CSS 205: Final Project Proof of Data

Women's Representation in Politics

Adeline Liem & Harley Clifton

2025-02-11

This assignment was completed by Adeline Liem and Harley Clifton, and we are both CSS M.S. students. We will be collaborating on this project together for the rest of the course.

```
# Read in dataset

filepath <- './data/PADD_Agreement_Level_Multiple_Positions.csv'
df <- read_csv(filepath)
```

```
## Rows: 116 Columns: 34
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr   (3): Con, Stage, StageSub
## dbl  (30): AgtId, GeWom, Mult_Pos_Weight, Del_N, NADel_N, FemDel_Bin, FemDel...
## date  (1): Dat
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
df
```

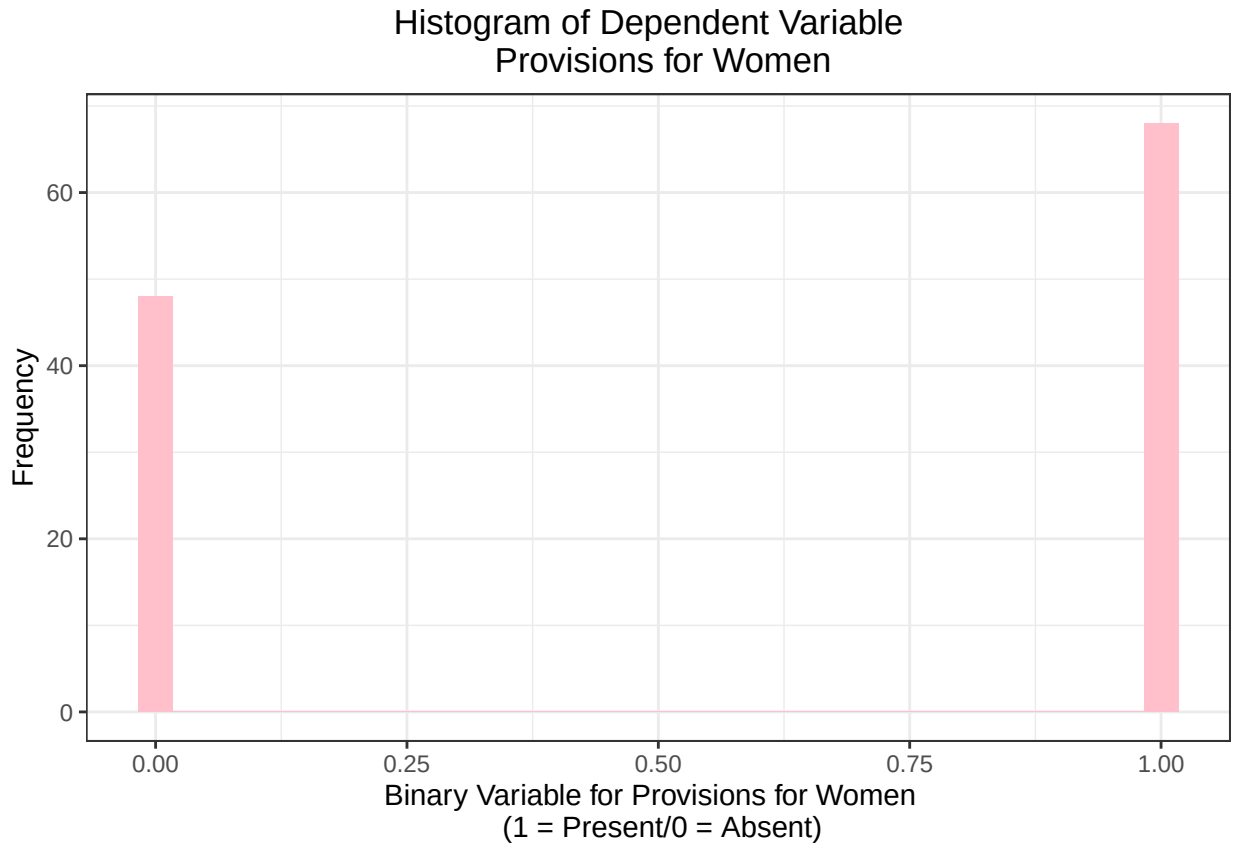
```
## # A tibble: 116 x 34
##    Con      AgtId Dat        Stage  StageSub GeWom Mult_Pos_Weight Del_N NADel_N
##    <chr>    <dbl> <date>     <chr>  <chr>    <dbl>           <dbl> <dbl>   <dbl>
##  1 Gabon        8 1994-10-07 SubCo~ FrAg         0               2    92       0
##  2 Colombia   147 1994-05-26 SubCo~ FrAg         0               1    18       0
##  3 Colombia   149 1991-02-15 SubCo~ FrAg         0               1    15       0
##  4 Colombia   150 1998-07-29 SubCo~ FrAg         0               1    15       0
##  5 Colombia   151 1991-05-27 SubCo~ FrAg         1               1    21       0
##  6 Colombia   152 1991-01-25 SubCo~ FrAg         0               1    11       0
##  7 Colombia   163 1991-07-01 SubCo~ FrCons       1               1    NA      NA
##  8 Colombia   168 1990-03-09 SubCo~ FrAg         0               1     7       0
##  9 Colombia   173 1994-04-09 SubCo~ FrAg         0               1     7       0
## 10 Djibouti   197 2000-02-07 SubCo~ FrAg         0               1     2       0
## # i 106 more rows
## # i 25 more variables: FemDel_Bin <dbl>, FemDel_N <dbl>, FemDel_P <dbl>,
## #   Sig_N <dbl>, FemSig_N <dbl>, FemSig_P <dbl>, Ob_N <dbl>, FemOb_N <dbl>,
## #   FemOb_P <dbl>, Neg_N <dbl>, FemNeg_N <dbl>, FemNeg_P <dbl>, Med_N <dbl>,
## #   FemMed_N <dbl>, FemMed_P <dbl>, Log_N <dbl>, FemLog_N <dbl>,
## #   FemLog_P <dbl>, Adv_N <dbl>, FemAdv_N <dbl>, FemAdv_P <dbl>,
## #   WomCom_N <dbl>, WomCom_P <dbl>, FemWomCom_N <dbl>, FemWomCom_P <dbl>
```

