sandbox-02-missing-data

Joshua Rosenberg

12/28/2017

** Race **

There are not statistically significant differences by race. Other includes white (n = 13), multiracial (n = 5), and Asian (n = 14).

```
## # A tibble: 4 x 2
##       race mean_response_rate
##     <fctr>              <dbl>
## 1    Black          0.6025861
## 2 Hispanic          0.6142559
## 3    Other          0.6918350
## 4     <NA>          0.7727273
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## race          2  0.178 0.08911   1.377  0.255
## Residuals   184 11.910 0.06473
## 2 observations deleted due to missingness
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = response_rate ~ race, data = ddd)
##
## $race
##                       diff         lwr       upr     p adj
## Hispanic-Black 0.01166976 -0.08472291 0.1080624 0.9559064
## Other-Black    0.08924889 -0.04165226 0.2201500 0.2435566
## Other-Hispanic 0.07757913 -0.04988610 0.2050444 0.3235459
```

** URM **

Students who are in a URM group have a response rate of 60.6% compared to a response rate of 74.1% for non-URM students (t = 2.622, p = .013, d = .56).

```
## # A tibble: 2 x 2
##     urm mean_response_rate
##   <dbl>              <dbl>
## 1     0          0.7409596
## 2     1          0.6060130
```

```
## [1] "mean in group 0  is  0.741"
## [1] "mean in group 1  is  0.606"
## [1] "Test statistic is  2.622"
## [1] "P-value is  0.013"
## [1] "Effect size is  0.56"
```

** Gender **

Female's response rate is 66.2% compared to male's, which is 58.4%; this is a significant difference (t = 2.093, p = .038, d = .31)

```
## # A tibble: 3 x 2
```

```
##   gender mean_response_rate
##    <chr>              <dbl>
## 1      F          0.6615822
## 2      M          0.5837968
## 3   <NA>          0.6363636
```

```
## [1] "mean in group F  is  0.662"
## [1] "mean in group M  is  0.584"
## [1] "Test statistic is  2.093"
## [1] "P-value is  0.038"
## [1] "Effect size is  0.31"
```

** Pre-interest **

Overall pre-interest and response rate are not correlated.

```
##
##  Pearson's product-moment correlation
##
## data:  ddd$overall_pre_interest and ddd$response_rate
## t = -0.032217, df = 167, p-value = 0.9743
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.1533953  0.1485229
## sample estimates:
##          cor
## -0.002493016
```