```r
# 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
```
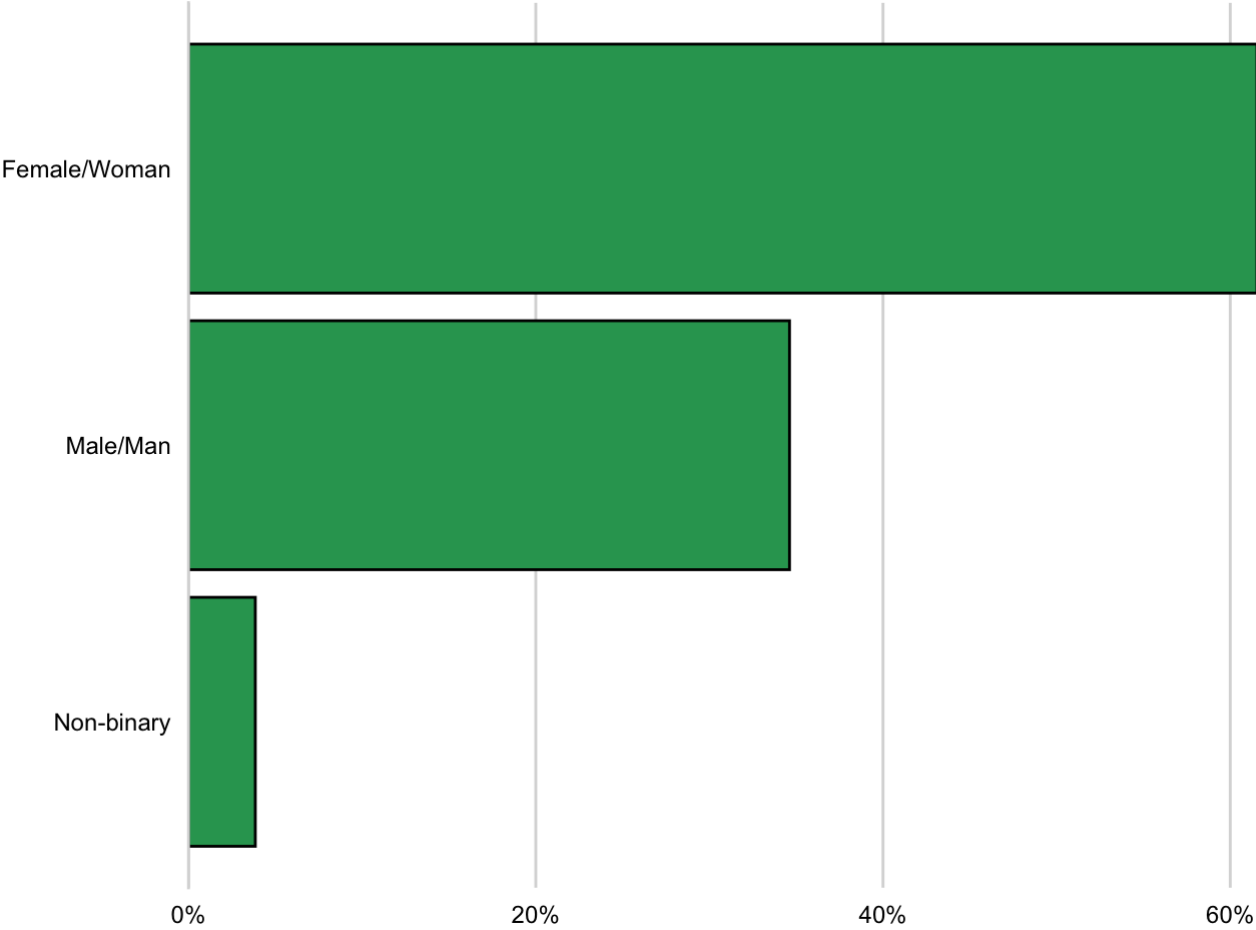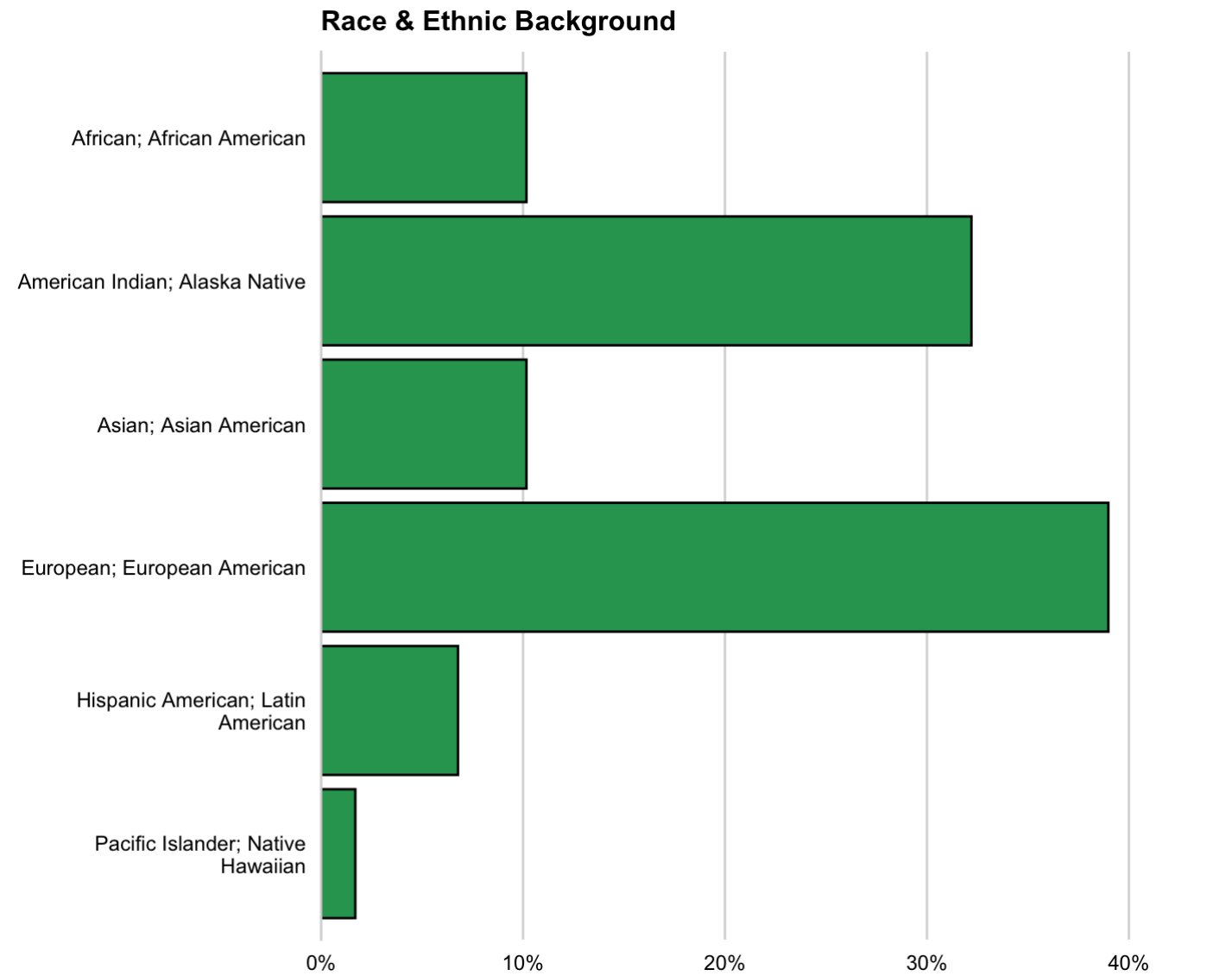

Participant Demographics

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.
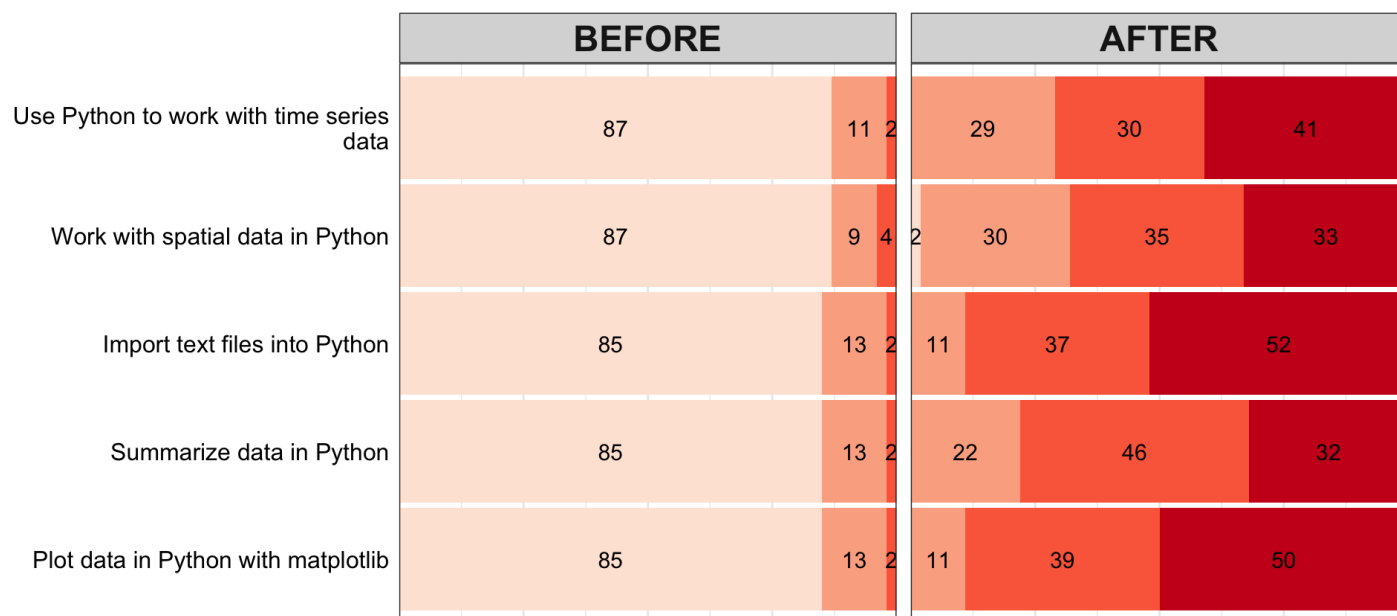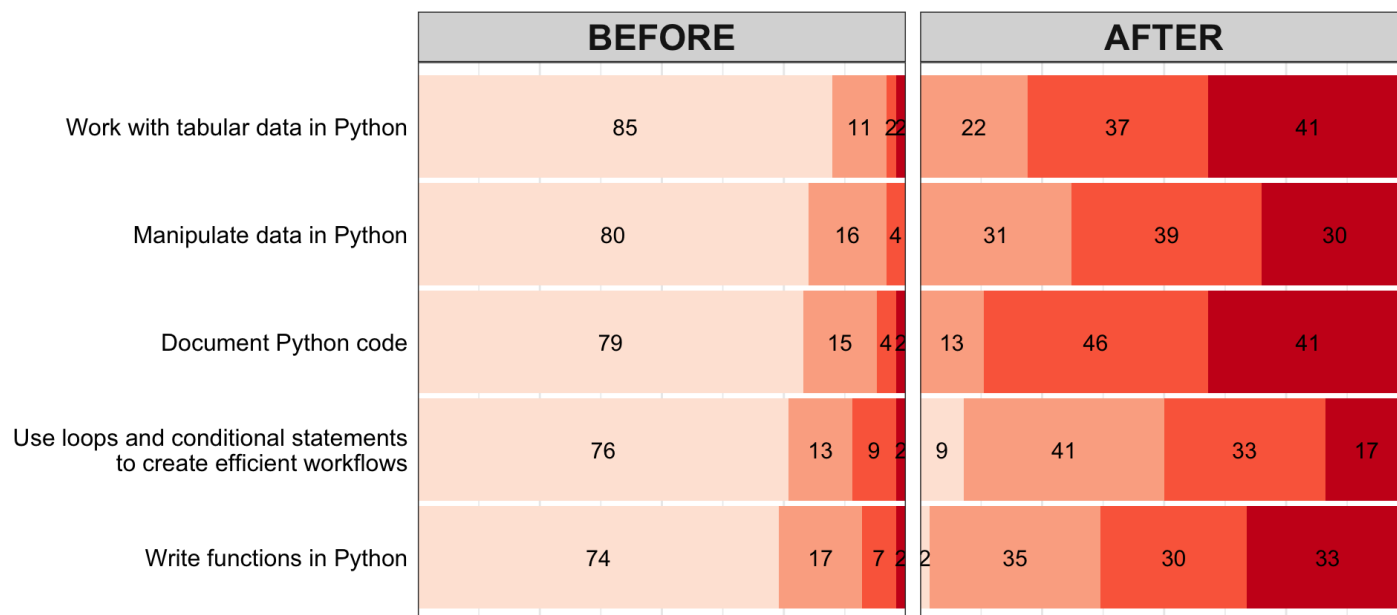
Race & Ethnic Background

Raw Responses (%)

