8.3 Model Fitting

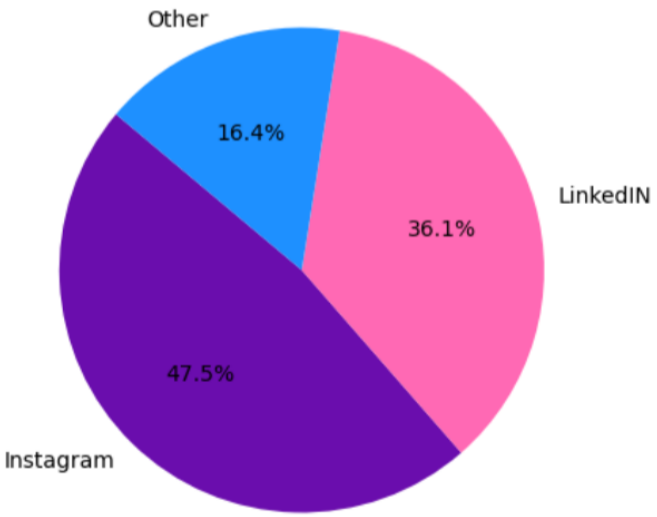
8.3.1 Generalized Linear Model

- Generalized Linear Models (GLMs) are a powerful class of regression models that extend the traditional linear regression framework.
- Unlike simple linear regression, GLMs allow for more flexible relationships by using different underlying statistical distributions.

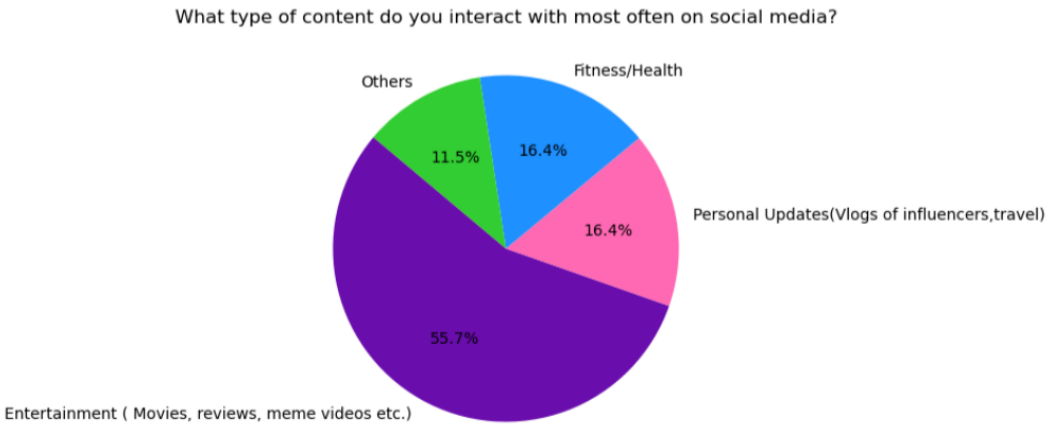
Research Question Does social media negatively affect self-esteem?
Response/Outcome Variable D1: Do you feel that social media negatively affects your self-esteem? (YES/NO)
Predictor Variables (Independent Variables)

- Q6: Age
- Q3: Do you feel pressure to conform to certain standards or trends because of what you see on social media? (Categorical variable based on pressure)
- Q1: How often do you feel like you're missing out on important events or experiences because of what you see on social media? (5-point Likert scale)
- Q4: How frequently do you compare yourself to others based on what you see on social media?
- Q8: How would you describe your level of engagement on social media?
- Q5: How satisfied are you with your overall experience on social media? (1-5-point continuous scale)
- Q2: Do you believe that social media has negatively affected your mental health?
- Q7: Which social media platform do you use the most?

