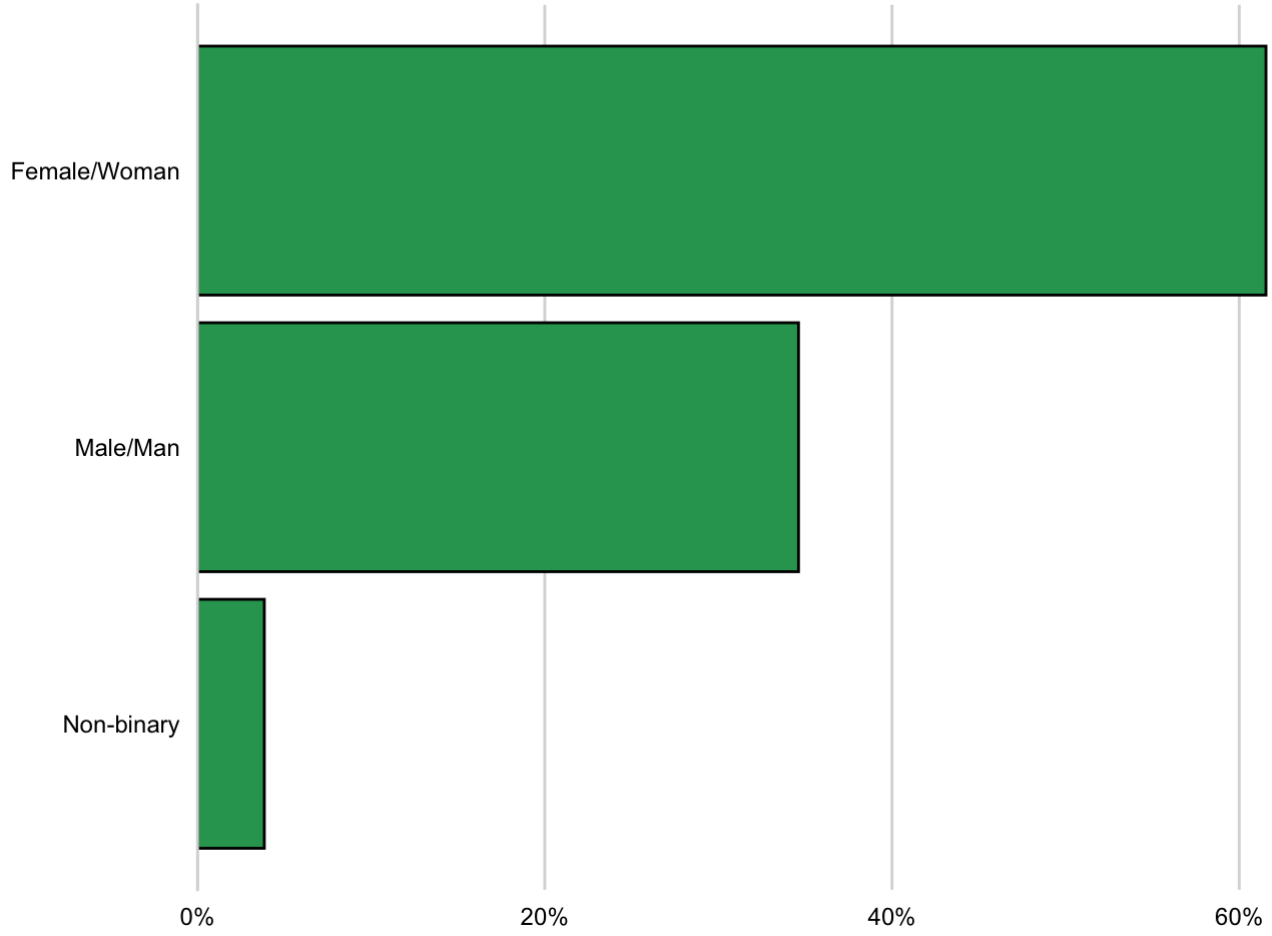
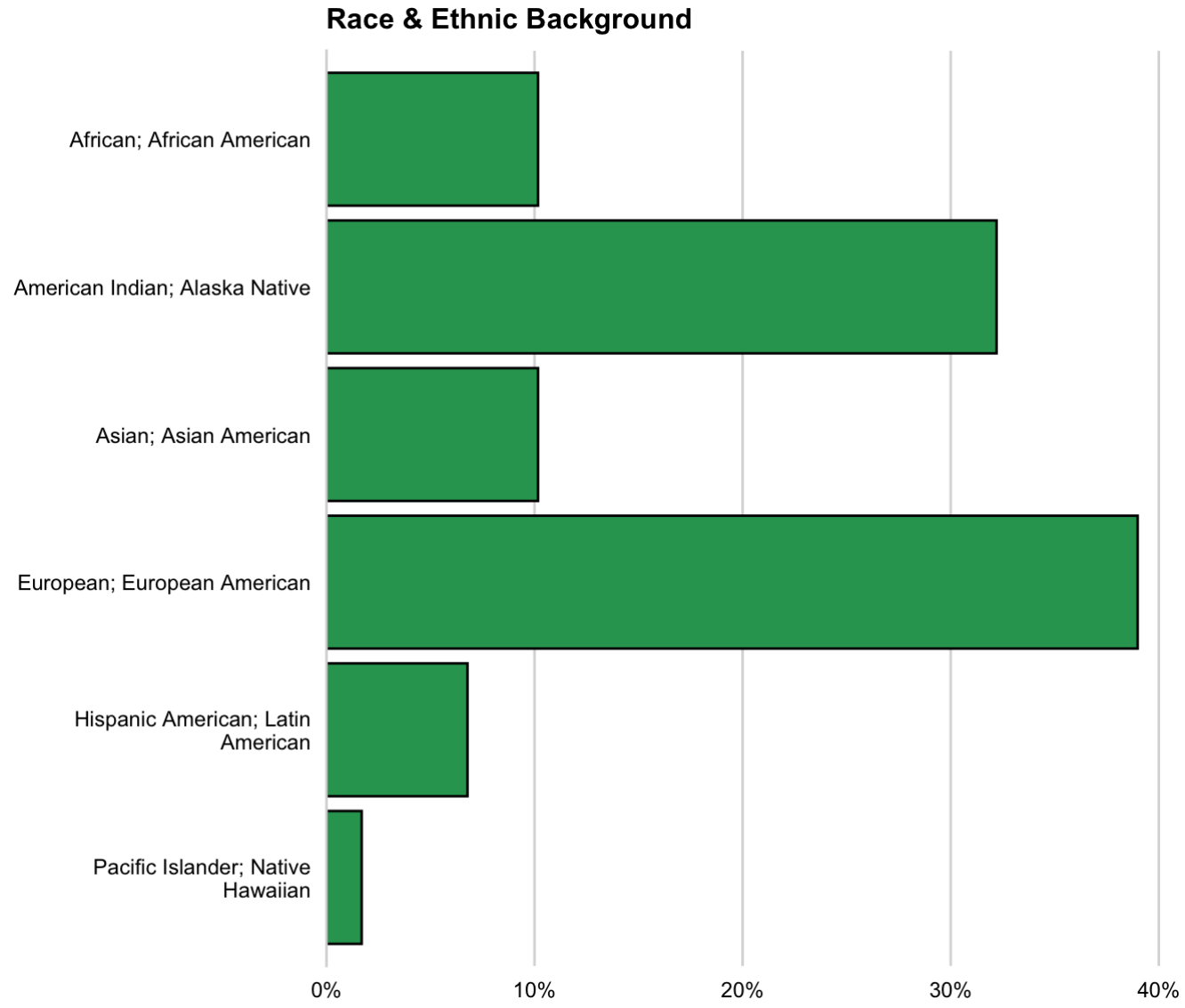
```
# Installing libraries
library(TAM) # for Rasch modeling
library(WrightMap) # to build item-person (Wright) maps
library(lsr)  #to calculate Cohen's d
library(DescTools)  #to calculate etasquared
library(Rmisc) # to calculate summary statistics
library(tidyverse)
library(cowplot)
library(knitr)
library(magick)


# Create logo banner for plots
earthlab_orig <- image_read(path = "earth-lab-logo-white.png") %>%
   image_scale("x80")

twitter_orig <- image_read(path = "plot-footer-twitter.png") %>%
   image_scale("x70")

black_banner <- image_read(path = "black-banner.png")

earthlab_logo <- image_composite(image_scale(black_banner, "1000x100"), earthlab_orig, o
ffset = "+30+10")
twitter_logo <- image_composite(image_scale(black_banner, "1000x100"), twitter_orig, off
set = "+540+15")

logo <- image_append(image_scale(c(earthlab_logo, twitter_logo)), stack = FALSE)

logo
```