Python Skills

Data Science Skills

Not Confident   Slightly Confident   Moderately Confident   Very Confident
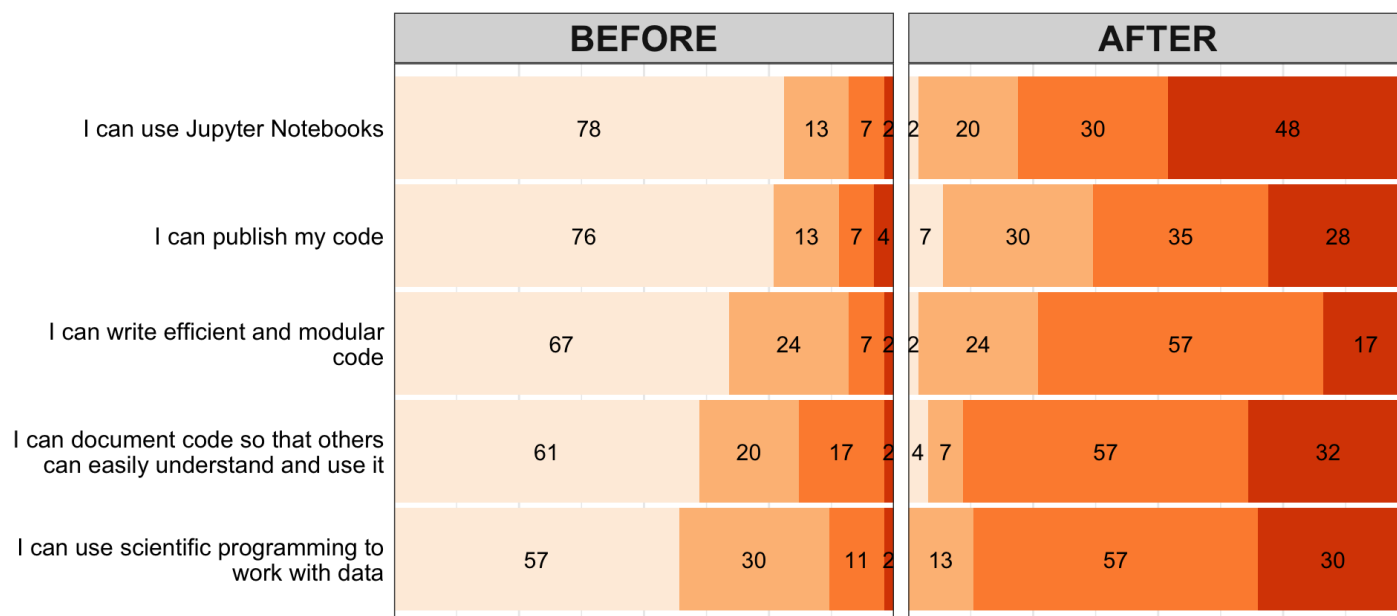
	BEFORE	AFTER
I can use Jupyter Notebooks	78 13 7 2	2 20 30 48
I can publish my code	76 13 7 4	7 30 35 28
I can write efficient and modular code	67 24 7 2	2 24 57 17
I can document code so that others can easily understand and use it	61 20 17 2	4 7 57 32
I can use scientific programming to work with data	57 30 11 2	13 57 30

Not Confident   Slightly Confident   Moderately Confident   Very Confident

	BEFORE	AFTER
I can use remote sensing data	55 24 17 4	9 28 43 20
I can create reproducible workflows	44 37 15 4	2 20 50 28
I can find and access data	28 33 32 7	9 30 61
I can draw analytical conclusions from data	24 37 30 9	9 50 41
I can visualize scientific data	24 37 32 7	15 37 48
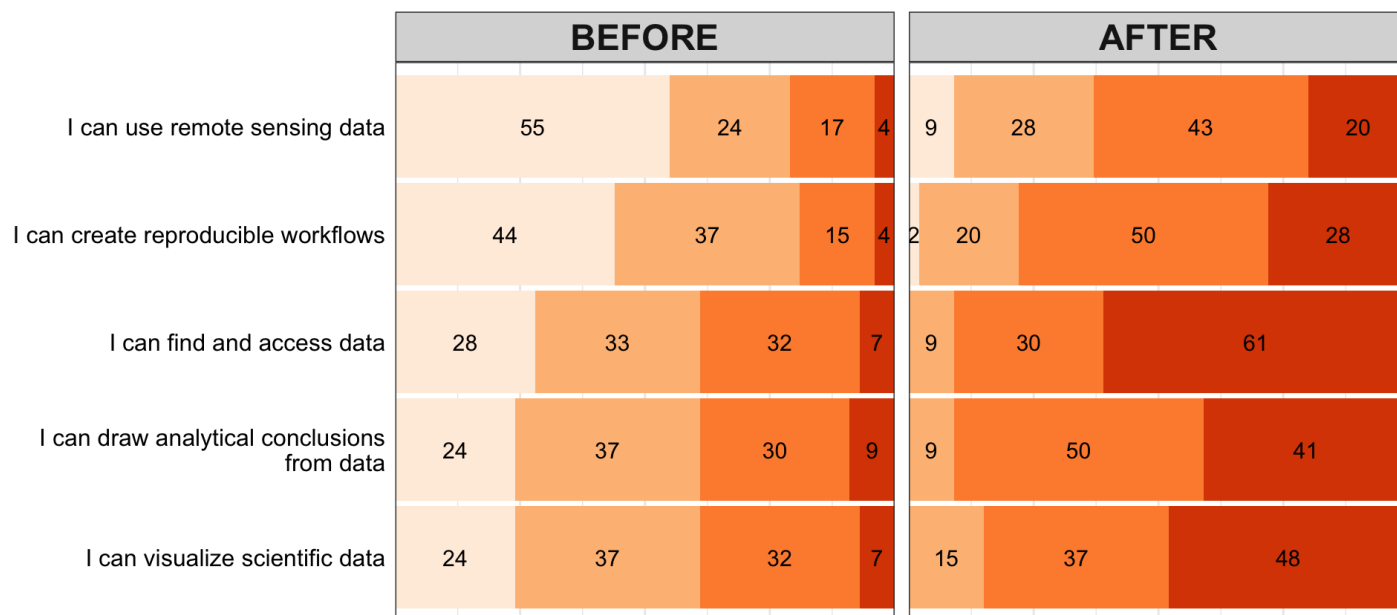
Data Science Communication
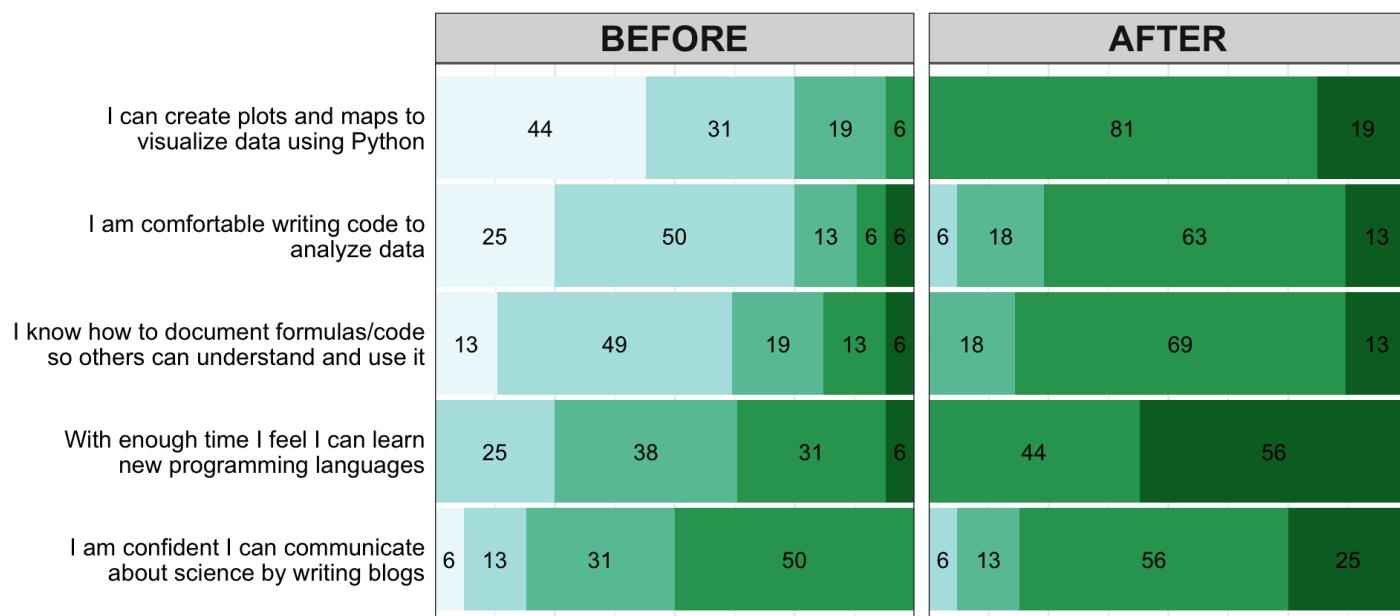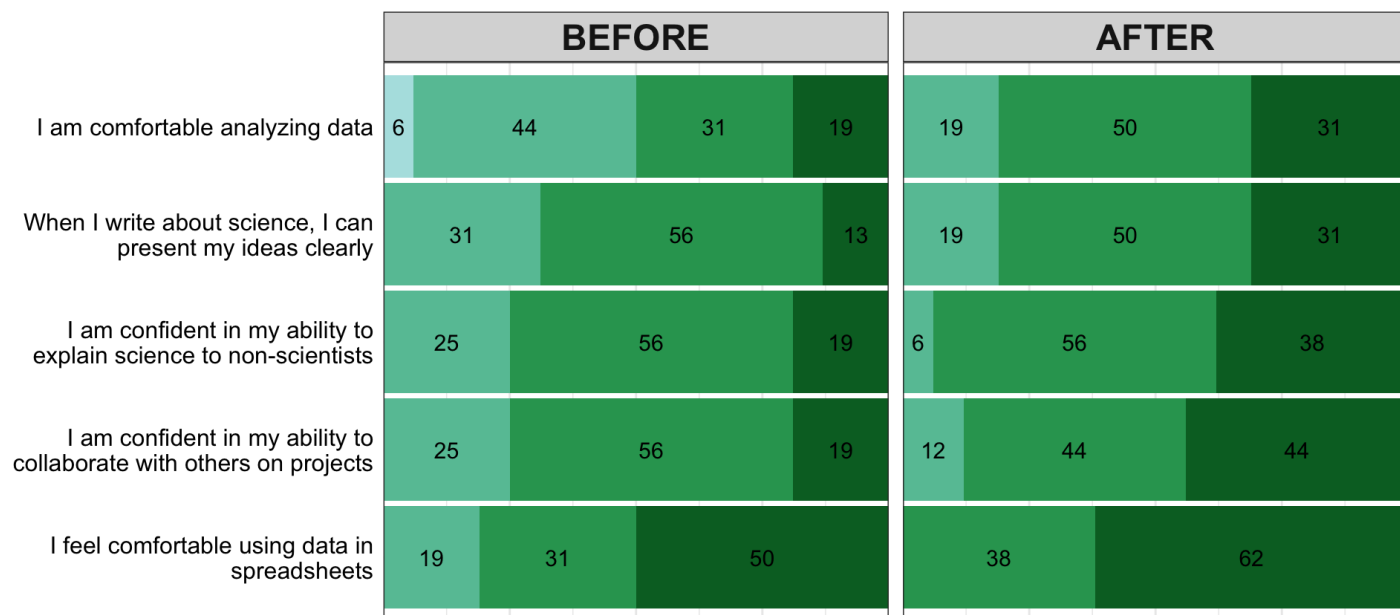
Data Science Practices

Not Confident　　Slightly Confident　　Moderately Confident　　Very Confident

	BEFORE				AFTER		
Use technical science skills (such as tools, instruments, and techniques)	25	44	27	4	9	58	33
Figure out how the data that you've downloaded for a project can be used to address your science question	25	36	33	6	2 8	46	44
Find the data that you need to address a science question	25	33	29	13	13	43	44
Communicate project results in a blog format	19	42	33	6	4 17	46	33
Describe the results of your data processing as it relates to your science question	18	42	27	13	8	46	46

Not Confident　　Slightly Confident　　Moderately Confident　　Very Confident

	BEFORE				AFTER		
Use scientific language and terminology	15	31	41	13	6	63	31
Relate results and explanations to the work of others	10	42	33	15	12	44	44
Communicate project results in an oral presentation	9	35	31	25	4	48	48
Use scientific literature and/or reports to guide research	8	31	40	21	4	50	46
Generate a research question	8	23	52	17	6	48	46

Science Identity

Rasch Modeling

Python Skills

Python Skills (BEFORE)

```
08.4
01.4 | 02.4 | 06.4 | 10.4
09.4
03.4
07.4
05.4 | 08.3
01.3 | 02.3 | 04.4 | 06.3 | 10.3
09.3
03.3
07.3
04.3 | 05.3
08.2
01.2 | 02.2 | 06.2 | 10.2
09.2
03.2
07.2

04.2 | 05.2
```

Person Distribution

Logits

Item Distriubtion

Python Skills (AFTER)

```
08.4
01.4 | 02.4 | 06.4 | 10.4
09.4
03.4
07.4
05.4 | 08.3
01.3 | 02.3 | 04.4 | 06.3 | 10.3
09.3
03.3
07.3
04.3 | 05.3
08.2
01.2 | 02.2 | 06.2 | 10.2
09.2
03.2
07.2
05.2
04.2
```