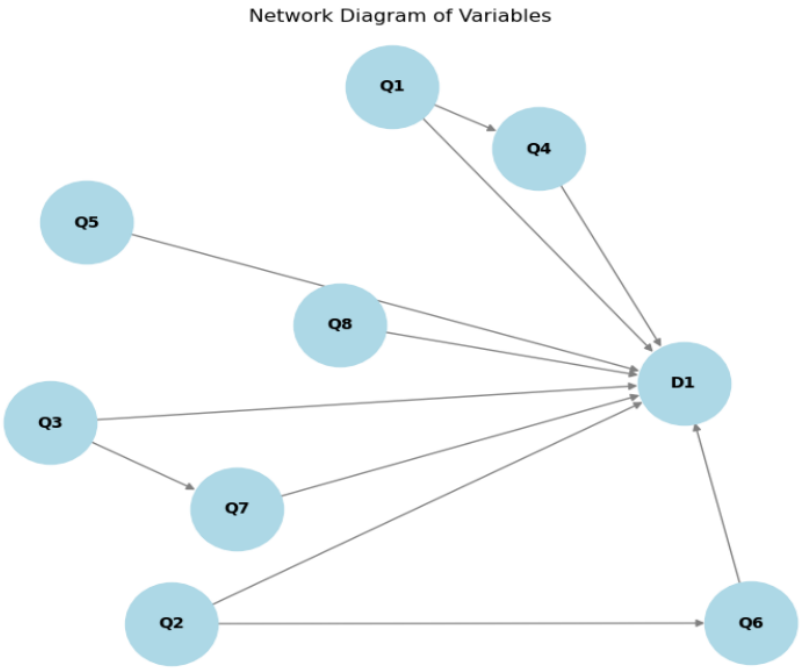
Visualizations



- Instagram has the largest segment, representing 47.5 percent of responses. This suggests that nearly half of the respondents believe Instagram has the most significant impact on their self image.LinkedIn follows, making up 36.1 percent of responses.



- The content most of them interact with is entertainment accounting for 55.7 percent followed by personal updates and fitness/health (16.4 percent).



The diagram presents a structural model of variables potentially influencing the perception that social media negatively affects self-esteem. This model employs a directed acyclic graph (DAG) to illustrate hypothesized causal relationships among constructs related to social media usage and its psychological impacts. Key Observations:

- **Outcome Variable:** D1, positioned centrally, likely represents the primary outcome: self-reported negative impact of social media on self-esteem.

- **Direct Predictors:** Seven variables (Q1, Q2, Q3, Q4, Q5, Q7, Q8) have direct paths to D1, suggesting they are hypothesized to have immediate effects on the outcome. These encompass various dimensions of social media engagement and perceived psychological effects.

- **Mediated Relationships:** Q6, presumed to represent age, exhibits an indirect relationship with D1 via Q2. This suggests a potential mediation effect, where age influences the outcome through its impact on the belief that social media affects mental health negatively.

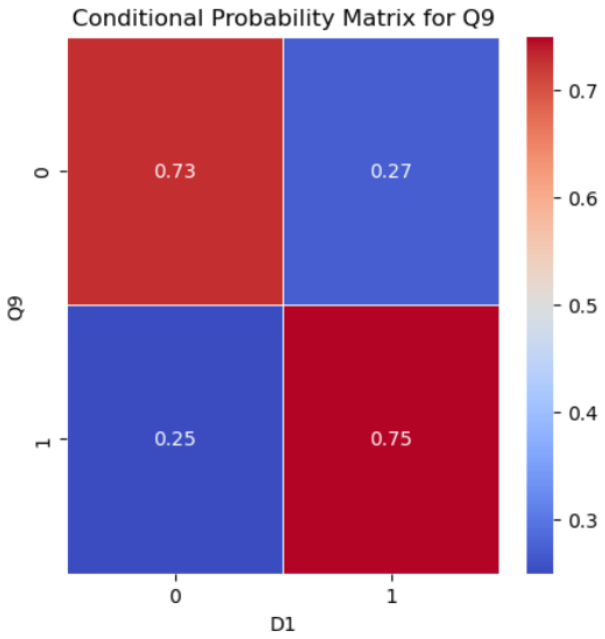
- **Variable Interrelationships:**
  Q1 (FOMO experiences) and Q4 (social comparison frequency) are interconnected, indicating a possible bidirectional relationship or shared underlying construct.
  Q3 (conformity pressure) and Q7 (primary social media platform) show a connection, suggesting platform-specific effects on perceived social pressures.

