in this research can you explain me the working of it how author perfrom data analysis i am confused

Reasoned about data analysis process for 12 seconds

The author’s data analysis is split into two main parts:

1. **Baseline Analysis:**
   * Participants first complete a questionnaire that records their social media use—such as the number of platforms they use, the average time spent on each, and how many friends they have.
   * The author uses descriptive statistics (like means and standard deviations) to summarize this baseline information and to profile the participants’ typical social media behavior.
2. **Outcome Analysis for Mental Health Measures:**
   * Participants’ scores on established psychological scales are then examined. These scales include:
     + The Kessler Psychological Distress Scale (measuring distress),
     + Rosenberg’s Self-Esteem Scale (measuring self-esteem), and
     + The Patient Health Questionnaire-9 (measuring depressive symptoms).
   * The study sets up several hypotheses (seven in total) comparing groups based on their social media use (e.g., Facebook users, Instagram users, and those not using social media).
   * For each hypothesis, a one-way ANOVA is performed. This test checks if there are overall significant differences among the groups for each mental health outcome.
   * If the ANOVA yields a significant omnibus F-test, indicating that at least one group differs from the others, the author then uses a Bonferroni post-hoc test. This follow-up test pinpoints exactly which groups differ from each other.

In summary, the analysis first characterizes how much and in what ways participants use social media, and then it statistically compares the mental health outcomes between different usage groups using one-way ANOVA and, if needed, Bonferroni tests to explore specific group differences. This method helps determine whether social media use is linked with increased psychological distress, lower self-esteem, or more depressive symptoms.

o3-mini

**You said:**

how is he scoring in the first baseline analysis on 3 questionaries , see in measures there are different weightage of points ranging 1-5 on each option but on 3quesrions thre is no weightage so how is her scoring it?

Reasoned about scoring for baseline analysis for 10 seconds

In the baseline analysis, the questionnaire isn’t scored using a weighted scale like the mental health measures. Instead, it simply collects descriptive data about each participant’s social media use—for example, which platforms they use, how much time they spend on them, and how many friends they have. This raw data is then summarized with descriptive statistics (averages, frequencies, etc.) rather than being converted into a weighted score.

In contrast, the later measures (such as the Kessler Psychological Distress Scale, Rosenberg’s Self-Esteem Scale, and the PHQ-9) use Likert-type scales (with weightages typically ranging from 1 to 5) because they are designed to quantify latent constructs like distress or self-esteem. The baseline questionnaire is simply about capturing usage behavior, so it doesn’t require that kind of scoring.

This approach allows the researcher to profile the participants’ social media habits before examining how those habits relate to the mental health outcomes.

o3-mini

**You said:**

so how is he measuring then if he does not have user data , explain me with a real example of a user

Reasoned about data measurement in analysis for 8 seconds

Imagine a participant named Alex who fills out the baseline questionnaire. Alex’s responses might look like this:

* **Social Media Accounts:** Alex uses Facebook and Instagram.
* **Time Spent Daily:** He reports spending 60 minutes on Facebook and 30 minutes on Instagram.
* **Number of Friends:** He indicates he has 150 Facebook friends and 100 Instagram followers.

Even though these questions don’t use a 1–5 weighted Likert scale (like the mental health measures do), Alex’s responses still provide valuable numerical data. The researcher collects these raw numbers from every participant. Then, using descriptive statistics, the researcher can calculate, for example, the average time spent on each platform across all users or the typical number of friends reported. This baseline data helps form a profile of social media behavior, which can later be related to the psychological scales that do use weighted scores.

So, while the baseline questionnaire isn’t “scored” in the same way as the mental health measures, it still captures essential user data that is analyzed quantitatively.

o3-mini

**You said:**

explain this example further in measures too

Reasoned about baseline data and measures for 5 seconds

Let's extend Alex’s example to include the standardized measures:

**Baseline Questionnaire**

Alex fills out the baseline questionnaire with questions like:

* **Social Media Accounts:** He indicates that he has accounts on Facebook and Instagram.
* **Time Spent Daily:** He reports spending 60 minutes on Facebook and 30 minutes on Instagram.
* **Number of Friends:** He states he has 150 Facebook friends and 100 Instagram contacts.

This information gives a descriptive snapshot of his social media behavior, but it isn’t “scored” with weights—it’s raw numerical data used to characterize his usage.

