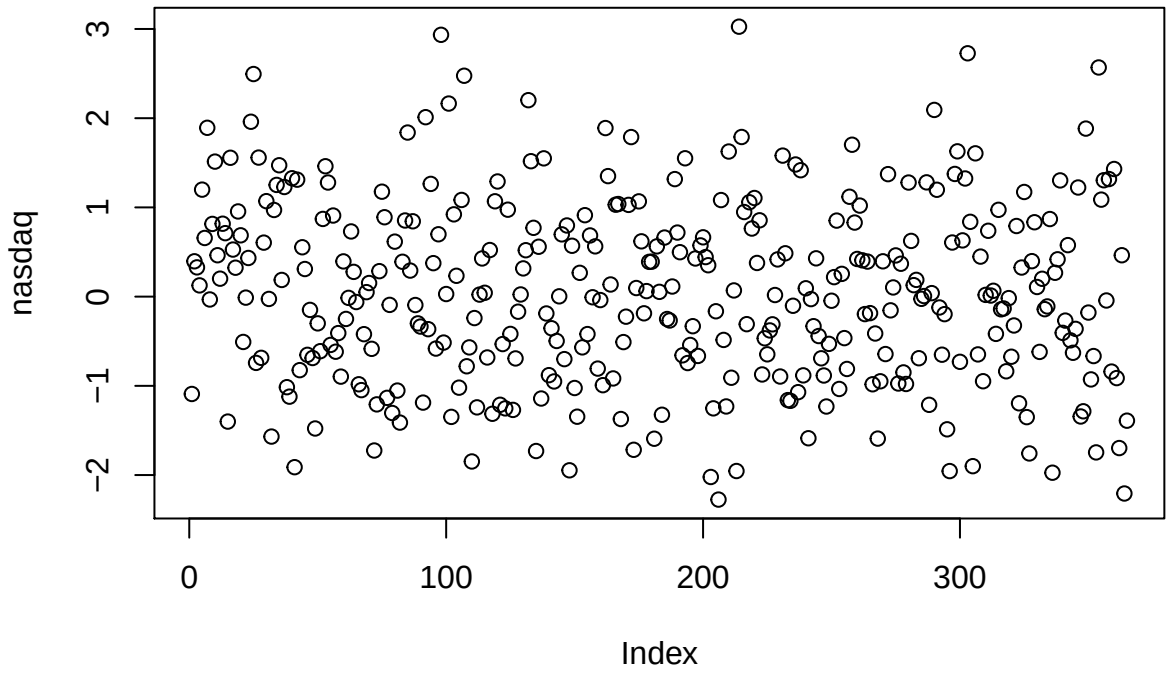
Zarate Marques R Camp Day 2 HW

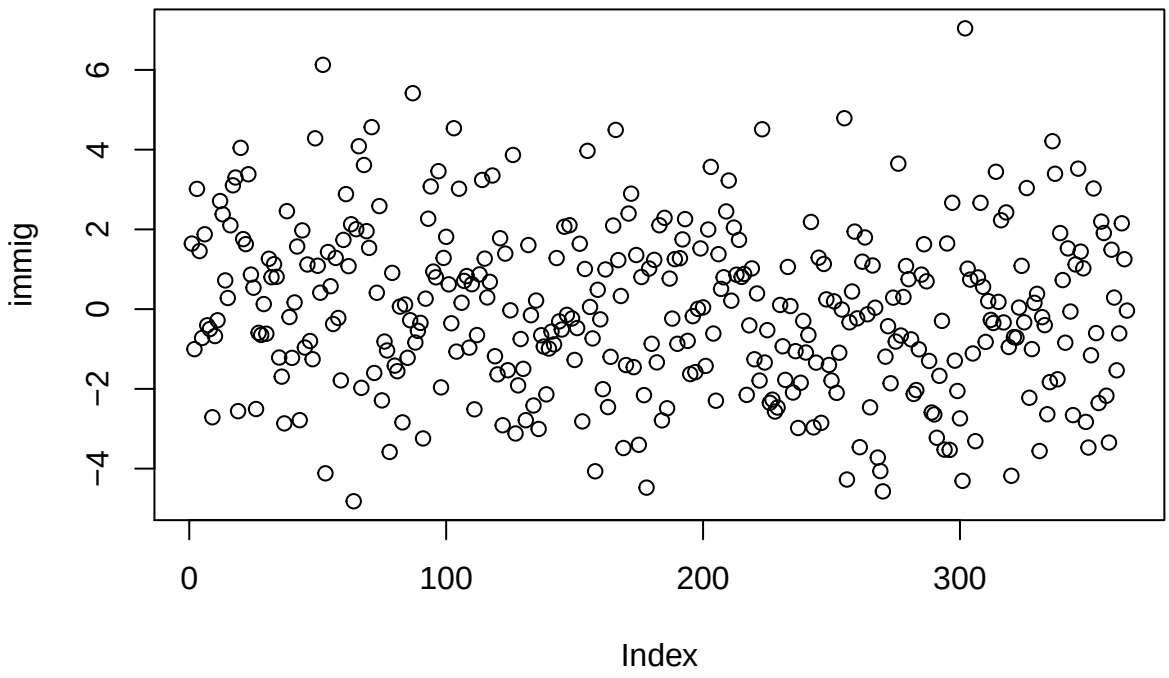
Problem 1

A major criticism of allowing a larger number of immigrants to enter the United States is that immigrants can have a negative impact on the economy. Others have responded by saying immigrants play a vital role in our economy by helping to work jobs in the secondary market that without immigrants would otherwise go unfilled. Corporations have the most to benefit from immigration because of the cost