CFG Spring 2025 Degree

Data Science

Group 2 Project

Investigating the Impact of Social Media on Teenagers' Mental Health

Antonia Badozi, Bethany Blakeley, Fariha Parvin, Helena Scotland, Humairaa Patel, Jodie Christian, Rebecca Webster
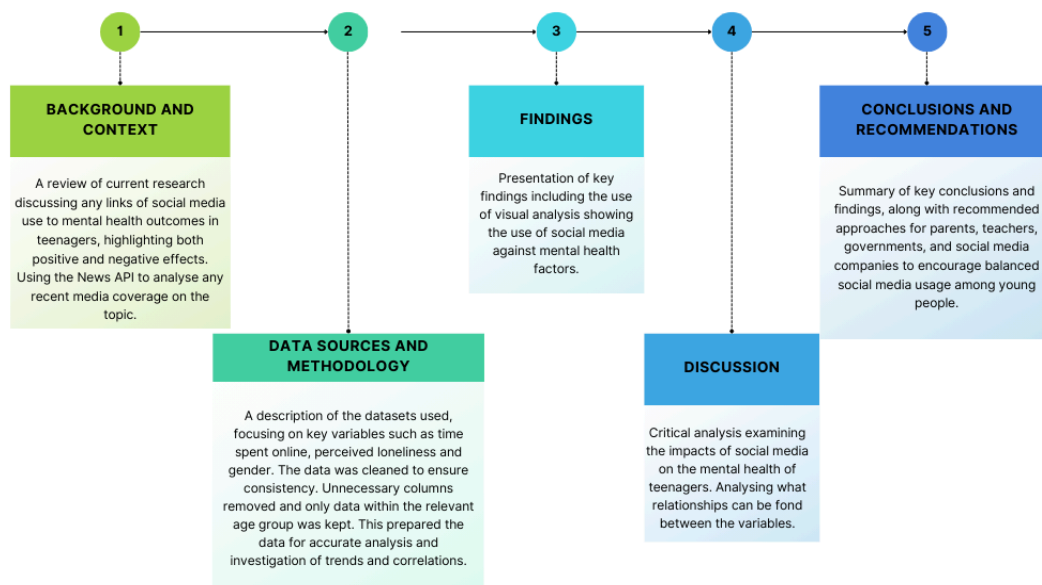
Introduction

We have seen the increase in the popularity of social media over recent years, and it is now a mainstay in modern society. Social media users worldwide in 2019 were 3.484 billion, a 9% increase year after year[1]. Now, in 2025, considering this rate of increase, we can assume the number of social media users is even greater. Social media is particularly popular with young people who have also experienced the use of various platforms, which have become relevant and offer unique features. The conversation that often follows the rise in the popularity of social media is whether it is having a positive or negative impact on us as humans, especially young people who haven't known anything different.

With "negative manifestations and symptoms of depression, anxiety, and stress"[2] often being reported, it's usually presumed that social media is unequivocally having a negative impact on young people. However, there have also been reports of numerous positive effects such as helping people stay connected, sharing ideas and access to educate themselves online[3].

This report investigates whether there is a relationship between social media and a variety of mental health factors, such as loneliness and anxiety. This report also explores whether these factors are impacted by age and gender. By analysing recent data, we hope to gain a broader understanding of the true impacts of social media on young people, allowing parents, teachers, healthcare professionals, government, and social media companies to make informed decisions about young people's access and use of social media in the future.

Project Roadmap

BACKGROUND AND CONTEXT

A review of current research discussing any links of social media use to mental health outcomes in teenagers, highlighting both positive and negative effects. Using the News API to analyse any recent media coverage on the topic.

DATA SOURCES AND METHODOLOGY

A description of the datasets used, focusing on key variables such as time spent online, perceived loneliness and gender. The data was cleaned to ensure consistency. Unnecessary columns removed and only data within the relevant age group was kept. This prepared the data for accurate analysis and investigation of trends and correlations.

FINDINGS

Presentation of key findings including the use of visual analysis showing the use of social media against mental health factors.

DISCUSSION

Critical analysis examining the impacts of social media on the mental health of teenagers. Analysing what relationships can be fond between the variables.

CONCLUSIONS AND RECOMMENDATIONS

Summary of key conclusions and findings, along with recommended approaches for parents, teachers, governments, and social media companies to encourage balanced social media usage among young people.

Background

The impact social media has on our lives is massive. Since the release of Facebook in 2004, it has become embedded in everyday life. Facebook alone has 3.07 billion monthly active users. With Instagram and TikTok also having over 1 billion users each, it is important that we study the impact social media can have on our mental health. As a team, we were especially interested in the impact it could have on children's mental health, as some of our team were teachers. It is also an important topic in popular culture during the completion of this report, Prince Harry and Meghan Markle called for stronger social media protections for children[4]. And, in 2024 Channel 4 produced a documentary about "Swiped: The school that banned Smartphones" also in 2024 the government held a debate on the "impact of smartphones and social media on children"[5]. News articles like these generally point out how social media is embedded in young people's

lives pointing to startling statistics like how "50% of 9 year olds own a smartphone and by age 11, 99% of children spend time online[6]. Studies in America have shown an increase in depression symptoms linked to social media use in young people[7].