Histogram of the Dependent Variable

Dependent Variable: Provisions for Women (`GeWom`): Women, girls, and gender
"This is a binary variable, taking the value of 1 if any of the peace agreement provisions are specifically addressing women, their inclusion, and their rights. This includes references to girls, widows, mothers, sexual violence (or forms thereof), gender violence, UNSC 1325 or CEDAW, lactating women. If no such provisions are present in the agreement, the value of the variable is 0" (Bell et al. 2021, 24).

```
# Generate histogram

hist <- ggplot(df, aes(x = GeWom)) +
  geom_histogram(fill = "pink") +
  labs(title = "Histogram of Dependent Variable  \n Provisions for Women",
       x = "Binary Variable for Provisions for Women  \n (1 = Present/0 = Absent)",
       y = "Frequency") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5),
        axis.title.x = element_text(size = 11),
        axis.title.y = element_text(size = 11))

hist
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Correlation Matrix

A correlation matrix for the DV and IVs that the original authors included in the model you are replicating.

Independent Variables: Women signatories (`FemSig_P`), women negotiators (`FemNeg_P`), women mediators (`FemMed_P`), women observers (`FemOb_P`)

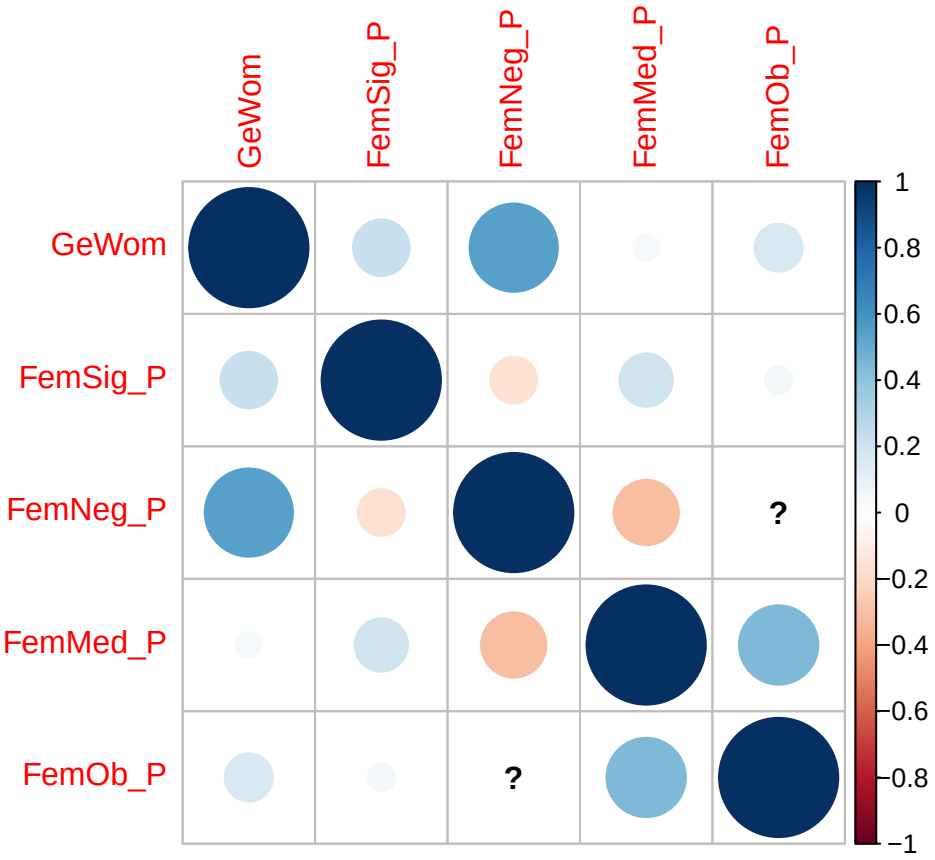
```
# Making subset of the data
dat <- df %>% select(GeWom, FemSig_P, FemNeg_P, FemMed_P, FemOb_P)

