SDS HW 9

Gianluca Bollo (gb25625) - https://github.com/gianlucabollo/HW9-SDS315

4/18/2024

How much more likely are GOTV call recipients to have voted in 1998? Preliminary Analysis:

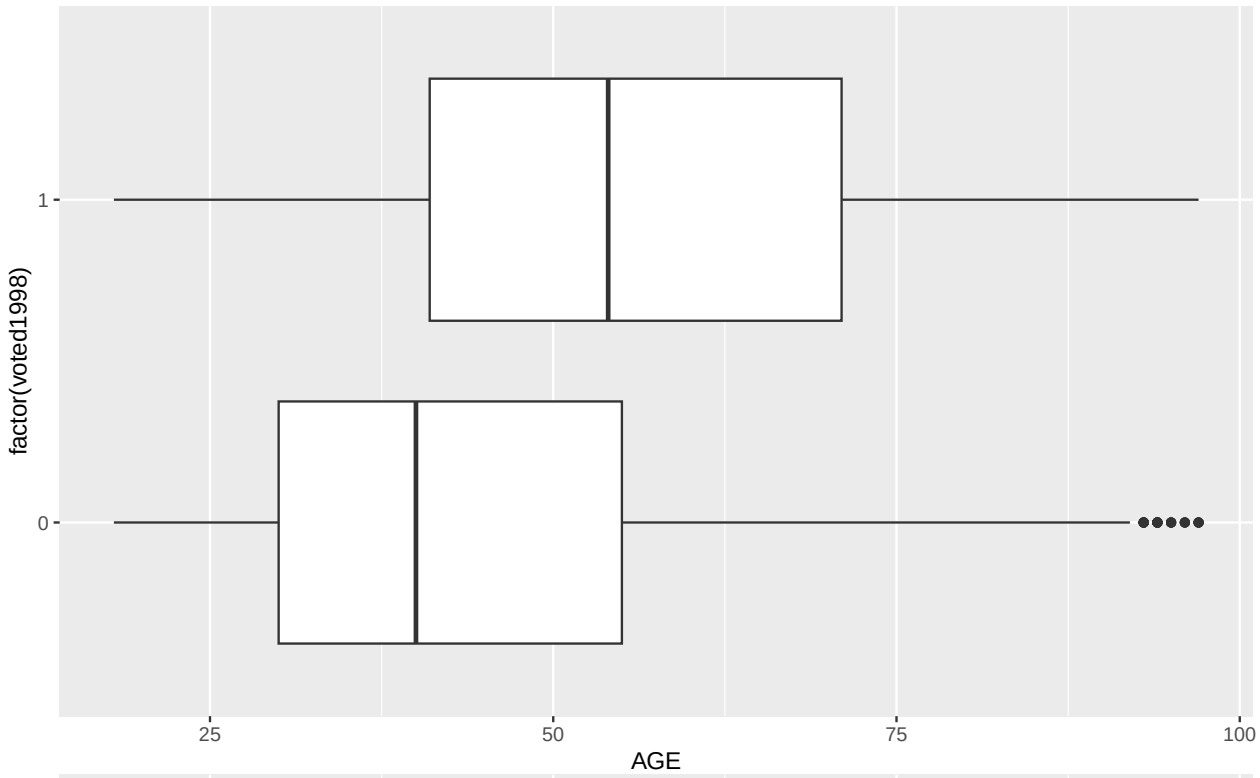
```
##  prop_1.0  prop_1.1
## 0.4442449 0.6477733
```