It is important to consider some positive impacts of social media, as research has shown "social connections are important for our well-being"[8] and loneliness has been declared a "global public health concern"[9] by the WHO. So we have endeavoured to compare a variety of sources to see what the truth is, looking for correlation between social media and the impact it is having on teenagers' lives.

Data collection

After deciding on our topic- the impact of being online on children ages 10-15- our group split up to research and collect data that we each thought would be useful. Everyone looked into different areas, such as online bullying, mental health, the role of parents and schools, at what age children should have their phones, and both the positive and negative aspects of online use.

We collected a wide range of information, including government reports, academic studies, news articles, and surveys. However, because our topic involves minors, it was difficult to access direct medical or psychological data, as much of this information is protected for privacy and ethical reasons. To work around this, we relied on secondary data from published studies, school-based reports and anonymised survey results that still gave us insights without violating privacy.

Our data sources included a wide range of websites, but we also used API data and CSV files, allowing us to organise our findings into clear tables and charts we could use in our notebook. On our shared Miro board, we went through the data and made links between related data sets, for example, connecting online time with academic performance or bullying with mental health. We categorised our data into API, CSV and general, helping us analyse patterns and draw a clear conclusion for our report.

We used NewsAPI to collect data on news headlines. After generating a free API key and installing NewsAPI in our notebook, we encountered a limitation: the free tier only allows access to articles from the past 30 days. To expand our dataset, we implemented a loop that ran multiple relevant query terms, enabling us to gather a broader and more representative set of results.

Specifications And Design

After collecting all our resources, we began putting it all together to analyse. To begin with, we looked at a news API and used sentiment analysis to see whether articles about social media are more positive or negative.

The findings from this sentiment analysis shows that there is a positive association between social media and how it's reported in the news (Figures 1 and 2). This is more apparent in the descriptions than in the headlines; however, a positive opinion is not universal.

From Dataset 1, we gathered information about the time teenagers spend on social
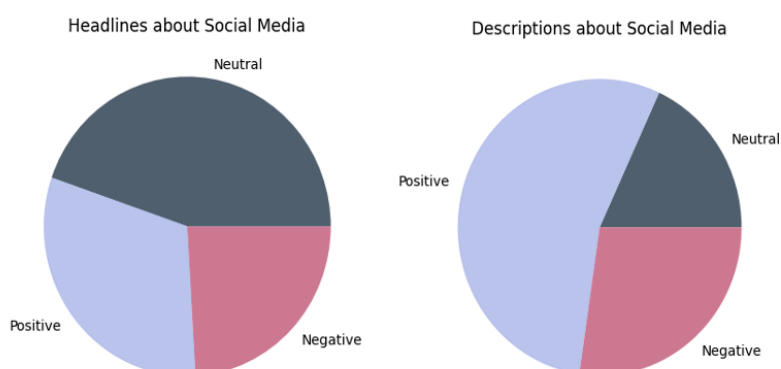
Figures 1 & 2 - News API and Sentiment Analysis

media and how that affects mental health-related variables. The categories we investigated were perceived loneliness, perceived anxiety, self-esteem and social media addiction as a representation of mental health.

There proves to be a slight increase in social media addiction with the more hours a day spent on these platforms, along with this, lower self-esteem is found (Figures 3 and 4). Furthermore, teenagers were less likely to feel anxious and lonely the more hours a day they spent on social media.
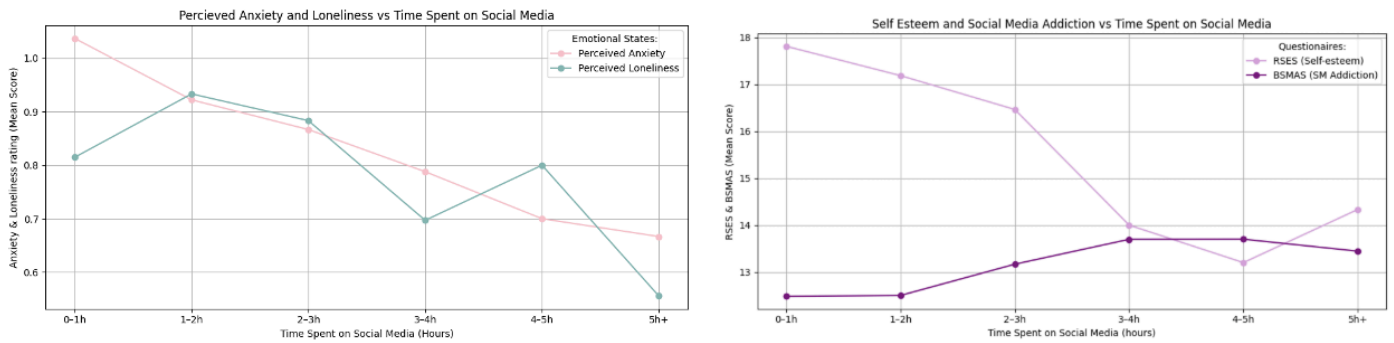

Figures 3 & 4 - Time Spent on SM & Mental Health Factors (Dataset 1)

