Exploring the Effects of Social Media Addiction on Student Well-Being

Amanda J. Reinhart

Western Governors University

Table of Contents

[A. Proposal Overview 4](#_Toc2137083979)

[A.1 Research Question or Organizational Need 4](#_Toc1815560889)

[A.2 Context and Background 4](#_Toc1384268585)

[A.3 and A3A Summary of Published Works and Their Relation to the Project 4](#_Toc2079458267)

[Review of Work 1 5](#_Toc694593955)

[Review of Work 2 5](#_Toc475195288)

[Review of Work 3 6](#_Toc1156669356)

[A.4 Project Deliverables 6](#_Toc374341712)

[A.5 Benefits and Support of Decision-Making Process 7](#_Toc1116633891)

[B. Data Analytics Project Plan 7](#_Toc1561262715)

[B.1 Goals, Objectives, and Deliverables 7](#_Toc1217623361)

[B.2 Scope of Project 10](#_Toc427690838)

B.2.A Included in Project Scope 10

[B.2.B Not included in Project Scope 11](#_Toc2036717847)

B.3 Standard Methodology 12

[B.4 Timeline and Milestones 13](#_Toc1049840195)

[B.5 Resources and Costs 13](#_Toc313275103)

[B.6 Criteria for Success 14](#_Toc1337049307)

[C. Design of Data Analytics Solution 15](#_Toc764513894)

[C.1 Hypothesis 15](#_Toc1169566000)

[C.2 and C.2.A Analytical Method 15](#_Toc1622461191)

[C.3 Tools and Environments 16](#_Toc53928856)

[C.4 and C.4.A Methods and Metrics to Evaluate Statistical Significance 16](#_Toc1386867513)

[C.5 Practical Significance 17](#_Toc157862960)

[C.6 Visual Communication 17](#_Toc1263386613)

[D. Description of Dataset 18](#_Toc34323025)

[D.1 Source of Data 18](#_Toc1867631323)

[D.2 Appropriateness of Dataset 18](#_Toc418252330)

[D.3 Data Collection Methods 19](#_Toc1665164903)

[D.4 Observations on Quality and Completeness of Data 19](#_Toc57902398)

[D.5 and D.5.A Data Governance, Privacy, Security, Ethical, Legal, and Regulatory Compliances 19](#_Toc1649406101)

[References 21](#_Toc1322159537)

# A. Proposal Overview

## A.1 Research Question or Organizational Need

## How do social media addiction levels vary across countries and academic levels, and how are higher addiction levels associated with students’ mental health, conflict frequency, and perceived academic impact?

## A.2 Context and Background

## Social Media is an ever present and growing force in the lives of students around the world. While social media provides an outlet for connection, self-expression, and information sharing, it can also be a negative influence. Excessive social media use can lead to addiction and can negatively impact mental health, relationships, and academic performance. Students are especially vulnerable as they navigate personal and educational development as young adults. Understanding the extent and nature of social media addiction in students is crucial for all who aim to support student well-being and academic success.

## This project seeks to explore the relationship between students’ social media addiction scores and various indicators of personal well-being, including mental health, sleep patterns, social conflicts, and perceived academic impact. By examining how these behavioral and emotional markers relate to addiction levels, the analysis aims to uncover meaningful trends. These insights could provide a valuable foundation for future efforts aimed at improving students’ relationships with social media and supporting healthier, more balanced usage patterns.

## A.3 and A3A Summary of Published Works and Their Relation to the Project

### Social media addiction: Its impact, mediation and intervention

This study examines the relationship between social media addiction, mental health, and academic performance among college students. While it found that viewing and editing one’s own profile may temporarily increase self-esteem, extended social media use negatively impacts other aspects of mental health, including overall self-esteem. Social media addiction was linked to higher rates of anxiety and depression and lower academic performance. The study consisted of two parts: the first explored the correlations between addiction and behavioral indicators, and the second implemented an experimental intervention based on those findings. The intervention, which involved cognitive restructuring, reminder cards, and diary-keeping, successfully reduced social media addiction and improved mental health outcomes (Hou, Xiong, Jiang, Song, & Wang, 2019).