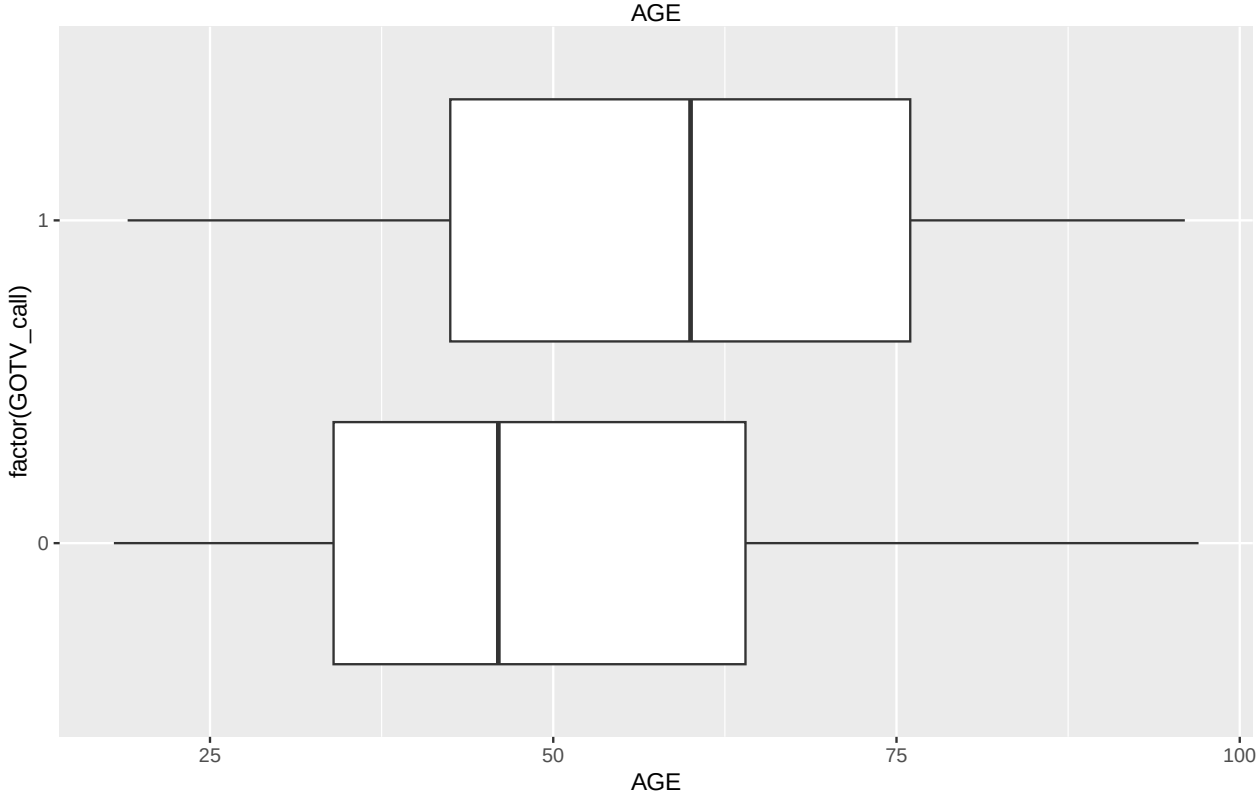
```
##       name     lower     upper level     method  estimate
## 1 diffprop 0.1416722 0.2626275  0.95 percentile 0.2035283
```

Consider the voted1996, AGE, and MAJORPTY variables. Provide evidence that at all three of these variables are confounders.

```
## # A tibble: 2 x 3
##   voted1996 prop_voted_1998 prop_gotv_call
##       <int>           <dbl>          <dbl>
## 1         0           0.229         0.0141
## 2         1           0.640         0.0304
```

```
## # A tibble: 2 x 3
##   MAJORPTY prop_voted_1998 prop_gotv_call
##      <int>           <dbl>          <dbl>
## 1        0           0.350         0.0178
## 2        1           0.482         0.0245
```

