title: "Homework#4, Lab#4, Econometrics B2000"

author: "Hassan Fayyaz"

date:"10/07/2021"

'Q#1: Names of study group '

'Hassan Fayyaz, Taulant Bega, Kyle Zhou, Joel Andrade'

'Q#2: Write up the results of your Lab 2 work.'

```
rm(list = ls(all = TRUE))
setwd("C:/Users/hassa/Desktop/Fall 2021/1_Econometrics_(B2000)/Homeworks/1_HW")
load("acs2017_ny_data.RData")

# The purpose of this research is to find the income wage of college graduates based on their race (Hispanic or Asian).

use_varb <- (AGE >= 30) & (AGE <= 55) & (LABFORCE == 2) & (WKSWORK2 > 4) & (UHRSWORK >= 35)
dat_use <- subset(acs2017_ny,use_varb)
detach()
attach(dat_use)
'The following objects are masked from dat_use (pos = 3):

    AfAm, AGE, Amindian, ANCESTR1, ANCESTR1D, ANCESTR2, ANCESTR2D, Asian,
below_150poverty, below_200poverty, below_povertyline, BPL, BPLD, BUILTYR2, CITIZEN,
```

CLASSWKR, CLASSWKRD, Commute_bus, Commute_car, Commute_other, Commute_rail, Commute_subway, COSTELEC, COSTFUEL, COSTGAS, COSTWATR, DEGFIELD, DEGFIELD2,

DEGFIELD2D, DEGFIELDD, DEPARTS, EDUC, educ_advdeg, educ_college, educ_hs, educ_nohs, educ_somecoll, EDUCD, EMPSTAT, EMPSTATD, FAMSIZE, female, foodstamps,

FOODSTMP, FTOTINC, FUELHEAT, GQ, has_AnyHealthIns, has_PvtHealthIns, HCOVANY, HCOVPRIV, HHINCOME, Hisp_Cuban, Hisp_DomR, Hisp_Mex, Hisp_PR, HISPAN, HISPAND,

Hispanic, in_Bronx, in_Brooklyn, in_Manhattan, in_Nassau, in_NYC, in_Queens, in_StatenI, in_Westchester, INCTOT, INCWAGE, IND, LABFORCE, LINGISOL, MARST,

MIGCOUNTY1, MIGPLAC1, MIGPUMA1, MIGRATE1, MIGRATE1D, MORTGAGE, NCHILD, NCHLT5, OCC, OWNCOST, OWNERSHP, OWNERSHPD, POVERTY, PUMA, PWPUMA00, RACE, race_oth, RACED,

RELATE, RELATED, RENT, ROOMS, SCHOOL, SEX, SSMC, TRANTIME, TRANWORK, UHRSWORK, UNITSSTR, unmarried, veteran, VETSTAT, VETSTATD, white, WKSWORK2, YRSUSA1

The following objects are masked from dat_use (pos = 4):

AfAm, AGE, Amindian, ANCESTR1, ANCESTR1D, ANCESTR2, ANCESTR2D, Asian, below_150poverty, below_200poverty, below_povertyline, BPL, BPLD, BUILTYR2, CITIZEN,

CLASSWKR, CLASSWKRD, Commute_bus, Commute_car, Commute_other, Commute_rail, Commute_subway, COSTELEC, COSTFUEL, COSTGAS, COSTWATR, DEGFIELD, DEGFIELD2,

DEGFIELD2D, DEGFIELDD, DEPARTS, EDUC, educ_advdeg, educ_college, educ_hs, educ_nohs, educ_somecoll, EDUCD, EMPSTAT, EMPSTATD, FAMSIZE, female, foodstamps,

FOODSTMP, FTOTINC, FUELHEAT, GQ, has_AnyHealthIns, has_PvtHealthIns, HCOVANY, HCOVPRIV, HHINCOME, Hisp_Cuban, Hisp_DomR, Hisp_Mex, Hisp_PR, HISPAN, HISPAND,

Hispanic, in_Bronx, in_Brooklyn, in_Manhattan, in_Nassau, in_NYC, in_Queens, in_StatenI, in_Westchester, INCTOT, INCWAGE, IND, LABFORCE, LINGISOL, MARST,