This work supports our hypothesis and exploration of how social media addiction affects students’ mental health and academic performance, which are two of the key outcomes we intend to evaluate. The structure of the study, with a focus on identifying behavioral indicators followed by testing an intervention, validates our approach of exploring these variables through correlation and visual analysis. Although our project does not implement an intervention, the first phase of this study closely mirrors our own goal of identifying trends and relationships between addiction scores and well-being metrics. The success of their multi-stage approach also suggests future opportunities for expanding on this project’s findings in practical settings.

### Investigating social media addiction and impact of social media addiction, loneliness, depression, life satisfaction, and problem-solving skills on academic self-efficacy and academic success among university students.

This study sought to determine the level of social media addiction and its effects on academic success among university students. It examined variables such as social media addiction, loneliness, depression, life satisfaction, problem-solving skills, and academic self-efficacy. The findings revealed that higher levels of social media addiction and depression were associated with lower academic performance and self-efficacy. Conversely, better problem-solving skills and greater life satisfaction correlated positively with academic success (Aslan & Polat, 2024).

This research supports our hypothesis by highlighting the negative impact of social media addiction on students’ academic performance. The study’s comprehensive approach, examining multiple psychological variables, aligns with our project’s aim to explore the multifaceted effects of social media addiction.

### Social media and mental health in students: a cross-sectional study during the Covid-19 pandemic

## This study looked at social media use, mental health, and academic performance among students during the COVID-19 pandemic. It found that while social media was sometimes used as a coping mechanism during isolation, overuse was linked to increased anxiety, depression, and poor academic outcomes. The results suggest that excessive reliance on social media in times of stress may amplify existing challenges rather than alleviate them (Wang et al., 2023).

This work supports the idea that social media addiction can have compounding effects on mental health and academic success, especially when students rely heavily on it during stressful periods. It strengthens the case for examining how addiction scores relate to overall well-being, and why those relationships matter in both day-to-day and high-stress environments.

## A.4 Project Deliverables

The deliverables of this analysis will include several components designed to clearly and effectively communicate the results. A Jupyter Notebook will be the primary working document, containing all code, analysis steps, visualizations, and accompanying commentary. This notebook will include cleaned and organized data, exploratory data analysis, statistical testing, and relevant visualizations such as bar graphs, scatter plots with regression lines, and histograms.

In support of the notebook, a written report will also be provided. This report will summarize the methodology, key findings, and conclusions in a more formal format suitable for broader distribution. This report will highlight major trends uncovered in the data, provide an interpretation of the statistical findings, and address the original research question.

Additionally, a video presentation will be provided to summarize the project and findings from the analysis. The video will outline the research question the analysis explored, demonstrate the functionality of the code utilized, and outline the findings and implications of the analysis.

Together, these deliverables will present a complete picture of how social media addiction scores relate to student well-being indicators and will offer insight for stakeholders interested in student mental health and academic success.

## A.5 Benefits and Support of the Decision-Making Process

This project will provide the data-based foundation for any future interventions needed for students and their mental well-being as it relates to social media addiction. The anticipated outcome of this project is a clearer understanding of how social media addiction levels are associated with students’ mental health, social relationships, and perceptions of academic performance. With the resulting data in hand, stakeholders such as educators, mental health counselors, and parents will be better equipped to handle interventions with the student(s) in their lives to counter-act the effects of social media addiction.