and availability of labor. Does the larger work force have an effect on the stock market? Does immigration help to improve the economy?

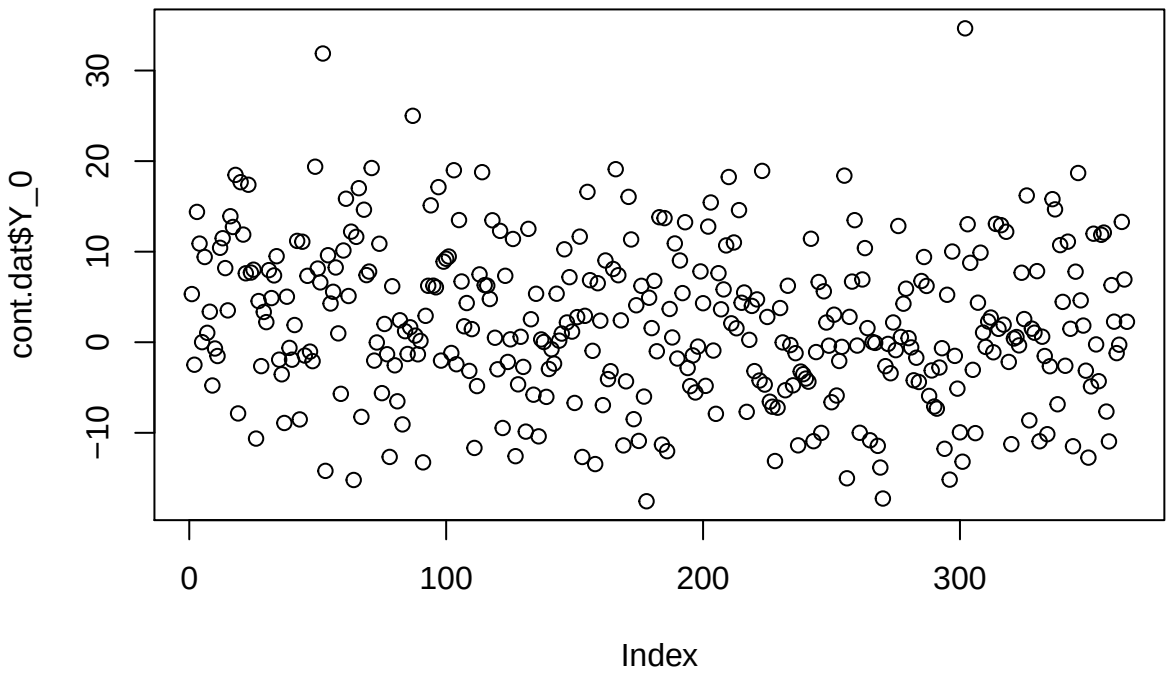
```
rm(list=ls())
set.seed(118234)

simN <- 365 # I am looking at the data from each day over the course of a year

nasdaq = rnorm(n=simN, mean = 0.03, sd =1) #this is the movement of the stock market
plot(nasdaq)
```




```
immig = rnorm(n=simN, mean = 0.05, sd =2) #This is the rate of immigration in terms of percentage leavi
plot(immig)
```