Person Distribution

Logits

Item Distriubtion
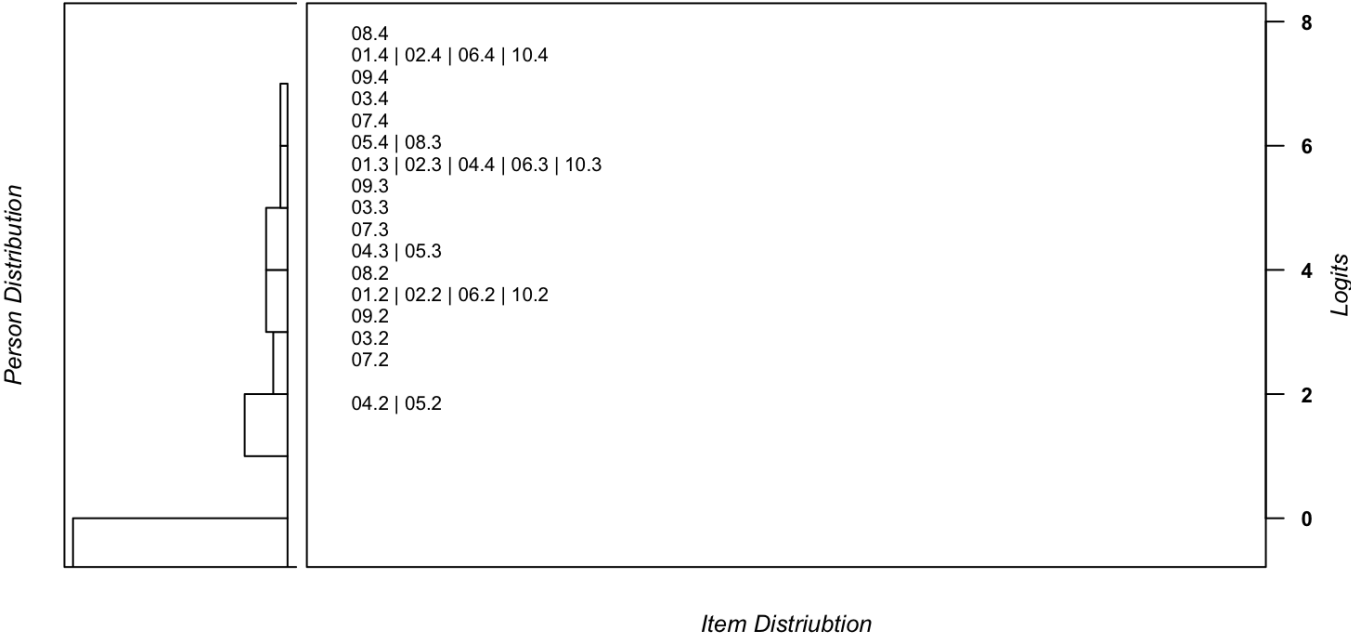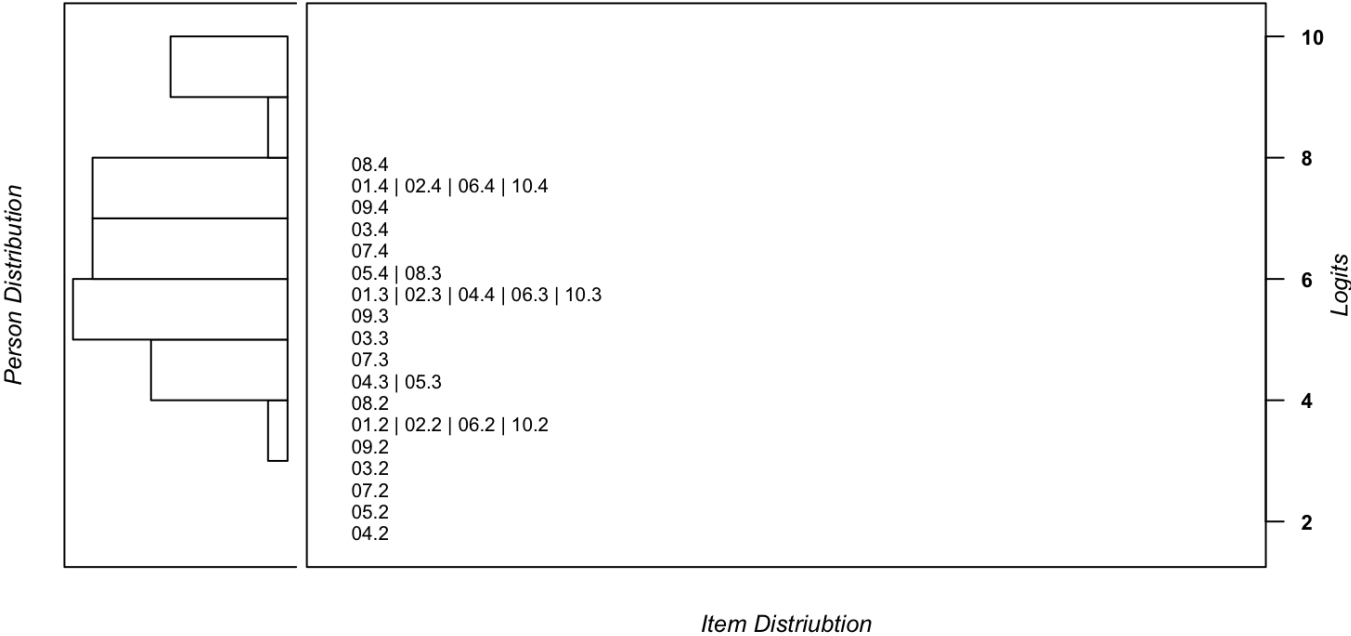
Data Science Skills

Data Science Skills (BEFORE)

```
10.4
04.4 | 09.4

03.4 | 08.4
02.4 | 05.4

10.3
01.4 | 04.3 | 06.4 | 07.4 | 09.3

03.3 | 08.3
02.3 | 05.3 | 10.2
04.2 | 09.2

01.3 | 03.2 | 06.3 | 07.3 | 08.2
02.2 | 05.2


01.2 | 06.2 | 07.2
```

Person Distribution

Logits

4

2

0

-2

Item Distriubtion

Data Science Skills (AFTER)

```
10.4
04.4 | 09.4

03.4 | 08.4
02.4 | 05.4

10.3
01.4 | 06.4 | 07.4 | 09.3
04.3
03.3 | 08.3
05.3 | 10.2
02.3 | 09.2
04.2
01.3 | 03.2 | 07.3 | 08.2
05.2 | 06.3
02.2

01.2
06.2 | 07.2
```

Person Distribution

Logits

6

4

2

0

Item Distriubtion
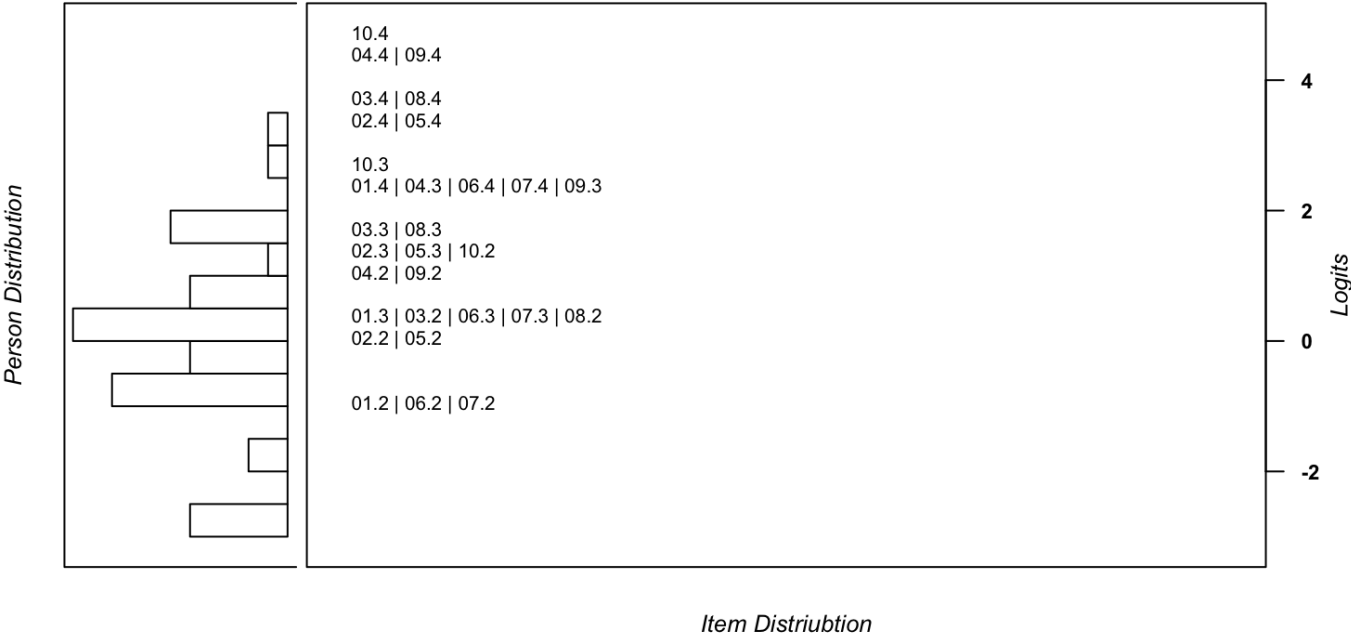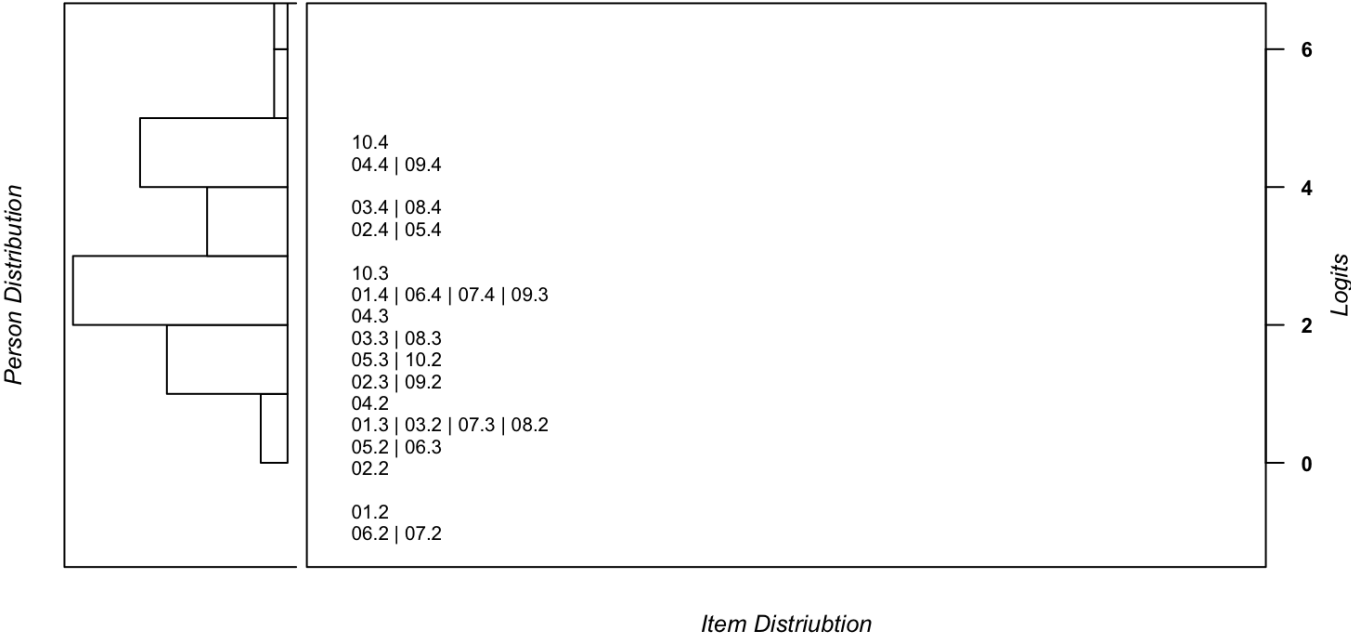
Data Science Communication

Data Science Communication (BEFORE)