MIGCOUNTY1, MIGPLAC1, MIGPUMA1, MIGRATE1, MIGRATE1D, MORTGAGE, NCHILD, NCHLT5, OCC, OWNCOST, OWNERSHP, OWNERSHPD, POVERTY, PUMA, PWPUMA00, RACE, race_oth, RACED,

RELATE, RELATED, RENT, ROOMS, SCHOOL, SEX, SSMC, TRANTIME, TRANWORK, UHRSWORK, UNITSSTR, unmarried, veteran, VETSTAT, VETSTATD, white, WKSWORK2, YRSUSA1

The following objects are masked from dat_NYC (pos = 5):

AfAm, AGE, Amindian, ANCESTR1, ANCESTR1D, ANCESTR2, ANCESTR2D, Asian, below_150poverty, below_200poverty, below_povertyline, BPL, BPLD, BUILTYR2, CITIZEN,

CLASSWKR, CLASSWKRD, Commute_bus, Commute_car, Commute_other, Commute_rail, Commute_subway, COSTELEC, COSTFUEL, COSTGAS, COSTWATR, DEGFIELD, DEGFIELD2,

DEGFIELD2D, DEGFIELDD, DEPARTS, EDUC, educ_advdeg, educ_college, educ_hs, educ_nohs, educ_somecoll, EDUCD, EMPSTAT, EMPSTATD, FAMSIZE, female, foodstamps,

FOODSTMP, FTOTINC, FUELHEAT, GQ, has_AnyHealthIns, has_PvtHealthIns, HCOVANY, HCOVPRIV, HHINCOME, Hisp_Cuban, Hisp_DomR, Hisp_Mex, Hisp_PR, HISPAN, HISPAND,

Hispanic, in_Bronx, in_Brooklyn, in_Manhattan, in_Nassau, in_NYC, in_Queens, in_StatenI, in_Westchester, INCTOT, INCWAGE, IND, LABFORCE, LINGISOL, MARST,

MIGCOUNTY1, MIGPLAC1, MIGPUMA1, MIGRATE1, MIGRATE1D, MORTGAGE, NCHILD, NCHLT5, OCC, OWNCOST, OWNERSHP, OWNERSHPD, POVERTY, PUMA, PWPUMA00, RACE, race_oth, RACED,

RELATE, RELATED, RENT, ROOMS, SCHOOL, SEX, SSMC, TRANTIME, TRANWORK, UHRSWORK, UNITSSTR, unmarried, veteran, VETSTAT, VETSTATD, white, WKSWORK2, YRSUSA1

The following objects are masked from dat_NYC (pos = 6):

AfAm, AGE, Amindian, ANCESTR1, ANCESTR1D, ANCESTR2, ANCESTR2D, Asian, below_150poverty, below_200poverty, below_povertyline, BPL, BPLD, BUILTYR2, CITIZEN,

CLASSWKR, CLASSWKRD, Commute_bus, Commute_car, Commute_other, Commute_rail, Commute_subway, COSTELEC, COSTFUEL, COSTGAS, COSTWATR, DEGFIELD, DEGFIELD2,

DEGFIELD2D, DEGFIELDD, DEPARTS, EDUC, educ_advdeg, educ_college, educ_hs, educ_nohs, educ_somecoll, EDUCD, EMPSTAT, EMPSTATD, FAMSIZE, female, foodstamps,

FOODSTMP, FTOTINC, FUELHEAT, GQ, has_AnyHealthIns, has_PvtHealthIns, HCOVANY, HCOVPRIV, HHINCOME, Hisp_Cuban, Hisp_DomR, Hisp_Mex, Hisp_PR, HISPAN, HISPAND,

Hispanic, in_Bronx, in_Brooklyn, in_Manhattan, in_Nassau, in_NYC, in_Queens, in_StatenI, in_Westchester, INCTOT, INCWAGE, IND, LABFORCE, LINGISOL, MARST,