We also looked at different social media platforms. As shown in Figure 5, WhatsApp and Instagram are shown to be the most popular platforms. Use of TikTok and Snapchat increases as time spent on social media increases.
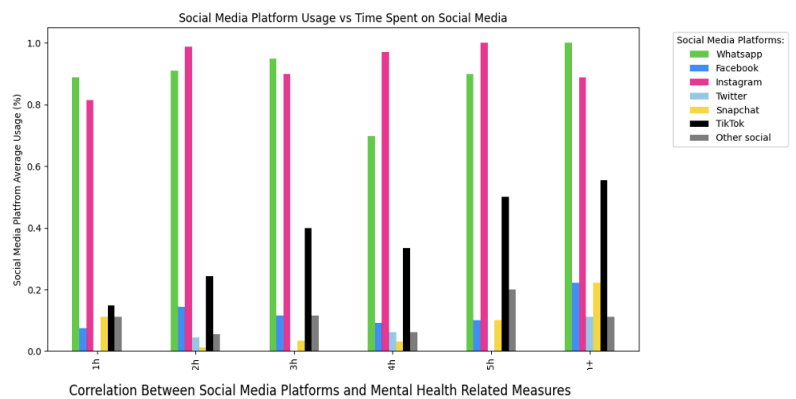
The correlation graph (Figure 6) suggests there are "other" platforms which cause the most effect on mental health variables. Twitter causes the least effect on mental health variables, which can be predicted as it is the least popular platform amongst teenagers. Another interesting finding is that there is a negative correlation between self-esteem and platforms related to sharing images and videos.
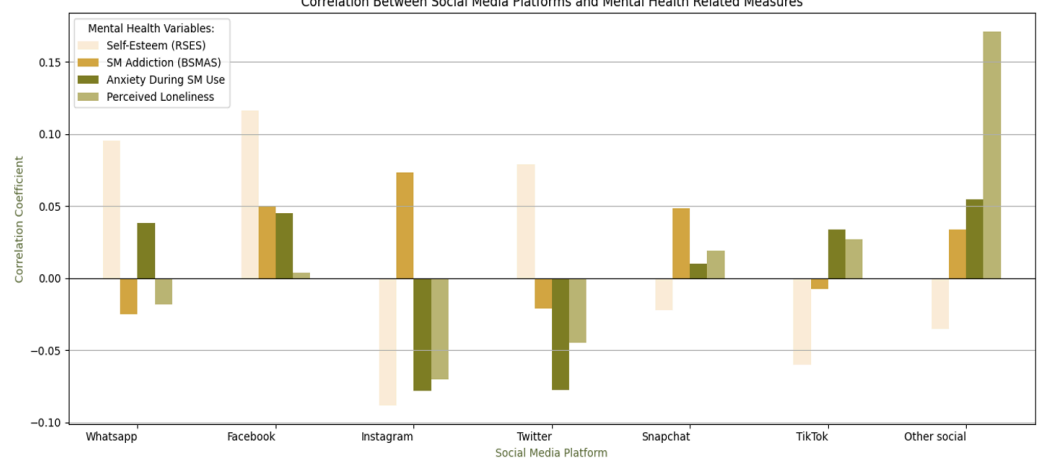
We then split these factors based on sex and age. It was found that the majority of males and females experienced the same level of anxiety and loneliness as always whilst using social media (Figures 7 and 8).
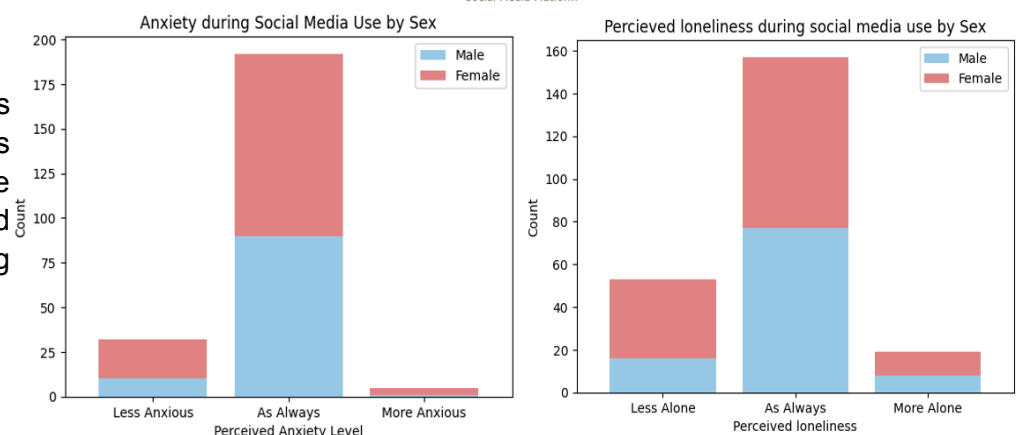

Figure 7 & 8 - Bar Charts showing anxiety & loneliness during SM use by Sex (Dataset 1)

In terms of age, we found that older teenagers generally experience slightly less anxiety but check their social media more (Figure 9).