Item Distriubtion
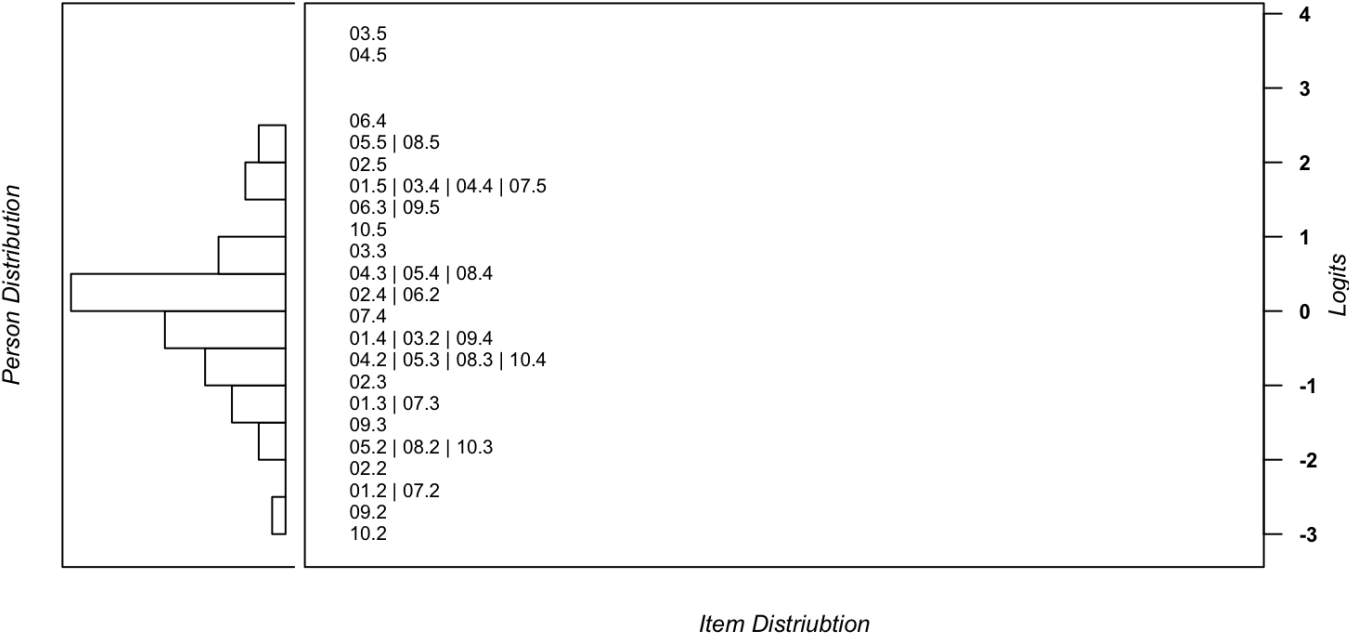
Data Science Communication (AFTER)

Item Distriubtion
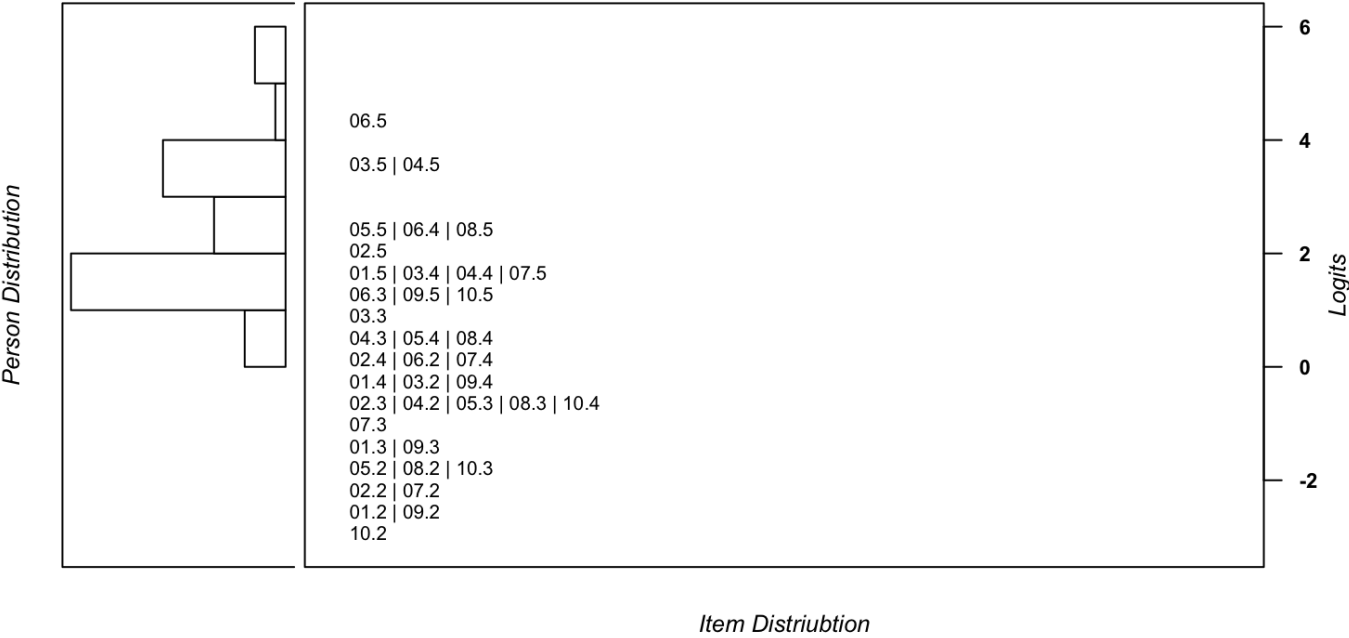
Data Science Practices

Data Science Practices (BEFORE)

Person Distribution

```
01.4
05.4
04.4 | 06.4 | 10.4
02.4 | 08.4
                                                                              2
03.4 | 07.4 | 09.4

01.3
05.3 | 10.3
04.3 | 06.3
02.3 | 08.3                                                                   0

03.3 | 07.3 | 09.3

01.2
04.2 | 05.2 | 10.2
06.2                                                                          -2
02.2 | 08.2

03.2 | 07.2 | 09.2

                                                                              -4
```

Logits

Item Distriubtion

Data Science Practices (AFTER)

Person Distribution

```
                                                                              4

01.4 | 05.4
04.4 | 06.4 | 10.4
02.4 | 08.4                                                                   2

03.4 | 07.4 | 09.4

01.3 | 05.3
04.3 | 06.3 | 10.3
02.3 | 08.3                                                                   0

03.3 | 07.3 | 09.3
01.2
05.2
04.2 | 06.2 | 10.2
02.2 | 08.2                                                                   -2

03.2 | 07.2 | 09.2
```

Logits

Item Distriubtion
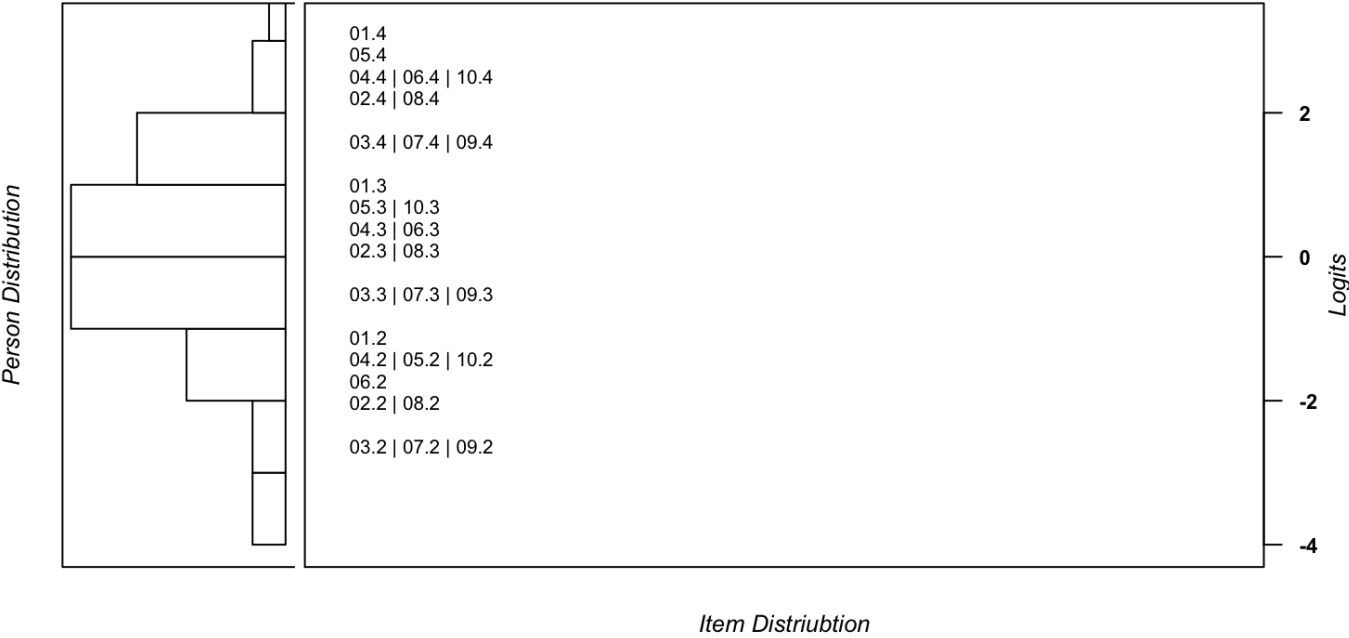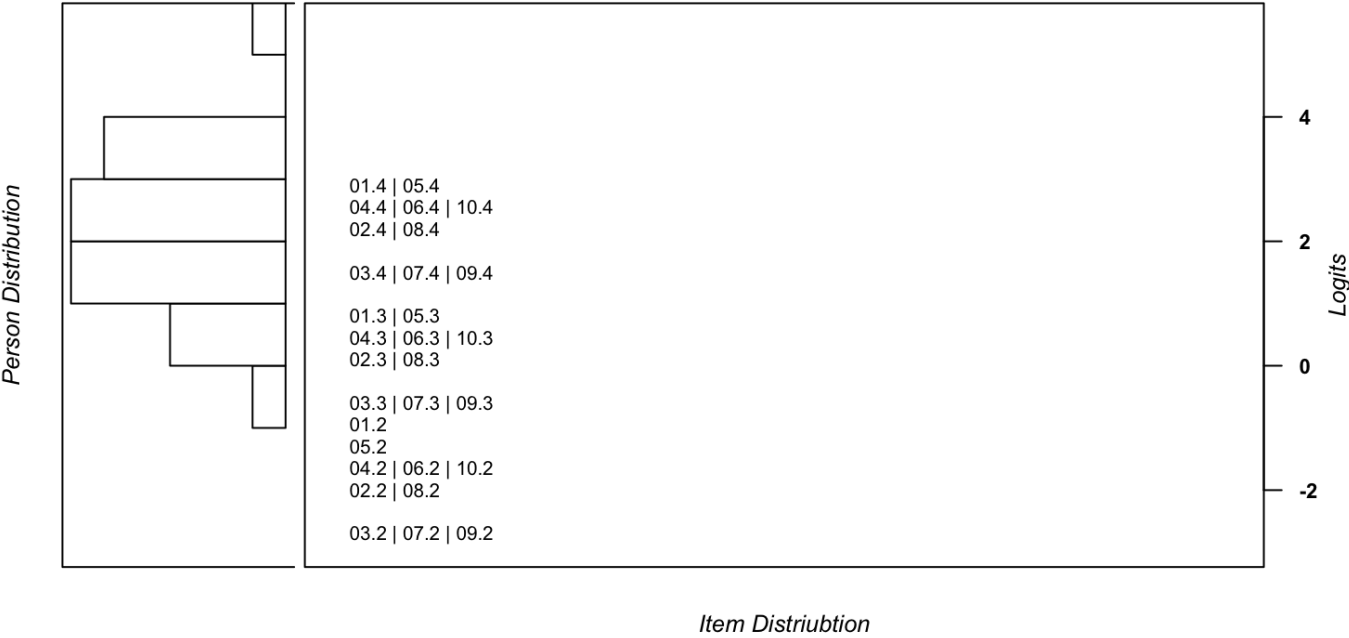
Science Identity

Science Identity (BEFORE)

Person Distribution

```
02.5
03.5 | 04.5
05.5
01.5 | 02.4
04.4
03.4 | 05.4
01.4

02.3
04.3
03.3 | 05.3
01.3

02.2
03.2 | 04.2
05.2
01.2
```

Logits

4
2
0
-2
-4

Item Distriubtion

Science Identity (AFTER)

Person Distribution

```
02.5
04.5
03.5 | 05.5
01.5
02.4
04.4
03.4 | 05.4
01.4

02.3
03.3 | 04.3
05.3
01.3

02.2
03.2 | 04.2
05.2
01.2
```