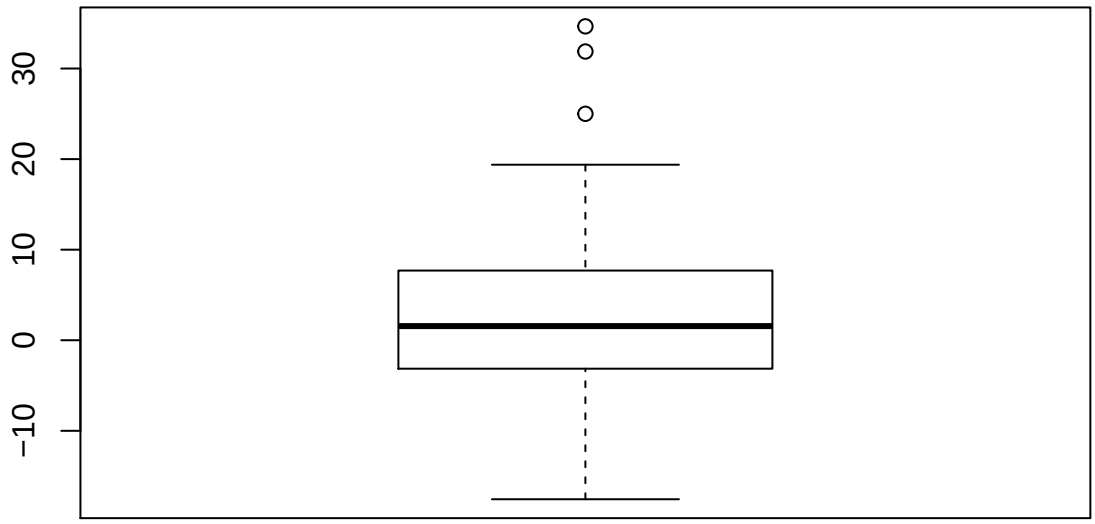

```
a = 2
b = 1 # coefficient for the effect of the stock market on economic growth, I chose 1 as I believe that
b2 = 4 #the effect of increased immigration on economic growth
error = rnorm(n=simN, mean =0, sd =2) #noise

Y_0 <- a + (b*nasdaq) + (b2*immig) + error # This represents economic growth, lets call it money spent

cont.dat <- data.frame(Y_0, nasdaq, immig)
plot(cont.dat$Y_0)
```




```
#First analysis of whether immigration benefits the stock market
mark.immig <-lm(nasdaq ~ immig, data = cont.dat)
summary(mark.immig)
```

```
##
## Call:
## lm(formula = nasdaq ~ immig, data = cont.dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.31651 -0.73889 -0.03335  0.71013  2.98574
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.048730   0.053159   0.917    0.360
## immig       -0.005081   0.025854  -0.197    0.844
##
## Residual standard error: 1.016 on 363 degrees of freedom
## Multiple R-squared:  0.0001064,  Adjusted R-squared:  -0.002648
## F-statistic: 0.03862 on 1 and 363 DF,  p-value: 0.8443
```

```
#looking at the economic growth
boxplot(cont.dat$Y_0)
```