- **Exogenous Variables:** Q5 (satisfaction with social media) appears as an exogenous variable, only directly influencing the main outcome without incoming paths from other variables.

- **Network Structure:** The model's structure implies a complex, multifaceted approach to understanding social media's impact on self-esteem, considering both direct and indirect pathways of influence.

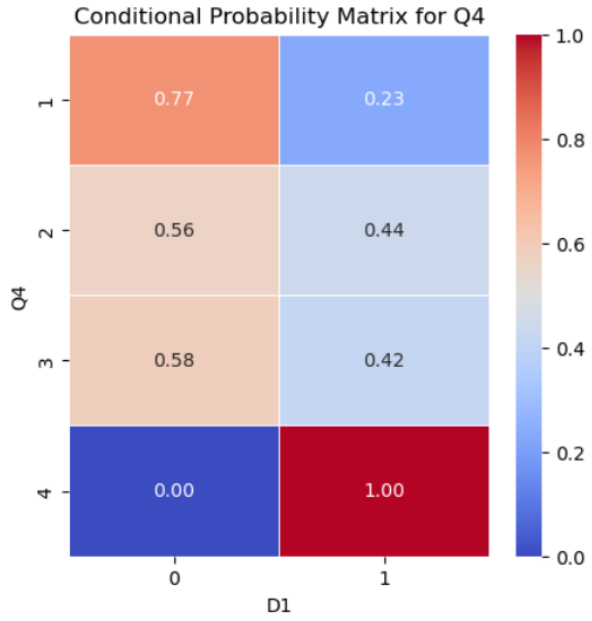
Implications: This network model suggests that the relationship between social media usage and self-esteem is multidimensional, involving various aspects of user experience, psychological processes, and demographic factors. The structure allows for the examination of both direct effects and potential mediating or moderating relationships among the variables.
Methodological Considerations: While this diagram provides insights into hypothesized relationships, further statistical analysis (e.g., path analysis, structural equation modeling) would be necessary to quantify the strength and significance of these relationships. Additionally, the directionality of some paths may require theoretical justification or empirical validation.
This model offers a comprehensive framework for investigating the complex dynamics between social media engagement and its psychological consequences, particularly focusing on self-esteem as a key outcome of interest.

Figure 4: Stress Heat Map



- The matrix shows the conditional probabilities of D1 given different outcomes of Q9.
- The matrix indicates that Q9 has a strong influence on D1.
- Specifically:
  If Q9 is 0, individuals are more likely to respond "NO" to D1, indicating that they do not feel social media negatively affects their self-esteem. If Q9 is 1, individuals are more likely to respond "YES" to D1, indicating that they do feel social media negatively affects their self-esteem.