Participant Demographics

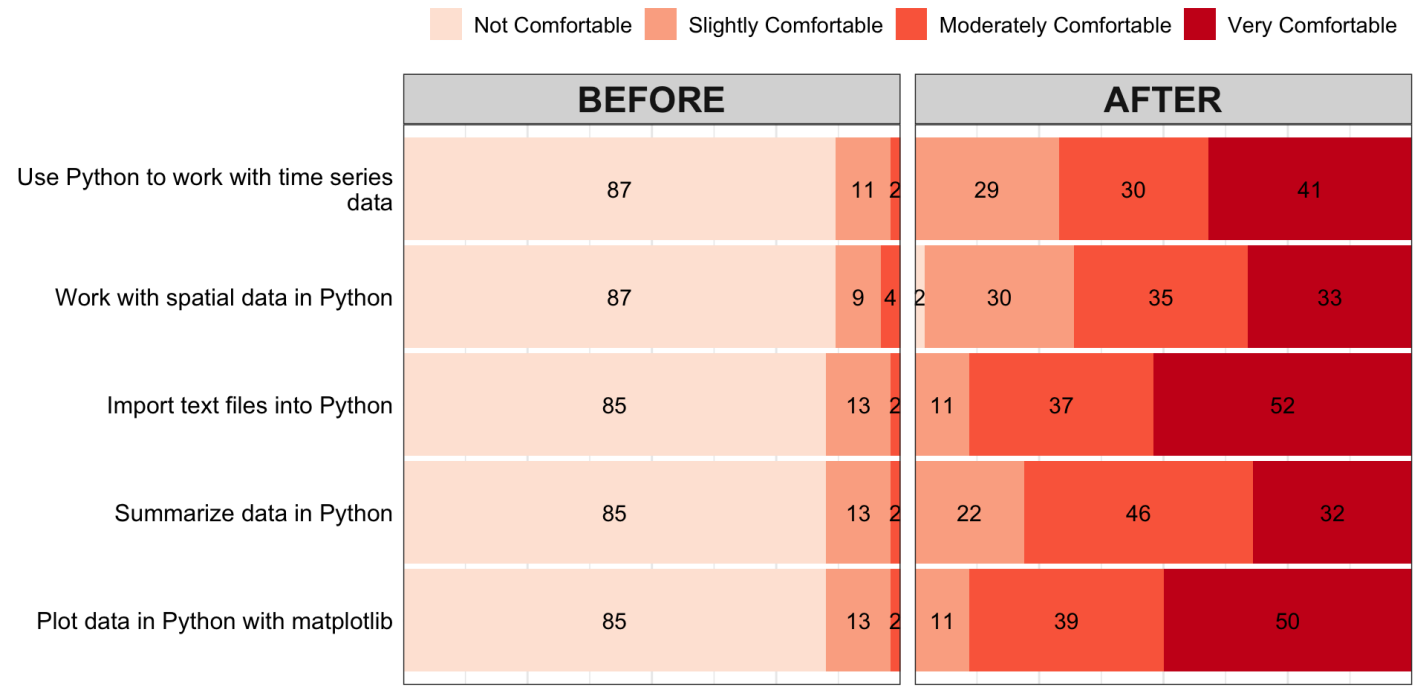
A total of 53 consenting participants provided demographic information related to gender, race & ethnicity through a series of items included on the pre-program survey instrument, administered prior to the start of the technical workshops.

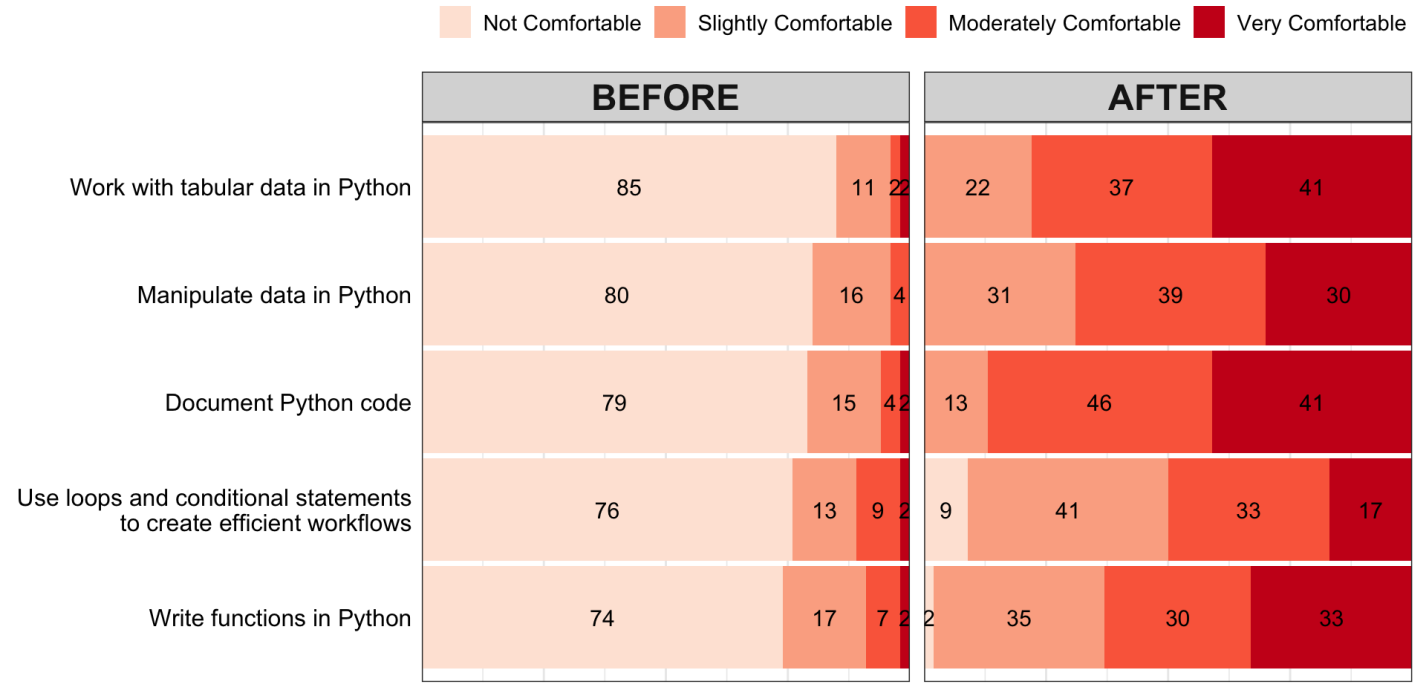

Female/Woman
Male/Man
Non-binary
0%
20%
40%
60%



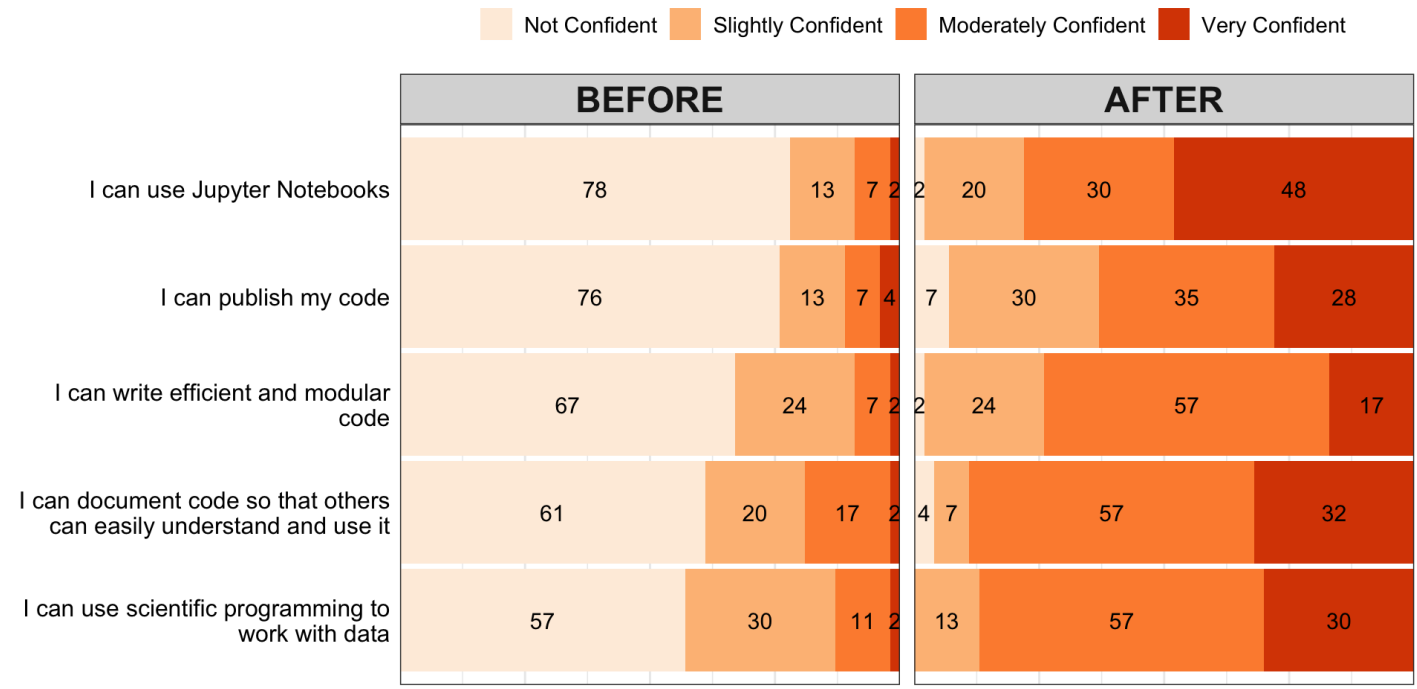
Raw Responses (%)