MIGCOUNTY1, MIGPLAC1, MIGPUMA1, MIGRATE1, MIGRATE1D, MORTGAGE, NCHILD, NCHLT5, OCC, OWNCOST, OWNERSHP, OWNERSHPD, POVERTY, PUMA, PWPUMA00, RACE, race_oth, RACED,

RELATE, RELATED, RENT, ROOMS, SCHOOL, SEX, SSMC, TRANTIME, TRANWORK, UHRSWORK, UNITSSTR, unmarried, veteran, VETSTAT, VETSTATD, white, WKSWORK2, YRSUSA1

The following objects are masked from dat_NYC (pos = 7):

AfAm, AGE, Amindian, ANCESTR1, ANCESTR1D, ANCESTR2, ANCESTR2D, Asian, below_150poverty, below_200poverty, below_povertyline, BPL, BPLD, BUILTYR2, CITIZEN,

CLASSWKR, CLASSWKRD, Commute_bus, Commute_car, Commute_other, Commute_rail, Commute_subway, COSTELEC, COSTFUEL, COSTGAS, COSTWATR, DEGFIELD, DEGFIELD2,

DEGFIELD2D, DEGFIELDD, DEPARTS, EDUC, educ_advdeg, educ_college, educ_hs, educ_nohs, educ_somecoll, EDUCD, EMPSTAT, EMPSTATD, FAMSIZE, female, foodstamps,

FOODSTMP, FTOTINC, FUELHEAT, GQ, has_AnyHealthIns, has_PvtHealthIns, HCOVANY, HCOVPRIV, HHINCOME, Hisp_Cuban, Hisp_DomR, Hisp_Mex, Hisp_PR, HISPAN, HISPAND,

Hispanic, in_Bronx, in_Brooklyn, in_Manhattan, in_Nassau, in_NYC, in_Queens, in_StatenI, in_Westchester, INCTOT, INCWAGE, IND, LABFORCE, LINGISOL, MARST,

MIGCOUNTY1, MIGPLAC1, MIGPUMA1, MIGRATE1, MIGRATE1D, MORTGAGE, NCHILD, NCHLT5, OCC, OWNCOST, OWNERSHP, OWNERSHPD, POVERTY, PUMA, PWPUMA00, RACE, race_oth, RACED,

RELATE, RELATED, RENT, ROOMS, SCHOOL, SEX, SSMC, TRANTIME, TRANWORK, UHRSWORK, UNITSSTR, unmarried, veteran, VETSTAT, VETSTATD, white, WKSWORK2, YRSUSA1'

require(stargazer)

'Loading required package: stargazer

Please cite as:

Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
R package version 5.2.2. https://CRAN.R-project.org/package=stargazer
'

require(AER)
'Loading required package: AER
Loading required package: car
Loading required package: carData
Loading required package: lmtest
Loading required package: zoo

Attaching package: ‘zoo’

The following objects are masked from ‘package:base’:

as.Date, as.Date.numeric

Loading required package: sandwich
Loading required package: survival

Attaching package: ‘survival’

The following object is masked from ‘dat_use’:

veteran

The following object is masked from ‘dat_NYC’:

veteran'

summary(dat_use)

"AGE female educ_nohs educ_hs educ_somecoll educ_college educ_advdeg SCHOOL EDUC

Min. : 0.00 Min. :0.0000 Min. :0.000 Min. :0.0000 Min. :0.000 Min. :0.0000 Min. :0.000 N/A : 5569 Grade 12 :55119

1st Qu.:22.00 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.000 No, not in school:144968 4 years of college :30802

Median :42.00 Median :1.0000 Median :0.000 Median :0.0000 Median :0.000 Median :0.0000 Median :0.000 Yes, in school : 46048 5+ years of college :23385

Mean :41.57 Mean :0.5156 Mean :0.271 Mean :0.2804 Mean :0.173 Mean :0.1567 Mean :0.119 Missing : 0 1 year of college :19947

3rd Qu.:60.00 3rd Qu.:1.0000 3rd Qu.:1.000 3rd Qu.:1.0000 3rd Qu.:0.000 3rd Qu.:0.0000 3rd Qu.:0.000 Nursery school to grade 4:14240