Logits

4
2
0
-2
-4

Item Distriubtion
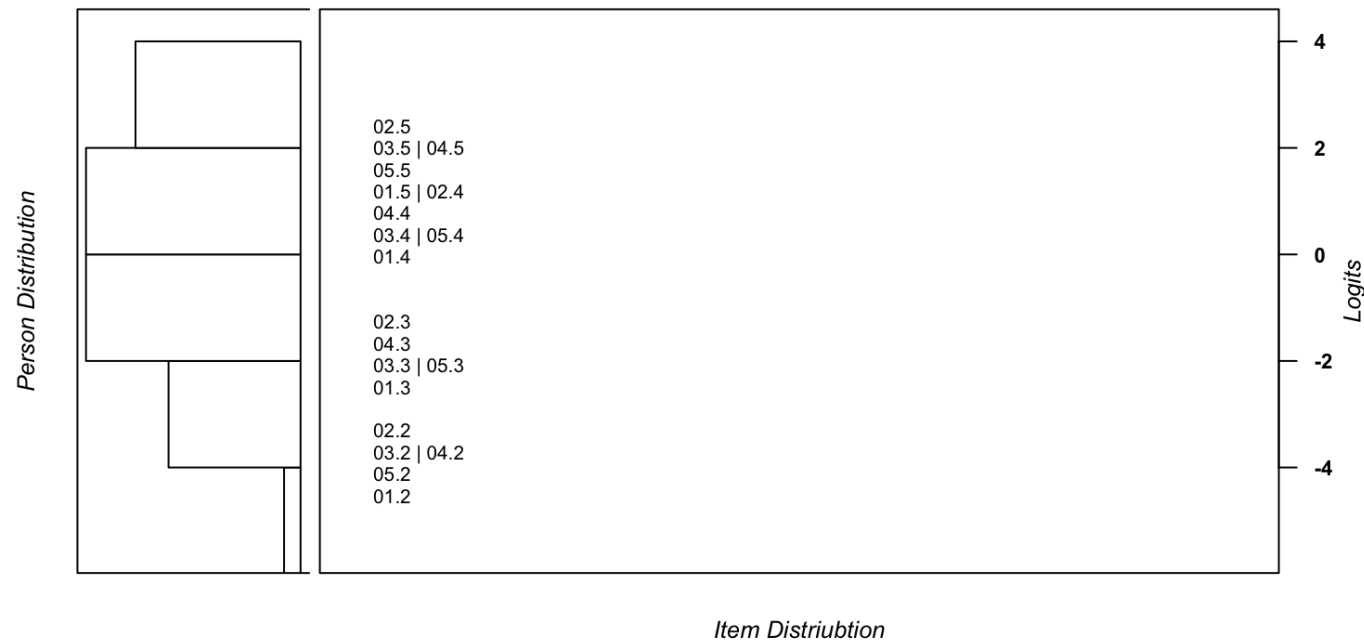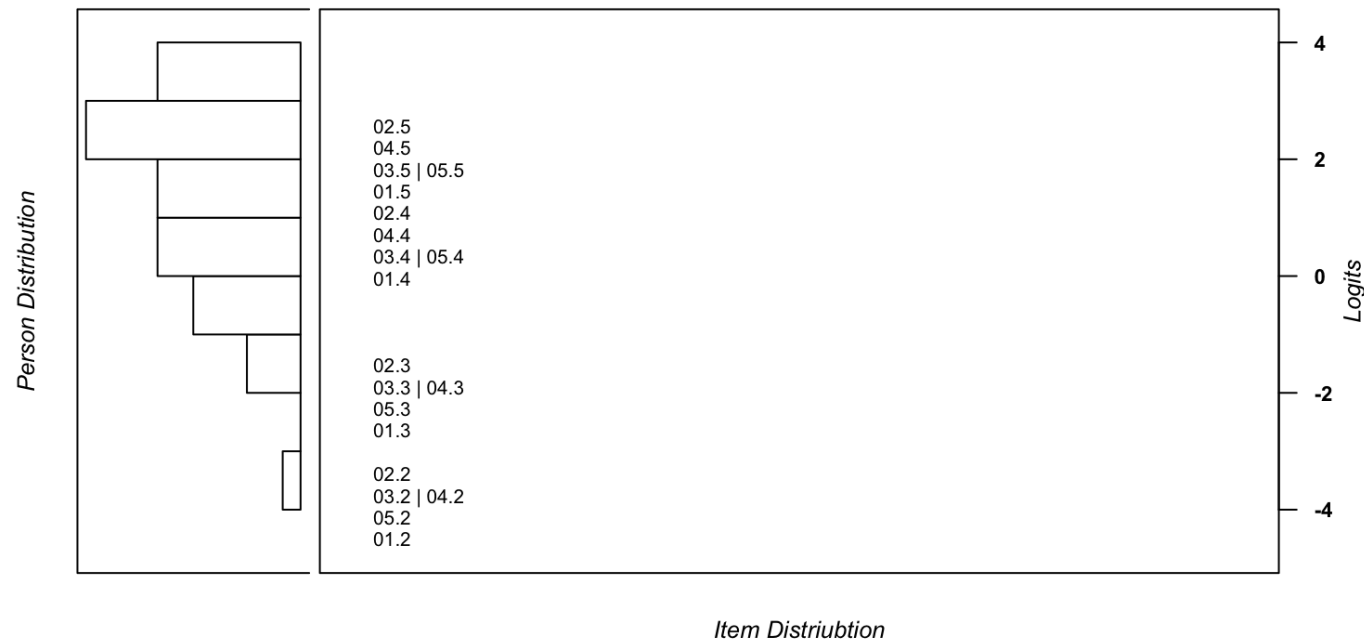
Analysis of Variance

```
##
## 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

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 Communication by Year

```
# science identity pre by year
ds_comms_pre1 <- filter(comfortpre, Cohort == "Year 1")
ds_comms_pre1
```

```
##          Ability                 Dimension Cohort    Trial
## 1    0.31876992 Data Science Communication Year 1 Before
## 2   -0.50742352 Data Science Communication Year 1 Before
## 3    1.77450789 Data Science Communication Year 1 Before
## 4   -0.37201079 Data Science Communication Year 1 Before
## 5    0.90986431 Data Science Communication Year 1 Before
## 6    0.60796709 Data Science Communication Year 1 Before
## 7   -0.50742352 Data Science Communication Year 1 Before
## 8    0.03831258 Data Science Communication Year 1 Before
## 9    0.03831258 Data Science Communication Year 1 Before
## 10  -0.09951847 Data Science Communication Year 1 Before
## 11   0.17763546 Data Science Communication Year 1 Before
## 12   2.46706048 Data Science Communication Year 1 Before
## 13   0.31876992 Data Science Communication Year 1 Before
## 14   0.46207106 Data Science Communication Year 1 Before
## 15   0.90986431 Data Science Communication Year 1 Before
## 16   0.75699922 Data Science Communication Year 1 Before
```

```
ds_comms_pre2 <- filter(comfortpre, Cohort == "Year 2")
ds_comms_pre2
```

```
##          Ability                 Dimension Cohort    Trial
## 1    0.03831258 Data Science Communication Year 2 Before
## 2   -0.50742352 Data Science Communication Year 2 Before
## 3   -1.65983969 Data Science Communication Year 2 Before
## 4   -1.65983969 Data Science Communication Year 2 Before
## 5    0.17763546 Data Science Communication Year 2 Before
## 6    0.90986431 Data Science Communication Year 2 Before
## 7   -0.23617675 Data Science Communication Year 2 Before
## 8   -1.34572147 Data Science Communication Year 2 Before
## 9   -0.50742352 Data Science Communication Year 2 Before
## 10  -0.23617675 Data Science Communication Year 2 Before
## 11   0.03831258 Data Science Communication Year 2 Before
## 12   1.58218261 Data Science Communication Year 2 Before
## 13  -0.50742352 Data Science Communication Year 2 Before
```

```
ds_comms_pre3 <- filter(comfortpre, Cohort == "Year 3")
ds_comms_pre3
```

```
##          Ability                   Dimension Cohort   Trial
## 1  -0.23617675 Data Science Communication Year 3 Before
## 2   0.03831258 Data Science Communication Year 3 Before
## 3   1.58218261 Data Science Communication Year 3 Before
## 4  -1.34572147 Data Science Communication Year 3 Before
## 5   0.03831258 Data Science Communication Year 3 Before
## 6   0.03831258 Data Science Communication Year 3 Before
## 7   0.46207106 Data Science Communication Year 3 Before
## 8  -0.09951847 Data Science Communication Year 3 Before
## 9  -0.23617675 Data Science Communication Year 3 Before
## 10 -0.50742352 Data Science Communication Year 3 Before
## 11  0.46207106 Data Science Communication Year 3 Before
## 12  0.17763546 Data Science Communication Year 3 Before
## 13 -0.37201079 Data Science Communication Year 3 Before
## 14 -1.19864112 Data Science Communication Year 3 Before
## 15  2.21085346 Data Science Communication Year 3 Before
## 16 -2.74709468 Data Science Communication Year 3 Before
## 17 -1.49876757 Data Science Communication Year 3 Before
## 18 -0.37201079 Data Science Communication Year 3 Before
## 19  0.03831258 Data Science Communication Year 3 Before
```

```
# science identity post by year
ds_comms_post1 <- filter(comfortpost, Cohort == "Year 1")
ds_comms_post1
```

```
##       Ability                   Dimension Cohort Trial
## 1  1.7657308 Data Science Communication Year 1 After
## 2  1.7657308 Data Science Communication Year 1 After
## 3  3.8788828 Data Science Communication Year 1 After
## 4  1.0660680 Data Science Communication Year 1 After
## 5  3.0255702 Data Science Communication Year 1 After
## 6  2.1895153 Data Science Communication Year 1 After
## 7  0.7564803 Data Science Communication Year 1 After
## 8  2.4334785 Data Science Communication Year 1 After
## 9  1.3983095 Data Science Communication Year 1 After
## 10 1.5765678 Data Science Communication Year 1 After
## 11 1.3983095 Data Science Communication Year 1 After
## 12 3.0255702 Data Science Communication Year 1 After
## 13 1.5765678 Data Science Communication Year 1 After
## 14 1.5765678 Data Science Communication Year 1 After
## 15 3.4039561 Data Science Communication Year 1 After
## 16 3.0255702 Data Science Communication Year 1 After
```

```
ds_comms_post2 <- filter(comfortpost, Cohort == "Year 2")
ds_comms_post2
```

```
##       Ability                  Dimension Cohort Trial
## 1  1.9687354 Data Science Communication Year 2 After
## 2  2.7083229 Data Science Communication Year 2 After
## 3  3.4039561 Data Science Communication Year 2 After
## 4  4.5387378 Data Science Communication Year 2 After
## 5  1.5765678 Data Science Communication Year 2 After
## 6  1.5765678 Data Science Communication Year 2 After
## 7  1.3983095 Data Science Communication Year 2 After
## 8  2.1895153 Data Science Communication Year 2 After
## 9  1.9687354 Data Science Communication Year 2 After
## 10 0.9090075 Data Science Communication Year 2 After
## 11 3.8788828 Data Science Communication Year 2 After
## 12 5.7939364 Data Science Communication Year 2 After
## 13 3.0255702 Data Science Communication Year 2 After
```

```
ds_comms_post3 <- filter(comfortpost, Cohort == "Year 3")
ds_comms_post3
```

```
##       Ability                  Dimension Cohort Trial
## 1  2.4334785 Data Science Communication Year 3 After
## 2  1.5765678 Data Science Communication Year 3 After
## 3  5.7939364 Data Science Communication Year 3 After
## 4  1.9687354 Data Science Communication Year 3 After
## 5  1.9687354 Data Science Communication Year 3 After
## 6  1.7657308 Data Science Communication Year 3 After
## 7  1.2287191 Data Science Communication Year 3 After
## 8  2.1895153 Data Science Communication Year 3 After
## 9  0.7564803 Data Science Communication Year 3 After
## 10 5.7939364 Data Science Communication Year 3 After
## 11 1.2287191 Data Science Communication Year 3 After
## 12 1.5765678 Data Science Communication Year 3 After
## 13 3.8788828 Data Science Communication Year 3 After
## 14 0.4618885 Data Science Communication Year 3 After
## 15 3.8788828 Data Science Communication Year 3 After
## 16 3.4039561 Data Science Communication Year 3 After
## 17 1.7657308 Data Science Communication Year 3 After
## 18 2.4334785 Data Science Communication Year 3 After
## 19 3.0255702 Data Science Communication Year 3 After
```

```
# t-testing data science practices by year
# Year 1
t.test(ds_comms_post1$Ability, ds_comms_pre1$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  ds_comms_post1$Ability and ds_comms_pre1$Ability
## t = 5.4751, df = 29.515, p-value = 6.412e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   1.040732 2.280408
## sample estimates:
## mean of x mean of y
## 2.1164297 0.4558599
```

```
cohensD(ds_comms_post1$Ability, ds_comms_pre1$Ability)
```

```
## [1] 1.93574
```

```
# Year 2
t.test(ds_comms_post2$Ability, ds_comms_pre2$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  ds_comms_post2$Ability and ds_comms_pre2$Ability
## t = 6.3928, df = 20.827, p-value = 2.549e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   2.015839 3.961171
## sample estimates:
##   mean of x   mean of y
##   2.6874496 -0.3010552
```

```
cohensD(ds_comms_post2$Ability, ds_comms_pre2$Ability)
```

```
## [1] 2.507472
```

```
# Year 3
t.test(ds_comms_post3$Ability, ds_comms_pre3$Ability)
```

```
##
##  Welch Two Sample t-test
##
## data:  ds_comms_post3$Ability and ds_comms_pre3$Ability
## t = 6.2996, df = 32.67, p-value = 4.182e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.806121 3.530194
## sample estimates:
##  mean of x  mean of y
##  2.4805006 -0.1876567
```

```
cohensD(ds_comms_post3$Ability, ds_comms_pre3$Ability)
```

```
## [1] 2.043861
```

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
```

