

Descriptive Report

William Murrah

4/9/2021

The purpose of this project is to demonstrate the use of Rmarkdown to generate reports integrating text and R code with output (Xie, Allaire, and Grolemund 2018).

```
source("code/cleanGSS2018.R")
```

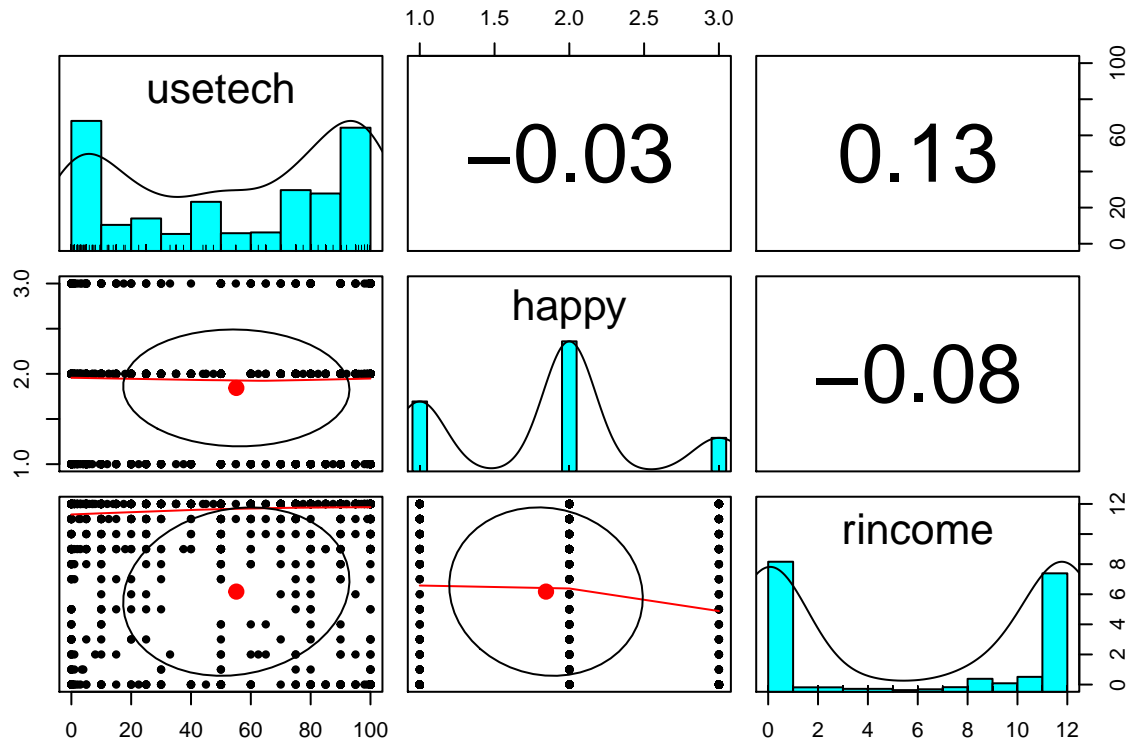
```
summary(gss)
```

```
##      ballot      usetech      happy      partyid
## Min.   :1.000   Min.    : 0.00   Min.    :1.000   independent  :414
## 1st Qu.:1.000   1st Qu.: 15.00   1st Qu.:1.000   strong democrat :379
## Median :2.000   Median : 60.00   Median :2.000   democrat       :351
## Mean   :2.002   Mean    : 55.15   Mean    :1.844   left independent:307
## 3rd Qu.:3.000   3rd Qu.: 90.00   3rd Qu.:2.000   republican      :271
## Max.    :3.000   Max.    :100.00   Max.    :3.000   (Other)         :513
##                      NA's    :936   NA's     :4   NA's          :110
##      rincome      race      sex      degree
## Min.    : 0.000   white:1692   male  :1051   Min.    :0.000
## 1st Qu.: 0.000   black: 383   female:1294   1st Qu.:1.000
## Median : 9.000   other: 270           Median :1.000
## Mean     : 6.176           Mean    :1.684
## 3rd Qu.:12.000           3rd Qu.:3.000
## Max.     :12.000           Max.    :4.000
## NA's     :134
##      educ      age      marital      hrs2
## less than hs : 87   Min.    :18.00   Min.    :1.00   Min.    : 6.00
## some hs      : 211  1st Qu.:34.00   1st Qu.:1.00   1st Qu.:40.00
## hs grad      : 656  Median :48.00   Median :2.00   Median :40.00
## some college :1312  Mean    :49.13   Mean    :2.67   Mean    :45.91
## college degree: 72  3rd Qu.:63.00   3rd Qu.:5.00   3rd Qu.:50.00
## NA's         : 7   Max.    :99.00   Max.    :9.00   Max.    :99.00
##                      NA's    :2291
##      hrs1      wrkstat      id      unhappy
## Min.    : 1.00   Min.    :1.000   Min.    : 1   Min.    :0.000
## 1st Qu.:35.00   1st Qu.:1.000   1st Qu.: 588  1st Qu.:0.000
## Median :40.00   Median :2.000   Median :1176  Median :0.000
## Mean    :41.87   Mean    :2.963   Mean    :1175  Mean    :1.039
## 3rd Qu.:50.00   3rd Qu.:5.000   3rd Qu.:1762  3rd Qu.:2.000
## Max.    :99.00   Max.    :9.000   Max.    :2348  Max.    :9.000
## NA's     :952
```