Max. :95.00 Max. :1.0000 Max. :1.000 Max. :1.0000 Max. :1.000 Max. :1.0000 Max. :1.000 2 years of college :14065

(Other) :39027

EDUCD DEGFIELD DEGFIELDD

Regular high school diploma :35689 N/A :142398 N/A :142398

Bachelor's degree :30802 Business : 9802 Psychology : 2926

1 or more years of college credit, no degree:19947 Education Administration and Teaching : 6708 Business Management and Administration: 2501

Master's degree :17010 Social Sciences : 4836 Accounting : 2284

Associate's degree, type not specified :14065 Medical and Health Sciences and Services: 3919 General Education : 2238

Some college, but less than 1 year : 9086 Fine Arts : 3491 English Language and Literature : 2202

(Other) :69986 (Other) : 25431 (Other) : 42036

DEGFIELD2 DEGFIELD2D PUMA GQ OWNERSHP OWNERSHPD

N/A :190425 N/A :190425 Min. : 100 Min. :1.000 Min. :0.000 Min. : 0.00

Business : 972 Psychology : 284 1st Qu.:1500 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:12.00

Social Sciences : 853 Economics : 260 Median :3201 Median :1.000 Median :1.000 Median :13.00

Education Administration and Teaching: 611 Political Science and Government : 243 Mean :2713 Mean :1.148 Mean :1.266 Mean :14.95

Fine Arts : 465 Business Management and Administration : 217 3rd Qu.:3902 3rd Qu.:1.000 3rd Qu.:2.000 3rd Qu.:22.00

Communications : 352 French, German, Latin and Other Common Foreign Language Studies: 205 Max. :4114 Max. :5.000 Max. :2.000 Max. :22.00

(Other) : 2907 (Other) : 4951

MORTGAGE OWNCOST RENT COSTELEC COSTGAS COSTWATR COSTFUEL HHINCOME FOODSTMP LINGISOL ROOMS

Min. :0.000 Min. : 0 Min. : 0 Min. : 0 Min. : 0 Min. : 0 Min. : 0 Min. : -11800 Min. :1.000 Min. :0.000 Min. : 0.000

1st Qu.:0.000 1st Qu.: 1208 1st Qu.: 0 1st Qu.: 960 1st Qu.: 840 1st Qu.: 320 1st Qu.:9993 1st Qu.: 41600 1st Qu.:1.000 1st Qu.:1.000 1st Qu.: 4.000

Median :1.000 Median : 2891 Median : 0 Median :1560 Median :2400 Median :1400 Median :9993 Median : 81700 Median :1.000 Median :1.000 Median : 6.000

Mean :1.453 Mean :38582 Mean : 393 Mean :2311 Mean :5032 Mean :4836 Mean :7935 Mean : 114902 Mean :1.147 Mean :1.002 Mean : 5.887

3rd Qu.:3.000 3rd Qu.:99999 3rd Qu.: 630 3rd Qu.:2520 3rd Qu.:9993 3rd Qu.:9993 3rd Qu.:9993 3rd Qu.: 140900 3rd Qu.:1.000 3rd Qu.:1.000 3rd Qu.: 8.000

Max. :4.000 Max. :99999 Max. :3800 Max. :9997 Max. :9997 Max. :9997 Max. :9997 Max. :2030000 Max. :2.000 Max. :2.000 Max. :16.000

NA's :10630

BUILTYR2 UNITSSTR FUELHEAT SSMC FAMSIZE NCHILD NCHLT5 RELATE RELATED MARST RACE

Min. : 0.000 Min. : 0.00 Min. :0.000 Min. :0.00000 Min. : 1.000 Min. :0.0000 Min. :0.00000 Min. : 1.000 Min. : 101.0 Min. :1.000 Min. :1.00

1st Qu.: 1.000 1st Qu.: 3.00 1st Qu.:2.000 1st Qu.:0.00000 1st Qu.: 2.000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.: 1.000 1st Qu.: 101.0 1st Qu.:1.000 1st Qu.:1.00

