

Lab 3: Individual 1

John Chandara <chandara@iastate.edu>

2/19/20

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --
## v ggplot2 3.2.1      v purrr 0.3.3
## v tibble 2.1.3       v dplyr 0.8.3
## v tidyr 1.0.0        v stringr 1.4.0
## v readr 1.3.1        v forcats 0.4.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()

dat <- readxl::read_xls('GSS.xls')

## Warning in read_fun(path = enc2native(normalizePath(path)), sheet_i = sheet, :
## Expecting numeric in A64817 / R64817C1: got 'Source'
```

GSS Data Exploration

Data Cleaning

```
indx <- sapply(dat, is.character)
dat[indx] <- lapply(dat[indx], function(x) as.factor(as.character(x)))
names(dat) <- c('year', 'happiness', 'party', 'residence', 'sex', 'education', 'children', 'marriage',
str(dat)

## Classes 'tbl_df', 'tbl' and 'data.frame': 64816 obs. of 10 variables:
## $ year : num 1972 1972 1972 1972 1972 ...
## $ happiness: Factor w/ 7 levels "Data collection: General Social Survey\nCase Selection: all cases"
## $ party : Factor w/ 10 levels "Don't know","Ind,near dem",...: 2 6 4 6 9 2 2 2 9 9 ...
## $ residence: Factor w/ 10 levels "E. nor. central",...: 4 1 1 3 1 1 1 4 10 10 ...
## $ sex : Factor w/ 2 levels "Female","Male": 1 2 1 1 1 2 2 2 1 1 ...
## $ education: Factor w/ 7 levels "Bachelor","Don't know",...: 1 6 4 1 4 4 4 1 4 4 ...
## $ children : Factor w/ 10 levels "0","1","2","3",...: 1 6 5 1 3 1 3 1 3 5 ...
## $ marriage : Factor w/ 6 levels "Divorced","Married",...: 3 2 2 2 2 3 1 3 3 2 ...
## $ id : num 1 2 3 4 5 6 7 8 9 10 ...
## $ ballot : Factor w/ 5 levels "Ballot a","Ballot b",...: 5 5 5 5 5 5 5 5 5 5 ...
```

Exploration

How does the happiness of a respondent relate to the political party affiliation?

Is political affiliation affected by education?

How does the political affiliation and the residence affect happiness?