Problem 2

Do Latino candidates mobilize Latino Voters? Latinx political participation rates are consistently lower than other racial/ethnic groups. Some have argued that the reason why Latinx participation rates are lower than other groups is because they have never been mobilized or asked to vote. I am looking at the likelihood of voting for both Latinx and non Latinx citizens when there is a Latinx candidate. I am assuming that the Latinx candidate helps to mobilize Latinx citizens which will increase their chances of voting.

```
rm(list=ls())
set.seed(4830389)

simN <- 50000 #Size of the district

latinx <- rbinom(n=simN, size=1, prob=0.18) # approximately 18% of the district is hispanic
# I chose this value because approximately 18% of the U.S. population is hispanic

mobilize0 <- rep(NA, simN)
mobilize1 <- ifelse(latinx==1 & runif(simN)<0.24, 1, mobilize0) #24 percent of hispanics have been mobi
mobilize2 <- ifelse(latinx==0 & runif(simN)<0.13, 1, mobilize1) #13 percent of the other population hav
mobilize <- ifelse(is.na(mobilize2)==T, 0, 1) #for those that were not mobilized
head(mobilize)
```

```
## [1] 0 0 0 0 0 0
```

```
sim.ethnicity <- latinx
#does having a coethnic candidate mobilize Latinos? I am including a binary variable that is a copy of
```

```r
sim.data <- data.frame(latinx, mobilize, sim.ethnicity)
head(sim.data)
```

```
##   latinx mobilize sim.ethnicity
## 1      1        0             1
## 2      0        0             0
## 3      0        0             0
## 4      0        0             0
## 5      0        0             0
## 6      0        0             0
```

```r
sum(latinx)
```

```
## [1] 9094
```

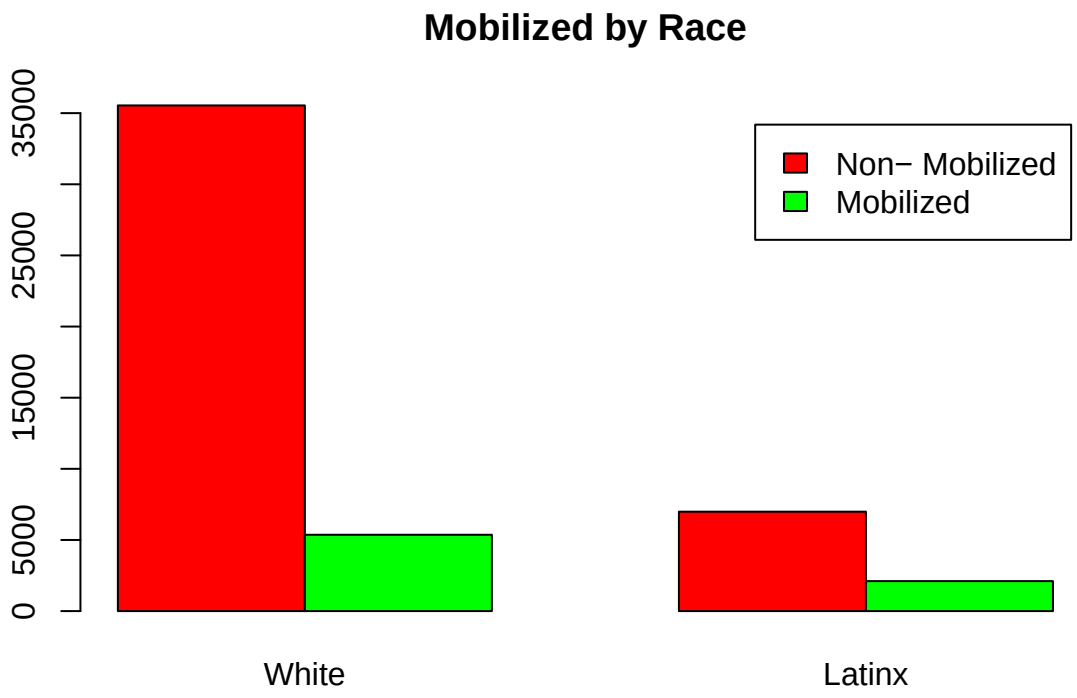
```r
sum(mobilize)
```

```
## [1] 7471
```

```r
counts <- t(table(sim.data$latinx, sim.data$mobilize))
counts
```

```
##
##         0     1
##   0 35542  6987
##   1  5364  2107
```