Median : 3.000 Median : 3.00 Median :2.000 Median :0.00000 Median : 3.000 Median :0.0000 Median :0.00000 Median : 2.000 Median : 201.0 Median :5.000 Median :1.00

Mean : 3.711 Mean : 4.39 Mean :2.959 Mean :0.01102 Mean : 3.087 Mean :0.5009 Mean :0.08441 Mean : 3.307 Mean : 335.6 Mean :3.742 Mean :2.03

3rd Qu.: 5.000 3rd Qu.: 6.00 3rd Qu.:4.000 3rd Qu.:0.00000 3rd Qu.: 4.000 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.: 3.000 3rd Qu.: 301.0 3rd Qu.:6.000 3rd Qu.:2.00

Max. :22.000 Max. :10.00 Max. :9.000 Max. :2.00000 Max. :19.000 Max. :9.0000 Max. :5.00000 Max. :13.000 Max. :1301.0 Max. :6.000 Max. :9.00

RACED HISPAN HISPAND BPL BPLD ANCESTR1 ANCESTR1D

Min. :100 Min. :0.0000 Min. : 0.00 New York :128517 New York :128517 Not Reported :32021 Not Reported :32021

1st Qu.:100 1st Qu.:0.0000 1st Qu.: 0.00 West Indies : 8481 China : 4116 Italian :20577 Italian (1990-2000, ACS, PRCS) :20577

Median :100 Median :0.0000 Median : 0.00 China : 4964 Dominican Republic: 3517 Irish, various subheads,:16388 Irish :15651

Mean :205 Mean :0.4153 Mean : 44.75 SOUTH AMERICA: 4957 Pennsylvania : 3303 German :12781 German (1990-2000, ACS/PRCS) :12605

3rd Qu.:200 3rd Qu.:0.0000 3rd Qu.: 0.00 India : 3476 New Jersey : 3127 African-American : 9559 African-American (1990-2000, ACS, PRCS): 9559

Max. :990 Max. :4.0000 Max. :498.00 Pennsylvania : 3303 Puerto Rico : 2272 United States : 8209 United States : 8209

(Other) : 42887 (Other) : 51733 (Other) :97050 (Other) :97963

```
      ANCESTR2                    ANCESTR2D      CITIZEN       YRSUSA1
HCOVANY       HCOVPRIV      SEX          EMPSTAT       EMPSTATD

 Not Reported:141487   Not Reported             :141487   Min.   :0.0000   Min.   : 0.000   Min.
:1.000   Min.   :1.000   Male  : 95222   Min.   :0.000   Min.   : 0.00

 German      : 9476   German (1990-2000, ACS, PRCS) :  9441   1st Qu.:0.0000   1st Qu.: 0.000
1st Qu.:2.000   1st Qu.:1.000   Female:101363   1st Qu.:1.000   1st Qu.:10.00

 Irish       : 9238   Irish                    : 8809   Median :0.0000   Median : 0.000   Median :2.000
Median :2.000             Median :1.000   Median :10.00

 English     : 4895   English                  : 4895   Mean   :0.4793   Mean   : 5.377   Mean
:1.951   Mean   :1.691             Mean   :1.514   Mean   :15.16

 Italian     : 4531   Italian (1990-2000, ACS, PRCS):  4531   3rd Qu.:0.0000   3rd Qu.: 0.000
3rd Qu.:2.000   3rd Qu.:2.000             3rd Qu.:3.000   3rd Qu.:30.00

 Polish      : 3113   Polish                   : 3113   Max.   :3.0000   Max.   :92.000   Max.   :2.000
Max.   :2.000             Max.   :3.000   Max.   :30.00

 (Other)     : 23845   (Other)                  : 24309

   LABFORCE          OCC            IND          CLASSWKR      CLASSWKRD
WKSWORK2       UHRSWORK       INCTOT         FTOTINC         INCWAGE
POVERTY

 Min.   :0.000   0      : 79987   0      :79987   Min.   :0.000   Min.   : 0.00   Min.   :0.000   Min.   :
0.00   Min.   : -7300   Min.   : -11800   Min.   :     0   Min.   :  0.0

 1st Qu.:1.000   2310   :  3494   7860   : 9025   1st Qu.:0.000   1st Qu.: 0.00   1st Qu.:0.000   1st
Qu.: 0.00   1st Qu.:  8000   1st Qu.:  35550   1st Qu.:     0   1st Qu.:159.0

 Median :2.000   5700   :  3235   8680   : 6354   Median :2.000   Median :22.00   Median :1.000
Median :12.00   Median :  25000   Median :  74000   Median :  10000   Median :351.0

 Mean   :1.331   430    :  3025   770    : 6279   Mean   :1.116   Mean   :13.03   Mean   :2.701
Mean   :19.77   Mean   :  45245   Mean   : 107111   Mean   :  33796   Mean   :318.7

 3rd Qu.:2.000   4720   :  2666   8190   : 5873   3rd Qu.:2.000   3rd Qu.:22.00   3rd Qu.:6.000
3rd Qu.:40.00   3rd Qu.:  56500   3rd Qu.: 132438   3rd Qu.:  47000   3rd Qu.:501.0

 Max.   :2.000   4760   :  2563   7870   : 4041   Max.   :2.000   Max.   :29.00   Max.   :6.000
Max.   :99.00   Max.   :1563000   Max.   :2030000   Max.   :638000   Max.   :501.0

                 (Other):101615   (Other):85026                                                  NA's
:31129   NA's   :10817   NA's   :33427

   MIGRATE1       MIGRATE1D      MIGPLAC1       MIGCOUNTY1      MIGPUMA1
VETSTAT        VETSTATD       PWPUMA00       TRANWORK       TRANTIME
```