Figure 9- Age and Mental Health Factors (Dataset 1)

We found a second dataset with the same age range and social media use, and merged both Dataset 1 and 2 to give us a larger collection of data. With this combined dataset, we could see that the data for 19-year-olds was skewing the graph (Figure 10), and so we removed the age group (Figure 11). By doing this, we were able to see that social media usage is consistent across all age ranges.
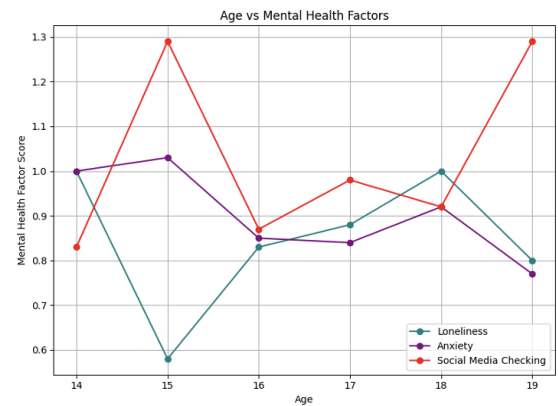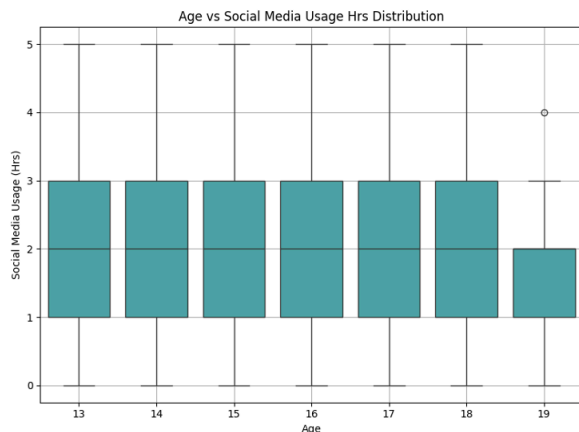

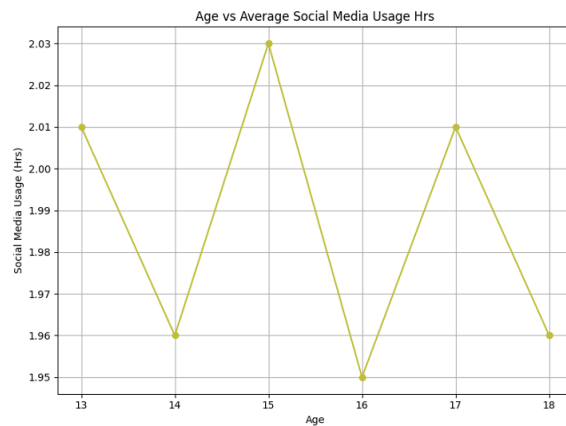Figure 10 – Social media usage based on age (merged data)

Figure 11 - SM usage & age after removing anomaly (merged data)

To see if these results are found only in teenagers, we analysed a third dataset looking at social media usage in adults. We chose to focus on an age range of 21-30. When calculating the mean time spent on social media (Figure 12), we can see that adults spend almost double the amount of time on social media than teenagers.

Mean time spent on social media (Hours)	Teenagers	Adults
Per Day	1.72	3.21
Per Week	12.04	22.47
Per Month	51.62	96.29

Figure 12 - Mean time on SM & age groups (Dataset 1 & 3)

To see how time spent on social media affects adults, a line graph was created (Figure 13). This line graph shows an increase in feelings of depression with the time spent on social media. This shows a similar trend to that felt by teenagers. It is important to note that Dataset 1 and Dataset 3 used different ways of asking how participants felt when using social media. In Dataset 1, they were asked about feeling anxious, whereas participants in Dataset 3 were asked about feeling depressed.
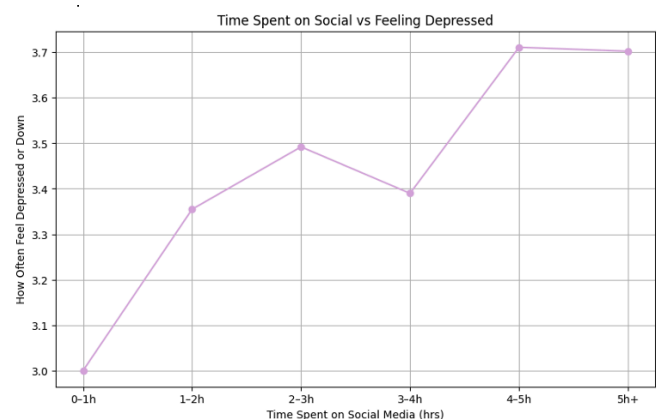

Figure 13 - Time spent on SM vs Mental Health (Dataset 3)

As we did in Dataset 1, we looked at different social media platforms and how they correlate with mental health (Figure 14). We were able to see that Facebook, Instagram and "Other" social media platforms were the most popular amongst adults. Snapchat had the highest positive correlation with feeling depressed, whereas Facebook had the least (Figure 15).
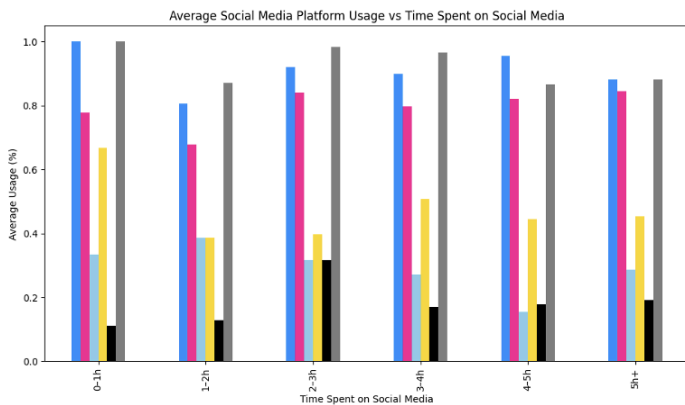
5