Figure 5: Comparison Heat Map



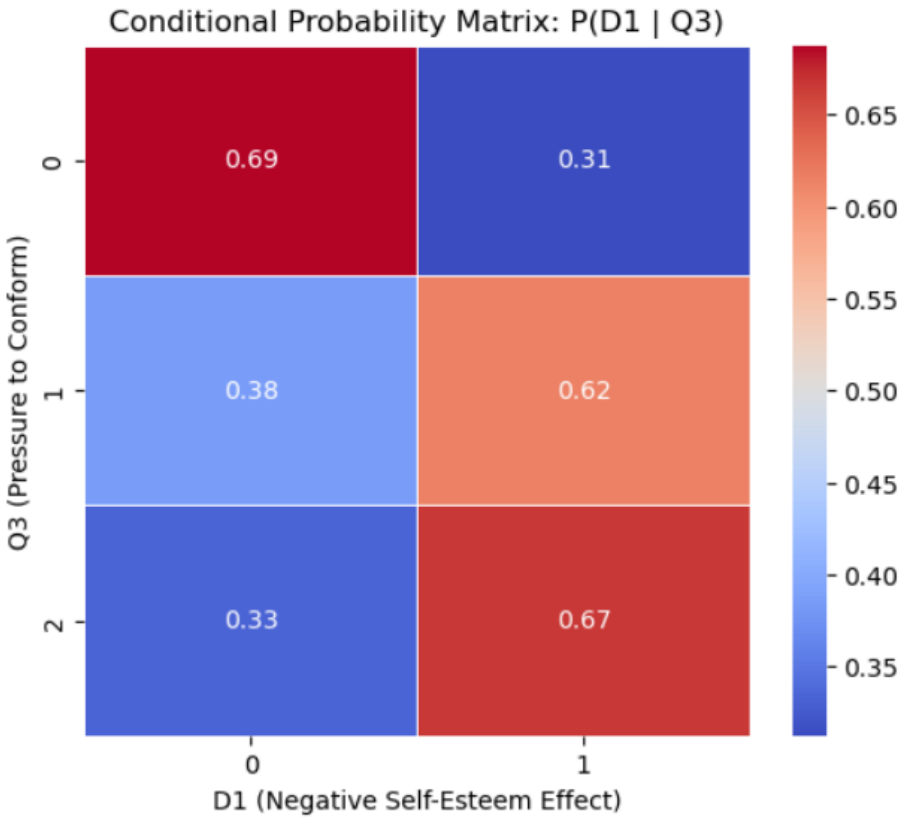
This matrix provides insights into how the variable Q4 influences the outcome D1:

- When Q4 = 1, there is a high probability (77 percent) that D1 will be 0, and a lower probability (23 percent) that D1 will be 1.
- When Q4 = 2, the probabilities are more balanced, with a 56 percent chance of D1 being 0 and a 44 percent chance of D1 being 1.
- When Q4 = 3, the probabilities are also relatively balanced, with a 58 percent chance of D1 being 0 and a 42 percent chance of D1 being 1.
- When Q4 = 4, D1 is almost certain to be 1 (100 percent), and there is no probability of D1 being 0.

The matrix indicates that Q4 has a varying influence on D1 depending on its value. Specifically:

- If Q4 is 1, individuals are more likely to respond "NO" to D1, indicating that they do not feel social media negatively affects their self-esteem.
- If Q4 is 4, individuals are almost certain to respond "YES" to D1, indicating that they do feel social media negatively affects their self-esteem.
- For Q4 values of 2 and 3, the probabilities are more balanced, indicating a less clear influence on D1.
- This suggests that Q4 is a significant predictor of D1, with clear patterns in the conditional probabilities.

Figure 6: Self-Esteem and Pressure Heat Map



The matrix indicates that Q3 has a varying influence on D1 depending on its value. Specifically:

- If Q3 is 0, individuals are more likely to respond "NO" to D1, indicating that they do not feel social media negatively affects their self-esteem.
- If Q3 is 2, individuals are more likely to respond "YES" to D1, indicating that they do feel social media negatively affects their self-esteem.
- For Q3 value of 1, the probabilities are more balanced, indicating a moderate influence on D1.

