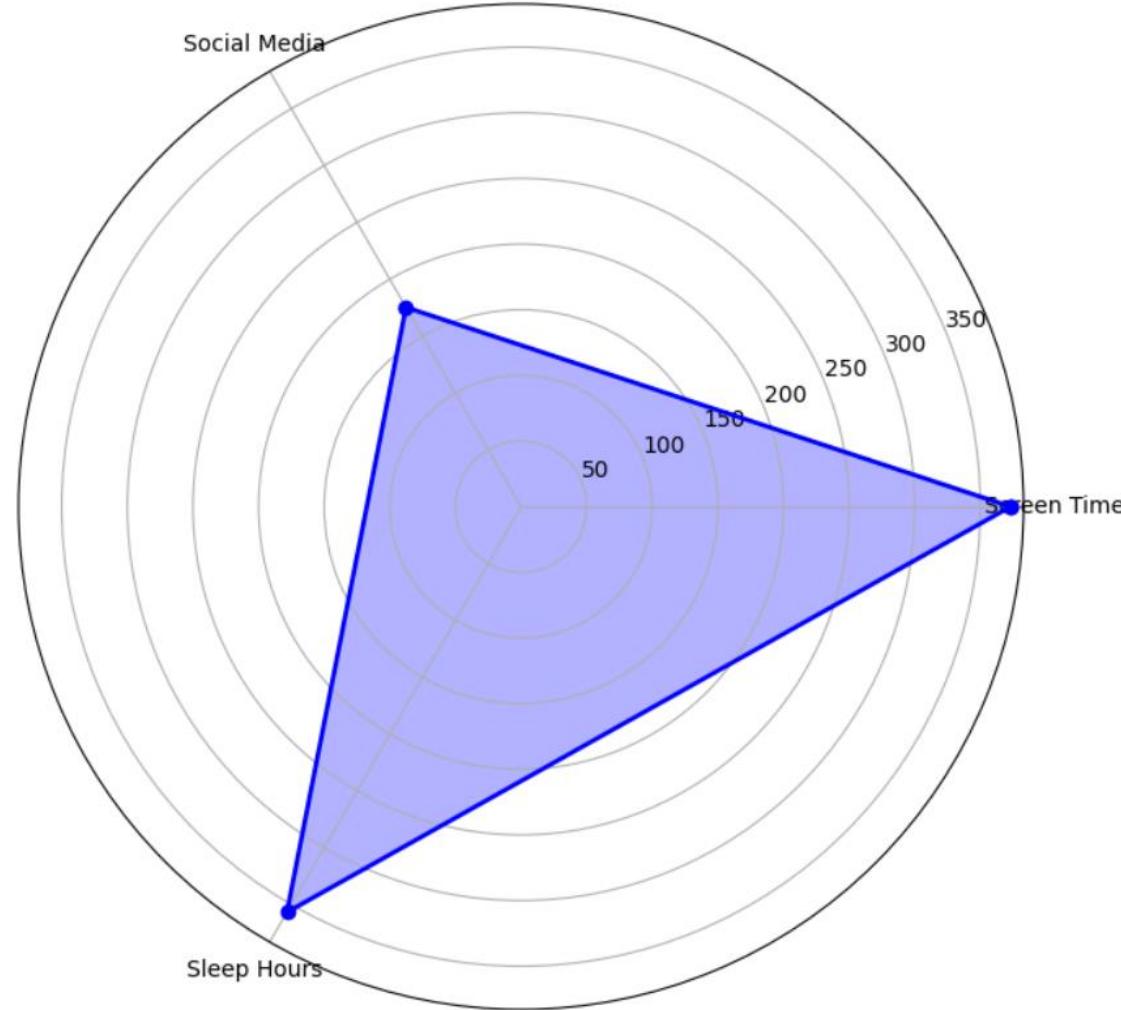
- Screen time: extends till the range of 350
- Sleep hours: extends till the range of 300
- Social media: extends till range of 200

Each axes radiates out from the center of the wheel. The farther a point is from the center the higher the value of the variable.

The blue shaded area connecting all three axes shows the proportion of these factors.

The triangle screen time reveals that screen time has the highest value followed by sleep hours. Social media usage has the lowest value.

Circular Area Chart



LICENCE

Slides are shared with the following license:

Mental Health Social Media Indicators.

<https://creativecommons.org/licenses/by-sa/4.0/>