Data Science Practices by Year

```
# science identity pre by year
ds_pract_pre1 <- filter(confpre, Cohort == "Year 1")
ds_pract_pre1
```

```
##          Ability                Dimension Cohort   Trial
## 1    3.01420897 Data Science Practices Year 1 Before
## 2    0.13092222 Data Science Practices Year 1 Before
## 3    0.73619507 Data Science Practices Year 1 Before
## 4   -0.06921477 Data Science Practices Year 1 Before
## 5    1.59756681 Data Science Practices Year 1 Before
## 6    1.15428309 Data Science Practices Year 1 Before
## 7   -1.56252692 Data Science Practices Year 1 Before
## 8   -1.32768269 Data Science Practices Year 1 Before
## 9   -0.88813267 Data Science Practices Year 1 Before
## 10   0.73619507 Data Science Practices Year 1 Before
## 11  -0.47277589 Data Science Practices Year 1 Before
## 12   0.94296217 Data Science Practices Year 1 Before
## 13   0.94296217 Data Science Practices Year 1 Before
## 14   1.83444032 Data Science Practices Year 1 Before
## 15   0.33123446 Data Science Practices Year 1 Before
## 16   1.15428309 Data Science Practices Year 1 Before
```

```
ds_pract_pre2 <- filter(confpre, Cohort == "Year 2")
ds_pract_pre2
```

```
##          Ability                Dimension Cohort   Trial
## 1   -0.88813267 Data Science Practices Year 2 Before
## 2   -0.06921477 Data Science Practices Year 2 Before
## 3   -0.88813267 Data Science Practices Year 2 Before
## 4   -2.38567830 Data Science Practices Year 2 Before
## 5    0.73619507 Data Science Practices Year 2 Before
## 6   -0.27011420 Data Science Practices Year 2 Before
## 7    1.59756681 Data Science Practices Year 2 Before
## 8   -3.77058166 Data Science Practices Year 2 Before
## 9   -0.67833205 Data Science Practices Year 2 Before
## 10  -0.27011420 Data Science Practices Year 2 Before
## 11   0.33123446 Data Science Practices Year 2 Before
## 12   0.13092222 Data Science Practices Year 2 Before
## 13  -1.56252692 Data Science Practices Year 2 Before
```

```
ds_pract_pre3 <- filter(confpre, Cohort == "Year 3")
ds_pract_pre3
```

```
##         Ability             Dimension Cohort  Trial
## 1    1.1542831 Data Science Practices Year 3 Before
## 2   -0.6783320 Data Science Practices Year 3 Before
## 3    1.1542831 Data Science Practices Year 3 Before
## 4   -2.0836556 Data Science Practices Year 3 Before
## 5    0.3312345 Data Science Practices Year 3 Before
## 6    0.1309222 Data Science Practices Year 3 Before
## 7    1.1542831 Data Science Practices Year 3 Before
## 8    0.3312345 Data Science Practices Year 3 Before
## 9   -0.2701142 Data Science Practices Year 3 Before
## 10   0.5326596 Data Science Practices Year 3 Before
## 11  -0.2701142 Data Science Practices Year 3 Before
## 12   1.1542831 Data Science Practices Year 3 Before
## 13   2.6630033 Data Science Practices Year 3 Before
## 14  -0.8881327 Data Science Practices Year 3 Before
## 15   2.6630033 Data Science Practices Year 3 Before
## 16  -3.7705817 Data Science Practices Year 3 Before
## 17  -1.3276827 Data Science Practices Year 3 Before
## 18  -1.1038583 Data Science Practices Year 3 Before
## 19  -1.5625269 Data Science Practices Year 3 Before
```

```r
# science identity post by year
ds_pract_post1 <- filter(confpost, Cohort == "Year 1")
ds_pract_post1
```

```
##          Ability             Dimension Cohort Trial
## 1     3.44744183 Data Science Practices Year 1 After
## 2     0.94296217 Data Science Practices Year 1 After
## 3     2.66300328 Data Science Practices Year 1 After
## 4     1.15428309 Data Science Practices Year 1 After
## 5     3.01420897 Data Science Practices Year 1 After
## 6     2.08638951 Data Science Practices Year 1 After
## 7     0.33123446 Data Science Practices Year 1 After
## 8     3.44744183 Data Science Practices Year 1 After
## 9     0.53265958 Data Science Practices Year 1 After
## 10    1.15428309 Data Science Practices Year 1 After
## 11    1.59756681 Data Science Practices Year 1 After
## 12    2.66300328 Data Science Practices Year 1 After
## 13   -0.06921477 Data Science Practices Year 1 After
## 14    2.08638951 Data Science Practices Year 1 After
## 15    2.66300328 Data Science Practices Year 1 After
## 16    3.44744183 Data Science Practices Year 1 After
```

```r
ds_pract_post2 <- filter(confpost, Cohort == "Year 2")
ds_pract_post2
```

```
##        Ability            Dimension Cohort Trial
## 1  1.1542831 Data Science Practices Year 2 After
## 2  2.6630033 Data Science Practices Year 2 After
## 3  3.4474418 Data Science Practices Year 2 After
## 4  3.4474418 Data Science Practices Year 2 After
## 5  2.6630033 Data Science Practices Year 2 After
## 6  1.1542831 Data Science Practices Year 2 After
## 7  2.6630033 Data Science Practices Year 2 After
## 8  2.3593798 Data Science Practices Year 2 After
## 9  1.5975668 Data Science Practices Year 2 After
## 10 1.1542831 Data Science Practices Year 2 After
## 11 2.6630033 Data Science Practices Year 2 After
## 12 5.2408556 Data Science Practices Year 2 After
## 13 0.9429622 Data Science Practices Year 2 After
```

```r
ds_pract_post3 <- filter(confpost, Cohort == "Year 3")
ds_pract_post3
```

```
##        Ability            Dimension Cohort Trial
## 1   3.4474418 Data Science Practices Year 3 After
## 2   1.1542831 Data Science Practices Year 3 After
## 3   5.2408556 Data Science Practices Year 3 After
## 4   1.1542831 Data Science Practices Year 3 After
## 5   1.3717866 Data Science Practices Year 3 After
## 6   3.0142090 Data Science Practices Year 3 After
## 7   1.1542831 Data Science Practices Year 3 After
## 8   2.0863895 Data Science Practices Year 3 After
## 9   0.5326596 Data Science Practices Year 3 After
## 10  3.0142090 Data Science Practices Year 3 After
## 11  0.7361951 Data Science Practices Year 3 After
## 12  2.6630033 Data Science Practices Year 3 After
## 13 -0.6783320 Data Science Practices Year 3 After
## 14  1.1542831 Data Science Practices Year 3 After
## 15  3.4474418 Data Science Practices Year 3 After
## 16  3.0142090 Data Science Practices Year 3 After
## 17  1.1542831 Data Science Practices Year 3 After
## 18  0.9429622 Data Science Practices Year 3 After
## 19  2.6630033 Data Science Practices Year 3 After
```

```r
# t-testing data science practices by year
# Year 1
t.test(ds_pract_post1$Ability, ds_pract_pre1$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  ds_pract_post1$Ability and ds_pract_pre1$Ability
## t = 3.4162, df = 29.98, p-value = 0.001844
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   0.5757861 2.2876111
## sample estimates:
## mean of x mean of y
## 1.9476311 0.5159325
```

```
cohensD(ds_pract_post1$Ability, ds_pract_pre1$Ability)
```

```
## [1] 1.207822
```

```
# Year 2
t.test(ds_pract_post2$Ability, ds_pract_pre2$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  ds_pract_post2$Ability and ds_pract_pre2$Ability
## t = 5.8855, df = 23.661, p-value = 4.777e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   1.954043 4.067099
## sample estimates:
##   mean of x  mean of y
##   2.3961931 -0.6143776
```

```
cohensD(ds_pract_post2$Ability, ds_pract_pre2$Ability)
```

```
## [1] 2.308496
```

```
# Year 3
t.test(ds_pract_post3$Ability, ds_pract_pre3$Ability)
```

```
##
##  Welch Two Sample t-test
##
## data:  ds_pract_post3$Ability and ds_pract_pre3$Ability
## t = 4.1292, df = 35.364, p-value = 0.0002114
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.015828 2.979252
## sample estimates:
##   mean of x   mean of y
##  1.96144469 -0.03609519
```

```
cohensD(ds_pract_post3$Ability, ds_pract_pre3$Ability)
```

```
## [1] 1.339705
```

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
```

Science Identity by Year

```
# science identity pre by year
idpre1 <- filter(idpre, Cohort == "Year 1")
idpre1
```

```
##         Ability          Dimension Cohort   Trial
## 1   1.8153924 Science Identity Year 1 Before
## 2  -1.2941506 Science Identity Year 1 Before
## 3   3.9602714 Science Identity Year 1 Before
## 4   0.4735946 Science Identity Year 1 Before
## 5   1.4568030 Science Identity Year 1 Before
## 6   1.1273724 Science Identity Year 1 Before
## 7   1.1273724 Science Identity Year 1 Before
## 8  -0.7836363 Science Identity Year 1 Before
## 9  -1.2941506 Science Identity Year 1 Before
## 10  2.2396942 Science Identity Year 1 Before
## 11  0.8055143 Science Identity Year 1 Before
## 12  1.8153924 Science Identity Year 1 Before
## 13  0.4735946 Science Identity Year 1 Before
## 14  0.4735946 Science Identity Year 1 Before
## 15 -0.3015740 Science Identity Year 1 Before
## 16  0.4735946 Science Identity Year 1 Before
```

```
idpre2 <- filter(idpre, Cohort == "Year 2")
idpre2
```

```
##         Ability          Dimension Cohort   Trial
## 1  -0.7836363 Science Identity Year 2 Before
## 2   2.2396942 Science Identity Year 2 Before
## 3  -2.5847635 Science Identity Year 2 Before
## 4  -4.2749159 Science Identity Year 2 Before
## 5  -2.9679977 Science Identity Year 2 Before
## 6  -2.9679977 Science Identity Year 2 Before
## 7  -0.7836363 Science Identity Year 2 Before
## 8  -0.3015740 Science Identity Year 2 Before
## 9  -3.7820238 Science Identity Year 2 Before
## 10  3.9602714 Science Identity Year 2 Before
## 11 -0.7836363 Science Identity Year 2 Before
## 12  3.9602714 Science Identity Year 2 Before
```

```
idpre3 <- filter(idpre, Cohort == "Year 3")
idpre3
```

```
##        Ability          Dimension Cohort   Trial
## 1  -0.7836363 Science Identity Year 3 Before
## 2  -0.7836363 Science Identity Year 3 Before
## 3  -0.7836363 Science Identity Year 3 Before
## 4   1.1273724 Science Identity Year 3 Before
## 5   0.8055143 Science Identity Year 3 Before
## 6  -1.7665849 Science Identity Year 3 Before
## 7  -3.7820238 Science Identity Year 3 Before
## 8  -3.3593292 Science Identity Year 3 Before
## 9  -2.1907468 Science Identity Year 3 Before
## 10  3.9602714 Science Identity Year 3 Before
## 11 -6.2108718 Science Identity Year 3 Before
## 12  2.8183989 Science Identity Year 3 Before
## 13  2.2396942 Science Identity Year 3 Before
## 14  0.1124107 Science Identity Year 3 Before
## 15 -3.7820238 Science Identity Year 3 Before
## 16  2.8183989 Science Identity Year 3 Before
## 17  3.9602714 Science Identity Year 3 Before
## 18 -0.7836363 Science Identity Year 3 Before
```

```
# science identity post by year
idpost1 <- filter(idpost, Cohort == "Year 1")
idpost1
```

```
##        Ability          Dimension Cohort Trial
## 1   2.2396950 Science Identity Year 1 After
## 2  -0.3015739 Science Identity Year 1 After
## 3   3.9602714 Science Identity Year 1 After
## 4   0.4735945 Science Identity Year 1 After
## 5   2.8183984 Science Identity Year 1 After
## 6   1.8153930 Science Identity Year 1 After
## 7   2.8183984 Science Identity Year 1 After
## 8   2.8183984 Science Identity Year 1 After
## 9  -0.7836289 Science Identity Year 1 After
## 10  2.2396950 Science Identity Year 1 After
## 11  0.8055138 Science Identity Year 1 After
## 12  2.2396950 Science Identity Year 1 After
## 13  0.1124107 Science Identity Year 1 After
## 14  0.4735945 Science Identity Year 1 After
## 15  1.1273719 Science Identity Year 1 After
## 16  1.8153930 Science Identity Year 1 After
```

```
idpost2 <- filter(idpost, Cohort == "Year 2")
idpost2
```

```
##        Ability          Dimension Cohort Trial
## 1    1.4568030 Science Identity Year 2 After
## 2    3.9602714 Science Identity Year 2 After
## 3    2.2396950 Science Identity Year 2 After
## 4   -1.2941469 Science Identity Year 2 After
## 5   -1.7665849 Science Identity Year 2 After
## 6   -0.7836289 Science Identity Year 2 After
## 7    3.9602714 Science Identity Year 2 After
## 8    1.8153930 Science Identity Year 2 After
## 9    2.2396950 Science Identity Year 2 After
## 10   3.9602714 Science Identity Year 2 After
## 11   0.8055138 Science Identity Year 2 After
## 12   3.9602714 Science Identity Year 2 After
```

```
idpost3 <- filter(idpost, Cohort == "Year 3")
idpost3
```

```
##        Ability          Dimension Cohort Trial
## 1    0.1124107 Science Identity Year 3 After
## 2   -0.7836289 Science Identity Year 3 After
## 3   -0.3015739 Science Identity Year 3 After
## 4    3.9602714 Science Identity Year 3 After
## 5    2.2396950 Science Identity Year 3 After
## 6   -1.7665849 Science Identity Year 3 After
## 7   -3.3593342 Science Identity Year 3 After
## 8    1.1273719 Science Identity Year 3 After
## 9   -0.3015739 Science Identity Year 3 After
## 10   3.9602714 Science Identity Year 3 After
## 11   1.4568030 Science Identity Year 3 After
## 12   2.8183984 Science Identity Year 3 After
## 13   2.2396950 Science Identity Year 3 After
## 14   0.4735945 Science Identity Year 3 After
## 15   0.8055138 Science Identity Year 3 After
## 16   2.8183984 Science Identity Year 3 After
## 17   3.9602714 Science Identity Year 3 After
## 18   1.4568030 Science Identity Year 3 After
```