**Standardized Measures (Scored with Weightage)**

1. **Kessler Psychological Distress Scale (K10):**  
   Alex is asked several questions about how often he felt certain negative emotions (e.g., “nervous” or “hopeless”) over the past month.
   * **Response Options:** These might range from 1 (indicating “never”) to 5 (indicating “always”).
   * **Scoring Example:**  
     Suppose Alex’s responses on 5 items are: 3, 4, 2, 4, and 3. His total distress score would be the sum: 3+4+2+4+3 = 16.
   * **Interpretation:** A higher total score suggests greater psychological distress.
2. **Rosenberg Self-Esteem Scale (RSES):**  
   This scale includes statements like “I feel that I have a number of good qualities.”
   * **Response Options:** Answers are given on a Likert scale (for example, 1 = “strongly disagree” to 5 = “strongly agree”).
   * **Scoring Example:**  
     If Alex answers 10 items with scores such as 4, 3, 4, 2, 3, 4, 3, 4, 3, and 4, his overall self-esteem score would be the sum of these responses. (Some items might be reverse-scored, but the process is similar.)
   * **Interpretation:** A higher score usually indicates higher self-esteem, although the specific interpretation depends on the scale’s guidelines.
3. **Patient Health Questionnaire-9 (PHQ-9):**  
   Alex answers questions about depressive symptoms like “Little interest or pleasure in doing things.”
   * **Response Options:** Each item is rated from 0 (not at all) to 3 (nearly every day).
   * **Scoring Example:**  
     If his responses are 1, 2, 1, 2, 1, 1, 0, 2, and 1, the total score would be 1+2+1+2+1+1+0+2+1 = 11.
   * **Interpretation:** A higher total score suggests more severe depressive symptoms.

**Putting It All Together**

* **Baseline Data:** Tells us about Alex’s social media habits (e.g., he spends a total of 90 minutes daily on social media, and he has a certain network size).
* **Standardized Measures:** Provide a quantifiable assessment of his mental health. For example, if Alex’s Kessler score is high, it indicates significant psychological distress; if his Rosenberg score is low, it suggests lower self-esteem; and his PHQ-9 score tells us about the severity of depressive symptoms.

By comparing these standardized scores across different groups (such as heavy social media users versus non-users), the researcher can investigate if and how social media usage correlates with mental health outcomes. Alex’s data, along with others’, would be analyzed statistically (using ANOVA and post-hoc tests) to see if differences in usage patterns translate into significant differences in mental health measures.

This comprehensive approach allows the study to link raw behavioral data (from the baseline questionnaire) with standardized psychological assessments to draw conclusions about the effects of social media on mental health.

o3-mini

**You said:**

how is her done all this , does he used some ML algo in that or what technical thing he done , or its just a research based thing only therory

Reasoned for a couple of seconds

The study isn’t based on machine learning algorithms; instead, it relies on traditional statistical methods grounded in research theory. Here’s how the author approaches it:

* **Baseline Data Analysis:**  
  The raw responses about social media habits (e.g., time spent, number of friends) are summarized using descriptive statistics. This means calculating averages, medians, or frequencies to create a profile of participant behavior.
* **Inferential Statistical Tests:**  
  For the mental health measures (like the Kessler Scale, Rosenberg’s Self-Esteem Scale, and PHQ-9), the author uses a one-way ANOVA to test each hypothesis. This analysis determines whether there are statistically significant differences in scores across the groups (Facebook users, Instagram users, and those not using social media).
  + **Bonferroni Post-Hoc Tests:** If the ANOVA indicates significant differences, the Bonferroni test is used to pinpoint exactly which groups differ from each other.

These methods are standard in psychological research and are used to test hypotheses rather than to train predictive models. There isn’t any machine learning component here; the focus is on comparing group differences through established statistical techniques.

o3-mini

**You said:**

so he analyse it one by one for each 200 user or he used some tech in that

Reasoned about data analysis approach for 5 seconds

The researcher doesn’t manually analyze each of the 200 users one by one. Instead, he collects all the data into a database or spreadsheet and then uses statistical software to run the analyses automatically. For example:

* **Descriptive Statistics:** The software calculates averages, medians, and frequencies for all users’ responses, providing a snapshot of overall social media usage.
* **Inferential Analysis:** Tools like SPSS, R, or similar are used to perform one-way ANOVA tests. These tests compare the groups (e.g., Facebook, Instagram, no social media) simultaneously, and if significant differences are found, Bonferroni post-hoc tests help pinpoint exactly which groups differ.

