

Week_6

Kalyani Cauwenberghs

2/14/2020

Part 1

Part 2

Part 3

Describe the distribution of the outcome variable, identify a main predictor that you're interested in studying its effect on the outcome

```
#read in dataset
data<-read.csv("hfi_cc_2019.csv")
data<-data[data$year=="2017",]
#predictor: political freedom
#outcome: economic freedom
#make sure both columns have no missing data
print("pf_score contains no missing data:")

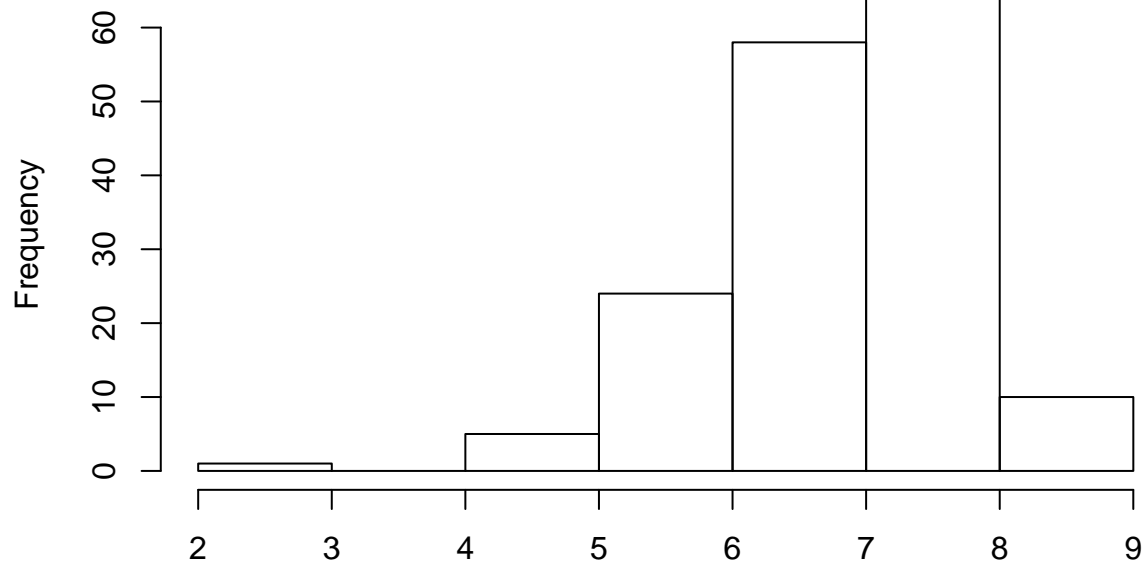
## [1] "pf_score contains no missing data:"
sum(as.character(data$pf_score)=="-")==0

## [1] TRUE
print("ef_score contains no missing data:")

## [1] "ef_score contains no missing data:"
sum(as.character(data$ef_score)=="-")==0

## [1] TRUE
#distribution of outcome aka EF
hist(as.numeric(as.character(data$ef_score)))
```