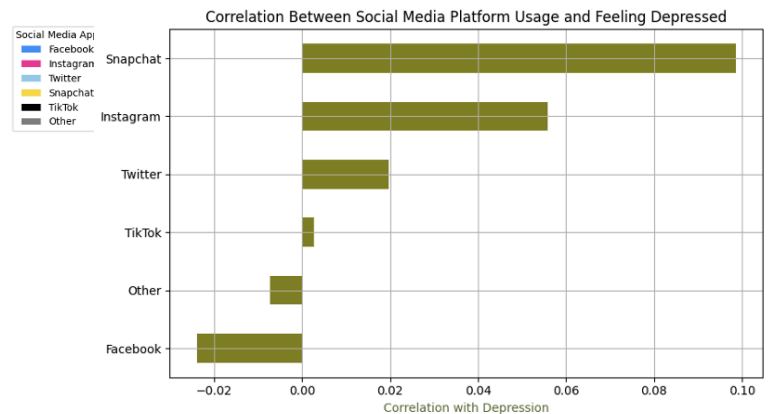Figure 14 - Platforms vs Time spent on SM (Dataset 3)

Figure 15 - Correlation between platforms & Depression (Dataset 3)

Figure 16 - Heat map for depression score by sex and time spent on SM (Dataset 3)

Finally, we investigated whether sex played a role in mental health and time spent on social media. We found that there was a 50% chance for both males and females to feel depressed in general.

The heat map, however, shows an increase in females feeling depressed the more hours they spend on social media.
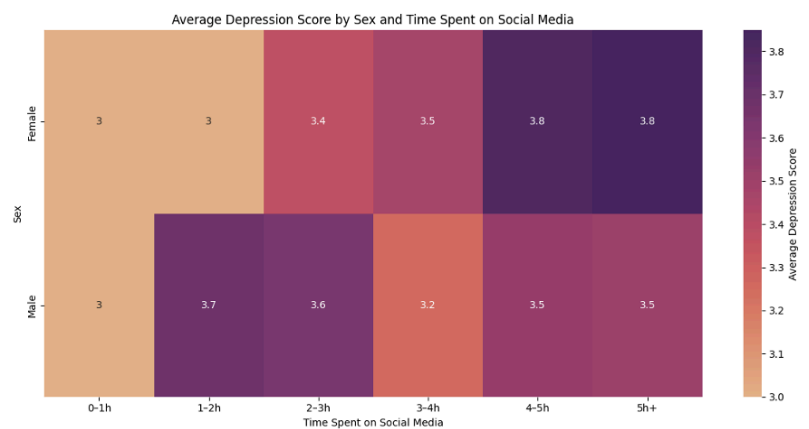
Implementation and Execution

Development approach and team member roles

For this project, we took a collaborative and flexible approach in assigning roles according to each team member's strengths. We conducted regular check-ins to adapt and assist each other where needed. Each member contributed to the coding and the analysis, and we utilised Miro, Zoom calls, Slack and pair programming to stay organised and support each other.

Bethany's role was to compile all of our code and make sure that everything ran smoothly and was easy to understand. She also worked on bar charts using Matplotlib to compare mental health differences by gender and teamed up with Rebecca to handle research of APIs and sentiment analysis.

Rebecca worked on researching and pulling data from the news API and helped to prepare it for analysis. She collaborated with Bethany on the sentiments and ensured that the articles we included were reputable and relevant to our topic.

Antonia focused on finding correlations in the datasets using Pandas, such as how time spent on social media relates to factors such as anxiety and self-esteem amongst teenagers.

Humairaa conducted descriptive analysis using Pandas. She calculated the mean time spent on social media per day, week, and month. She also analysed average mental health ratings—such as anxiety and loneliness—by grouping the data by age and gender to explore potential differences across demographic groups..

Helena contributed to the visualisation analysis and merging of datasets. She created line plots using Matplotlib to explore relationships between age and mental health scores, such as anxiety and loneliness

Fariha looked closely at the time spent on social media and how this affects various mental health indicators using line graphs with Matplotlib. She also used bar charts and line graphs to discover which social media apps are the most popular among teenagers.

Jodie focused on cleaning and preparing the datasets. She ensured that missing or inconsistent data were handled appropriately and that the datasets were ready for accurate analysis. She helped Rebecca to research and find possible APIs. She also managed the GitHub repository for the team.

Though each member had specific areas to focus on, we all supported each other throughout the process.

Tools and Libraries

We utilised a range of Python libraries to carry out our analysis. Pandas and NumPy allowed us to clean and process the data, and to calculate averages and group together variables such as age and gender. For the visualisations, we used Matplotlib and Seaborn to help us create line graphs, bar charts and box plots. We also used Textblob to run sentiment analysis on social media-related articles collected via the NewsAPI.

Our coding was conducted in Google Colab, which allowed us to collaborate in real time and comment on each other's work. We used GitHub to manage our files and track changes. For communication, we utilised Slack and Miro, which were useful for visual planning and organisation. We also used Excel early on in the project for initial data exploration.