Overview of the model

```
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:                     D1   No. Observations:                   60
Model:                            GLM   Df Residuals:                       51
Model Family:                Binomial   Df Model:                            8
Link Function:                  Logit   Scale:                          1.0000
Method:                          IRLS   Log-Likelihood:                -16.903
Date:                Wed, 28 Aug 2024   Deviance:                       33.806
Time:                        13:24:14   Pearson chi2:                     51.7
No. Iterations:                     7   Pseudo R-squ. (CS):             0.5564
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -2.9429      4.418     -0.666      0.505     -11.602       5.716
Q1             2.0511      0.748      2.741      0.006       0.584       3.518
Q2             3.7846      1.180      3.207      0.001       1.472       6.098
Q3             0.0767      0.815      0.094      0.925      -1.521       1.675
Q4             0.2915      0.545      0.535      0.593      -0.776       1.359
Q5            -1.1476      0.595     -1.930      0.054      -2.313       0.018
Q6            -0.1156      0.115     -1.007      0.314      -0.341       0.109
Q7             0.8852      0.730      1.213      0.225      -0.545       2.316
Q8            -0.3680      0.813     -0.452      0.651      -1.962       1.226
```

Model Summary

- Dependent Variable (D1): This is a binary outcome variable.
- No. Observations: 60
- Df Residuals: 51
- Df Model: 8 (indicating 8 predictors in the model)
- Log-Likelihood: -16.903
- Deviance: 33.806
- Pseudo R-squared (CS): 0.5564, which indicates that approximately 55.64 percent of the variability in D1 is explained by the model. This is a good level of explanatory power.

Coefficients and Interpretations

- Intercept: -2.9429 (p = 0.505)
  This is the log odds of D1 being 1 when all predictors are zero. The intercept is not statistically significant (p ¿ 0.05), suggesting it does not provide significant information about D1 on its own.
- Q1: 2.0511 (p = 0.006)
  For every one-unit increase in Q1, the log odds of D1 being 1 increase by approximately 2.0511. This predictor is statistically significant (p ¡ 0.05), indicating a strong positive effect on D1. In practical terms, higher values of Q1 are associated with a higher likelihood of D1 being 1.

- Q2: 3.7846 (p = 0.001)
  For every one-unit increase in Q2, the log odds of D1 being 1 increase by approximately 3.7846. This predictor is statistically significant (p ¡ 0.05), suggesting a strong positive effect. Higher values of Q2 are associated with a higher probability of D1 being 1.

- Q3: 0.0767 (p = 0.925)
  The coefficient for Q3 is close to zero and not statistically significant (p ¿ 0.05). This predictor does not have a meaningful effect on D1.

- Q4: 0.2915 (p = 0.593)
  The coefficient for Q4 is also close to zero and not statistically significant (p ¿ 0.05). This predictor does not appear to significantly impact D1.

- Q5: -1.1476 (p = 0.054)
  For every one-unit increase in Q5, the log odds of D1 being 1 decrease by approximately 1.1476. Although this predictor is close to being statistically significant (p 0.054), it is not quite below the 0.05 threshold. This suggests a potential negative effect, but it's not statistically robust given the p-value.

- Q6: -0.1156 (p = 0.314)
  The coefficient for Q6 is not statistically significant (p ¿ 0.05). This predictor does not have a significant effect on D1

- Q7: 0.8852 (p = 0.225)
  The coefficient for Q7 is not statistically significant (p ¿ 0.05). This predictor does not significantly affect D1.

- Q8: -0.3680 (p = 0.651)
  The coefficient for Q8 is also not statistically significant (p ¿ 0.05). This predictor does not have a meaningful impact on D1.

Summary of Model

- **Significant Predictors:** Q1 and Q2 have significant positive effects on the log odds of D1 being 1. Q5 has a potential negative effect but is not statistically significant.

- **Non-significant Predictors:** Q3, Q4, Q6, Q7, and Q8 are not statistically significant, suggesting they do not have a meaningful impact on the response variable D1 in this model.

Overall, the model provides a good fit with a pseudo R-squared value of 0.5564, explaining over half of the variability in the response variable. The significant predictors (Q1 and Q2) should be considered key factors influencing D1, while the non-significant predictors might need further investigation or could be excluded from the model.