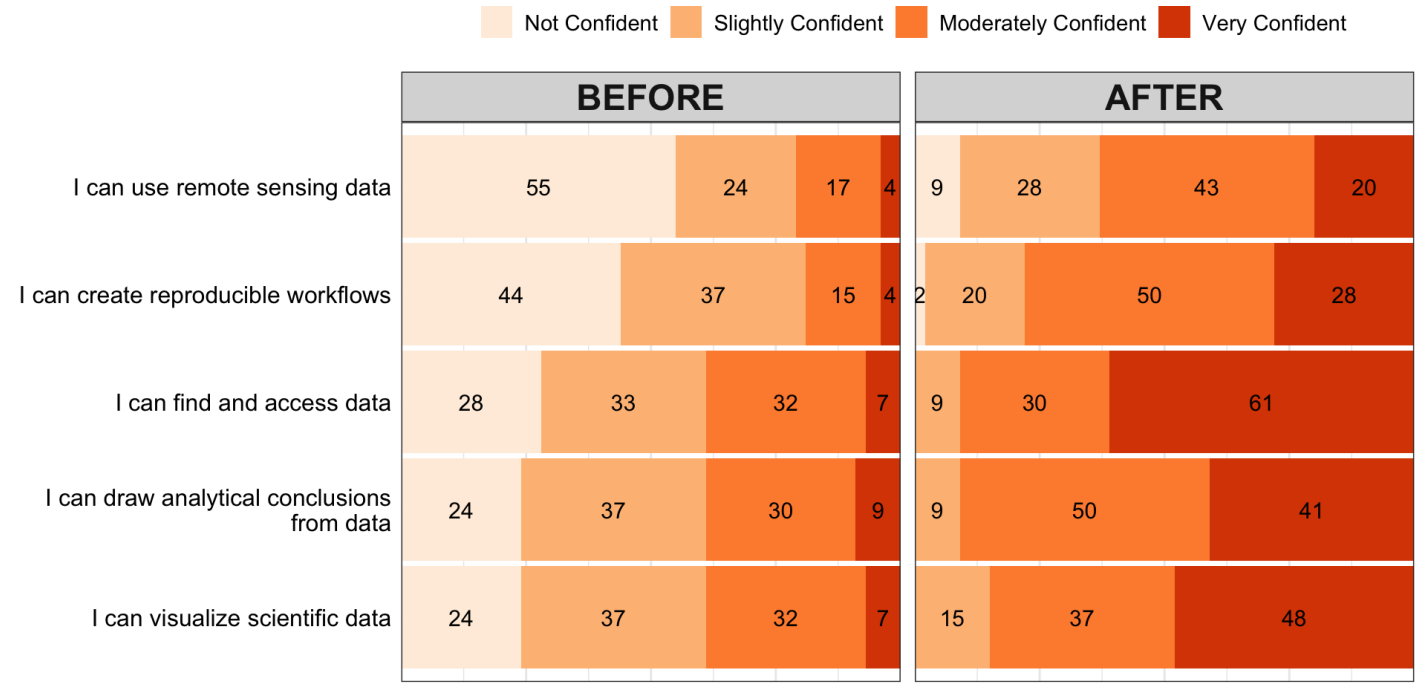
Python Skills



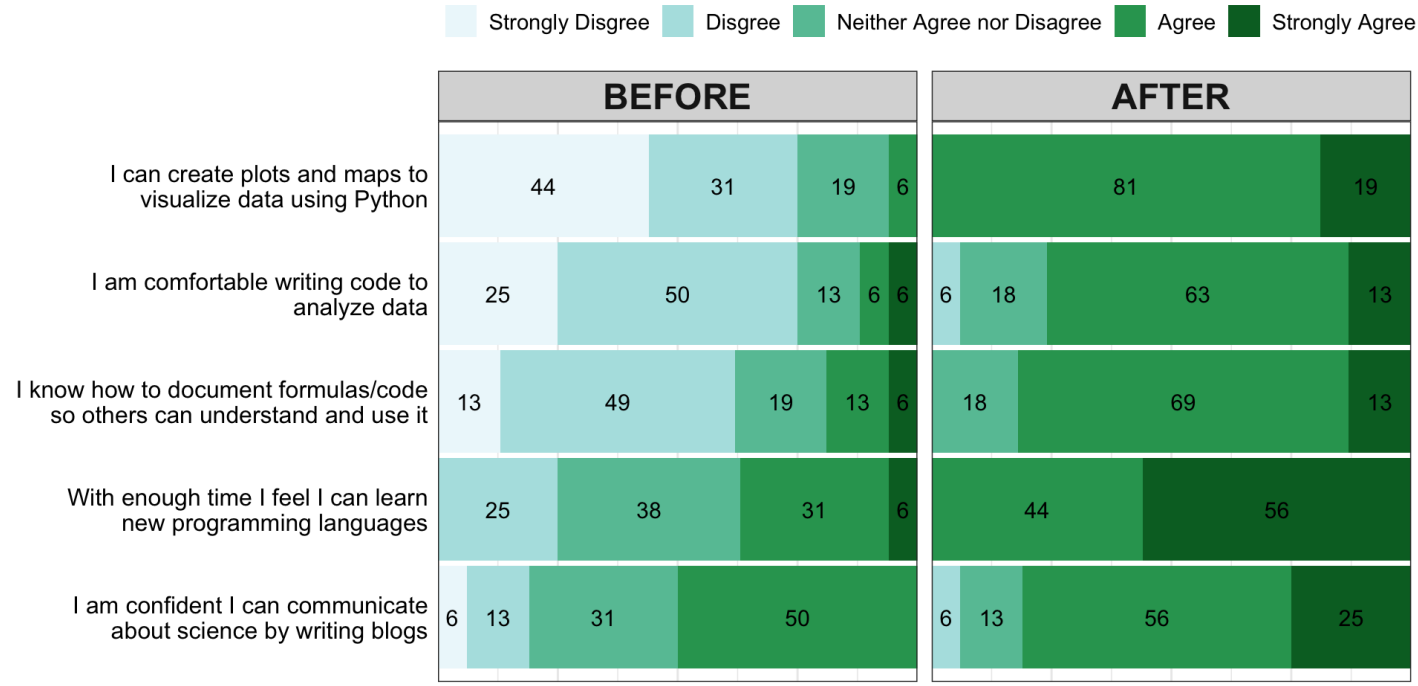


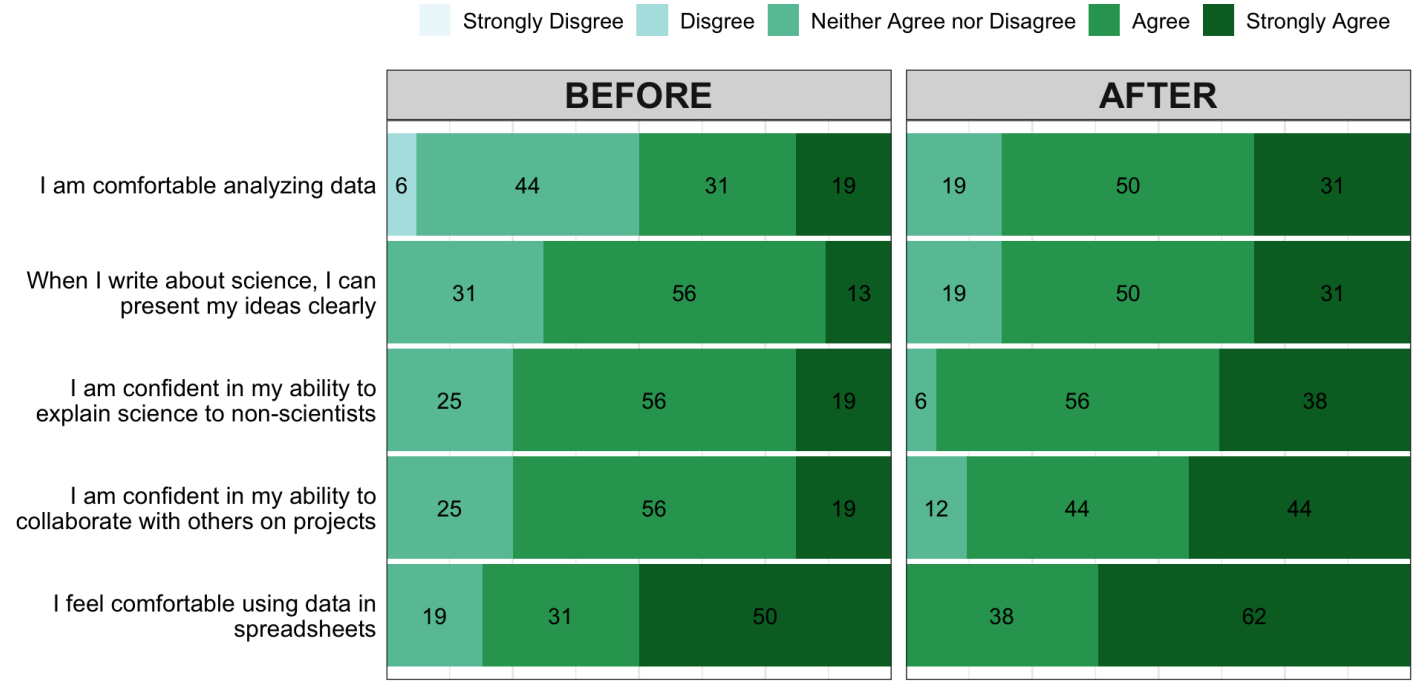
Data Science Skills