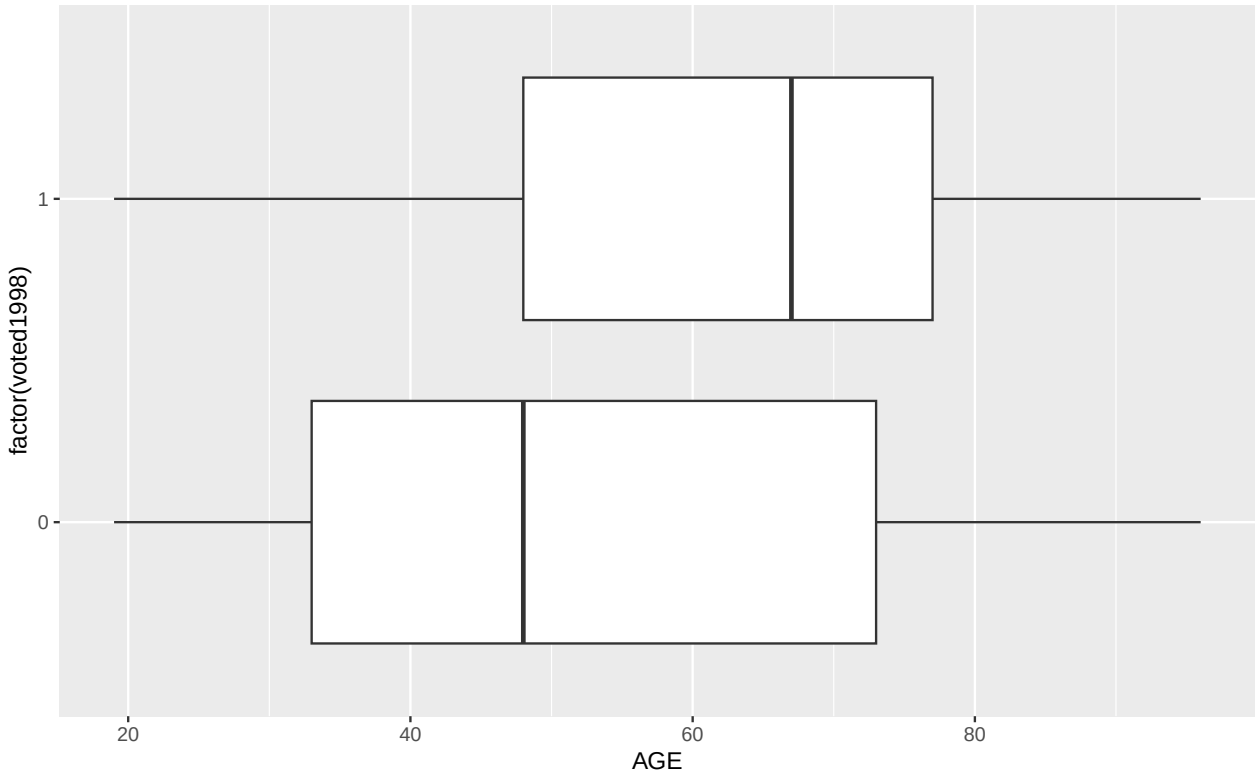

factor(voted1998)
1
0
25
50
75
100
AGE

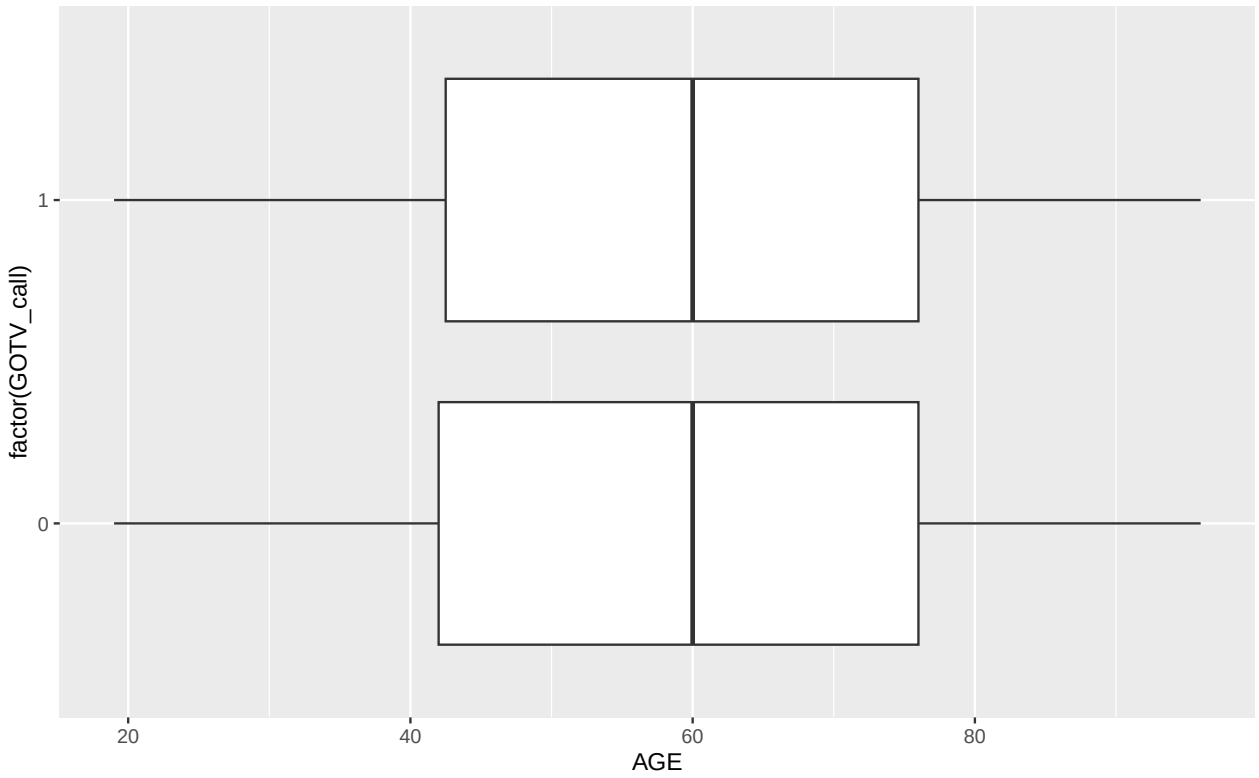

factor(GOTV_call)
1
0
25
50
75
100
AGE

Use matching to construct a data set with GOTV_call as our treatment variable, and with voted1996, AGE,and MAJORPTY as our "matching" or "balancing" variables. Use 5 control cases for each treated case in your matching (ratio=5). Provide evidence that your "matched" data set is, indeed, balanced with respect to the three confounders of voted1996, AGE, and MAJORPTY. (That is, show that these variables are no longer confounders for the matched data.) Then repeat your analysis from Part A, except using the matched data only.

```
## # A tibble: 2 x 3
##   voted1996 prop_voted_1998 prop_gotv_call
##       <int>           <dbl>          <dbl>
## 1         0           0.265          0.167
## 2         1           0.716          0.167
```

```
## # A tibble: 2 x 3
##   MAJORPTY prop_voted_1998 prop_gotv_call
##      <int>           <dbl>          <dbl>
## 1        0           0.460          0.171
## 2        1           0.617          0.166
```