# B. Data Analytics Project Plan

* Goal: The goal of this project is to analyze how students’ social media addiction scores correlate with key behavioral and well-being indicators to assess the broader impact of social media on student life.
  + Objective 1.1: Determine the relationship between Social Media Addiction and Average Sleep Per Night
    - Deliverable 1.1.1: A scatterplot with a line of best fit to visually represent the relationship between the two variables in our high-usage dataset
    - Deliverable 1.1.2: OLS (Ordinary Least Squares) Regression Results to statistically describe the relationship between the two variables in our high-usage dataset
    - Deliverable 1.1.3: A scatterplot with a line of best fit to visually represent the relationship between the two variables in our low-usage dataset
    - Deliverable 1.1.4: OLS Regression Results to statistically describe the relationship between the two variables in our low-usage dataset
  + Objective 1.2: Determine the relationship between Social Media Addiction and Mental Health Score
    - Deliverable 1.2.1: A scatterplot with a line of best fit to visually represent the relationship between the two variables in our high-usage dataset
    - Deliverable 1.2.2: OLS Regression Results to statistically describe the relationship between the two variables in the same group
    - Deliverable 1.2.3: A scatterplot with a line of best fit to visually represent the relationship between the two variables in our low-usage dataset
    - Deliverable 1.2.4: OLS Regression Results to statistically describe the relationship between the two variables in the same group
  + Objective 1.3: Determine the relationship between Social Media Addiction and Number of Conflicts on Social Media
    - Deliverable 1.3.1: A scatterplot with a line of best fit to visually represent the relationship between the two variables in our high-usage dataset
    - Deliverable 1.3.2: OLS Regression Results to statistically describe the relationship between the two variables in the same group
    - Deliverable 1.3.3: A scatterplot with a line of best fit to visually represent the relationship between the two variables in our low-usage dataset
    - Deliverable 1.3.4: OLS Regression Results to statistically describe the relationship between the two variables in the same group
  + Objective 1.4: Determine the relationship between Social Media Addiction and perceived Affect on Academic Performance
    - Deliverable 1.4.1: A bar chart representing the average social media addiction score for students who believe that their social media usage affects their academic performance and for those who do not in our high-usage dataset
    - Deliverable 1.4.2: A bar chart representing the average social media addiction score for students who believe that their social media usage affects their academic performance and for those who do not in our low-usage dataset
  + Objective 1.5: Determine the relationship between Social Media Addiction and Academic Level
    - Deliverable 1.5.1: a bar chart representing the average social media addiction score for students at each academic level surveyed in our high-usage dataset
    - Deliverable 1.5.2: a bar chart representing the average social media addiction score for students at each academic level surveyed in our low-usage dataset
* Goal 2: Communicate findings from the analysis in a clear and actionable format for stakeholder decision making
  + Objective 2.1: Prepare a report that summarizes the project’s background, methodology, results, and recommendations
    - Deliverable 2.1.1: A written report, provided in .docx format
  + Object 2.2: Create a visual presentation to clearly convey the findings and recommendations to stakeholders
    - Deliverable 2.2.1: A recorded presentation summarizing the project and results

## B.2 Scope of Project

## B.2.A Included in Project Scope

This project includes the analysis of social media addiction scores in students in relation to behavioral and well-being indicators that include: average hours of sleep per night, number of conflicts on social media, mental health score, perceived effect on academic performance, and addiction level. The analysis will be split into two groups based on average daily usage to allow for comparative analysis between high-usage and low-usage student groups.

The analysis will be completed within JupyterLab using Python, Pandas, Matplotlib, Seaborn, and StatsModels. The dataset utilized will be manually downloaded from Kaggle.com and loaded into JupyterLab. The dataset will be cleaned and evaluated to complete an analysis in line with the goal for the project, as outlined in B1.

Based on the completed analysis, a written report will be provided for stakeholders along with a video presentation. The report and presentation will summarize the project and its execution, the methodologies utilized during the analysis, the results of the analysis, and recommend two courses of action based on the results of the analysis.

### B.2.B Not included in Project Scope

Developing and delivering the recommended courses of action based on the analysis is outside of the scope of this project. This project is intended to gather the necessary data and make suggestions for stakeholders to take action on the impacts of social media addiction, but not to deliver those action items directly.

