HW1

Homework #1 Group 2 Econ B2000, MA Econometrics Amira, Muhibul, Minghao

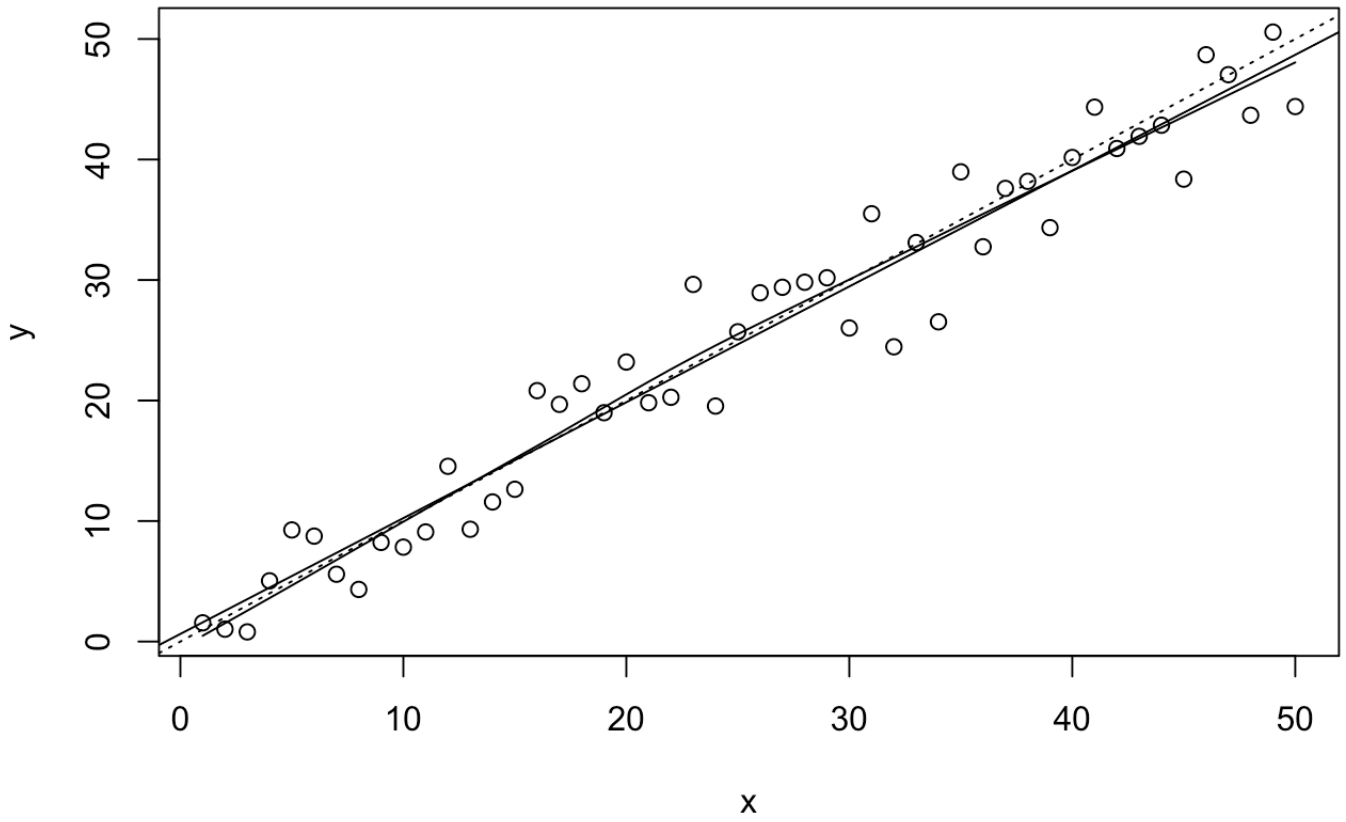
```
x <- 1:50
w <- 1 + sqrt(x)/2
example1 <- data.frame(x=x, y= x + rnorm(x)*w)
attach(example1)
```

```
## The following object is masked _by_ .GlobalEnv:
##
##     x
```

```
fm <- lm(y ~ x)
summary(fm)
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9389 -2.4671  0.0262  2.6228  6.8909
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.63244    0.94770   0.667    0.508
## x            0.96137    0.03234  29.723   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.3 on 48 degrees of freedom
## Multiple R-squared:  0.9485, Adjusted R-squared:  0.9474
## F-statistic: 883.4 on 1 and 48 DF,  p-value: < 2.2e-16
```

```
lrf <- lowess(x, y)
plot(x, y)
lines(x, lrf$y)
abline(0, 1, lty=3)
abline(coef(fm))
```




```
detach()
```

```
load("Household_Pulse_data.RData")
#glimpse(acs2017_ny) try this later
Household_Pulse_data[1:10,1:7]
```

```
##       RHISPANIC RRACE      EEDUC      MS EGENID_BIRTH GENID_DESCRIBE
## 1  Not Hispanic White   bach deg      NA       female         female
## 2  Not Hispanic White HS diploma married       female         female
## 3  Not Hispanic White   bach deg widowed       female         female
## 4  Not Hispanic White    adv deg   never       female         female
## 5  Not Hispanic Black  some coll   never       female         female
## 6      Hispanic White   bach deg   never       female         female
## 7  Not Hispanic White    adv deg widowed       female         female
## 8  Not Hispanic White  some coll widowed       female         female
## 9  Not Hispanic White  assoc deg married       female         female
## 10 Not Hispanic White    adv deg married       female         female
##    SEXUAL_ORIENTATION
## 1            straight
## 2            straight
## 3            straight
## 4            straight
## 5            straight
## 6            straight
## 7            straight
## 8            straight
## 9            straight
## 10           straight
```

```
attach(Household_Pulse_data)
summary(Household_Pulse_data)
```

```
##         RHISPANIC       RRACE                EEDUC                MS
##  Not Hispanic:62660   White:56938   less than hs:  411   NA       :  881
##  Hispanic    : 6454   Black: 5412   some hs     :  936   married  :40036
##                       Asian: 3561   HS diploma  : 7857   widowed  : 3872
##                       Other: 3203   some coll   :14596   divorced :10310
##                                     assoc deg   : 7508   separated: 1214
##                                     bach deg    :20075   never    :12801
##                                     adv deg     :17731
##  EGENID_BIRTH       GENID_DESCRIBE       SEXUAL_ORIENTATION
##  male  :27592   NA         : 1131   NA            : 1506
##  female:41522   male       :26796   gay or lesbian: 2343
##                 female     :40263   straight      :61238
##                 transgender:  202   bisexual      : 2288
##                 other      :  722   something else:  871
##                                     dont know     :  868
##
##                       KIDS_LT5Y                         KIDS_5_11Y
##  NA                        :62342   NA                       :58467
##  Yes children under 5 in HH: 6772   Yes children 5 - 11 in HH:10647
```

```
##
##
##
##
##
##                      KIDS_12_17Y                                    ENROLLNONE
##  NA                          :58046   NA                                   :64285
##  Yes children 12 - 17 in HH:11068   children not in any type of school: 4829
##
##
##
##
##
##
##                RECVDVACC                            DOSESRV
##  NA                  :  851   NA                         : 9105
##  yes got vaxx        :60326   yes got all doses          :57762
##  no did not get vaxx: 7937   yes plan to get all doses: 1993
##                               no will not get all doses:  254
##
##
##