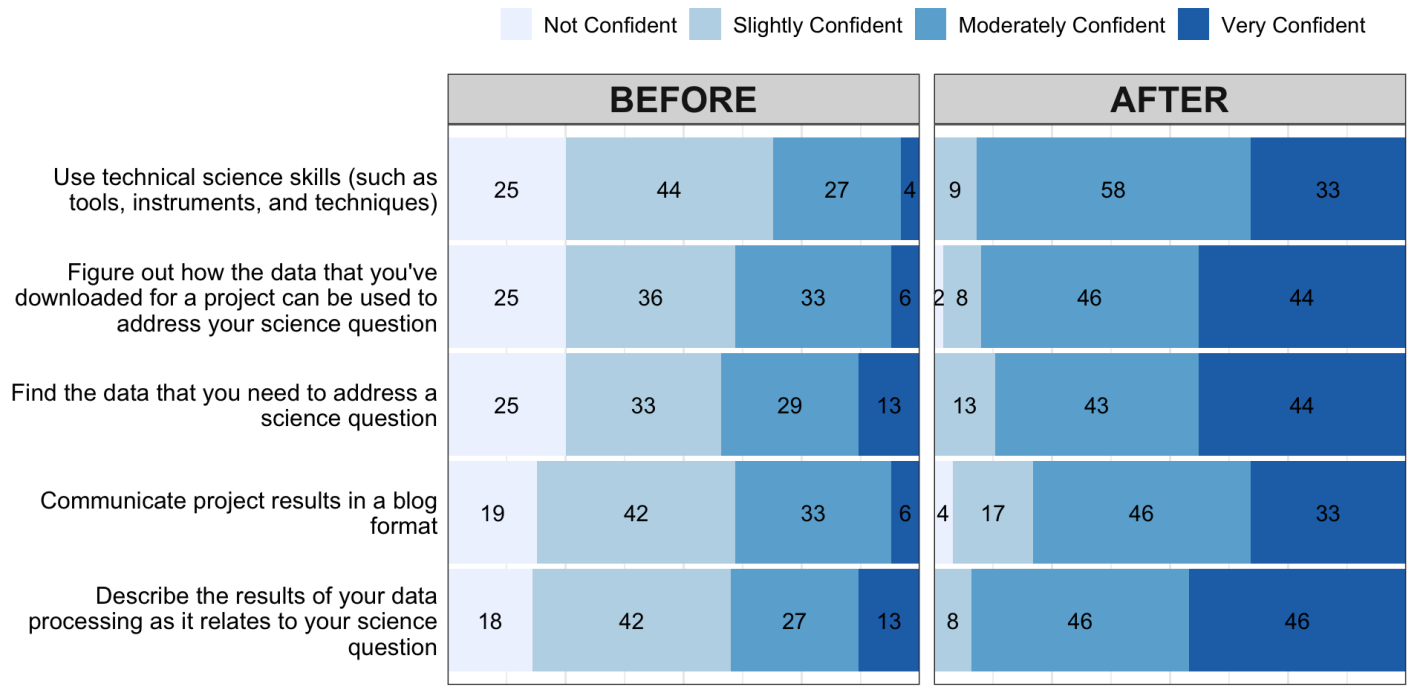


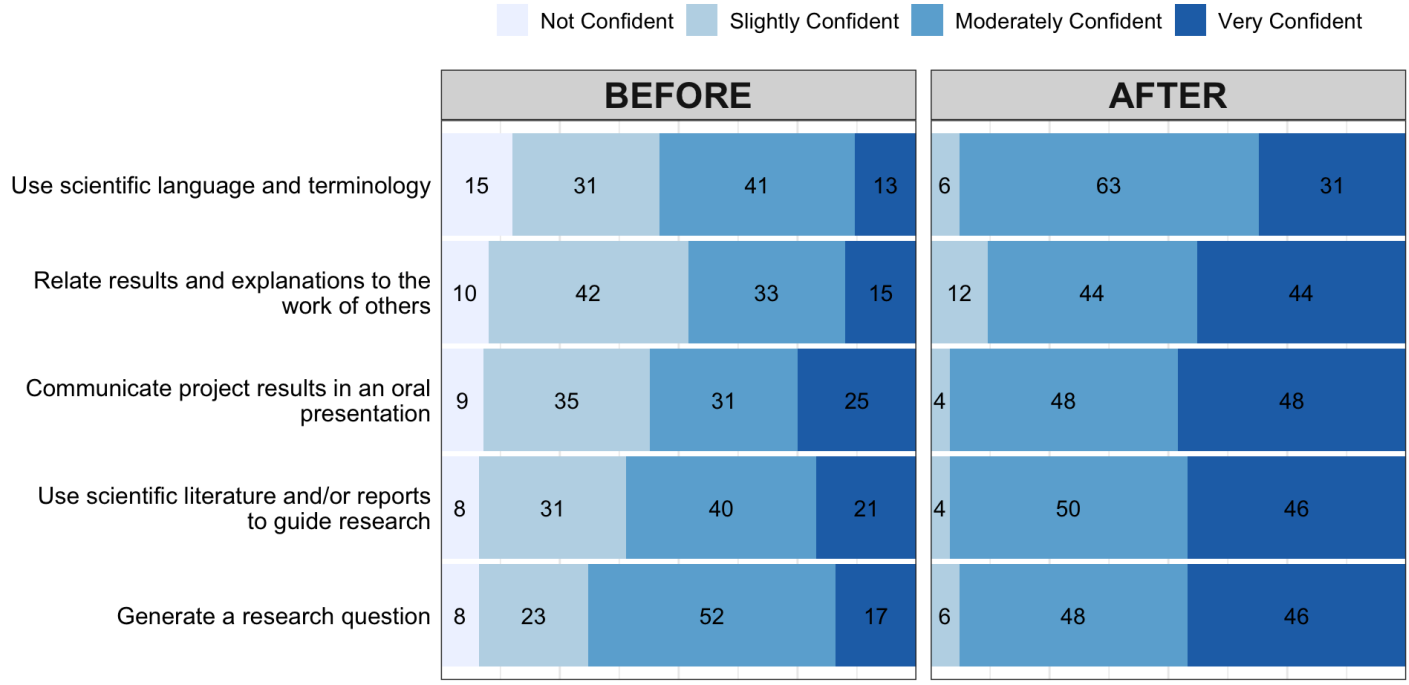
Data Science Communication



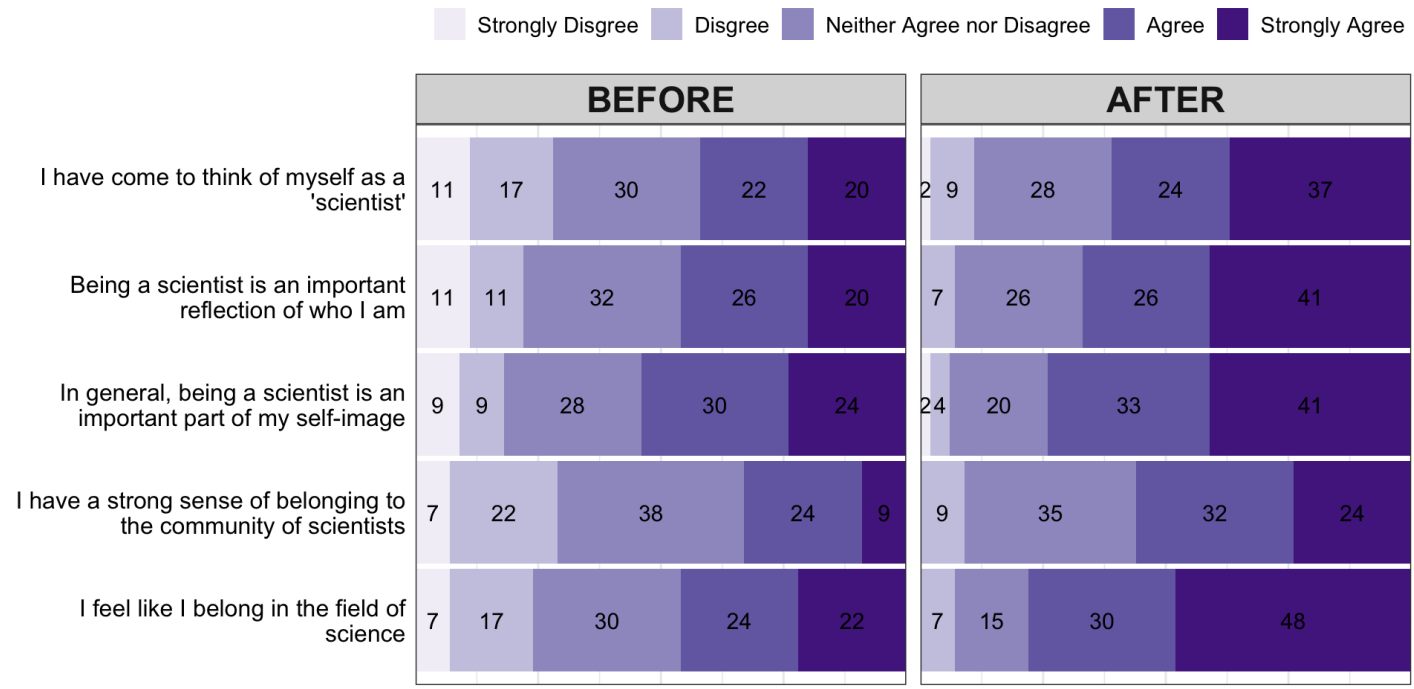