## B.3 Standard Methodology

This project will follow the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology. This approach provides a clear, flexible framework that aligns well with the structure of this analysis and the tools being used.

* Business Understanding:  
  This phase will involve identifying the research question and defining the goal of the project: to examine how social media addiction scores relate to indicators of student well-being. Key objectives and deliverables will be outlined to support this goal, as detailed in section B1.
* Data Understanding:  
  The dataset will be manually downloaded from Kaggle, imported into JupyterLab, and explored to understand the structure of the data and the meaning of each variable. This step will include assessing what cleaning will be necessary.
* Data Preparation**:**  
  The dataset will be cleaned and transformed to make it usable for analysis. Unnecessary columns will be dropped and categorical variables will be standardized. The dataset will also be split into two groups based on average daily usage to allow for comparison between students who use social media more frequently and those who use it less frequently.
* Modeling:  
  OLS Regression will be used to explore the relationships between addiction scores and each behavioral or well-being indicator. Scatterplots with lines of best fit will be created to visually represent each relationship. These steps will be repeated for both the high-usage and low-usage datasets to allow for meaningful comparison.
* Evaluation:  
  Regression output will be evaluated using p-values and R-squared values to determine statistical significance and the strength of each relationship. These results, combined with visualizations and group comparisons, will inform the overall conclusions drawn from the analysis.
* Deployment:  
  A final report and video presentation will be created to summarize the project, share the results of the analysis, and recommend two possible actions based on those results. These materials will be intended for stakeholders and will serve as the final deliverables of the project.

## B.4 Timeline and Milestones

|  |  |  |  |
| --- | --- | --- | --- |
| **Milestone or deliverable** | **Duration**  **(hours or days)** | **Projected start date** | **Anticipated end date** |
| Establish Requirements for Analysis Project | 2 days | 6/1/2025 | 6/2/2025 |
| Gather and Clean Data | 1 day | 6/3/2025 | 6/3/2025 |
| Evaluate and Analyze Data | 2 days | 6/4/2025 | 6/5/2025 |
| Compile Written Analysis Report | 2 days | 6/6/2025 | 6/7/2025 |
| Record Analysis Presentation | 1 day | 6/8/2025 | 6/8/2025 |

## B.5 Resources and Costs

* Python: Free
  + Associated libraries (i.e. Panadas, NumPy, MatplotLib, etc): Free
* JupyterLab: Free
* Word Processor: Free through WGU
* Panopto Video Recording Software: Free through WGU
* Labor: Free

The resources utilized through this project are made up of free and open-source software languages, free versions of software, and software with access included in WGU admission. The labor hours for this project are my own, and will be completed in my own time, and therefore, free. Completion of this project will not incur any costs that would not have already been covered by other needs.

**B.6 Criteria for Success**

The success of this project will be determined by the completion and accuracy of the planned statistical analyses. The analysis will be considered successful if all planned visualizations and statistical tests are correctly executed for both the high-usage and low-usage datasets, including scatterplots, bar charts, and OLS regression models. Additionally, success will be measured by whether or not the statistical tests return results that can be clearly interpreted using an established alpha level. Ultimately, if the analysis is carried out as planned and gives us a solid basis for interpreting the relationships between variables, the project will be considered successful.

# C. Design of Data Analytics Solution

## C.1 Hypothesis

## Students with higher social media addiction scores will report greater negative impacts on overall well-being, including lower mental health scores, more frequent social conflicts, and a higher likelihood of perceiving social media as detrimental to their academic performance. Additionally, higher addiction levels are expected to be more prevalent among students at lower academic levels.

Null Hypothesis: There is no relationship between students’ social media addiction scores and their behavioral or well-being indicators, including sleep, mental health, social conflicts, or perceived academic impact. Academic level will have no effect on reported addiction levels.