##                       GETVACRV                                    KIDDOSES
##  NA                          :61159   NA                                   :58318
##  definitely will get vaxx:  609   Yes kids got or will get all doses: 7135
##  probably will get vaxx  :  731   no kids did not or will not        : 3661
##  unsure about vaxx       : 1584
##  probably not            : 1599
##  definitely not          : 3432
##
##                      KIDGETVAC                            HADCOVID
##  NA                          :65384   NA                         : 1363
##  definitely will get vaxx:  487   yes doctor told had covid: 9122
##  probably will get vaxx  :  439   no did not                 :58221
##  unsure about vaxx       :  736   not sure                   :  408
##  probably not            :  593
##  definitely not          : 1036
##  dont know yet           :  439
##                    WRKLOSSRV                                ANYWORK
##  NA                        : 1961   NA                             : 2135
##  yes recent HH job loss: 8058   yes employment in last 7 days:39237
##  no recent HH job loss :59095   no employment in last 7 days :27742
##
##
##
##
##                  KINDWORK                 RSNNOWRKRV
##  NA                      :30540   NA                  :42659
```

```
##  work for govt      : 6378   retired          :15024
##  work for private co:21370   other            : 4027
##  work for nonprofit : 5055   sick or disabled: 1451
##  self employed      : 4966   caring for kids  : 1329
##  work in family biz :  805   laid off         : 1164
##                              (Other)          : 3460
##                                       CHLDCARE
##  NA                                       :58419
##  yes impacts to childcare because pandemic: 2566
##  no                                       : 8129
##
##
##
##
##                            CURFOODSUF
##  NA                             : 6770
##  had enough food                :49234
##  had enough but not what wanted: 9947
##  sometimes not enough food      : 2486
##  often not enough food          :  677
##
##
##                                                CHILDFOOD
##  NA                                                 :64258
##  often kids not eating enough because couldnt afford:  271
##  sometimes kids not eating enough                   : 1191
##  kids got enough food                               : 3394
##
##
##
##                                              ANXIOUS
##  NA                                              : 7946
##  no anxiety over past 2 wks                      :26611
##  several days anxiety over past 2 wks            :19794
##  more than half the days anxiety over past 2 wks: 6140
##  nearly every day anxiety                        : 8623
##
##
##                                               WORRY
##  NA                                              : 8016
##  no worry over past 2 wks                        :31876
##  several days worried over past 2 wks            :17936
##  more than half the days worried over past 2 wks: 4979
##  nearly every day worry                          : 6307
##
##
##                              TENURE
```

```
##  NA                            :11103
##  housing owned free and clear :16738
##  housing owned with mortgage  :28016
##  housing rented               :12579
##  housing occupied without rent:  678
##
##
##                                LIVQTRRV                RENTCUR
##  live in detached 1 family          :41348   NA              :56572
##  NA                                 :11336   current on rent:11239
##  live in bldg w 5+ apts             : 6731   behind on rent : 1303
##  live in 1 family attached to others: 4628
##  live in mobile home                : 1781
##  live in building with 3-4 apts     : 1737
##  (Other)                            : 1553
##                MORTCUR                                                EVICT
##  NA                  :41200   