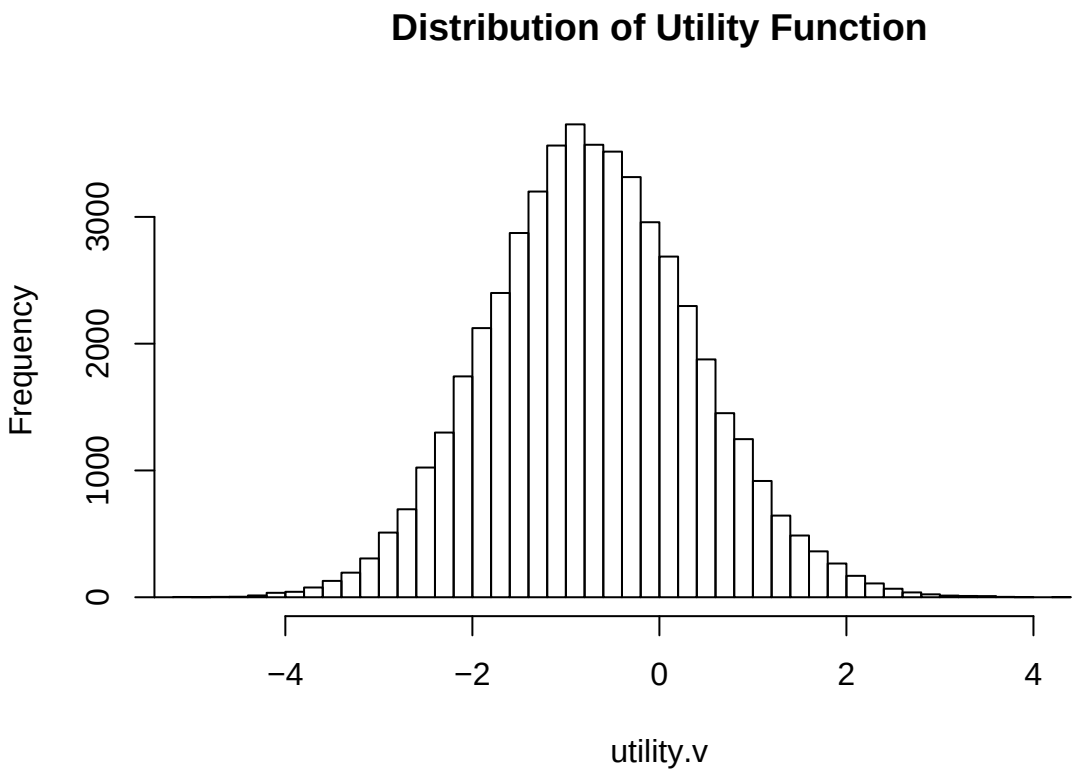
```r
par(mfrow=c(1,1))
barplot(counts, main="Mobilized by Race", xlab="", beside=T,
        col=c("red", "green"), legend=c("Non- Mobilized", "Mobilized"),
        axisnames=T, names.arg=c("White", "Latinx"))
```




```
rho <- cor(latinx, mobilize)
rho   #correlation between race and mobilization
```

```
## [1] 0.1088103
```

```
#generating the likelihood of voting
error      <- rnorm(n=simN, mean=-1, sd=1) #error
b.latinx   <- runif(n=simN, min=0.24, max=0.24) #likelihood of voting if Latinx
b.mobilize <- runif(n=simN, min=0.82, max=0.82) #Likelihood of voting if mobilized
b.sim.ethnicity <- runif(n=simN, min=0.64, max=0.64) #mobilizing factor of shared ethnicity
#voting utility function
utility.v <- b.latinx*latinx + b.mobilize*mobilize + b.sim.ethnicity*sim.ethnicity + error

hist(utility.v, breaks=50, xlim =c(-5,5),
     main="Distribution of Utility Function")
```




```
vote <- ifelse(utility.v > 0, 1, 0)
sim.data$vote <- vote


poll <- sample(1:100, size = nrow(sim.data), replace = TRUE)
sim.data$poll <- poll

agg.vote <- c()
agg.mobilize <- c()
agg.latinx <- c()
agg.total <- c()

for(i in 1:100) {
agg.vote[i]     <- sum(sim.data$vote[poll==i])
agg.mobilize[i] <- sum(sim.data$mobilize[poll==i])
agg.latinx[i]   <- sum(sim.data$latinx[poll==i])
agg.total[i]    <- length(sim.data$poll[poll==i])
}

latinx <- agg.latinx
other <- agg.total - agg.latinx
vote.m <- agg.vote # voting if mobilized
vote.nm <- agg.total - agg.vote # did not vote
nmobilize <- agg.mobilize # number of people that have been mobilized
nmobilizen <- agg.total - agg.mobilize # those who were not mobilized
total <- latinx + other #total population
```

```r
agg.data <- data.frame(poll, latinx, other, vote.m, vote.nm, nmobilize, nmobilizen, total, sim.ethnicity

head(agg.data)
```

```
##   poll latinx other vote.m vote.nm nmobilize nmobilizen total
## 1   18     77   405    119     363        80        402   482
## 2   50     92   407    136     363        64        435   499
## 3   24     93   389    134     348        71        411   482
## 4   17    100   410    135     375        71        439   510
## 5    6     94   418    137     375        97        415   512
## 6   85    100   388    123     365        76        412   488
##   sim.ethnicity
## 1             1
## 2             0
## 3             0
## 4             0
## 5             0
## 6             0
```

```r
#Analysis
anal <- lm(vote ~ latinx + nmobilize +sim.ethnicity, data=agg.data)
summary(anal)
```

```
##
## Call:
## lm(formula = vote ~ latinx + nmobilize + sim.ethnicity, data = agg.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5261 -0.1976 -0.1940  0.4756  0.8152
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.2069680  0.0236561   8.749   <2e-16 ***
## latinx         0.0002516  0.0002073   1.214   0.2249
## nmobilize     -0.0004649  0.0002454  -1.894   0.0582 .
## sim.ethnicity  0.3213432  0.0048346  66.467   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.417 on 49996 degrees of freedom
## Multiple R-squared:  0.08129,    Adjusted R-squared:  0.08123
## F-statistic:  1475 on 3 and 49996 DF,  p-value: < 2.2e-16
```

```r
logistic <- glm(vote ~ latinx + nmobilize + sim.ethnicity, data=agg.data)
summary(logistic)
```