Min. :0.000 Min. : 0.00 Min. : 0.000 Min. : 0.000 Min. : 0 Min. :0.0000 Min. : 0.000 Min. : 0 Min. : 0.000 Min. : 0.00

1st Qu.:1.000 1st Qu.:10.00 1st Qu.: 0.000 1st Qu.: 0.000 1st Qu.: 0 1st Qu.:1.0000 1st Qu.:11.000 1st Qu.: 0 1st Qu.: 0.000 1st Qu.: 0.00

Median :1.000 Median :10.00 Median : 0.000 Median : 0.000 Median : 0 Median :1.0000 Median :11.000 Median : 0 Median : 0.000 Median : 0.00

Mean :1.122 Mean :11.51 Mean : 6.184 Mean : 4.117 Mean : 277 Mean :0.8621 Mean : 9.412 Mean : 1255 Mean : 9.725 Mean : 14.75

3rd Qu.:1.000 3rd Qu.:10.00 3rd Qu.: 0.000 3rd Qu.: 0.000 3rd Qu.: 0 3rd Qu.:1.0000 3rd Qu.:11.000 3rd Qu.: 3100 3rd Qu.:10.000 3rd Qu.: 20.00

Max. :4.000 Max. :40.00 Max. :900.000 Max. :810.000 Max. :70100 Max. :2.0000 Max. :20.000 Max. :59300 Max. :70.000 Max. :138.00

DEPARTS in_NYC in_Bronx in_Manhattan in_StatenI in_Brooklyn in_Queens in_Westchester in_Nassau Hispanic

Min. : 0.0 Min. :0.0000 Min. :0.0000 Min. :0.00000 Min. :0.00000 Min. :0.000 Min. :0.0000 Min. :0.00000 Min. :0.00000 Min. :0.0000

1st Qu.: 0.0 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.0000

Median : 0.0 Median :0.0000 Median :0.0000 Median :0.00000 Median :0.00000 Median :0.000 Median :0.0000 Median :0.00000 Median :0.00000 Median :0.0000

Mean : 373.3 Mean :0.3615 Mean :0.0538 Mean :0.04981 Mean :0.02084 Mean :0.126 Mean :0.1111 Mean :0.04413 Mean :0.07032 Mean :0.1387

3rd Qu.: 732.0 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.000 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.0000