Data Science Practices



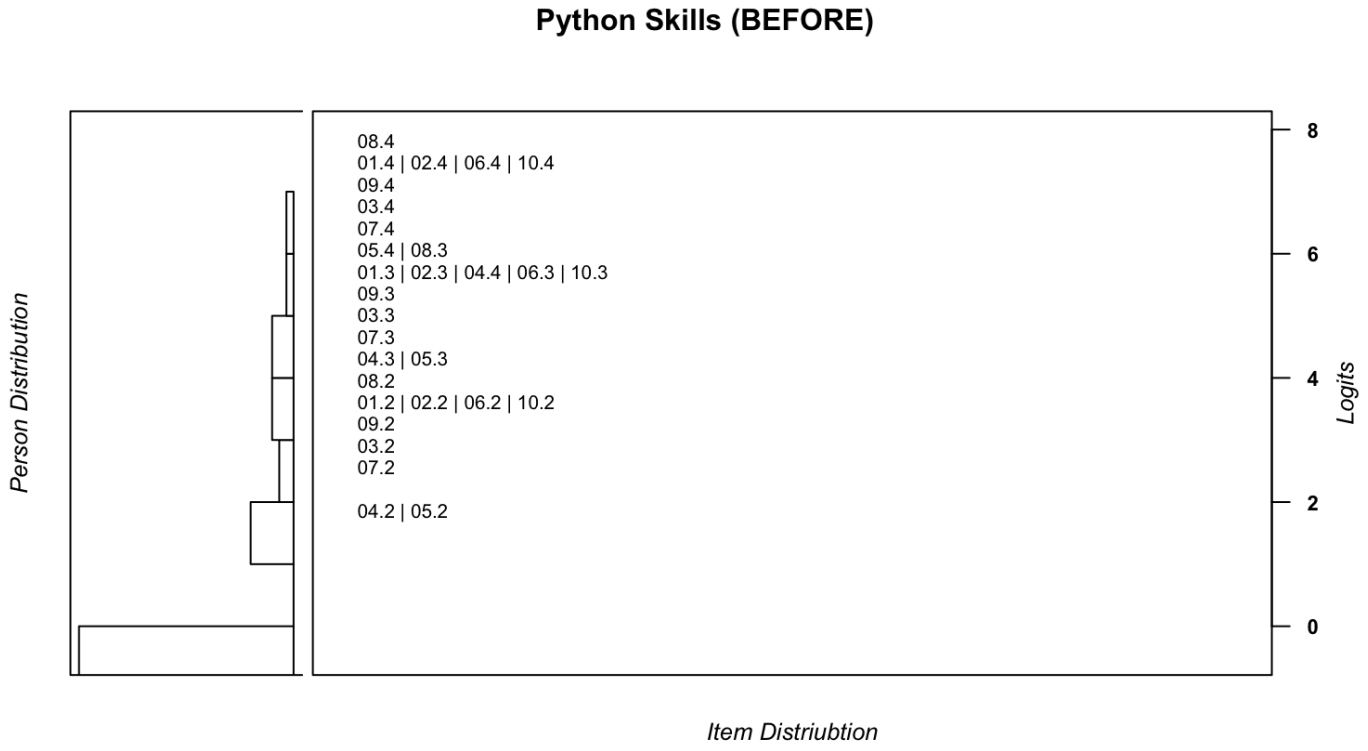


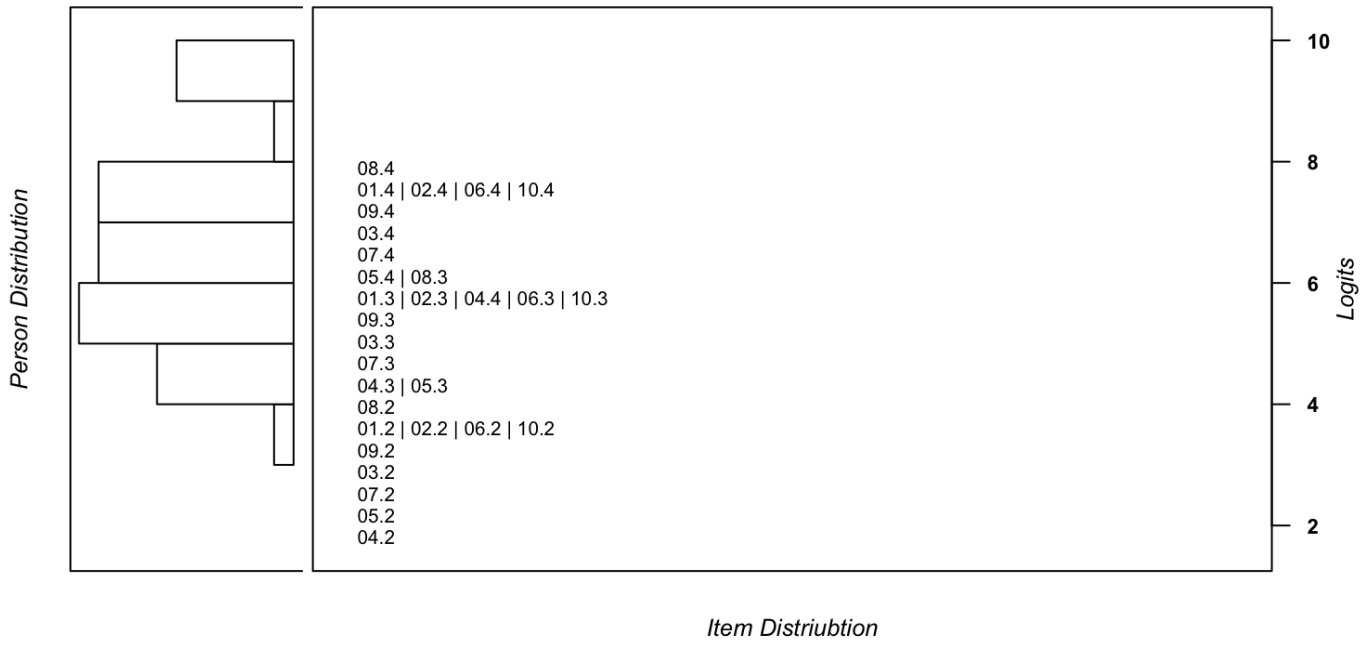
Science Identity



Rasch Modeling (Wright Maps)