## C.2 and C.2.A Analytical Method

This project will use a combination of OLS regression and scatterplot and bar chart visualizations to analyze the relationship between students’ social media addiction scores and various well-being indicators. OLS regression and the scatterplot visualizations will be used to evaluate the strength and direction of the relationship between variables. Bar charts will be used to compare addiction scores by perceived impact on academic performance and by academic level, allowing for a clear visual comparison of average scores across categorical responses. For all of these, the analysis will be run twice, once on our high-usage dataset and once on the low-usage dataset to compare the relationships between groups.

OLS regression is an appropriate choice for this analysis as it provides a clear summary of how much variation in the dependent variable can be explained by each independent variable. It also provides p-values, which will help determine if the relationships are statistically significant. Bar charts are appropriate for the categorical variables analyzed in this dataset because they clearly illustrate average differences in addiction scores across these groupings.

## C.3 Tools and Environments

The primary environment in which the analysis will be completed is JupyterLabs. A Jupyter notebook is ideal for this project as it provides a single location to house all code, visualizations, and commentary. This clean and efficient vehicle for completing our analysis also has the added perk of being available free to the public, keeping project costs low.

Python is the language of choice utilized for this project. Python is frequently used for data analysis due to its flexibility and extensive range of open-source libraries. Of the available libraries, this project utilizes Pandas for working with dataframes, Matplotlib and Seaborn for creating our visualizations, and Statsmodels for completing our OLS Regression.

## C.4 and C.4.A Methods and Metrics to Evaluate Statistical Significance

For each relationship explored in this analysis, the null hypothesis assumes no connection between a student’s social media addiction score and the well-being or behavioral indicator being evaluated. For the continuous variables (average sleep per night, mental health score, and number of conflicts on social media), OLS regression will be used to determine whether a statistically significant relationship exists. Each regression will produce a p-value, and we’ll use an alpha level of 0.05 to determine significance. If the p-value is below 0.05, we’ll reject the null hypothesis and consider the relationship statistically meaningful.

OLS regression is a good fit here because we’re examining linear relationships between one numerical outcome and several other numeric variables. This method provides both visual and statistical insight into whether changes in well-being indicators are tied to changes in addiction levels. Using a standard alpha level of 0.05 keeps the significance threshold consistent and clear.

For the bar chart comparisons (academic level and perceived impact on academic performance), visual differences in average addiction scores will be the focus. These won’t include statistical testing, but will still help us identify useful patterns worth noting in the context of other findings.

**C.5 Practical Significance**

While statistical significance tells us whether relationships are likely due to chance, practical significance focuses on whether those relationships actually matter in real life. In this project, practical significance will be assessed by looking at the strength of the relationships between addiction scores and each well-being factor. Stronger relationships, especially those with consistent patterns across both high- and low-usage datasets, will be considered more impactful.

For example, if students with higher addiction scores consistently report lower mental health scores or fewer hours of sleep, that insight becomes meaningful for school counselors, educators, or policymakers aiming to support student well-being. The goal is not just to confirm if relationships exist, but to provide findings that can be acted on, whether that’s starting conversations around digital boundaries, building awareness campaigns, or prioritizing support services. If the results help highlight clear trends that could drive real decisions, the analysis will be considered practically significant.

## C.6 Visual Communication

In this analysis, various visuals will be utilized throughout the project to help make decisions and identify trends. Initially, a histogram grouping average daily social media usage by country will be utilized to give us our two data groups for comparative analysis.

Throughout the primary portion of the analysis, two main visualizations will be utilized. Scatterplots with lines of best fit will be utilized alongside our OLS regression calculations to assist in determining the strength and direction of relationships between our dependent variable and various independent variables. Bar charts will be used to show the differences in average addiction score across several categorical variables to help identify trends worth noting in the context of our other findings.

All visualizations will be created alongside our code and commentary in the JupyterLab notebook that will be utilized for our analysis. Python libraries Matplotlib and Seaborn will be utilized in the creation of these visuals.

