Name: Anusha Mittal
EID: am98969
HOMEWORK 1

CODE:

```
{smcl}
{com}{sf}{ul off}{txt}{.-}
      name:  {res}<unnamed>
       {txt}log:  {res}\\cnsdisk.austin.utexas.edu\home\am98968\Desktop\hw1.smcl
  {txt}log type:  {res}smcl
 {txt}opened on:  {res}31 Aug 2023, 23:58:01

{com}. describe

{txt}Contains data from {res}\\cnsdisk.austin.utexas.edu\home\am98968\Desktop\metrics stata\BWGHT.DTA
{txt}  obs:{res}        1,388
{txt}  vars:{res}           14                          3 Jun 1997 13:47
{txt}  size:{res}       49,968
{txt}{hline}
               storage   display    value
variable name   type    format     label      variable label
{hline}
{p 0 48}{res}{bind:faminc         }{txt}{bind: float   }{bind:{txt}%9.0g     }{space 1}{bind:        }{bind:   }{res}{res}1988 family income, $1000s{p_end}
{p 0 48}{bind:cigtax        }{txt}{bind: float   }{bind:{txt}%9.0g     }{space 1}{bind:        }{bind:   }{res}{res}cig. tax in home state, 1988{p_end}
{p 0 48}{bind:cigprice      }{txt}{bind: float   }{bind:{txt}%9.0g     }{space 1}{bind:        }{bind:   }{res}{res}cig. price in home state, 1988{p_end}
{p 0 48}{bind:bwght         }{txt}{bind: int     }{bind:{txt}%8.0g     }{space 1}{bind:        }{bind:   }{res}{res}birth weight, ounces{p_end}
{p 0 48}{bind:fatheduc      }{txt}{bind: byte    }{bind:{txt}%8.0g     }{space 1}{bind:        }{bind:   }{res}{res}father's yrs of educ{p_end}
{p 0 48}{bind:motheduc      }{txt}{bind: byte    }{bind:{txt}%8.0g     }{space 1}{bind:        }{bind:   }{res}{res}mother's yrs of educ{p_end}
{p 0 48}{bind:parity        }{txt}{bind: byte    }{bind:{txt}%8.0g     }{space 1}{bind:        }{bind:   }{res}{res}birth order of child{p_end}
{p 0 48}{bind:male          }{txt}{bind: byte    }{bind:{txt}%8.0g     }{space 1}{bind:        }{bind:   }{res}{res}=1 if male child{p_end}
{p 0 48}{bind:white         }{txt}{bind: byte    }{bind:{txt}%8.0g     }{space 1}{bind:        }{bind:   }{res}{res}=1 if white{p_end}
{p 0 48}{bind:cigs          }{txt}{bind: byte    }{bind:{txt}%8.0g     }{space 1}{bind:        }{bind:   }{res}{res}cigs smked per day while preg{p_end}
{p 0 48}{bind:lbwght        }{txt}{bind: float   }{bind:{txt}%9.0g     }{space 1}{bind:        }{bind:   }{res}{res}log of bwght{p_end}
{p 0 48}{bind:bwghtlbs      }{txt}{bind: float   }{bind:{txt}%9.0g     }{space 1}{bind:        }{bind:   }{res}{res}birth weight, pounds{p_end}
{p 0 48}{bind:packs         }{txt}{bind: float   }{bind:{txt}%9.0g     }{space 1}{bind:        }{bind:   }{res}{res}packs smked per day while preg{p_end}
{p 0 48}{bind:lfaminc       }{txt}{bind: float   }{bind:{txt}%9.0g     }{space 1}{bind:        }{bind:   }{res}{res}log(faminc){p_end}
{txt}{hline}
Sorted by:

{com}. summarize

{txt}    Variable {c |}        Obs        Mean    Std. Dev.       Min        Max
{hline 13}{c +}{hline 57}
{space 6}faminc {c |}{res}       1,388    29.02666    18.73928         .5         65
{txt}{space 6}cigtax {c |}{res}       1,388    19.55295    7.795598          2         38
{txt}{space 4}cigprice {c |}{res}       1,388     130.559    10.24448      103.8      152.5
{txt}{space 7}bwght {c |}{res}       1,388    118.6996    20.35396         23        271
{txt}{space 4}fatheduc {c |}{res}       1,192    13.18624    2.745985          1         18
{txt}{hline 13}{c +}{hline 57}
{space 4}motheduc {c |}{res}       1,387    12.93583    2.376728          2         18
{txt}{space 6}parity {c |}{res}       1,388    1.632565    .8940273          1          6
{txt}{space 8}male {c |}{res}       1,388    .5208934    .4997433          0          1
{txt}{space 7}white {c |}{res}       1,388    .7845821    .4112601          0          1
{txt}{space 8}cigs {c |}{res}       1,388    2.087176    5.972688          0         50
{txt}{hline 13}{c +}{hline 57}
{space 6}lbwght {c |}{res}       1,388    4.760031    .1906622    3.135494    5.602119
{txt}{space 4}bwghtlbs {c |}{res}       1,388    7.418723    1.272123      1.4375     16.9375
{txt}{space 7}packs {c |}{res}       1,388    .1043588    .2986344          0        2.5