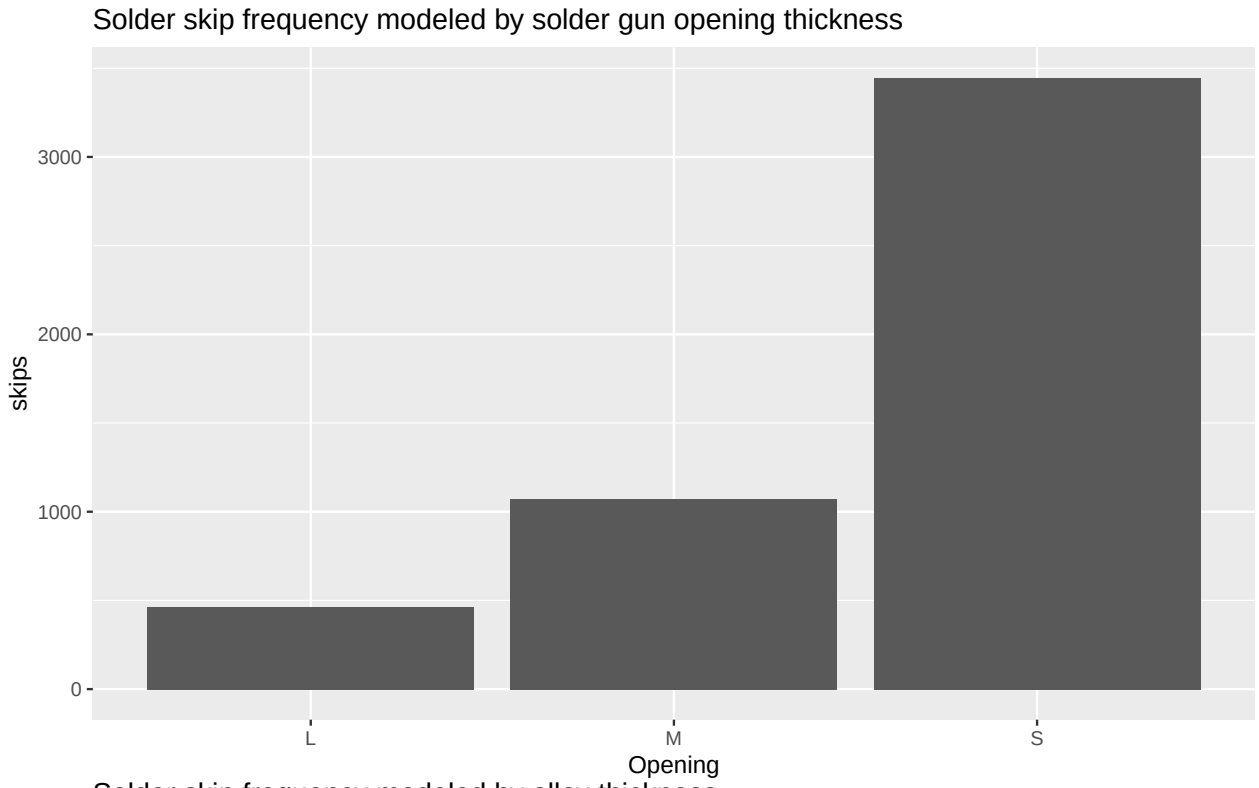


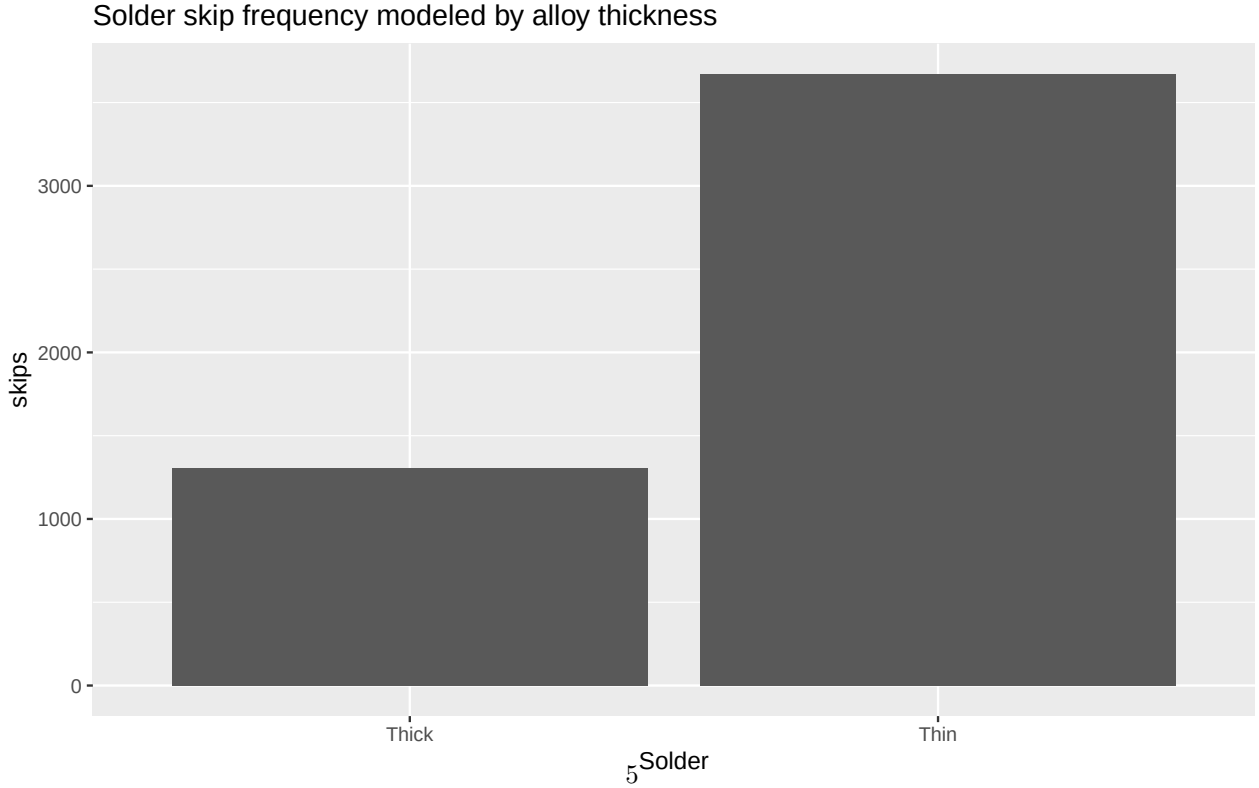

```
##  prop_1.0  prop_1.1
## 0.5740891 0.6477733
```

```
##       name       lower     upper level     method   estimate
## 1 diffprop 0.008403442 0.1378708  0.95 percentile 0.07368421
```