# D. Description of Dataset

## D.1 Source of Data

## The necessary dataset will be derived from the complete dataset found [here](https://www.kaggle.com/datasets/adilshamim8/social-media-addiction-vs-relationships), from Kaggle. The title of the data is Students' Social Media Addiction. The dataset consists of 1 file, with 13 columns and over 700 student responses

## D.2 Appropriateness of Dataset

The dataset selected for this project will be well-suited to support the goals of the analysis. It includes a wide range of variables directly tied to student behavior, well-being, and social media use, all of which align with the research question. The inclusion of a standardized social media addiction score will allow for consistent measurement of addiction levels across respondents, while additional fields such as mental health, sleep, conflict, and academic experience will provide the necessary context to explore how that addiction may show up in day-to-day life.

The dataset will also be large enough to support meaningful comparison between groups (i.e., high vs. low daily usage), making it possible to evaluate whether patterns hold across different levels of user behavior. The structure of the data will support the planned regression and visual analysis methods without requiring extensive transformation. Overall, this dataset will align well with the scope of the project and allow for a focused, well-rounded approach to meeting the stated objectives.

**D.3 Data Collection Methods**

## The data will be downloaded directly from Kaggle.com in CSV format. The CSV file will then be manually loaded into the analysis environment for cleaning, exploration, and modeling. D.4 Observations on Quality and Completeness of Data

The dataset is complete, with no null values, no missing records, etc., making it ideal for analysis for this project. The dataset consists of survey results of students enrolled in high school, undergraduate, or graduate programs in multiple countries across the globe. The questions were based on valid social media and relationship-conflict scales. Participants voluntarily opted into the survey. Additionally, there are multiple data quality control checks on the dataset, including data validation, removal of duplicate entries, and anonymization (Shamim, 2025).

## D.5 and D.5.A Data Governance, Privacy, Security, Ethical, Legal, and Regulatory Compliance

This project will use a publicly available dataset downloaded from Kaggle. The dataset will not be altered beyond necessary cleaning and preparation steps required for analysis. I will follow consistent documentation practices, apply clear file naming conventions, and keep all files and outputs organized in a structured JupyterLab workspace to maintain good data governance throughout the project.

Since the dataset is fully anonymized and contains no personally identifiable information, there are no direct privacy concerns. No new data will be collected, and no individual participants will be contacted during the project. Still, the data will be stored locally on a password-protected device to ensure that it remains secure and access is limited.

All analysis will be completed in an offline environment. The dataset will not be uploaded to cloud-based services or shared publicly beyond what is required for the capstone submission, minimizing exposure to any external threats. These precautions are appropriate given the low sensitivity of the data.

This dataset does not require IRB approval or special regulatory clearance, but ethical handling of the data remains a priority. The analysis and resulting conclusions will stay within the limits of what the data supports, avoiding overgeneralization. All data use will follow Kaggle’s terms of service, and findings will be communicated responsibly to reflect respect for the populations represented.

# References

Shamim, A. (2025). *Students’ Social Media Addiction*. Kaggle.com. https://www.kaggle.com/datasets/adilshamim8/social-media-addiction-vs-relationships

Hou, Y., Xiong, D., Jiang, T., Song, L., & Wang, Q. (2019). Social media addiction: Its impact, mediation, and intervention. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 13*(1), Article 4. <https://cyberpsychology.eu/article/view/11562>

Aslan, I., & Polat, H. (2024). Investigating social media addiction and impact of social media addiction, loneliness, depression, life satisfaction and problem-solving skills on academic self-efficacy and academic success among university students. *Frontiers in Public Health*, *12*, 1359691. <https://doi.org/10.3389/fpubh.2024.1359691>

Nazari, A., Hosseinnia, M., Torkian, S. *et al.* Social media and mental health in students: a cross-sectional study during the Covid-19 pandemic. *BMC Psychiatry* 23, 458 (2023). <https://doi.org/10.1186/s12888-023-04859-w>