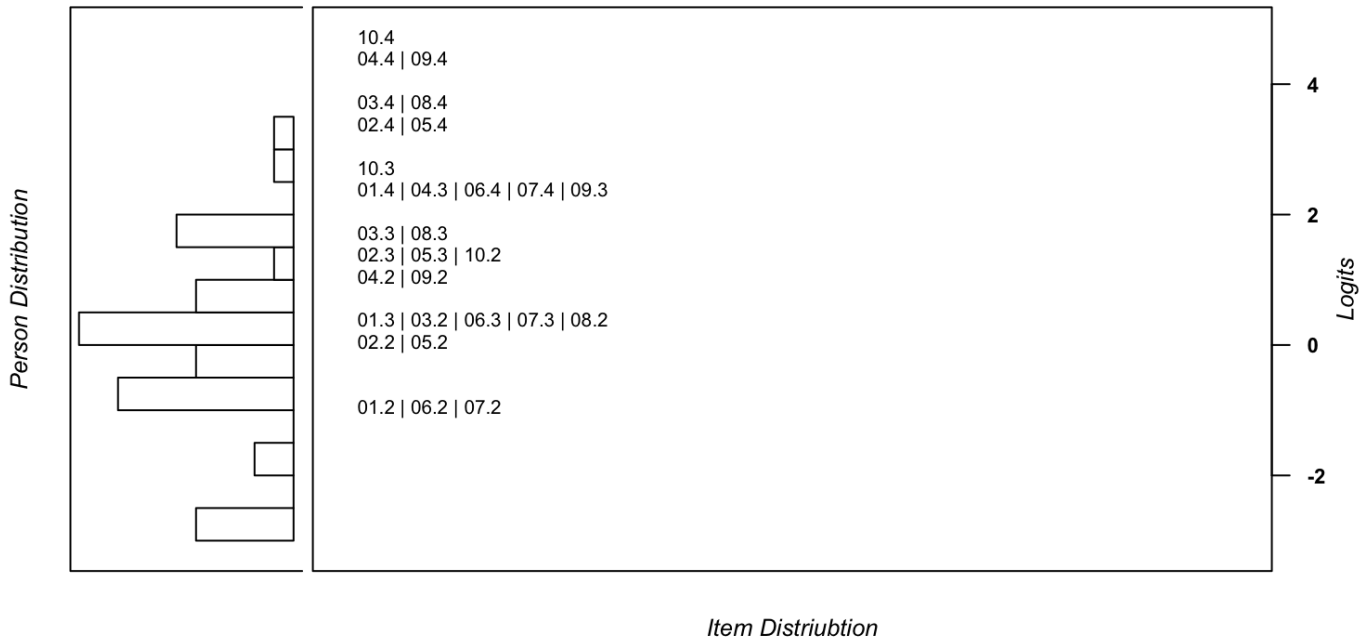
Python Skills



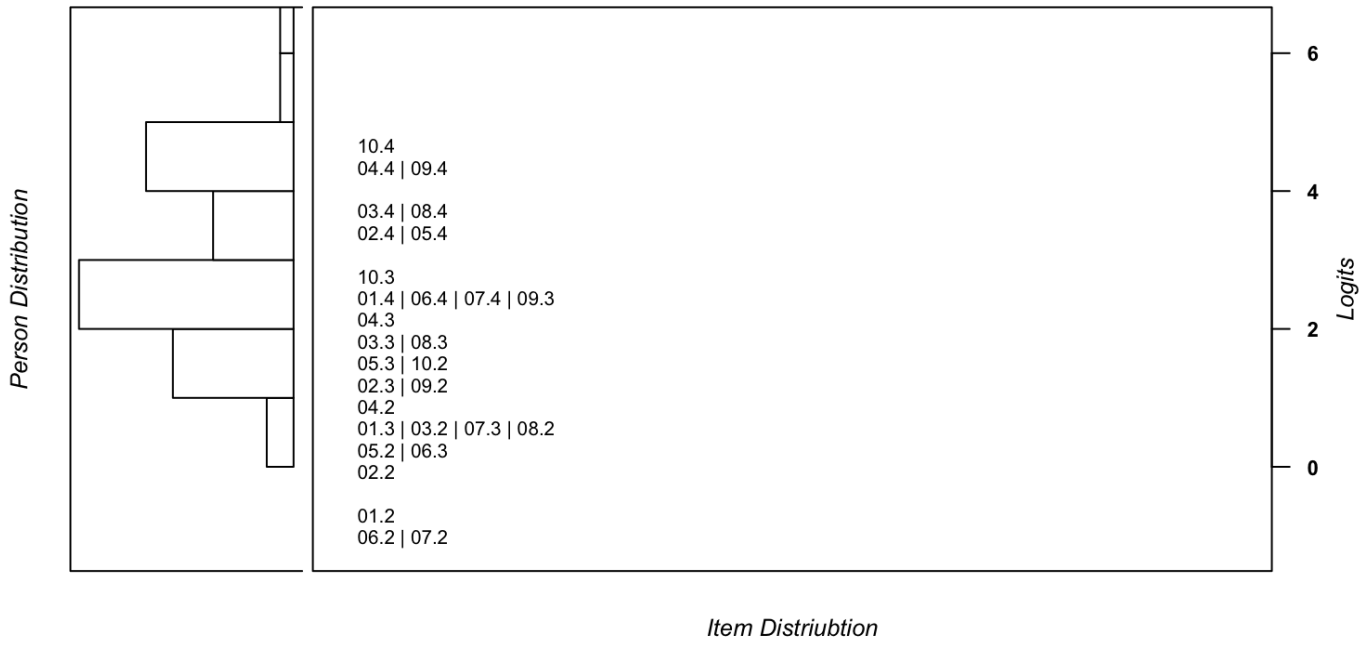


Data Science Skills

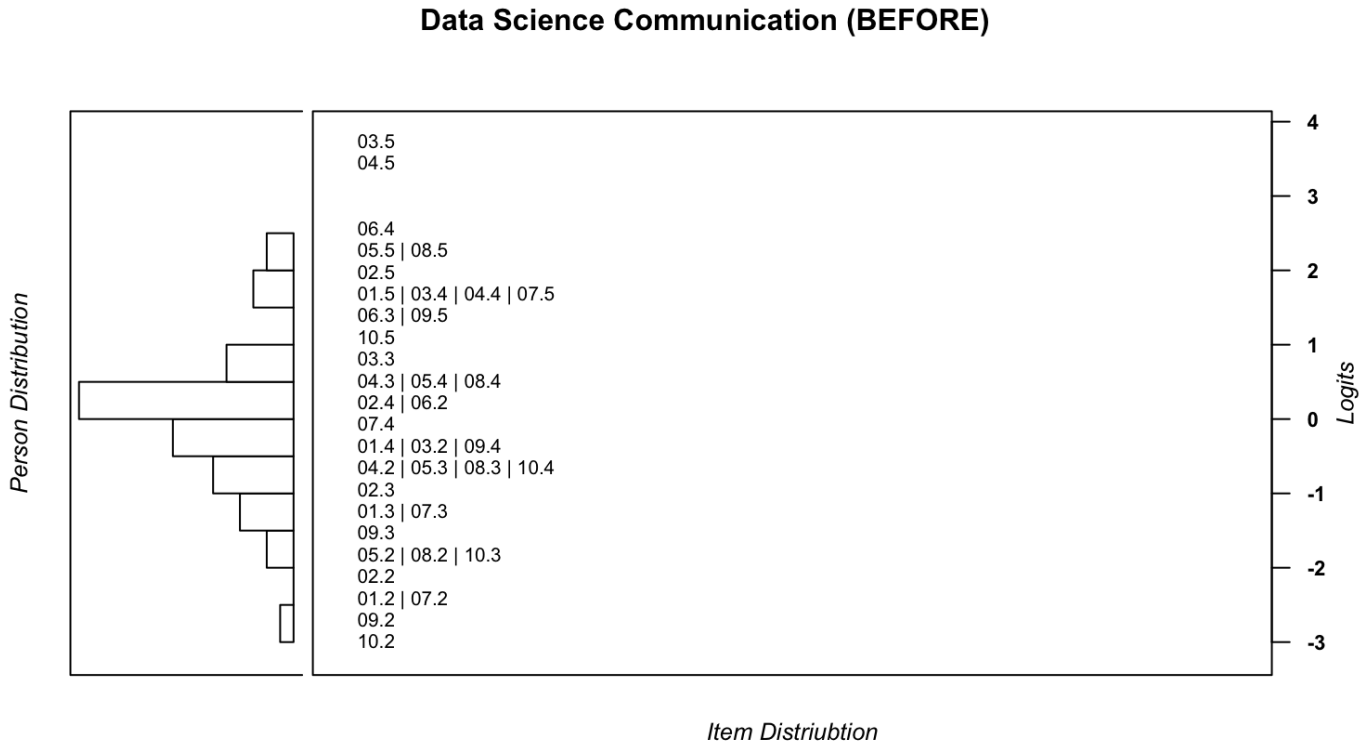
Data Science Skills (BEFORE)

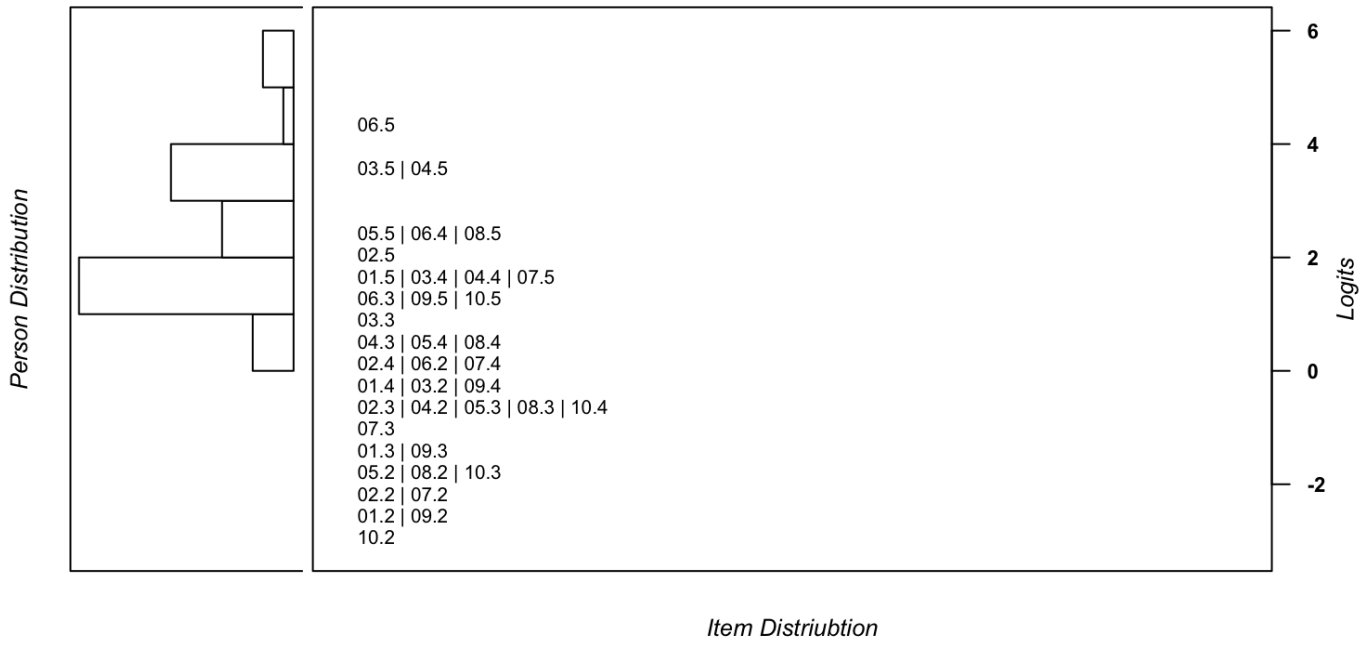


Data Science Skills (AFTER)