This approach ensures efficiency, accuracy, and consistency in handling data from all 200 participants rather than processing each response manually.

o3-mini

**You said:**

The Effects of Social Media on Mental Health\_ A Proposed Study.pdf

PDF

does he provided the dataset in his pdf?

Reasoned about availability of dataset for 5 seconds

No, the actual dataset isn’t provided in the PDF. This document is a proposed study where the author outlines the methodology, including how data would be collected (through a baseline questionnaire and standardized psychological scales) and the planned statistical analyses (using one-way ANOVA and Bonferroni post-hoc tests). Instead of including real user data, the paper details the procedure for gathering data from 200 participants and provides predicted results based on prior research, rather than presenting an actual dataset.

o3-mini

**You said:**

DataSet\_1.xlsx

Spreadsheet

i want to use his research and method in my research project can i do it could you help me in that , i have a dataset with me can it be used for same work i have applied it check it and tell me

3/3

Reasoned about using existing research methods for 9 seconds

Yes, you can definitely build on his methods for your research project. His study outlines a structured approach that you can adapt if your dataset contains similar variables. Here’s how you might proceed:

**1. Mapping Your Dataset to the Study’s Framework**

* **Baseline Social Media Usage Data:**  
  In the proposed study, participants completed a short baseline questionnaire that gathered information such as:
  + **Social Media Accounts:** Which platforms (e.g., Facebook, Instagram) they use.
  + **Time Spent:** The average amount of time per day on each platform.
  + **Social Network Size:** The number of friends or connections they have.

Check if your dataset includes these or similar variables. For instance, do you have columns like “platform,” “daily\_usage\_minutes,” or “number\_of\_connections”? If so, you can calculate descriptive statistics (means, medians, frequencies) to profile your sample’s social media behavior.

* **Mental Health Measures:**  
  The study uses standardized scales such as:
  + **Kessler Psychological Distress Scale**
  + **Rosenberg Self-Esteem Scale**
  + **Patient Health Questionnaire-9 (PHQ-9)**

Your dataset should have comparable outcome variables, such as scores from depression, self-esteem, or distress questionnaires. Verify that your data includes these or that you can compute these scores from survey responses.

**2. Analytical Methods**

Based on his method, here’s a step-by-step guide on how you might use your dataset:

* **Descriptive Analysis for Baseline Data:**  
  Use statistical software (like R, SPSS, or Python) to calculate:
  + **Averages and Distributions:** For time spent on social media, number of social media accounts, and friend counts.
  + **Group Profiles:** Identify if participants naturally cluster into groups (e.g., heavy vs. light users).
* **Inferential Analysis:**  
  Following his approach, you can:
  + **Compute Total Scores:** For each of the mental health measures (e.g., sum up the Likert-scale responses for the PHQ-9, RSES, and Kessler scales).
  + **Group Comparison:** If your dataset allows you to classify participants into groups (e.g., Facebook users, Instagram users, and non-users), run a one-way ANOVA to check if there are significant differences in mental health outcomes between these groups.
  + **Post-hoc Testing:** If your ANOVA is significant, use a Bonferroni post-hoc test to identify which specific groups differ from one another.

**3. Next Steps and Technical Implementation**

* **Data Cleaning and Preparation:**  
  Ensure your dataset is cleaned and that the variables are correctly formatted. For example, ensure that responses for standardized scales are numerically coded as expected.
* **Using Statistical Software:**
  + **R Example:** You can use functions like aov() for ANOVA and pairwise.t.test() with the Bonferroni correction for post-hoc analysis.
  + **Python Example:** Libraries such as pandas for data manipulation, scipy.stats for ANOVA, and statsmodels can be very helpful.
* **Interpreting Your Findings:**  
  Compare the descriptive and inferential statistics from your analysis with the predictions and rationale outlined in the proposed study. This can help you determine whether your findings are consistent with or extend the previous research.

**4. Getting Specific Help**

If you share more details about the structure of your dataset (e.g., variable names, types of measurements), I can guide you through the specific steps or even provide sample code to get you started.