{txt}{space 5}lfaminc {c |}{res}       1,388    3.071271    .9180645   -.6931472   4.174387

{com}. save "\\cnsdisk.austin.utexas.edu\home\am98968\Desktop\metrics stata\BWGHT.DTA", replace
{txt}file \\cnsdisk.austin.utexas.edu\home\am98968\Desktop\metrics stata\BWGHT.DTA saved

{com}. gen smoker = (cigs > 0 )

. gen smoker = 0
{err}variable {bf}smoker{sf} already defined
{txt}{search r(110), local:r(110);}

{com}. drop smoker

. gen smoker = 0

. replace smoker = 1 if cigs>0
{txt}(212 real changes made)

{com}. tab smoker bwght
{err}too many values
{txt}{search r(134), local:r(134);}

{com}. mean bwt if smoker == 1
{err}variable {bf}bwt{sf} not found
{txt}{search r(111), local:r(111);}

{com}. mean bwght if smoker == 1
{res}
{txt}Mean estimation{col 35}Number of obs{col 51}= {res}         212

{txt}{hline 13}{c TT}{hline 11}{hline 11}{hline 14}{hline 12}
{col 14}{c |}        Mean{col 26}   Std. Err.{col 38}     [95% Con{col 51}f. Interval]
{hline 13}{c +}{hline 11}{hline 11}{hline 14}{hline 12}
{space 7}bwght {c |}{col 14}{res}{space 2} 111.1462{col 26}{space 2} 1.317384{col 37}{space 5} 108.5493{col 51}{space 3} 113.7431
{txt}{hline 13}{c BT}{hline 11}{hline 11}{hline 14}{hline 12}

{com}. mean bwght if smoker == 0
{res}
{txt}Mean estimation{col 35}Number of obs{col 51}= {res}       1,176

{txt}{hline 13}{c TT}{hline 11}{hline 11}{hline 14}{hline 12}
{col 14}{c |}        Mean{col 26}   Std. Err.{col 38}     [95% Con{col 51}f. Interval]
{hline 13}{c +}{hline 11}{hline 11}{hline 14}{hline 12}
{space 7}bwght {c |}{col 14}{res}{space 2} 120.0612{col 26}{space 2} .5910411{col 37}{space 5} 118.9016{col 51}{space 3} 121.2208
{txt}{hline 13}{c BT}{hline 11}{hline 11}{hline 14}{hline 12}

{com}. 120.0612  - 111.1462
{bf}{err}120.0612{sf} is not a valid command name
{txt}{search r(199), local:r(199);}

{com}. gen diff_bwt = 120.0612  - 111.1462

. diff_bwt
{err}command {bf}diff_bwt{sf} is unrecognized
{txt}{search r(199), local:r(199);}

{com}. display 120.0612  - 111.1462
{res}8.915

{com}. display " the difference in average birthweight between smoking mothers and non- smoking mothers is 8.915"
{res} the difference in average birthweight between smoking mothers and non- smoking mothers is 8.915

{com}. correlate bwght faminc motheduc cigs
{txt}(obs=1,387)

             {c |}     bwght   faminc motheduc      cigs
{hline 13}{c +}{hline 36}
       bwght {c |}{res}   1.0000
    {txt}faminc {c |}{res}   0.1085    1.0000
  {txt}motheduc {c |}{res}   0.0691    0.4559    1.0000
      {txt}cigs {c |}{res}  -0.1509   -0.1734   -0.2139    1.0000


{com}. display "the highest correlation in magintude is shown by number of cigarettes smoked per day. The signs also make sense, since the more the mother smokes the less healthy the babe is likely to be hence, the lesser birth weight. Moreover, with greater income and education a mother is likely able to give the best resources for child nourishment, so the child is healthy and of a higher weight"
{res}the highest correlation in magnitude is shown by number of cigarettes smoked per day. The signs also make sense, since the more the mother smokes the less healthy the babe is likely to be hence, the lesser birth weight. Moreover, with greater income and education a mother is likely able to give the best resources for child nourishment, so the child is healthy and of a higher weight

{com}. save "\\cnsdisk.austin.utexas.edu\home\am98968\Desktop\metrics stata\BWGHT.DTA", replace
{txt}file \\cnsdisk.austin.utexas.edu\home\am98968\Desktop\metrics stata\BWGHT.DTA saved

{com}. save "\\cnsdisk.austin.utexas.edu\home\am98968\Desktop\metrics stata\BWGHT.DTA", replace
{txt}file \\cnsdisk.austin.utexas.edu\home\am98968\Desktop\metrics stata\BWGHT.DTA saved

{com}. log close
      {txt}name:  {res}<unnamed>
       {txt}log:  {res}\\cnsdisk.austin.utexas.edu\home\am98968\Desktop\hw1.smcl
  {txt}log type:  {res}smcl
 {txt}closed on:  {res} 1 Sep 2023, 20:51:13
{txt}{.-}
{smcl}
{txt}{sf}{ul off}
```