Max. :2345.0 Max. :1.0000 Max. :1.0000 Max. :1.00000 Max. :1.00000 Max. :1.000 Max. :1.0000 Max. :1.00000 Max. :1.00000 Max. :1.0000

Hisp_Mex Hisp_PR Hisp_Cuban Hisp_DomR white AfAm Amindian Asian race_oth unmarried

Min. :0.00000 Min. :0.0000 Min. :0.000000 Min. :0.00000 Min. :0.0000 Min. :0.000 Min. :0.000000 Min. :0.00000 Min. :0.0000 Min. :0.00

1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.000000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.000000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.00

Median :0.00000 Median :0.0000 Median :0.000000 Median :0.00000 Median :1.0000 Median :0.000 Median :0.000000 Median :0.00000 Median :0.0000 Median :0.00

Mean :0.01626 Mean :0.0436 Mean :0.003403 Mean :0.02827 Mean :0.6997 Mean :0.125 Mean :0.003779 Mean :0.08656 Mean :0.1324 Mean :0.45

3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.000000 3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:0.000 3rd Qu.:0.000000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:1.00

Max. :1.00000 Max. :1.0000 Max. :1.000000 Max. :1.00000 Max. :1.0000 Max. :1.000 Max. :1.000000 Max. :1.00000 Max. :1.0000 Max. :1.00

veteran has_AnyHealthIns has_PvtHealthIns Commute_car Commute_bus Commute_subway Commute_rail Commute_other below_povertyline below_150poverty

Min. :0.00000 Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.000 Min. :0.0000

1st Qu.:0.00000 1st Qu.:1.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.000 1st Qu.:0.0000

Median :0.00000 Median :1.0000 Median :1.0000 Median :0.0000 Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000 Median :0.000 Median :0.0000

Mean :0.04443 Mean :0.9513 Mean :0.6906 Mean :0.2997 Mean :0.02162 Mean :0.07468 Mean :0.01332 Mean :0.05506 Mean :0.122 Mean :0.1965

3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.000 3rd Qu.:0.0000

Max. :1.00000 Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.00000 Max. :1.000 Max. :1.0000