In summary, you can indeed use his research methods as a framework for your project. With your dataset, you’ll be able to replicate the baseline analysis and statistical tests to explore how social media usage correlates with mental health outcomes. Let me know how you’d like to proceed or if you need further technical assistance with your analysis!

o3-mini

**You said:**

Processed\_SocialMediaData.xlsx

Spreadsheet

it is the dataset i have check it

Reasoned about using dataset in research for 6 seconds

I can’t directly view the contents of your Excel file. Could you please share a brief overview of its structure? For example:

* What are the column names (e.g., social media platform, daily usage time, number of friends, scores on mental health scales)?
* How many rows/participants does it include?
* Are there any specific variables that resemble those used in the proposed study (like baseline social media usage and standardized mental health measures)?

Once I know more about your dataset’s structure, I can help you map it to the methods in the research and guide you on how to proceed with your analysis.

o3-mini

**You said:**

what file can you directly see?

Reasoned about dataset analysis options for a few seconds

I can directly see the PDF file ("The Effects of Social Media on Mental Health: A Proposed Study.pdf") that you uploaded. However, while I acknowledge that you’ve also uploaded an Excel file ("Processed\_SocialMediaData.xlsx"), I can’t directly view its content. Could you provide a brief overview of its structure or list the column names? That will help me guide you on how to use it with the research methods.

o3-mini

**You said:**

columsn name are 1. What is your age? 2. Gender 3. Relationship Status 4. Occupation Status 5. What type of organizations are you affiliated with? 6. Do you use social media? 7. What social media platforms do you commonly use? 8. What is the average time you spend on social media every day? 9. How often do you find yourself using Social media without a specific purpose? 10. How often do you get distracted by Social media when you are busy doing something? 11. Do you feel restless if you haven't used Social media in a while? 12. On a scale of 1 to 5, how easily distracted are you? 13. On a scale of 1 to 5, how much are you bothered by worries? 14. Do you find it difficult to concentrate on things? 15. On a scale of 1-5, how often do you compare yourself to other successful people through the use of social media? 16. Following the previous question, how do you feel about these comparisons, generally speaking? 17. How often do you look to seek validation from features of social media? 18. How often do you feel depressed or down? 19. On a scale of 1 to 5, how frequently does your interest in daily activities fluctuate? 20. On a scale of 1 to 5, how often do you face issues regarding sleep? and total 484 rows are there

Reasoned about dataset structure and analysis for 32 seconds

Below is a step-by-step guide on how you can adapt the methods from the “Effects of Social Media on Mental Health” proposed study to **your** dataset, even though you don’t have the exact standardized scales (Kessler, Rosenberg, PHQ-9). You can still capture similar ideas—psychological distress, social comparison, etc.—using the questions you do have.

**1. Identify Variables That Parallel the Proposed Study**

**a. Social Media Usage (Baseline-Like Data)**

* **Average time on social media** (Column 8)
* **Platforms used** (Column 7)
* **Possible grouping**: “No Social Media,” “Light Use,” “Moderate Use,” “Heavy Use,” etc.

**b. Mental Health / Well-Being Indicators**

Although your dataset does not contain standardized measures like the PHQ-9 or Rosenberg Self-Esteem Scale, you can approximate certain constructs:

1. **Psychological Distress**
   * “How much are you bothered by worries?” (Column 13)
   * “Do you find it difficult to concentrate on things?” (Column 14)
   * “How often do you feel depressed or down?” (Column 18)
   * “How frequently does your interest in daily activities fluctuate?” (Column 19)
   * “How often do you face issues regarding sleep?” (Column 20)

You could sum or average some of these to create an **ad-hoc “distress” index**. For instance, if each question is on a 1–5 scale, you can sum them so that higher totals indicate greater distress.

1. **Social Comparison / Validation**
   * “How often do you compare yourself to other successful people?” (Column 15)
   * “How do you feel about these comparisons?” (Column 16)
   * “How often do you look to seek validation from features of social media?” (Column 17)

You might treat these items individually or combine them into a **“comparison/validation” index** if they’re all measured on a similar scale (e.g., 1–5).

1. **Distraction / Restlessness**
   * “How often do you find yourself using Social Media without a specific purpose?” (Column 9)
   * “How often do you get distracted by Social Media when busy?” (Column 10)
   * “Do you feel restless if you haven’t used Social Media in a while?” (Column 11)
   * “On a scale of 1 to 5, how easily distracted are you?” (Column 12)

These could combine into a **“compulsive use” or “distraction” index** if they’re similarly scaled.

