Project Proposals

Research Question 1: How does the number of hours spent with friends influence life satisfaction?

Variables and Suggested Visualizations:

- Life satisfaction (Ordinal and measured from 1–10): A histogram of life satisfaction scores would be useful to see the distribution. This will also show if there is any skewness or multimodality present.
- Hours spent with friends in the past 7 days: A box plot would show the range, spread, and outliers in the data for hours spent with friends. This can reveal if the data is highly variable or concentrated around certain values.

Assumptions of Simple Linear Regression:

- Linearity: The relationship between hours with friends and life satisfaction should be approximately linear. This can be assessed by creating a scatter plot with a trend line to visualize the relationship.
- Independence of observations: Each participant's data should be independent of others.
- Homoscedasticity: The spread of residuals should be consistent across values of the independent variable (hours spent with friends).
- Normality of Residuals: The residuals (errors) should follow an approximately normal distribution.

Null Hypothesis:

- There is no significant relationship between hours spent with friends and life satisfaction, this implies that the amount of time with friends does not predict life satisfaction.

Analysis Plan:

- We will perform a simple linear regression with life satisfaction as the dependent variable and hours with friends as the independent variable. If the assumptions are met, this model will allow us to determine whether more time with friends is significantly associated with higher life satisfaction scores.

Possible Outcomes:

- Significant positive relationship: A significant positive relationship would suggest that more hours spent with friends correlates with higher life satisfaction, implying that social time positively impacts well-being.
- Non-significant relationship: If the relationship is non-significant, it would indicate that hours spent with friends might not directly influence life satisfaction, suggesting that other factors might be more influential in determining well-being.

How Each Result Answers the Research Question:

- A significant positive result would support the idea that social engagement with friends has a beneficial impact on life satisfaction, highlighting the importance of social time for mental well-being.
- A non-significant result would suggest that time spent with friends alone may not play a significant role in affecting life satisfaction, potentially indicating that the quality of connections or other factors might be more critical for well-being.

Research Question 2: How does time spent with family relate to the likelihood of having a depression diagnosis?

Variables and Suggested Visualizations:

- Depression diagnosis (Binary: Yes/No): A bar chart showing the counts of participants with and without a depression diagnosis would offer a quick overview of the proportions in each category. This is helpful for understanding the baseline prevalence of depression within the sample.
- Hours with family in the past 7 days: A box plot or histogram to show the distribution of time spent with family can reveal the central tendency and range, indicating if most participants spend similar or highly variable amounts of time with family.

Assumptions of Logistic Regression:

- Linearity of the logit: The relationship between hours with family and the log-odds of depression diagnosis should be linear.
- Independence of observations: Each participant's data should be independent.
- Sufficient sample size: Logistic regression requires a sufficient sample size to ensure valid results, especially when the outcome is relatively rare.

Null Hypothesis:

- There is no significant relationship between hours spent with family and the likelihood of a depression diagnosis, meaning family time does not affect depression outcomes.

Analysis Plan:

- We will use logistic regression to analyze whether hours spent with family predicts the likelihood of a depression diagnosis. The odds ratio obtained from this analysis will show whether more family time increases or decreases the odds of experiencing depression.

Possible Outcomes:

- Significant negative association: A significant negative association (odds ratio < 1) would indicate that more hours with family reduces the odds of a depression diagnosis.
- Non-significant association: If the relationship is non-significant, it would suggest that family time may not directly reduce the risk of depression, and other factors may play a more prominent role.

How Each Result Answers the Research Question:

- A significant negative relationship would support the idea that family time contributes to mental health by potentially lowering depression risk, emphasizing family's role in mental well-being.
- A non-significant result would imply that time with family alone might not strongly influence depression outcomes, suggesting the need to explore other social or personal factors that could impact mental health.

Research Question 3: How does the amount of time someone spends on social media each day relate to their self-rated mental health?

Variables and Suggested Visualizations:

- Self-rated mental health (Ordinal, 5 Categories: Poor, Fair, Good, Very Good and Excellent): A bar chart displaying the mental health categories would provide a clear overview of the distribution, showing how participants rate their mental health across the five levels.
- Social media time per day (Ordinal, 6 Categories: Less than 10 minutes, 10-30 minutes, 31-60 minutes, 1-2 hours, 2-3 hours and More than 3 hours) A bar chart that displays the counts of participants in each range of social media time per day would give a good overview of the proportion of participants in each range.

Assumptions of the Chi-Square Test of Independence:

- Independence of observations: Each participant's data should be independent of others.
- Expected frequencies: Each cell in the contingency table created by cross-tabulating social media time per day and mental health categories should have an expected frequency of at least 5 to ensure the validity of the chi-square test.

Null Hypothesis:

- There is no association between the amount of time spend on social media each day and self-rated mental health categories, indicating that social media usage does not significantly influence self-rated mental health.

Analysis Plan:

- We will perform a chi-square test of independence to assess whether there's a significant association between the time spent on social media (grouped into ranges) and self-rated mental health categories. This test will determine if self-rated mental health varies across groups with different amounts of time spent on social media.

Possible Outcomes:

- Significant association: A significant chi-square result would indicate that self-rated mental health differs across levels of social media usage, suggesting that using social media more each day might be linked to worse mental health ratings.

- Non-significant association: If the chi-square test is non-significant, it would imply that the amount of social media usage may not be strongly associated with self-rated mental health.

How Each Result Answers the Research Question:

- A significant association would support the idea that social media usage is linked to mental health, indicating that increased social media usage may decrease mental health.
- A non-significant result would suggest that the quantity of social media usage alone may not have a strong impact on mental health ratings, perhaps indicating that there are other factors that may be more critical to mental well-being.