Implementation process and challenges

We followed a week-by-week structure over 6 weeks with regular check-ins after each session (around 3-4 times a week). In week 1, we finalised our project idea, assigned roles and learned each other's strengths and weaknesses. Week 2 involved acquiring datasets and navigating which tools to use, such as Pandas, NumPy and Matplotlib. By week 3, we had begun cleaning the data and assigning coding tasks, with the remainder of our time spent coding, visualising, writing up findings and preparing for submission.

The entire process did involve a few challenges. Early on, we struggled to find an API that produced mental health data which focused on teenagers - most options were either too limited or required permissions, which were impractical considering our limited time frame. Upon group consultation, we decided to switch to NewsAPI, which allowed us to collect articles related to social media and mental health. Rebecca and Jodie handled the setup, and Rebecca worked on looping multiple queries and compiling the results into one CSV to get a broader range of sources.

We also had to rethink some of our visualisations. Some of our early ideas, such as scatter plots for certain comparisons, did not show the data enough. In other cases, we ran plots that technically worked, but were not visually useful. We had to be more flexible and choose better-fitting options, such as line and bar charts. These trials and errors were a good way for us to understand what works best for certain situations.

On a more technical side, some members of the team ran into issues with data transformations not saving, or code not performing the way we had expected, without throwing an error. GitHub was another learning curve we encountered, where merges and pulls were becoming convoluted. However, despite all this, the team worked exceptionally well together by assisting each other in debugging and using pair programming throughout. In many ways, our process and approach were agile in practice - we worked in iterations, adapted as we went along, and kept on improving based on what was working well and what wasn't. We also used agile principles as we were fluid when working. We held stand-ups after class as regular

meetings to keep up to date with what all members of the team were doing and to ensure we were on track to deliver value to the end project and users.

Conclusion

This report has investigated the complex relationship between social media usage and teenage mental health. It was highlighted that most young people report no considerable differences in anxiety or loneliness. The data shows that a higher daily use of social media is connected to higher social media addiction scores and lower self-esteem.

The data also presented that reasonable and higher levels of social media usage are connected to some users feeling less lonely and anxious. This suggests that digital connectivity provides mental support and comfort for some users.

Notable findings indicate that:
- 15 year olds reported the highest levels of anxiety and the lowest perceived loneliness, whilst being the most active social media users
- Females have a higher addiction score reported compared to males and have lower self-esteem scores. However, males reported higher levels of anxiety and perceived loneliness
- There is a correlation between social media usage on platforms which focus on image and video sharing and negative mental health reporting compared to text-based platforms
- When combining two datasets, time spent on social media in 14 to 18-year-olds was relatively consistent, typically varying by only ±0.05 to 2 hours

To understand the effects of social media on teenage mental health in greater depth, we recommend involving psychology and mental health professionals in future analysis so they can provide science-based tests when measuring mental health factors rather than participant-based self-assessments.

Recommendations for key stakeholders:
- Parents - Promote health screen time usage and form open discussions surrounding the impact of social media on mental health and confidence. Enforce usage limits to aid teenagers in managing social media usage.
- Teachers - Schedule discussions on social media well-being within the school curriculum to enable students to develop introspection, awareness and resilience.
- Governments - Allocate resources to education and research into digital media skills and literacy focusing on current as well as future platforms. Implement and enforce regulations for social media companies to manage addiction and create social media usage limits.
- Social Media Companies - Provide documentation and education on best practices for use of their respective platforms. Create design features which can allow users to limit usage as stated in proposed government regulations.

A balanced and well-informed approach will empower young people to navigate the current and future digital world whilst being able to benefit from its opportunities.

Appendices

Appendix A: Social media addiction and self-esteem score by sex and age (Dataset 1)

```
Average Mental Health Scores by Gender:

        Social Media Addiction Score  Self-Esteem Score
Gender
Female                        13.41              14.96
Male                          12.33              18.06

Average Mental Health Scores by Age:

      Social Media Addiction Score  Self-Esteem Score
Age
14                          14.17              19.83
15                          11.87              16.03
16                          12.61              16.91
17                          12.22              15.90
18                          13.17              17.45
19                          14.74              13.91
```

Appendix B: average mental health scores by sex (Dataset 1)

```
Average Mental Health Scores by Sex:

        Perceived anxiety during SM use  Perceived loneliness
Sex
Female                            0.86                  0.80
Male                              0.91                  0.92

Difference in Mental Health Scores between Male and Female:

Perceived anxiety during SM use    0.05
Perceived loneliness               0.12
dtype: float64
```

```
      RSES  BSMAS  Perceived loneliness
Sex
1    18.06  12.33                  0.92
2    14.96  13.41                  0.80
```

Appendix C: Average mental health scores by age (Dataset 1)

```
      Perceived loneliness  Perceived anxiety during SM use  SM checking
Age
14                    1.00                             1.00         0.83
15                    0.58                             1.03         1.29
16                    0.83                             0.85         0.87
17                    0.88                             0.84         0.98
18                    1.00                             0.92         0.92
19                    0.80                             0.77         1.29
```