(viii) If birthweight is measured in pounds rather than ounces (bwghtlbs in the data), how would your answer from (vi) change? how would the correlation values in (vii) change? how would the covariances of the various variables in (vii) with bwght change? (You should be able to answer these without doing it in Stata (though you might want to check your answers with Stata))

- My answer to (vi) would change where the answer would be multiplied by 0.0625, based on that the covariance would also change in the dataset. The correlation however would remain the same.

WRITTEN ANSWERS:

A justification for job training programs is that they improve worker productivity. Suppose that you are asked to evaluate whether more job training makes workers more productive. However, rather than having data on individual workers, you have access to data on manufacturing firms in Ohio. In particular, for each firm, you have information on hours of job training per worker (training) and number of nondefective items produced per worker hour (output).

(i) Carefully state the ceteris paribus thought experiment underlying this policy question.
Every worker recieves the training equally effectively. So we're trying to evaluate the effect of training on output keeping all other factors constant and not affectng the output.

(ii) Does it seem likely that a firm's decision to train its workers will be independent of worker characteristics? What are some of those measurable and unmeasurable worker characteristics?
No I think the firms will definetly care about how much effort workers are putting in and also how much the training will really help the output. I think measureable characteristics would be their performance and numbers which would be the output along with their previous job experience. The unmeasureable characteristics would be how they recieve the training and if they're paying attention. Also how much they are really interested in the company to see if the training will be an investment to the company or will they just leave the company eventually.

(iii) Name a factor other than worker characteristics that can affect worker productivity.
The machines, external management and administration, the schedules and overall manufacturing plant productivity.

(iv) If you find a positive correlation between output and training, would you have convincingly established that job training makes workers more productive? Explain.

I could've definietly made a good argument that training does help with output but of course there are confounding factors to this since correlation does not necessarily mean causation. Factors such as 1) the workers are more productive becuase they see that the company is investing time and money in them so they feel valued and secure in their job so they put in more effort 2) It could also be that the training brings co-workers together and encourgaes for a better work envionrment which would make the output more productive.

CHAPTER 2, Question 4

The data set BWGHT contains data on births to women in the United States. Two variables of interest are the dependent variable, infant birth weight in ounces (bwght), and an explanatory variable, average number of cigarettes the mother smoked per day during pregnancy (cigs). The following simple regression was estimated using data on n 5 1,388 births:
bwght = 119.77 − 0.514 cigs

(i) What is the predicted birth weight when cigs = 0? What about when cigs = 20 (one pack per day)? Comment on the difference.

When cigs=0 the bwght = 119.77 which is the intercept. When cigs = 20 it will be 109.49. The difference shows that as more number of cigarettes are consumed by the mother, the bright weight does indeed descrease.

(ii) Does this simple regression necessarily capture a causal relationship between the child's birth weight and the mother's smoking habits? Explain.
I would say it could, since there is negative relationship but it is also good to keep in mind that we are probably holding other external factors constant which could be the reason for this relationship.

(iii) To predict a birth weight of 125 ounces, what would cigs have to be? Comment.
125 = 119.77 − 0.514c, c = − 10.18, which doesn't make logical sense since consuming negative amounts of cigarettes isn't possible in the real world.

(iv) The proportion of women in the sample who do not smoke while pregnant is about .85. Does this help reconcile your finding from part (iii)?

Around 85% of the women in our sample do not smoke while pregnant, and this fact plays a role in how we interpret the model's predictions. Given the high proportion of non-smokers in our dataset, the model might not accurately predict outcomes for heavy smokers who are less represented. So, when considering the model's predictions for extreme values like a birth weight of 125 ounces, we can see that the model is not necessarily completely reliable.