NA                                       :67859
##  current on mortgage:26462   very likely evicted in next 2 months      :  207
##  behind on mortgage : 1452   somewhat likely evicted in next 2 months  :  377
##                              not very likely evicted in next 2 months  :  345
##                              not at all likely evicted in next 2 months:  326
##
##
##                                     FORCLOSE               EST_ST
##  NA                                     :67695   California   : 5359
##  very likely forclosed in next 2 months :   65   Texas        : 3766
##  somewhat likely forclosed in next 2 months:  218   Florida      : 2728
##  not very likely forclosed in next 2 months:  474   Washington   : 2634
##  not at all forclosed in next 2 months  :  662   Massachusetts: 1965
##                                                  Oregon       : 1934
##                                                  (Other)      :50728
##                   PRIVHLTH                          PUBHLTH            REGION
##  has private health ins:46869   has public health ins:23346   Northeast:10478
##  no private health ins :11275   no public health ins :33381   South    :22680
##  NA                    :10970   NA                   :12387   Midwest  :13651
##                                                              West     :22305
##
##
##
##                    INCOME       TBIRTH_YEAR   Num_kids_Pub_School
##  NA                    :14637   Min.   :1933   Min.   :0.00
##  HH income $100k - 149:10117   1st Qu.:1955   1st Qu.:1.00
##  HH income $50k - 74.9: 9330   Median :1967   Median :2.00
##  HH income $75 - 99.9 : 7830   Mean   :1968   Mean   :1.71
##  HH income $200k +    : 6117   3rd Qu.:1981   3rd Qu.:2.00
##  HH income $35k - 49.9: 5805   Max.   :2003   Max.   :4.00
##  (Other)               :15278                  NA's   :55108
```

```
##  Num_kids_Priv_School Num_kids_homeschool        Works_onsite
##  Min.   :0.00         Min.   :0.00         NA           : 6350
##  1st Qu.:0.00         1st Qu.:0.00         worked onsite:34918
##  Median :1.00         Median :1.00         no           :27846
##  Mean   :1.03         Mean   :0.87
##  3rd Qu.:2.00         3rd Qu.:2.00
##  Max.   :2.00         Max.   :2.00
##  NA's   :66430        NA's   :67421
##           works_remote            Shop_in_store
##  NA             : 8022   NA              : 6873
##  worked remotely:22863   shopped in store:53576
##  no             :38229   no              : 8665
##
##
##
##
##                  eat_in_restaurant
##  NA                        : 7217
##  eat at restaurant indoors:32405
##  no                        :29492
##
##
##
##
```

```
summary(TBIRTH_YEAR[GENID_DESCRIBE == "female"])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1933    1956    1968    1969    1981    2003
```

```
summary(TBIRTH_YEAR[GENID_DESCRIBE == "male"])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1933    1953    1965    1967    1980    2003
```

```
summary(TBIRTH_YEAR[GENID_DESCRIBE == "transgender"])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1933    1975    1988    1983    1995    2003
```

```
summary(TBIRTH_YEAR[GENID_DESCRIBE == "other"])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1933    1959    1973    1973    1988    2003
```

```
summary(TBIRTH_YEAR[GENID_DESCRIBE == "NA"])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1933    1952    1962    1964    1977    2003
```

```
mean(TBIRTH_YEAR[GENID_DESCRIBE == "female"])
```

```
## [1] 1968.666
```

```
sd(TBIRTH_YEAR[GENID_DESCRIBE == "female"])
```

```
## [1] 15.45378
```

```
mean(TBIRTH_YEAR[GENID_DESCRIBE == "male"])
```

```
## [1] 1966.549
```

```
sd(TBIRTH_YEAR[GENID_DESCRIBE == "male"])
```

```
## [1] 16.29105
```

```
summary(EEDUC)
```

```
## less than hs      some hs   HS diploma    some coll    assoc deg     bach deg
##          411          936         7857        14596         7508        20075
##      adv deg
##        17731
```