Histogram of `as.numeric(as.character(data$ef_score))`

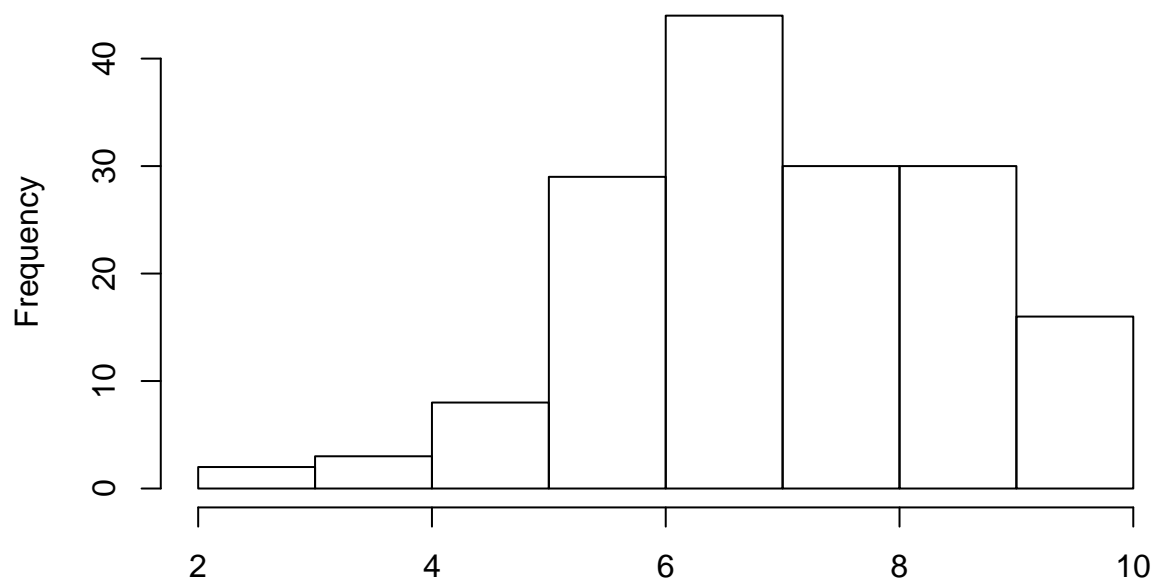


`as.numeric(as.character(data$ef_score))`

Our outcome (EF score) is normal but left skewed, mean around 6.5, median around 7. Our main predictor is PF score.

```
hist(as.numeric(as.character(data$pf_score)))
```

Histogram of `as.numeric(as.character(data$pf_score))`



`as.numeric(as.character(data$pf_score))`

Identify other variables (i.e. predictors, often called covariates) that might be related to the outcome or the

main predictor, discuss these variables in the context of part 2 above of this assignment

political freedom variables can be our predictors.

pf_rol

pf_ss

pf_movement

pf_religion

pf_association

pf_expression

pf_identity

Carry out univariate logistic regression of the outcome on each of the predictors including the main predictor, interpret the results in terms of odds ratio etc.

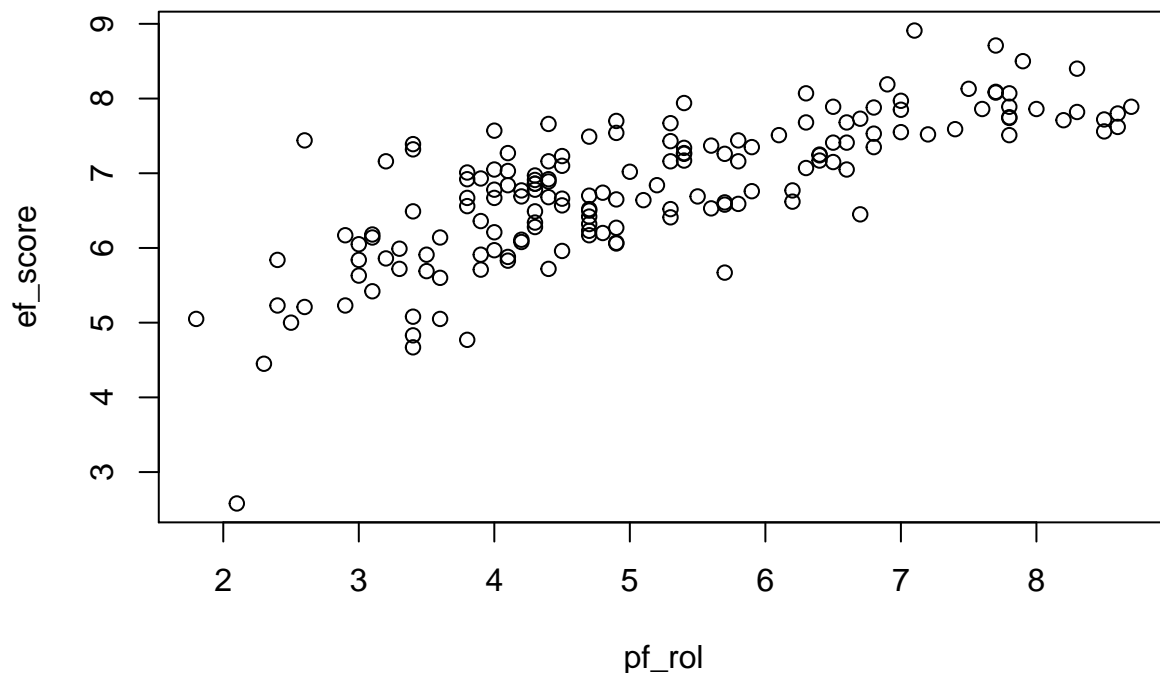
```
#predictor: pf_rol, outcome: ef_score  
print("pf_rol contains no missing data:")
```

```
## [1] "pf_rol contains no missing data:"
```

```
sum(as.character(data$pf_rol)=="-")==0
```

```
## [1] TRUE
```

```
pf_rol<-as.numeric(as.character(data$pf_rol))  
ef_score<-as.numeric(as.character(data$ef_score))  
plot(pf_rol,ef_score)
```



Fit a multiple logistic regression model by including more than one predictors, interpret the results in terms of conditional odds ratio etc.