```
# t-testing science identity by year
# Year 1
t.test(idpost1$Ability, idpre1$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  idpost1$Ability and idpre1$Ability
## t = 1.5967, df = 29.977, p-value = 0.1208
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.2111635  1.7241560
## sample estimates:
## mean of x mean of y
## 1.5420388 0.7855425
```

```
cohensD(idpost1$Ability, idpre1$Ability)
```

```
## [1] 0.5645036
```

```
# Year 2
t.test(idpost2$Ability, idpre2$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  idpost2$Ability and idpre2$Ability
## t = 2.4198, df = 20.336, p-value = 0.025
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.3428631 4.5944318
## sample estimates:
##  mean of x  mean of y
##  1.7128187 -0.7558287
```

```
cohensD(idpost2$Ability, idpre2$Ability)
```

```
## [1] 0.9878965
```

```
# Year 3
t.test(idpost3$Ability, idpre3$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  idpost3$Ability and idpre3$Ability
## t = 1.8263, df = 30.586, p-value = 0.07758
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1780179  3.2114174
## sample estimates:
##  mean of x  mean of y
##  1.1620446 -0.3546552
```

```
cohensD(idpost3$Ability, idpre3$Ability)
```

```
## [1] 0.60876
```

```
# Year 3 vs. 1
t.test(idpost3$Ability, idpost1$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  idpost3$Ability and idpost1$Ability
## t = -0.65286, df = 29.458, p-value = 0.5189
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.5696000  0.8096115
## sample estimates:
## mean of x mean of y
##   1.162045  1.542039
```

```
cohensD(idpost3$Ability, idpre1$Ability)
```

```
## [1] 0.2151985
```

```
# Looking at differences in growth across the three cohorts
sci_id_growth <- read.csv("sci-id-growth.csv")
sci_id_growth
```

```
##     Cohort          Diff
## 1   Year 1   0.424302606
## 2   Year 1   0.992576732
## 3   Year 1  -0.000000036
## 4   Year 1  -0.000000144
## 5   Year 1   1.361595374
## 6   Year 1   0.688020626
## 7   Year 1   1.691026016
## 8   Year 1   3.602034707
## 9   Year 1   0.510521677
## 10  Year 1   0.000000788
## 11  Year 1  -0.000000502
## 12  Year 1   0.424302606
## 13  Year 1  -0.361183961
## 14  Year 1  -0.000000144
## 15  Year 1   1.428945868
## 16  Year 1   1.341798371
## 17  Year 2   2.240439321
## 18  Year 2   1.720577168
## 19  Year 2   4.824458528
## 20  Year 2   2.980769085
## 21  Year 2   1.201412741
## 22  Year 2   2.184368750
## 23  Year 2   4.743907684
## 24  Year 2   2.116966986
## 25  Year 2   6.021718816
## 26  Year 2  -0.000000036
## 27  Year 2   1.589150093
## 28  Year 2  -0.000000036
## 29  Year 3   0.896046985
## 30  Year 3   0.000007380
## 31  Year 3   0.482062438
## 32  Year 3   2.832898993
## 33  Year 3   1.434180709
## 34  Year 3   0.000000001
## 35  Year 3   0.422689603
## 36  Year 3   4.486701108
## 37  Year 3   1.889172951
## 38  Year 3  -0.000000036
## 39  Year 3   7.667674865
## 40  Year 3  -0.000000490
## 41  Year 3   0.000000788
## 42  Year 3   0.361183817
## 43  Year 3   4.587537605
## 44  Year 3  -0.000000490
## 45  Year 3  -0.000000036
## 46  Year 3   2.240439321
```

```
sci_id_growth1 <- filter(sci_id_growth, Cohort == "Year 1")
sci_id_growth1
```

```
##     Cohort        Diff
## 1  Year 1  0.424302606
## 2  Year 1  0.992576732
## 3  Year 1 -0.000000036
## 4  Year 1 -0.000000144
## 5  Year 1  1.361595374
## 6  Year 1  0.688020626
## 7  Year 1  1.691026016
## 8  Year 1  3.602034707
## 9  Year 1  0.510521677
## 10 Year 1  0.000000788
## 11 Year 1 -0.000000502
## 12 Year 1  0.424302606
## 13 Year 1 -0.361183961
## 14 Year 1 -0.000000144
## 15 Year 1  1.428945868
## 16 Year 1  1.341798371
```

```
sci_id_growth2 <- filter(sci_id_growth, Cohort == "Year 2")
sci_id_growth2
```

```
##     Cohort        Diff
## 1  Year 2  2.240439321
## 2  Year 2  1.720577168
## 3  Year 2  4.824458528
## 4  Year 2  2.980769085
## 5  Year 2  1.201412741
## 6  Year 2  2.184368750
## 7  Year 2  4.743907684
## 8  Year 2  2.116966986
## 9  Year 2  6.021718816
## 10 Year 2 -0.000000036
## 11 Year 2  1.589150093
## 12 Year 2 -0.000000036
```

```
sci_id_growth3 <- filter(sci_id_growth, Cohort == "Year 3")
sci_id_growth3
```

```
##      Cohort        Diff
## 1   Year 3   0.896046985
## 2   Year 3   0.000007380
## 3   Year 3   0.482062438
## 4   Year 3   2.832898993
## 5   Year 3   1.434180709
## 6   Year 3   0.000000001
## 7   Year 3   0.422689603
## 8   Year 3   4.486701108
## 9   Year 3   1.889172951
## 10  Year 3  -0.000000036
## 11  Year 3   7.667674865
## 12  Year 3  -0.000000490
## 13  Year 3   0.000000788
## 14  Year 3   0.361183817
## 15  Year 3   4.587537605
## 16  Year 3  -0.000000490
## 17  Year 3  -0.000000036
## 18  Year 3   2.240439321
```

```
# t-testing for differences in growth
# Year 2 vs. Year 1
t.test(sci_id_growth2$Diff, sci_id_growth1$Diff)
```

```
##
##  Welch Two Sample t-test
##
## data:  sci_id_growth2$Diff and sci_id_growth1$Diff
## t = 2.8688, df = 15.509, p-value = 0.01141
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.4436788 2.9806235
## sample estimates:
## mean of x mean of y
## 2.4686474 0.7564963
```

```
# Year 3 vs. Year 1
t.test(sci_id_growth3$Diff, sci_id_growth1$Diff)
```

```
##
##   Welch Two Sample t-test
##
## data:  sci_id_growth3$Diff and sci_id_growth1$Diff
## t = 1.349, df = 24.446, p-value = 0.1897
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.401722  1.922129
## sample estimates:
## mean of x mean of y
## 1.5166998 0.7564963
```

```
cohensD(sci_id_growth3$Diff, sci_id_growth1$Diff)
```

```
## [1] 0.4455407
```

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
```

Data Science Skills by Year

```
# science identity pre by year
ds_skills_pre1 <- filter(techconfpre, Cohort == "Year 1")
ds_skills_pre1
```

```
##          Ability           Dimension Cohort   Trial
## 1   1.78498502 Data Science Skills Year 1 Before
## 2  -2.89555118 Data Science Skills Year 1 Before
## 3   0.76338580 Data Science Skills Year 1 Before
## 4   0.19183423 Data Science Skills Year 1 Before
## 5   0.39200645 Data Science Skills Year 1 Before
## 6  -0.51889635 Data Science Skills Year 1 Before
## 7   0.58165000 Data Science Skills Year 1 Before
## 8   3.40743841 Data Science Skills Year 1 Before
## 9   0.39200645 Data Science Skills Year 1 Before
## 10 -0.02249625 Data Science Skills Year 1 Before
## 11 -0.25643576 Data Science Skills Year 1 Before
## 12  0.39200645 Data Science Skills Year 1 Before
## 13 -0.51889635 Data Science Skills Year 1 Before
## 14  0.39200645 Data Science Skills Year 1 Before
## 15 -0.02249625 Data Science Skills Year 1 Before
## 16 -0.82595605 Data Science Skills Year 1 Before
```

```
ds_skills_pre2 <- filter(techconfpre, Cohort == "Year 2")
ds_skills_pre2
```

```
##         Ability           Dimension Cohort   Trial
## 1   0.5816500 Data Science Skills Year 2 Before
## 2   1.7849850 Data Science Skills Year 2 Before
## 3  -2.8955512 Data Science Skills Year 2 Before
## 4  -0.2564358 Data Science Skills Year 2 Before
## 5   1.7849850 Data Science Skills Year 2 Before
## 6   1.6156635 Data Science Skills Year 2 Before
## 7   0.1918342 Data Science Skills Year 2 Before
## 8   0.3920065 Data Science Skills Year 2 Before
## 9   1.1109643 Data Science Skills Year 2 Before
## 10 -0.8259560 Data Science Skills Year 2 Before
## 11 -1.7612494 Data Science Skills Year 2 Before
## 12 -0.8259560 Data Science Skills Year 2 Before
```

```
ds_skills_pre3 <- filter(techconfpre, Cohort == "Year 3")
ds_skills_pre3
```

```
##        Ability           Dimension Cohort   Trial
## 1    0.7633858 Data Science Skills Year 3 Before
## 2   -2.8955512 Data Science Skills Year 3 Before
## 3   -0.5188964 Data Science Skills Year 3 Before
## 4    2.5077364 Data Science Skills Year 3 Before
## 5   -1.7612494 Data Science Skills Year 3 Before
## 6    0.7633858 Data Science Skills Year 3 Before
## 7   -0.5188964 Data Science Skills Year 3 Before
## 8   -0.8259560 Data Science Skills Year 3 Before
## 9    0.1918342 Data Science Skills Year 3 Before
## 10   1.6156635 Data Science Skills Year 3 Before
## 11  -2.8955512 Data Science Skills Year 3 Before
## 12  -0.2564358 Data Science Skills Year 3 Before
## 13  -2.8955512 Data Science Skills Year 3 Before
## 14   1.9571955 Data Science Skills Year 3 Before
## 15   0.3920065 Data Science Skills Year 3 Before
## 16  -0.5188964 Data Science Skills Year 3 Before
## 17   0.3920065 Data Science Skills Year 3 Before
## 18   0.3920065 Data Science Skills Year 3 Before
```

```
# science identity post by year
ds_skills_post1 <- filter(techconfpost, Cohort == "Year 1")
ds_skills_post1
```

```
##       Ability           Dimension Cohort Trial
## 1  3.4074384 Data Science Skills Year 1 After
## 2  4.8202632 Data Science Skills Year 1 After
## 3  2.7092572 Data Science Skills Year 1 After
## 4  1.9571955 Data Science Skills Year 1 After
## 5  0.3920064 Data Science Skills Year 1 After
## 6  3.1553096 Data Science Skills Year 1 After
## 7  2.1338550 Data Science Skills Year 1 After
## 8  4.0010730 Data Science Skills Year 1 After
## 9  2.5077364 Data Science Skills Year 1 After
## 10 4.3679136 Data Science Skills Year 1 After
## 11 2.9240041 Data Science Skills Year 1 After
## 12 2.7092572 Data Science Skills Year 1 After
## 13 0.7633858 Data Science Skills Year 1 After
## 14 2.5077364 Data Science Skills Year 1 After
## 15 2.9240041 Data Science Skills Year 1 After
## 16 1.7849850 Data Science Skills Year 1 After
```

```
ds_skills_post2 <- filter(techconfpost, Cohort == "Year 2")
ds_skills_post2
```

```
##       Ability           Dimension Cohort Trial
## 1  4.367914 Data Science Skills Year 2 After
## 2  4.001073 Data Science Skills Year 2 After
## 3  1.784985 Data Science Skills Year 2 After
## 4  4.820263 Data Science Skills Year 2 After
## 5  2.709257 Data Science Skills Year 2 After
## 6  3.686306 Data Science Skills Year 2 After
## 7  5.446170 Data Science Skills Year 2 After
## 8  2.133855 Data Science Skills Year 2 After
## 9  2.316706 Data Science Skills Year 2 After
## 10 1.110964 Data Science Skills Year 2 After
## 11 3.407438 Data Science Skills Year 2 After
## 12 3.686306 Data Science Skills Year 2 After
```

```
ds_skills_post3 <- filter(techconfpost, Cohort == "Year 3")
ds_skills_post3
```

```
##       Ability           Dimension Cohort Trial
## 1  1.957195 Data Science Skills Year 3 After
## 2  1.615663 Data Science Skills Year 3 After
## 3  2.924004 Data Science Skills Year 3 After
## 4  6.657984 Data Science Skills Year 3 After
## 5  1.615663 Data Science Skills Year 3 After
## 6  1.615663 Data Science Skills Year 3 After
## 7  1.110964 Data Science Skills Year 3 After
## 8  2.507736 Data Science Skills Year 3 After
## 9  3.407438 Data Science Skills Year 3 After
## 10 4.001073 Data Science Skills Year 3 After
## 11 4.367914 Data Science Skills Year 3 After
## 12 2.709257 Data Science Skills Year 3 After
## 13 2.507736 Data Science Skills Year 3 After
## 14 2.924004 Data Science Skills Year 3 After
## 15 4.367914 Data Science Skills Year 3 After
## 16 4.820263 Data Science Skills Year 3 After
## 17 2.133855 Data Science Skills Year 3 After
## 18 4.367914 Data Science Skills Year 3 After
```

```
# t-testing science identity by year
# Year 1
t.test(ds_skills_post1$Ability, ds_skills_pre1$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  ds_skills_post1$Ability and ds_skills_pre1$Ability
## t = 5.6845, df = 29.684, p-value = 3.503e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   1.594562 3.384041
## sample estimates:
## mean of x mean of y
## 2.6915888 0.2022869
```

```
cohensD(ds_skills_post1$Ability, ds_skills_pre1$Ability)
```

```
## [1] 2.009757
```

```
# Year 2
t.test(ds_skills_post2$Ability, ds_skills_pre2$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  ds_skills_post2$Ability and ds_skills_pre2$Ability
## t = 5.6842, df = 21.699, p-value = 1.077e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   2.040757 4.388293
## sample estimates:
## mean of x mean of y
##   3.289270   0.074745
```

```
cohensD(ds_skills_post2$Ability, ds_skills_pre2$Ability)
```

```
## [1] 2.320549
```

```
# Year 3
t.test(ds_skills_post3$Ability, ds_skills_pre3$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  ds_skills_post3$Ability and ds_skills_pre3$Ability
## t = 6.558, df = 33.632, p-value = 1.725e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.28937 4.34663
## sample estimates:
##  mean of x  mean of y
##  3.0895691 -0.2284313
```

```
cohensD(ds_skills_post3$Ability, ds_skills_pre3$Ability)
```

```
## [1] 2.185985
```