I want to see if there's a trend associated with higher education vs vaccination status

```
all_doses <- data.frame(matrix(ncol=1+length(summary(EEDUC)),nrow=0))
colnames(all_doses) <- levels(unique(EEDUC))
for (i in 1:length(summary(EEDUC))){
  all_doses[1,i] <- summary(EEDUC[DOSESRV=='yes got all doses' | DOSESRV == 'yes plan
to get all doses'])[i] / summary(EEDUC)[i]
}
all_doses
```

```
##   less than hs   some hs HS diploma some coll assoc deg  bach deg   adv deg NA
## 1    0.6885645 0.6826923  0.7650503 0.8144697 0.8251199 0.9030635 0.9367774 NA
```

I want to also include people who “claim” they are going to get vaccinated. Some people might say vaccination is a good thing but never do it. Let’s see what percentage of people actually got vaccinated.

```
for (i in 1:length(summary(EEDUC))){
  all_doses[2,i] <- summary(EEDUC[DOSESRV=='yes got all doses'])[i] / summary(EEDUC)[
i]
}
all_doses
```

```
##   less than hs   some hs HS diploma some coll assoc deg  bach deg   adv deg NA
## 1    0.6885645 0.6826923  0.7650503 0.8144697 0.8251199 0.9030635 0.9367774 NA
## 2    0.6131387 0.6111111  0.7146494 0.7760345 0.7904901 0.8812951 0.9231854 NA
```

```
all_doses[3,] <- all_doses[1,] - all_doses[2,] #for percentage difference
```

Some cleaning

```
all_doses[8] <- NULL
all_doses_t <- t(all_doses) #For better viewing
colnames(all_doses_t) <- c('Included','Not Included','percentage_difference')
all_doses_t
```

```
##               Included Not Included percentage_difference
## less than hs 0.6885645    0.6131387            0.07542579
## some hs      0.6826923    0.6111111            0.07158120
## HS diploma   0.7650503    0.7146494            0.05040092
## some coll    0.8144697    0.7760345            0.03843519
## assoc deg    0.8251199    0.7904901            0.03462973
## bach deg     0.9030635    0.8812951            0.02176837
## adv deg      0.9367774    0.9231854            0.01359201
```

It’s interesting that for each cases, there’s a clear trend that with higher level of education, people tend to get

vaccinated. It's also interesting to note that the difference between the percentage of people who "claim" they are going to get vaccinated and the percentage of people who actually get vaccinated also has a decreasing trend with higher education. This could suggest that higher education people genuinely believe in getting vaccinated.