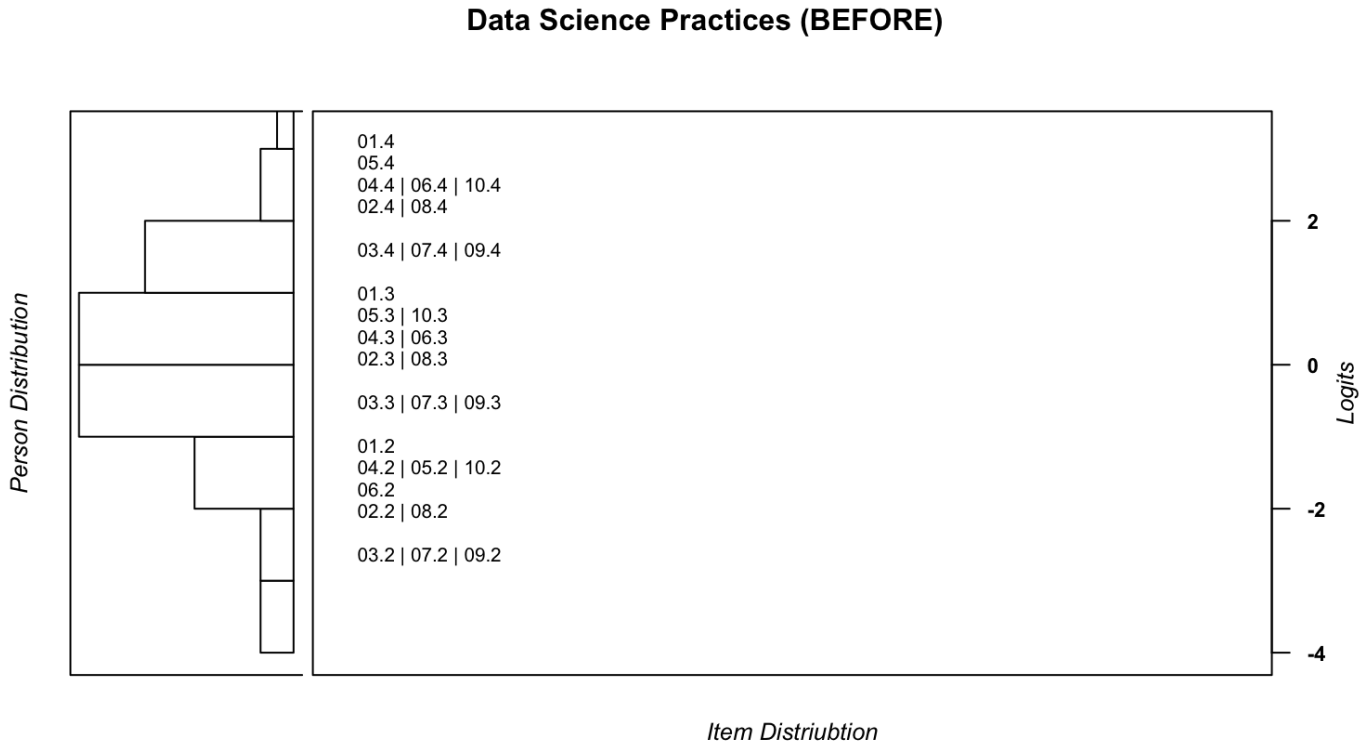


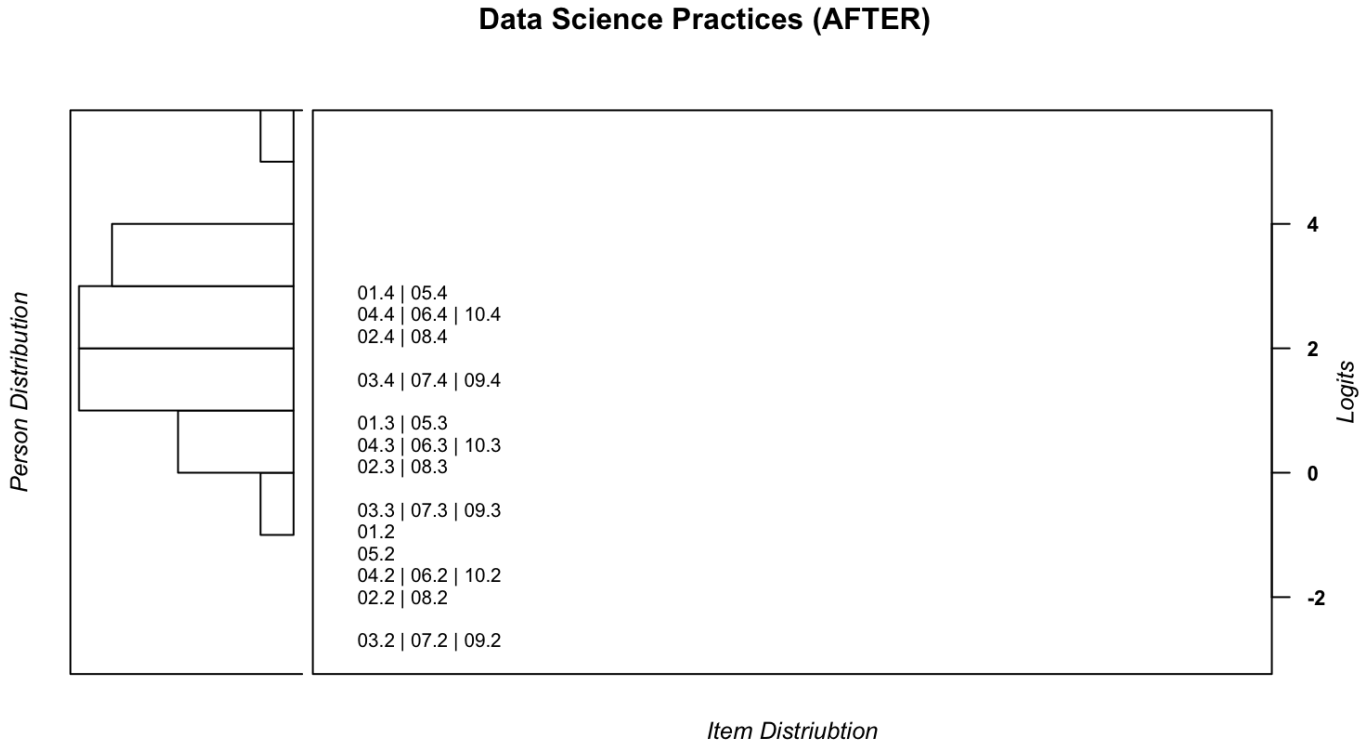
Data Science Communication



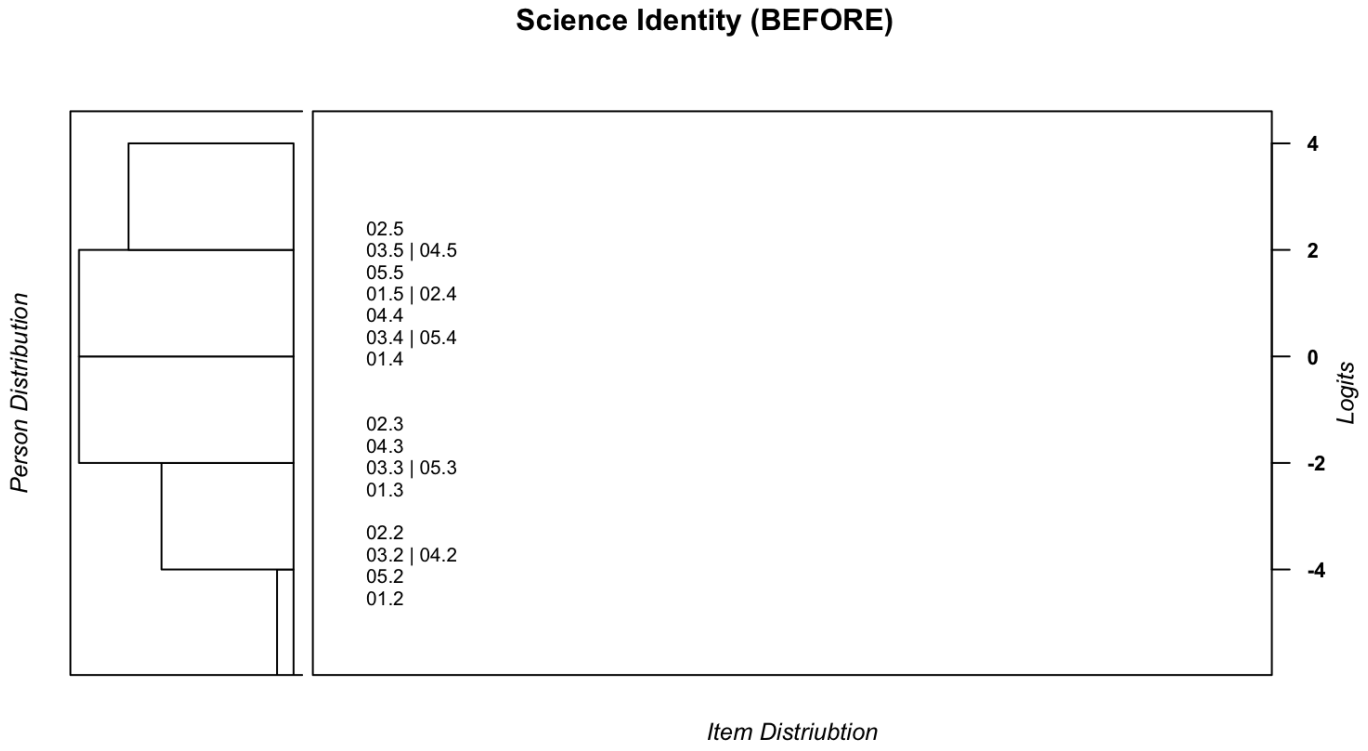


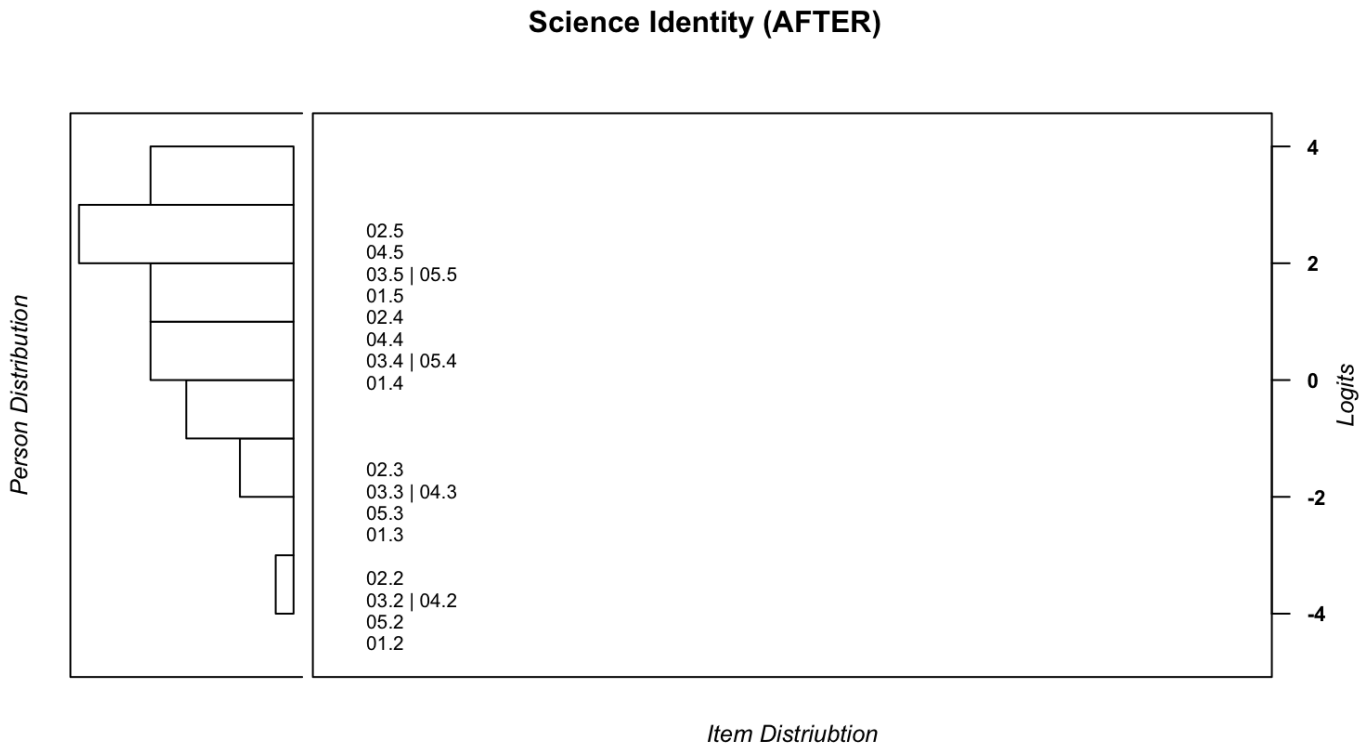
Data Science Practices





Science Identity





Analysis of Variance

```
##
## Call:
## lm(formula = Ability ~ Trial + Dimension + Cohort + Cohort *
##     Dimension, data = abil_trial_dimension_all6)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.228 -1.128  0.032  0.982  5.022
##
## Coefficients:
##                                             Estimate Std. Error t value
## (Intercept)                                  2.75115    0.32182   8.549
## TrialBefore                                 -2.93001    0.16283 -17.994
## DimensionData Science Practices             -0.05436    0.44032  -0.123
## DimensionData Science Skills                 0.16079    0.44032   0.365
## DimensionPython Skills                       2.11823    0.44032   4.811
## DimensionScience Identity                   -0.12235    0.44032  -0.278
## CohortYear 2                                -0.09295    0.46503  -0.200
## CohortYear 3                                -0.13972    0.42259  -0.331
## DimensionData Science Practices:CohortYear 2 -0.24793    0.65766  -0.377
## DimensionData Science Skills:CohortYear 2    0.32802    0.66517   0.493
## DimensionPython Skills:CohortYear 2          0.72306    0.66517   1.087
## DimensionScience Identity:CohortYear 2      -0.59235    0.66517  -0.891
## DimensionData Science Practices:CohortYear 3 -0.12938    0.59763  -0.216
## DimensionData Science Skills:CohortYear 3    0.12335    0.60141   0.205
## DimensionPython Skills:CohortYear 3          0.70361    0.60141   1.170
## DimensionScience Identity:CohortYear 3      -0.62037    0.60141  -1.032
##                                             Pr(>|t|)
## (Intercept)                                  < 2e-16 ***
## TrialBefore                                  < 2e-16 ***
## DimensionData Science Practices                0.902
## DimensionData Science Skills                   0.715
## DimensionPython Skills                      2.05e-06 ***
## DimensionScience Identity                      0.781
## CohortYear 2                                   0.842
## CohortYear 3                                   0.741
## DimensionData Science Practices:CohortYear 2   0.706
## DimensionData Science Skills:CohortYear 2      0.622
## DimensionPython Skills:CohortYear 2            0.278
## DimensionScience Identity:CohortYear 2         0.374
## DimensionData Science Practices:CohortYear 3   0.829
## DimensionData Science Skills:CohortYear 3      0.838
## DimensionPython Skills:CohortYear 3            0.243
## DimensionScience Identity:CohortYear 3         0.303
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.761 on 452 degrees of freedom