below_200poverty foodstamps

Min. :0.0000 Min. :0.0000

1st Qu.:0.0000 1st Qu.:0.0000

Median :0.0000 Median :0.0000

Mean :0.2676 Mean :0.1465

3rd Qu.:1.0000 3rd Qu.:0.0000

Max. :1.0000 Max. :1.0000 "

```
##############################################################################
#######
#INCOME WAGE OF Asian College Graduates Between The Ages Of 30-55
##############################################################################
#######

age_wage <- lm(INCWAGE ~ AGE + educ_college + Asian)
summary(age_wage)
'Call:
lm(formula = INCWAGE ~ AGE + educ_college + Asian)

Residuals:
   Min     1Q Median     3Q    Max
-57820 -28592 -23183  11554 614963

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  36340.46     453.22  80.183   <2e-16 ***
AGE           -146.19       8.37 -17.467   <2e-16 ***
educ_college 23907.06     415.09  57.595   <2e-16 ***
Asian          350.47     577.07   0.607    0.544
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 65420 on 163154 degrees of freedom
  (33427 observations deleted due to missingness)
Multiple R-squared:  0.02259,           Adjusted R-squared:  0.02257
F-statistic:  1257 on 3 and 163154 DF,  p-value: < 2.2e-16
'
```

```
plot(age_wage)
#Hit <Return> to see next plot: #Hit <Return> to see next plot:
#Hit <Return> to see next plot:#maybe get fancy
#maybe get fancy
require(stargazer)
#Hit <Return> to see next plot: require(stargazer)
#Hit <Return> to see next plot:
stargazer(age_wage, type = "text")

'===============================================
                  Dependent variable:
              ---------------------------
                       INCWAGE
-----------------------------------------------
AGE                    -146.195***
                       (8.370)

educ_college           23,907.060***
                       (415.090)

Asian                  350.467
                       (577.068)

Constant               36,340.460***
                       (453.221)
```

```
-----------------------------------------------
Observations                163,158
R2                          0.023
Adjusted R2                 0.023
Residual Std. Error   65,418.710 (df = 163154)
F Statistic         1,257.048*** (df = 3; 163154)
===============================================
Note:                *p<0.1; **p<0.05; ***p<0.01'

# subset in order to plot...
NNobs <- length(INCWAGE)
set.seed(12345) # just so you can replicate and get same "random" choices
graph_obs <- (runif(NNobs) < 0.1) # so something like just 1/10 as many obs
dat_graph <-subset(dat_use,graph_obs)

plot(INCWAGE ~ jitter(AGE, factor = 2), pch = 16, col = rgb(0.5, 0.5, 0.5, alpha = 0.2), data = dat_graph)
# ^^ that looks like crap since Wages are sooooooooo skew!  So try to find some sensible ylim = c(0, ??)
plot(INCWAGE ~ jitter(AGE, factor = 2), pch = 16, col = rgb(0.5, 0.5, 0.5, alpha = 0.2), ylim = c(0,150000), data = dat_graph)
# discus what you see in this plot

# change this line to fit your regression
to_be_predicted2 <- data.frame(AGE = 30:55, educ_college = 1, Asian = 0)
to_be_predicted2$yhat <- predict(age_wage, newdata = to_be_predicted2)

detach()
```

```
##############################################################################
#######
#INCOME WAGE OF Hispanci College Graduates Between The Ages Of 30-55
##############################################################################
#######

age_wage <- lm(INCWAGE ~ AGE + educ_college + Hispanic)
summary(age_wage)
'Call:
lm(formula = INCWAGE ~ AGE + educ_college + Hispanic)

Residuals:
   Min     1Q Median     3Q    Max
-58944 -28818 -21035  11678 623147

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   38848.981    461.679   84.15   <2e-16 ***
AGE            -167.272      8.398  -19.92   <2e-16 ***
educ_college  23273.352    415.146   56.06   <2e-16 ***
Hispanic     -10279.419    485.262  -21.18   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 65330 on 163154 degrees of freedom
  (33427 observations deleted due to missingness)
Multiple R-squared:  0.02527,           Adjusted R-squared:  0.02525
F-statistic:  1410 on 3 and 163154 DF,  p-value: < 2.2e-16'
```

```
plot(age_wage)
#Hit <Return> to see next plot:
# maybe get fancy
require(stargazer)
# <Return> to see next plot: require(stargazer)
#Hit <Return> to see next plot:
stargazer(age_wage, type = "text")

'================================================
                    Dependent variable:
                ----------------------------
                          INCWAGE
------------------------------------------------
AGE                       -167.272***
                          (8.398)

educ_college              23,273.350***
                          (415.146)

Hispanic                  -10,279.420***
                          (485.262)

Constant                  38,848.980***
                          (461.679)

------------------------------------------------
Observations                  163,158
```

```
R2                        0.025
Adjusted R2               0.025
Residual Std. Error   65,329.000 (df = 163154)
F Statistic        1,409.956*** (df = 3; 163154)
================================================
Note:              *p<0.1; **p<0.05; ***p<0.01'
```

```
# subset in order to plot...
NNobs <- length(INCWAGE)
set.seed(12345) # just so you can replicate and get same "random" choices
graph_obs <- (runif(NNobs) < 0.1) # so something like just 1/10 as many obs
dat_graph <-subset(dat_use,graph_obs)

plot(INCWAGE ~ jitter(AGE, factor = 2), pch = 16, col = rgb(0.5, 0.5, 0.5, alpha = 0.2), data = dat_graph)
# ^^ that looks like crap since Wages are sooooooooo skew!  So try to find some sensible ylim = c(0, ??)
plot(INCWAGE ~ jitter(AGE, factor = 2), pch = 16, col = rgb(0.5, 0.5, 0.5, alpha = 0.2), ylim = c(0,150000), data = dat_graph)
# discus what you see in this plot

# change this line to fit your regression
to_be_predicted2 <- data.frame(AGE = 30:55, educ_college = 1, Hispanic = 1)
to_be_predicted2$yhat <- predict(age_wage, newdata = to_be_predicted2)

detach()
```