# Correlation Matrix
cor(dat, use = 'pairwise.complete.obs')
```

```
## Warning in cor(dat, use = "pairwise.complete.obs"): the standard deviation is
## zero
```

```
##               GeWom    FemSig_P   FemNeg_P    FemMed_P    FemOb_P
## GeWom    1.00000000  0.22615740  0.5473703  0.04776698 0.16359421
## FemSig_P 0.22615740  1.00000000 -0.1545483  0.20064415 0.05230759
## FemNeg_P 0.54737034 -0.15454826  1.0000000 -0.30089443         NA
## FemMed_P 0.04776698  0.20064415 -0.3008944  1.00000000 0.44321486
## FemOb_P  0.16359421  0.05230759         NA  0.44321486 1.00000000
```

```
# Correlation Plot
corrplot(cor(dat, use = 'pairwise.complete.obs'))
```

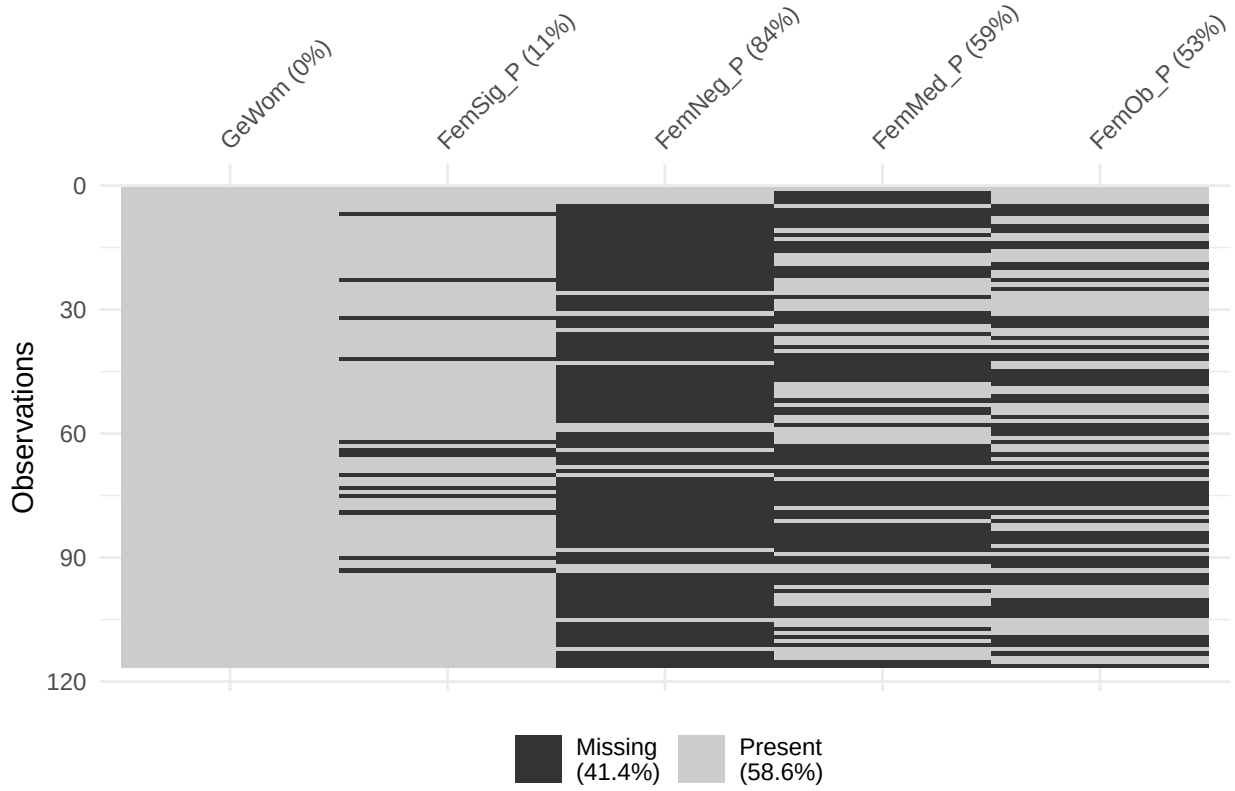
```
## Warning in cor(dat, use = "pairwise.complete.obs"): the standard deviation is
## zero
```



Data Missingness

A visual or tabular depiction of the missingness in the data from part (2); see p. 251-255 of the text.

```
# Generate missingness visualization

vis_miss(dat)
```



Appendix

"I certify that we did not use any LLM or generative AI tool in this assignment"