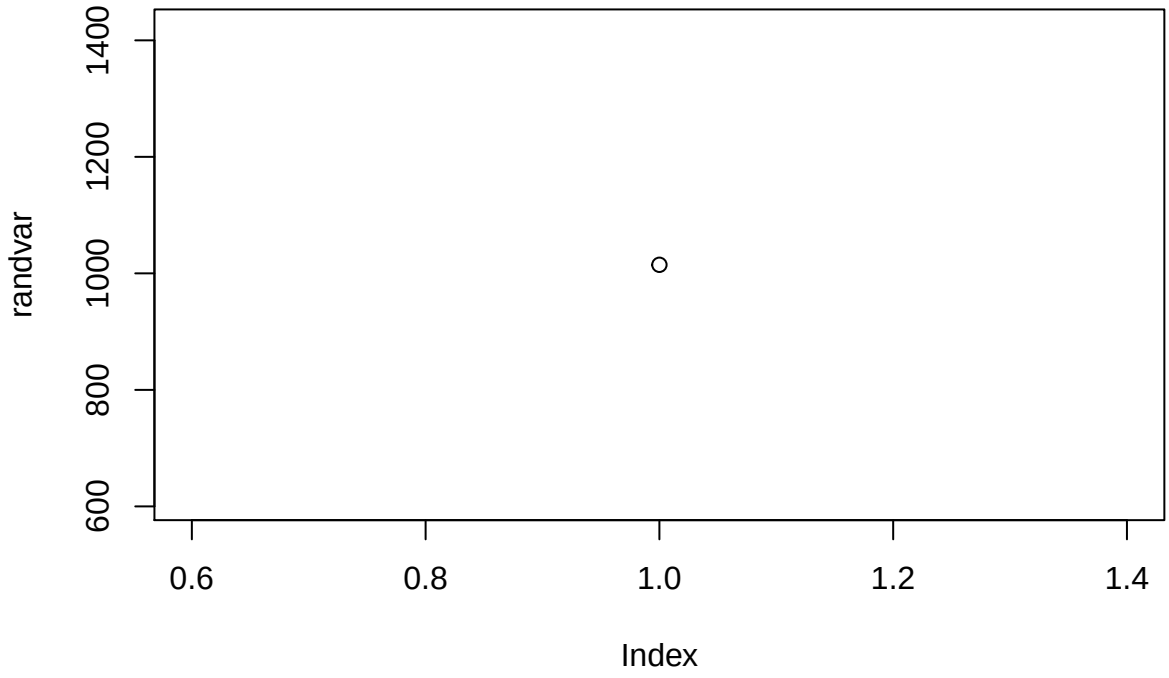
```
##
## Call:
## glm(formula = vote ~ latinx + nmobilize + sim.ethnicity, data = agg.data)
##
## Deviance Residuals:
```

```
##     Min       1Q   Median       3Q      Max
## -0.5261  -0.1976  -0.1940   0.4756   0.8152
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.2069680  0.0236561   8.749   <2e-16 ***
## latinx         0.0002516  0.0002073   1.214   0.2249
## nmobilize     -0.0004649  0.0002454  -1.894   0.0582 .
## sim.ethnicity  0.3213432  0.0048346  66.467   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.1738966)
##
##     Null deviance: 9463.4  on 49999  degrees of freedom
## Residual deviance: 8694.1  on 49996  degrees of freedom
## AIC: 54435
##
## Number of Fisher Scoring iterations: 2
```