Explore Distributions

Tukey and others (1977) said exploring data is really important.

```
pairs.panels(gss[, c("usetech", "happy", "rincome")])
```



```
describe(gss[, c("usetech", "happy", "rincome")])
```

```
##      vars   n  mean    sd median trimmed  mad min max range  skew kurtosis
## usetech   1 1409 55.15 37.83     60   56.42 51.89   0 100   100 -0.23   -1.52
## happy     2 2341  1.84  0.65      2    1.81  0.00   1   3     2  0.16   -0.67
## rincome   3 2211  6.18  5.60      9    6.22  4.45   0  12    12 -0.09   -1.89
##
##      se
## usetech 1.01
## happy   0.01
## rincome 0.12
```

```
aggregate(rincome ~ race + sex, data = gss,
          FUN = function(x) c(M = mean(x), SD = sd(x), n = length(x)))
```

```
##    race    sex rincome.M rincome.SD rincome.n
## 1 white  male   6.646648   5.705549  716.000000
## 2 black  male   6.405405   5.615622  148.000000
## 3 other  male   7.991228   5.208053  114.000000
## 4 white female  5.626561   5.551836  881.000000
## 5 black female  6.133028   5.473498  218.000000
## 6 other female  5.552239   5.320647  134.000000
```

```
library(tables)
M <- function(x) mean(x, na.rm = TRUE)
SD <- function(x) sd(x, na.rm = TRUE)
n <- function(x) round(length(x))
tabular((rincome)*(M + SD) + 1 ~ (race * sex) + 1, data = gss)
```

	Model 1	Model 2	Model 3
(Intercept)	6.18*** (0.12)	9.55*** (0.40)	9.89*** (0.62)
happy		-0.51** (0.18)	-0.69* (0.31)
usetech		0.01*** (0.00)	0.01 (0.01)
happy:usetech			0.00 (0.00)
R ²	0.00	0.02	0.02
Adj. R ²	0.00	0.02	0.02
Num. obs.	2211	1314	1314

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 1: Statistical models

		white		race		other		
		sex		black		sex		
		male	female	male	female	male	female	All
rincome	M	6.647	5.627	6.405	6.133	7.991	5.552	6.176
	SD	5.706	5.552	5.616	5.473	5.208	5.321	5.597
	All	768.000	924.000	157.000	226.000	126.000	144.000	2345.000

Models

```
mod0 <- lm(rincome ~ 1, data = gss)
mod1 <- lm(rincome ~ happy + usetech, data = gss)
mod1x <- lm(rincome ~ happy*usetech, data = gss)

texreg(list(mod0, mod1, mod1x))
```

References

- Tukey, John W, and others. 1977. *Exploratory Data Analysis*. Vol. 2. Reading, Mass.
- Xie, Yihui, Joseph J Allaire, and Garrett Grolemond. 2018. *R Markdown: The Definitive Guide*. CRC Press.