Appendix D: Correlation scores between mental health factors and time spent on social media (Dataset 1)

```
RSES                             -0.25
Perceived loneliness             -0.13
Perceived anxiety during SM use  -0.23
BSMAS                             0.10
Name: Time spent on social media, dtype: float64
```
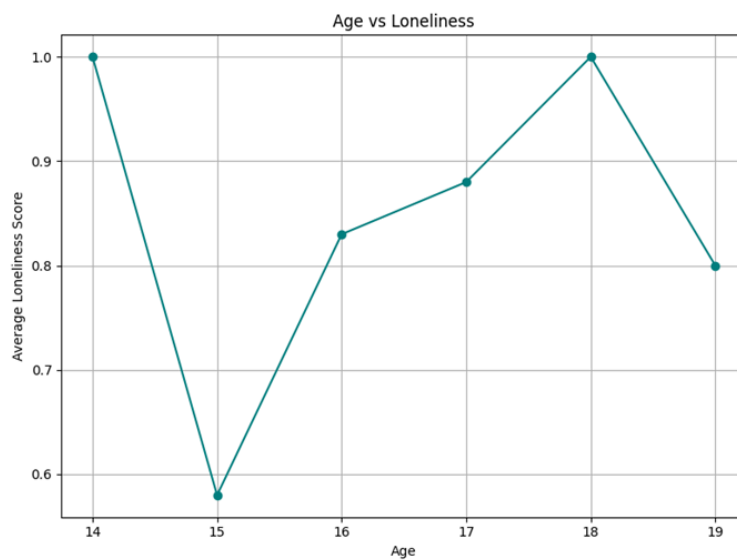
Appendix E: Number of teenagers in each time group (Dataset 1)

```
   Time Spent (hours/day)  Number of Teens  Percentage of Teens (%)
0                       0               27                    11.79
1                       1               90                    39.30
2                       2               60                    26.20
3                       3               33                    14.41
4                       4               10                     4.37
5                       5                9                     3.93
```
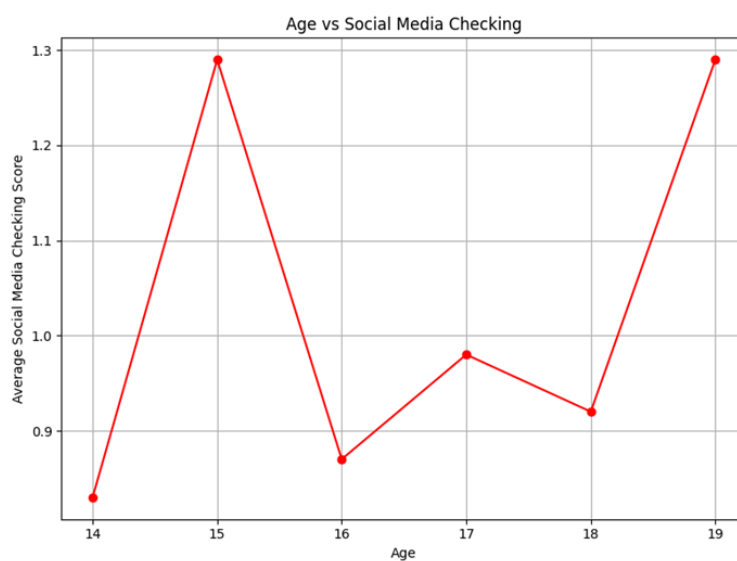
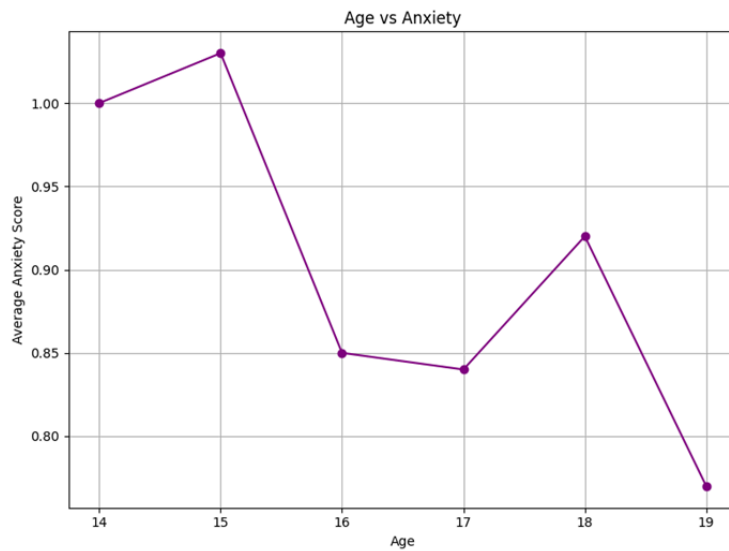### Appendix F: Levels of anxiety (Dataset 1)
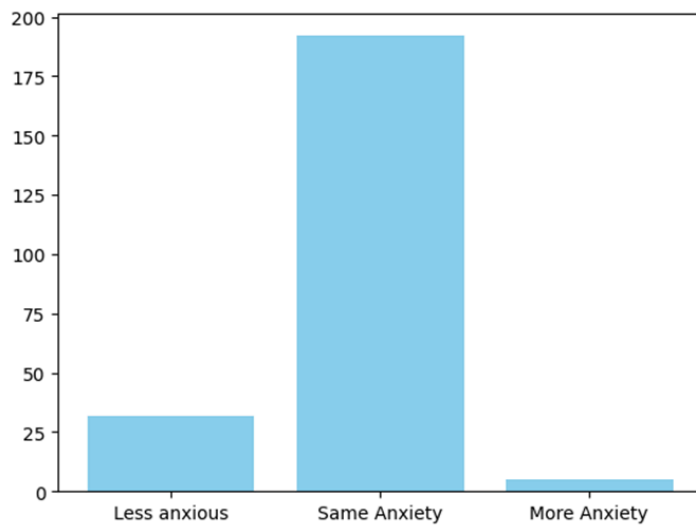
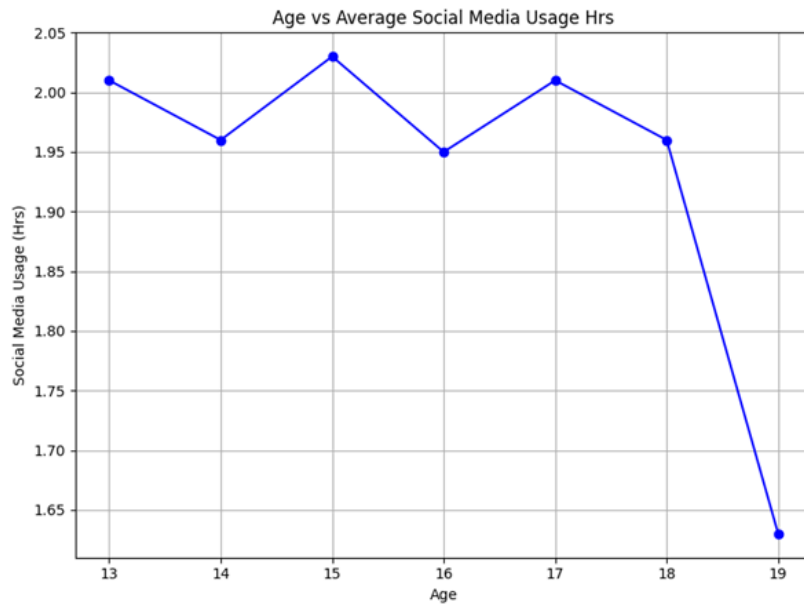Appendix G: Age vs average loneliness score (Dataset 1)