## Multiple R-squared:  0.5314, Adjusted R-squared:  0.5158
## F-statistic: 34.17 on 15 and 452 DF,  p-value: < 2.2e-16
```

```
## Analysis of Variance Table
##
## Response: Ability
##                   Df  Sum Sq Mean Sq  F value Pr(>F)
## Trial              1 1004.44 1004.44 323.7867 <2e-16 ***
## Dimension          4  563.71  140.93  45.4291 <2e-16 ***
## Cohort             2    1.34    0.67   0.2158 0.8060
## Dimension:Cohort   8   20.59    2.57   0.8295 0.5769
## Residuals        452 1402.18    3.10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##                        eta.sq  eta.sq.part
## Trial            0.3356796615 0.4173656492
## Dimension        0.1883110526 0.2866604141
## Cohort           0.0004473933 0.0009538292
## Dimension:Cohort 0.0068797627 0.0144690270
```

t-testing & Cohen's *d*

Data Science Communication

```
##
##  Welch Two Sample t-test
##
## data:  comfortpost$Ability and comfortpre$Ability
## t = 10.27, df = 87.882, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.950954 2.887157
## sample estimates:
##    mean of x    mean of y
##  2.415192354 -0.003863267
```

```
## [1] 2.096379
```

Data Science Practices

```
##
##  Welch Two Sample t-test
##
## data:  confpost$Ability and confpre$Ability
## t = 7.5006, df = 92.239, p-value = 3.837e-11
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.531676 2.634901
## sample estimates:
##    mean of x    mean of y
##  2.074584526 -0.008704105
```

```
## [1] 1.531062
```

Science Identity

```
##
##  Welch Two Sample t-test
##
## data:  idpost$Ability and idpre$Ability
## t = 3.3174, df = 82.505, p-value = 0.001353
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.6008267 2.4004040
## sample estimates:
##   mean of x   mean of y
##  1.43789668 -0.06271865
```

```
## [1] 0.6917187
```

Data Science Skills

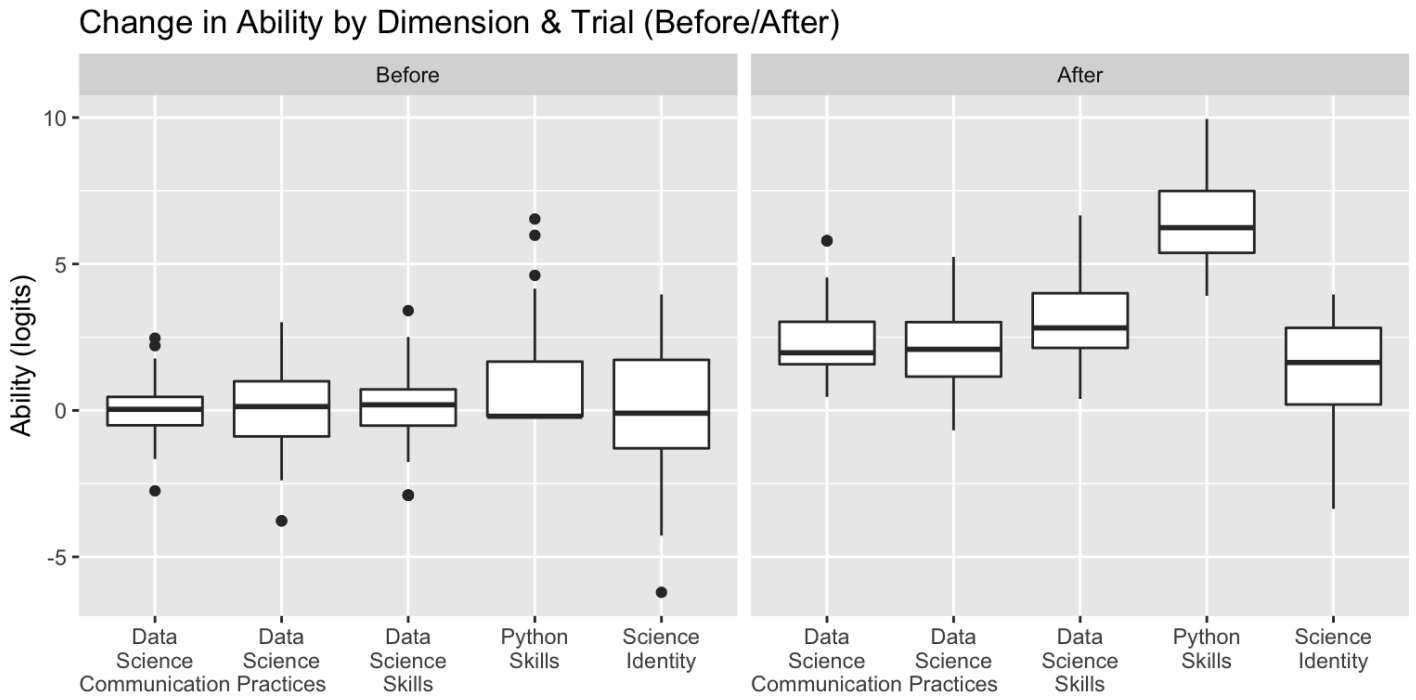
```
##
##  Welch Two Sample t-test
##
## data:  techconfpost$Ability and techconfpre$Ability
## t = 10.45, df = 89.127, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.431845 3.573682
## sample estimates:
##    mean of x    mean of y
## 3.0032369769 0.0004732148
```

```
## [1] 2.179052
```

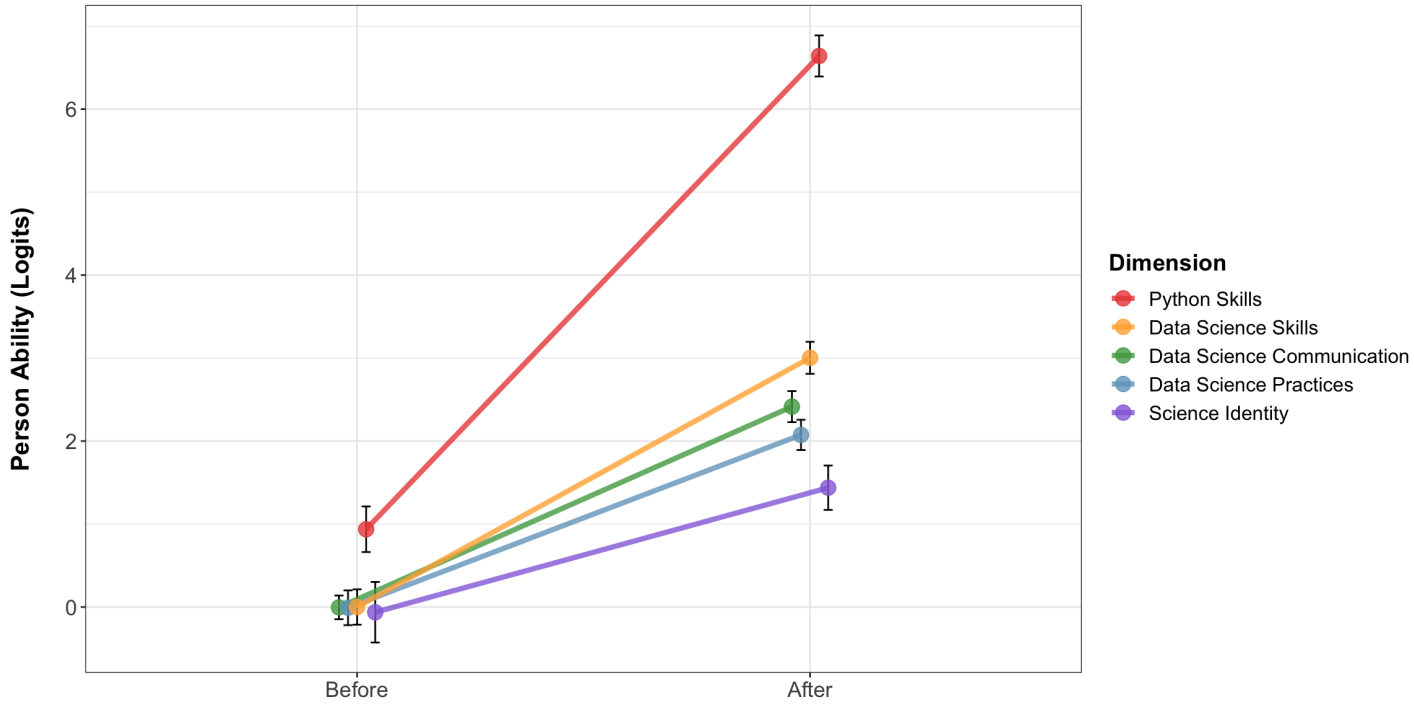
Python Skills

```
##
##  Welch Two Sample t-test
##
## data:  techcomfortpost$Ability and techcomfortpre$Ability
## t = 15.426, df = 89.051, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  4.968754 6.437964
## sample estimates:
## mean of x mean of y
##  6.641079  0.937720
```

```
## [1] 3.216642
```



Growth across the five dimensions



Splitting by cohort