You won’t get a perfect match to the original study’s standardized scales, but these items can approximate **mental health** (distress, depressive mood, restlessness) and **social comparison** (validation, comparison frequency).

**2. Clean and Code Your Data**

1. **Numeric Coding**:
   * For yes/no or free-text responses, convert them to numeric codes where possible (e.g., 0 = “No,” 1 = “Yes”).
   * For 1–5 scale questions, confirm they are consistently coded (e.g., 1 = “never,” 5 = “always”).
2. **Check Missing Data**:
   * Decide how to handle participants with incomplete responses (e.g., remove them or use mean substitution).
3. **Group Participants by Usage**:
   * You might create usage categories from **Column 8** (average time on social media). For example:
     + “No Social Media” = 0 minutes
     + “Light Use” = 1–60 minutes
     + “Moderate Use” = 61–120 minutes
     + “Heavy Use” = 121+ minutes
   * Alternatively, if you’re especially interested in platform differences, you could split by **Column 7** (e.g., “Facebook-only,” “Instagram-only,” “Multiple Platforms,” etc.), though that requires careful grouping logic.

**3. Create Ad-Hoc Indices (If Desired)**

* **Distress Index** (e.g., sum Columns 13, 14, 18, 19, 20 if they’re on a 1–5 scale).
  + A higher total = higher distress.
* **Comparison/Validation Index** (e.g., sum Columns 15, 16, 17 if each is 1–5; if Column 16 is open-ended text, you may need a different approach).
* **Distraction/Compulsive Use Index** (e.g., sum Columns 9, 10, 11, 12).

Make sure each question is on the same scale or that you standardize them before summing.

**4. Statistical Analyses**

**a. Descriptive Statistics**

* **Means, Medians, Frequencies**:
  + Average time spent on social media
  + Distribution of usage categories
  + Means of any ad-hoc indices (distress, comparison, etc.)

**b. Group Comparisons (ANOVA)**

If you want to follow the original study’s structure, do a **one-way ANOVA** comparing your usage groups on each outcome:

1. **Group = Usage Category** (No use, Light, Moderate, Heavy)
2. **Dependent Variables =**
   * Distress Index
   * Comparison/Validation Index
   * Distraction/Compulsive Use Index

**Example ANOVA Model (in R)**:

r

CopyEdit

# Suppose “usage\_group” is your categorical variable (No, Light, Moderate, Heavy)

# and “distress\_index” is a numeric score.

model\_distress <- aov(distress\_index ~ usage\_group, data = yourData)

summary(model\_distress)

# If significant, do pairwise comparisons with Bonferroni:

pairwise.t.test(yourData$distress\_index, yourData$usage\_group, p.adjust="bonferroni")

Repeat for any other outcomes (e.g., comparison\_index, distraction\_index).

**c. Correlation or Regression**

If you want a more continuous approach:

* **Correlation**: Check if “average time spent” (Column 8) is correlated with your indices (distress, comparison, etc.).
* **Multiple Regression**: Include demographic variables (age, gender, relationship status) as predictors to see if they explain additional variance in mental health outcomes.

**5. Interpret the Findings**

* **Check for Significant Differences**: Does heavy usage correlate with higher distress? Do users who compare themselves more frequently also report feeling more depressed?
* **Compare to the Proposed Study**: Although you’re not using the exact scales, you can see if your results mirror the general patterns: e.g., more social media use → higher distress.

**6. Caveats and Recommendations**

1. **Non-Standardized Items**: Because your questions aren’t official scales like Kessler or Rosenberg, your results will be more exploratory.
2. **Multiple Comparisons**: If you run many ANOVAs, remember to correct for multiple testing (e.g., Bonferroni or false discovery rate).
3. **Self-Report Bias**: All data are self-reported; participants might under- or over-report usage or mental health symptoms.

**Putting It All Together**

1. **Categorize** participants by usage (or by platform).
2. **Aggregate** relevant questions into indices (distress, comparison, etc.).
3. **Run ANOVAs** (or correlations/regressions) to see how usage level relates to those indices.
4. **Conclude** whether your findings support the general idea that heavier social media use is linked with lower well-being.

By following these steps, you’ll be applying a similar logic to the original “Effects of Social Media on Mental Health” study—just with your own questions and scales. If you need more specific help with code or interpretation, let me know, and I can walk you through a more concrete example!