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
```

Python Skills by Year

```
# science identity pre by year
python_pre1 <- filter(techcomfortpre, Cohort == "Year 1")
python_pre1
```

```
##          Ability        Dimension Cohort   Trial
## 1    1.6668819 Python Skills Year 1 Before
## 2    2.8461207 Python Skills Year 1 Before
## 3   -0.1996241 Python Skills Year 1 Before
## 4   -0.1996241 Python Skills Year 1 Before
## 5    1.0266831 Python Skills Year 1 Before
## 6   -0.1996241 Python Skills Year 1 Before
## 7    5.9798356 Python Skills Year 1 Before
## 8    1.6668819 Python Skills Year 1 Before
## 9   -0.1996241 Python Skills Year 1 Before
## 10  -0.1996241 Python Skills Year 1 Before
## 11  -0.1996241 Python Skills Year 1 Before
## 12  -0.1996241 Python Skills Year 1 Before
## 13  -0.1996241 Python Skills Year 1 Before
## 14   1.0266831 Python Skills Year 1 Before
## 15  -0.1996241 Python Skills Year 1 Before
## 16  -0.1996241 Python Skills Year 1 Before
```

```
python_pre2 <- filter(techcomfortpre, Cohort == "Year 2")
python_pre2
```

```
##          Ability        Dimension Cohort  Trial
## 1    3.1439720 Python Skills Year 2 Before
## 2    4.1577919 Python Skills Year 2 Before
## 3   -0.1996241 Python Skills Year 2 Before
## 4    1.0266831 Python Skills Year 2 Before
## 5   -0.1996241 Python Skills Year 2 Before
## 6   -0.1996241 Python Skills Year 2 Before
## 7   -0.1996241 Python Skills Year 2 Before
## 8   -0.1996241 Python Skills Year 2 Before
## 9    4.6112586 Python Skills Year 2 Before
## 10   2.5150264 Python Skills Year 2 Before
## 11  -0.1996241 Python Skills Year 2 Before
## 12  -0.1996241 Python Skills Year 2 Before
```

```
python_pre3 <- filter(techcomfortpre, Cohort == "Year 3")
python_pre3
```

```
##       Ability       Dimension Cohort   Trial
## 1   3.6764147 Python Skills Year 3 Before
## 2  -0.1996241 Python Skills Year 3 Before
## 3  -0.1996241 Python Skills Year 3 Before
## 4   6.5363310 Python Skills Year 3 Before
## 5  -0.1996241 Python Skills Year 3 Before
## 6   4.1577919 Python Skills Year 3 Before
## 7  -0.1996241 Python Skills Year 3 Before
## 8  -0.1996241 Python Skills Year 3 Before
## 9  -0.1996241 Python Skills Year 3 Before
## 10 -0.1996241 Python Skills Year 3 Before
## 11 -0.1996241 Python Skills Year 3 Before
## 12 -0.1996241 Python Skills Year 3 Before
## 13 -0.1996241 Python Skills Year 3 Before
## 14  3.4186043 Python Skills Year 3 Before
## 15 -0.1996241 Python Skills Year 3 Before
## 16 -0.1996241 Python Skills Year 3 Before
## 17 -0.1996241 Python Skills Year 3 Before
## 18  1.6668819 Python Skills Year 3 Before
```

```
# science identity post by year
python_post1 <- filter(techcomfortpost, Cohort == "Year 1")
python_post1
```

```
##      Ability       Dimension Cohort Trial
## 1  7.217052 Python Skills Year 1 After
## 2  7.812953 Python Skills Year 1 After
## 3  5.950066 Python Skills Year 1 After
## 4  6.141463 Python Skills Year 1 After
## 5  6.748451 Python Skills Year 1 After
## 6  5.950066 Python Skills Year 1 After
## 7  7.217052 Python Skills Year 1 After
## 8  4.580920 Python Skills Year 1 After
## 9  6.748451 Python Skills Year 1 After
## 10 6.141463 Python Skills Year 1 After
## 11 5.760429 Python Skills Year 1 After
## 12 4.366900 Python Skills Year 1 After
## 13 5.380136 Python Skills Year 1 After
## 14 5.760429 Python Skills Year 1 After
## 15 5.186832 Python Skills Year 1 After
## 16 5.760429 Python Skills Year 1 After
```

```
python_post2 <- filter(techcomfortpost, Cohort == "Year 2")
python_post2
```

```
##      Ability      Dimension Cohort Trial
## 1  7.491356 Python Skills Year 2 After
## 2  9.952048 Python Skills Year 2 After
## 3  4.788244 Python Skills Year 2 After
## 4  7.491356 Python Skills Year 2 After
## 5  6.748451 Python Skills Year 2 After
## 6  6.336577 Python Skills Year 2 After
## 7  9.952048 Python Skills Year 2 After
## 8  5.186832 Python Skills Year 2 After
## 9  6.972738 Python Skills Year 2 After
## 10 4.788244 Python Skills Year 2 After
## 11 5.570921 Python Skills Year 2 After
## 12 7.491356 Python Skills Year 2 After
```

```
python_post3 <- filter(techcomfortpost, Cohort == "Year 3")
python_post3
```

```
##      Ability      Dimension Cohort Trial
## 1  4.989882 Python Skills Year 3 After
## 2  5.186832 Python Skills Year 3 After
## 3  6.141463 Python Skills Year 3 After
## 4  9.952048 Python Skills Year 3 After
## 5  4.989882 Python Skills Year 3 After
## 6  5.380136 Python Skills Year 3 After
## 7  3.914069 Python Skills Year 3 After
## 8  6.141463 Python Skills Year 3 After
## 9  7.812953 Python Skills Year 3 After
## 10 9.952048 Python Skills Year 3 After
## 11 9.952048 Python Skills Year 3 After
## 12 7.812953 Python Skills Year 3 After
## 13 7.217052 Python Skills Year 3 After
## 14 7.491356 Python Skills Year 3 After
## 15 6.748451 Python Skills Year 3 After
## 16 9.952048 Python Skills Year 3 After
## 17 4.145073 Python Skills Year 3 After
## 18 8.216612 Python Skills Year 3 After
```

```
# t-testing science identity by year
# Year 1
t.test(python_post1$Ability, python_pre1$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  python_post1$Ability and python_pre1$Ability
## t = 10.946, df = 23.525, p-value = 1.036e-10
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   4.284734 6.278547
## sample estimates:
## mean of x mean of y
## 6.0451932 0.7635529
```

```
cohensD(python_post1$Ability, python_pre1$Ability)
```

```
## [1] 3.870099
```

```
# Year 2
t.test(python_post2$Ability, python_pre2$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  python_post2$Ability and python_pre2$Ability
## t = 7.6856, df = 21.854, p-value = 1.193e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   4.180347 7.271787
## sample estimates:
## mean of x mean of y
##   6.897514   1.171447
```

```
cohensD(python_post2$Ability, python_pre2$Ability)
```

```
## [1] 3.137618
```

```
# Year 3
t.test(python_post3$Ability, python_pre3$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  python_post3$Ability and python_pre3$Ability
## t = 8.8424, df = 33.996, p-value = 2.471e-10
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   4.669601 7.456560
## sample estimates:
## mean of x mean of y
## 6.9997981 0.9367173
```

```
cohensD(python_post3$Ability, python_pre3$Ability)
```

```
## [1] 2.94747
```

```
# Year 3 vs. 1 (PRE)
t.test(python_pre3$Ability, python_pre1$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  python_pre3$Ability and python_pre1$Ability
## t = 0.26877, df = 31.783, p-value = 0.7898
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   -1.139565  1.485894
## sample estimates:
## mean of x mean of y
## 0.9367173 0.7635529
```
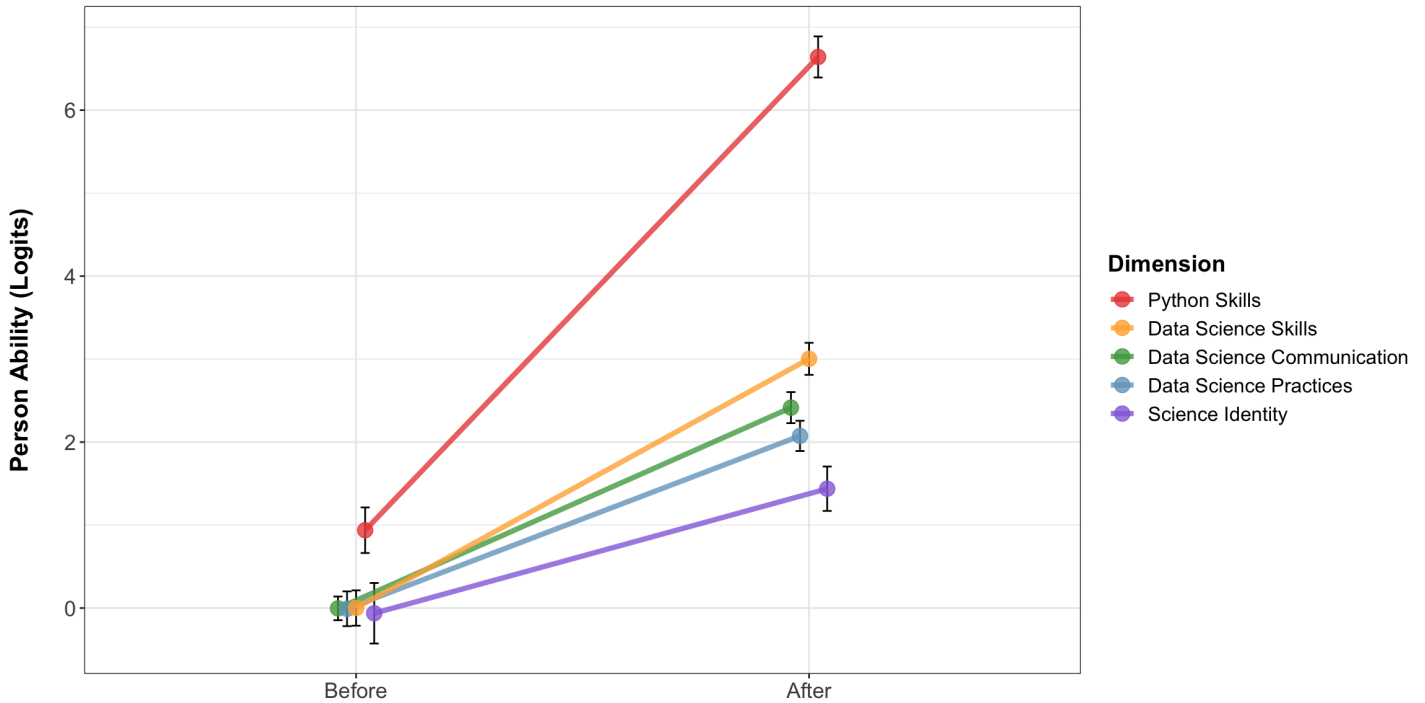
```
t.test(python_pre2$Ability, python_pre1$Ability)
```

```
##
##   Welch Two Sample t-test
##
## data:  python_pre2$Ability and python_pre1$Ability
## t = 0.59017, df = 22.169, p-value = 0.561
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   -1.024813  1.840601
## sample estimates:
## mean of x mean of y
## 1.1714470 0.7635529
```

```
cohensD(idpost3$Ability, idpre1$Ability)
```

```
## [1] 0.2151985
```

Growth across the five dimensions

Splitting by cohort