Appendix H: Age vs social media checking (Dataset 1)

Appendix I: Age vs anxiety score (Dataset 1)

Appendix J: Age vs average social media usage line graph including anomaly (Merged Dataset 1 & 2)


Appendix K: Average depression scores by age and sex (Dataset 3)

```
Average 'How Often Feel Depressed or Down' by Age:
Age
21.0    3.57
22.0    3.36
23.0    3.60
24.0    3.50
25.0    3.82
26.0    3.65
26.7    4.00
27.0    3.36
28.0    4.20
29.0    3.83
30.0    3.50
Name: How Often Feel Depressed or Down, dtype: float64

Average 'How Often Feel Depressed or Down' by Sex:
Sex
Female    3.58
Male      3.49
Other     2.50
Name: How Often Feel Depressed or Down, dtype: float64
```

```
Average 'How Often Feel Depressed or Down' by Sex:
Sex
Female    3.58
Male      3.49
Other     2.50
Name: How Often Feel Depressed or Down, dtype: float64

 Difference in Mental Health Scores between Male and Female:

-0.09
```

Appendix L: Pie chart of average depression scores by sex (Dataset 3)

Average Depression Score by Sex

Male 49.4% Female 50.6%
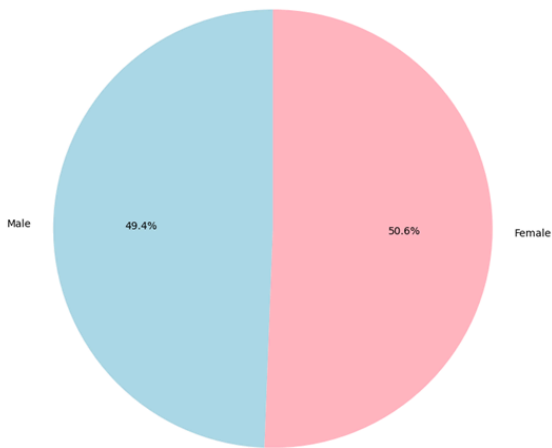
References

1. Kim, H.H., 2017. The impact of online social networking on adolescent psychological well-being (WB): A population-level analysis of Korean school-aged children. International Journal of Adolescence and Youth, 22, pp.364–376.
2. Khalaf, A.M., Alubied, A.A., Khalaf, A.M. and Rifaey, A.A., 2023. The impact of social media on the mental health of adolescents and young adults: A systematic review. Cureus, 15(8), p.e43210. Available at: https://doi.org/10.7759/cureus.43210
3. Pantic, I., 2014. Online social networking and mental health. Cyberpsychology, Behavior, and Social Networking, 17, pp.652–657.
4. https://www.bbc.co.uk/news/articles/cjewne81lq4o
5. https://commonslibrary.parliament.uk/research-briefings/cdp-2024-0103/
6. https://www.channel4.com/press/news/channel-4-orders-two-part-special-swiped-school-banned-smartphones-boldprint-studios
7. https://cls.ucl.ac.uk/heavy-social-media-use-linked-to-depression-in-young-teens-new-study-shows/
8. https://ourworldindata.org/social-connections-and-loneliness
9. https://www.theguardian.com/global-development/2023/nov/16/who-declares-loneliness-a-global-public-health-concern