Upon using a matched dataset and removing the effects of identified confounders, I proceeded to investigate the effect of a GOTV call on the likelihood of voting in the 1998 election. To do this, I calculated the confidence interval for the proportions of voting in 1998 (voted1998==1) for those who received a GOTV call versus those who didn't. As shown above, the interval does not contain 0, which deem the results significantly significant. In terms of practicality and making my own conclusion, I believe to a small degree that people receiving GOTV calls were more likely to go vote than those who didnt recieve the call. However, I would not be so sure to conclude that this observed difference is purely due to the call, as there may be other confounders or reasons that contribute to this difference not accounted for in this analysis.

Make two plots. The first plot should provide evidence that the size of the opening on the solder gun is related to the number of skips. The second should provide evidence that the thickness of the alloy used for soldering is related to the number of skips. Give each plot an informative caption describing what is shown in the plot.





Build a regression model and table for the given outcome and predictor variables.

```
##
## Call:
## lm(formula = skips ~ Opening + Solder + Opening:Solder, data = solder)
##
## Coefficients:
##         (Intercept)             OpeningM             OpeningS
##              0.3933               2.4067               5.1267
##          SolderThin  OpeningM:SolderThin  OpeningS:SolderThin
##              2.2800              -0.7400               9.6533
```

```
## # A tibble: 6 x 7
##   term                  estimate std_error statistic p_value lower_ci upper_ci
##   <chr>                    <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
## 1 intercept                 0.39      0.52      0.76    0.45    -0.63     1.41
## 2 Opening: M                2.41      0.74      3.27    0        0.96     3.85
## 3 Opening: S                5.13      0.74      6.97    0        3.68     6.57
## 4 Solder: Thin              2.28      0.74      3.1     0        0.84     3.72
## 5 Opening: M:SolderThin    -0.74      1.04     -0.71    0.48    -2.78     1.3
## 6 Opening: S:SolderThin     9.65      1.04      9.28    0        7.61    11.7
```

- The baseline number of skips for circuit boards that were manufactured with large-opening solder gun and thick alloy is .39.
- The main effect for a medium opening in a solder gun is 2.41 skips. This is the effect of a medium-opening solder gun in isolation.
- The main effect for a small opening in a solder gun is 5.13 skips. This is the effect of a small-opening solder gun in isolation.
- The main effect for a using thin alloy when soldering is 2.28 skips. This is the effect of a thin-alloy in isolation.
- The interaction effect for a medium-opening solder gun and thin alloy is -0.74 skips. This means that circuit boards constructed from thin alloy and a medium-opening solder gun yield 0.74 less skips than if you were to sum both of the individual effects of the variables.
- The interaction effect for a small-opening solder gun and thin alloy is 9.65 skips. This means that circuit boards constructed from thin alloy and a small-opening solder gun yield 9.65 more skips than if you were to sum both of the individual effects of the variables.

If you had to recommend a combination of Opening size and Solder thickness to AT&T based on this analysis, which one would it be, and why? (Remember, the goal is to minimize the number of skips in the manufacturing process.)

If I had to recommend a combination of Opening size and Solder thickness to AT&T based on this analysis, I would recommend the combination of using thick alloy and a large-opening solder gun. As seen from the above table, this combination is the baseline, and all other combinations yield a higher amount of skips. Due to the primarily positive estimated results and confidence intervals of the main and interaction effects, it is evident that the combination of thick alloy and a large-opening solder gun is the best combination when trying to minimize solder skips.