Discussion: My simulation does not truly represent the question that I was trying to answer. I had trouble thinking about the question in terms of an aggregate and so I don't believe that my model is correct. I was able to successfully simulate data as appropriate for the goal of this homework assignment.

Playing around with creating continuous random variables

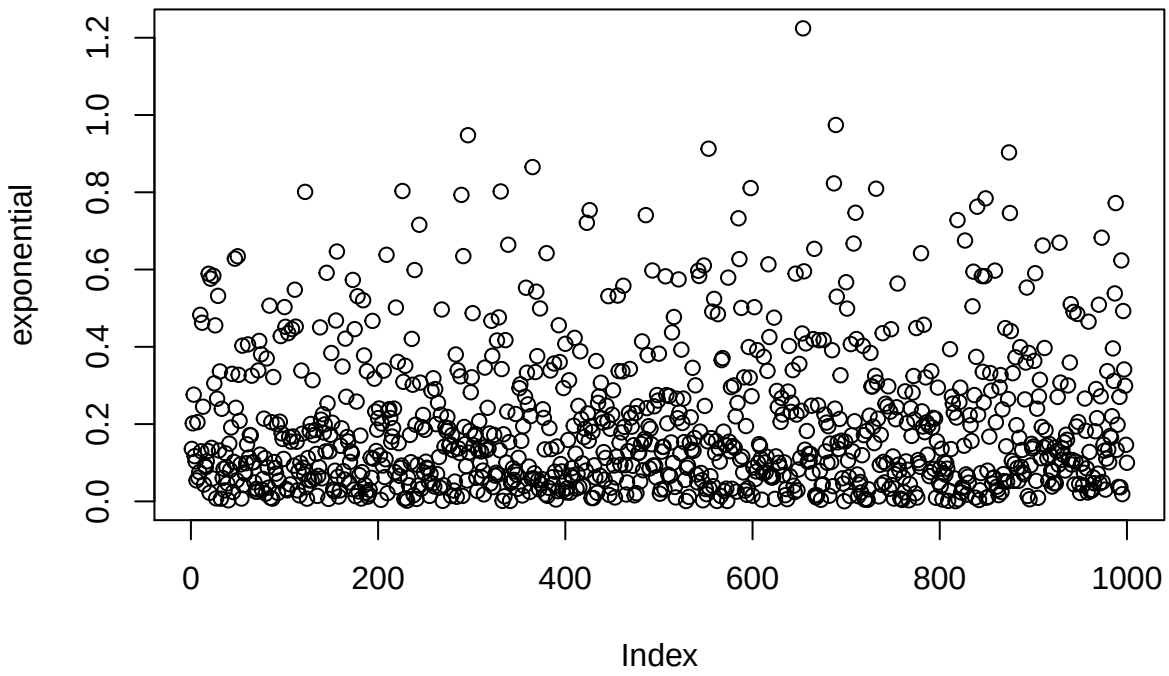
```
#These were other attempts at creating a function random continuous variable
rm(list=ls())

simN <- 1000

lambda = 1
randvar = (-1/lambda) * sum(log(runif(n=simN, min =0, max =1)))
plot(randvar)
```




```
U = runif(n=simN,min = 0, max = 1)
inflation = -log(U)


#vote <-rbern(n=simN, prob = 0.3)



exponential <- rexp(n=simN, rate=5)
summary(exponential)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.0003507 0.0614964 0.1408444 0.1976188 0.2732808 1.2242853
```

```
plot(exponential)
```


exponential
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
200
400
600
800
1000
Index