**Study**: In this study, participants were given various news articles to read that discussed the use of chemical weapons in Syria. After reading these articles, they were then instructed to write how they felt about the use of chemical weapons for five to ten minutes. Those writings were analyzed for different markers of linguistic styles (described below). Instead of doing a traditional ANOVA to see if there are group differences between men and women’s writing, we will be predicting gender from their writing patterns.

**Important:** You will find that the predictors are likely not significant. This analysis is practice, so interpret the predictors in the write up/analysis as if they were significant. Remember to paste your output in this document and upload your R script to blackboard to complete this assignment.

**Binominal Simultaneous Log Regression**

**IV X-Variables:**

All variables are z-scored ranges calculated from the raw word frequencies. A person who scores a 0 on any of these scales uses around an average range of complex thinking/honesty/status, etc. A score that is negative indicates that they use less honest/categorical/complex language, while positive scores indicate they use more honest/categorical/complex language.

* Honesty – a marker of how much a person uses deceptive or misleading language.
* Status – a marker of how much a person uses inclusive pronouns (we – high scores indicate more inclusive status) versus exclusive pronouns (I – low scores indicate a more individualistic status).
* Complex thinking – a marker of how much complexity exists in the writing: high scores are more complex, low scores are less complex.
* Cognitive processing – a marker of how much a person wrote about the relationships between things: high scores indicate more causal explanations, while low scores are more simplistic thinking.
* Psychological distancing – a marker of how much a person is relating to an event: high psychological distancing is more abstract thinking, while low scores are more personal thought processes.

**DV – Y-variable:**

* Gender: Men and Women

**Data Screening:**

**Accuracy and Ratio of Cases:**

1. Include summary output and comment if we have a good ratio of cases in the DV.
2. You can assume the IV data is accurate.

**Additivity:**

1. Include the symnum table of the IVs.
2. Are any of the variables too highly correlated?

**Power:**

1. Calculate the number of participants you would need for this analysis with a large effect size.
   1. Include a screen shot or summary of the numbers you typed into G\*Power, so we can give you partial credit if you get a different sample size than us.

**Analysis:**

1. Include the regression output for your model.
2. Is the overall model significant?
   1. Include the chi-square change, df change, and significance level.
3. What is the effect size for the overall model?
   1. Include the PseduoR2 output for your model.
4. Are the individual predictors significant?
   1. Note, they are likely not significant. Include which way they WOULD predict if they were significant by completing a table like the one in the notes.

|  |  |  |
| --- | --- | --- |
| **Predictor** |  |  |
|  |  |  |
|  |  |  |

1. Include the classification percentages:
   1. Men
   2. Women
   3. Total
2. Include a dot-plot (graph) of the classifications.
3. Write up:
   1. Include a short description of variables/analysis.
   2. Include a brief section on the data screening/assumptions.
      1. Ratio of cases
      2. Additivity
   3. Model summary chi square change statistics.
   4. Coefficients:
      1. Table of coefficient values (you can edit the one from the example).
      2. Interpret the coefficients, describe their effects as if they were significant.
   5. Include the classification percentages.