```
-------------------------------------------------------------------------------
      name:  PK
      log:  /Users/priyakoirala/Desktop/school/econometrics
/koirala_project1.log
      log type:  text
      opened on:  27 Feb 2023, 19:16:40

.
/*=================================================================
> (Q1): Open the data set in Stata (via the use command).
=================================================================*/

. use "/Users/priyakoirala/Desktop/school/econometrics/bwght.dta"
```

```
/*==============================================================================
> (Q2): Summarize the data in Stata. What is the average birthweight of
infants in the sample?
==============================================================================*/

. summarize

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
       bwght |      1,180    119.0551    20.19077         23        271
        cigs |      1,180    2.208475    6.222801          0         50


The average birthweight of infants in the sample is 119.0551.
```

```
/*==========================================================================
> (Q3): Construct a binary variable (using the "gen" command) named
anycig that equals one if the mother smoked at least one cigarette per
day and equals zero otherwise.
==========================================================================*/

. gen anycig = .
(1,180 missing values generated)

. replace anycig = 1 if cig >= 1
(187 real changes made)

. replace anycig = 0 if cig < 1
(993 real changes made)
```

```
/*==============================================================================
> (Q4): Use the "tab" command to determine the share (percent) of moms
who smoked during pregnancy.
==============================================================================*/

. tab anycig

     anycig |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        993       84.15       84.15
          1 |        187       15.85      100.00
------------+-----------------------------------
      Total |      1,180      100.00
```

15.85% of mothers smoked during pregnancy.

```
/*==================================================================
> (Q5): We want to test the hypothesis that the average birthweight of
infants born to smokers vs. non-smokers is the same. Use the "reg"
command to estimate the following model using OLS: bwghti = β0 +
β1anycigi + ui
==================================================================*/

. reg bwght anycig

Source   |       SS           df       MS        Number of obs  =   1,180
---------+----------------------------------    F(1, 1178)     =   31.54
Model    |  12531.7427         1  12531.7427     Prob > F       =  0.0000
Residual |  468107.677     1,178  397.374938     R-squared      =  0.0261
---------+----------------------------------    Adj R-squared  =  0.0252
Total    |  480639.419     1,179  407.667022     Root MSE       =  19.934

------------------------------------------------------------------------------
bwght    | Coefficient  Std. err.      t    P>|t|    [95% conf.interval]
---------+--------------------------------------------------------------------
anycig   |   -8.92383    1.58908    -5.62   0.000   -12.04157  -5.806087
_cons    |   120.4693   .6325948   190.44   0.000    119.2281   121.7104
------------------------------------------------------------------------------
```

/*==
> (Q6): How do you interpret β1 in this regression?
==*/

β1 is the difference in the average birthweights of infants born to mothers who smoked and the mothers who did not smoke during pregnancy.

β1 is -8.92383.

/*==========
> (Q7): How do you interpret β0?
==========*/

β0 represents the average (constant) birthweight of infants, whether they were born to mothers who smoked during pregnancy or mothers who did not.

β0 is 120.4693

```
/*==========================================================================
> (Q8): Rather than looking at the effect of smoking versus not smoking,
we want to determine how smoking intensity affects birthweight. Use the
"reg" command to estimate: bwghti = β0+ β1cigsi + ui
==========================================================================*/

. reg bwght cigs

      Source |       SS           df       MS      Number of obs   =
1,180
-------------+----------------------------------   F(1, 1178)      =
29.80
       Model |   11857.097         1   11857.097   Prob > F        =
0.0000
    Residual |  468782.322     1,178  397.947642   R-squared       =
0.0247
-------------+----------------------------------   Adj R-squared   =
0.0238
       Total |  480639.419     1,179  407.667022   Root MSE        =
19.949

------------------------------------------------------------------------------
       bwght | Coefficient  Std. err.      t    P>|t|     [95% conf.
interval]
-------------+----------------------------------------------------------------
        cigs |  -.5096199    .093362    -5.46   0.000    -.6927942   -
.3264455
       _cons |   120.1806   .6162443   195.02   0.000     118.9715
121.3896
------------------------------------------------------------------------------
```

```
/*=============================================================================
> (Q9): What is the marginal effect of smoking an additional cigarette on
> birthweight?
=============================================================================*/
```

For every additional cigarette smoked by pregnant mothers, an infant's birthweight is estimated to decrease by 0.5096199.

/*==========
> (Q10): Using your regression results in question (8). If a mother smokes 5 cigarettes per day, what is the effect on birthweight?
==========*/

Based on the model bwghti = β0 + β1cigsi + ui, if we were to assign cig to equal to 5 and β1 is approximated to -0.509619, then that gives us the algebraic product of -2.548095. Therefore at 5 cigarettes a day the effect on birthweight is that it decreases by 2.548095.

```
/*==================================================================
> (Q11): Does the regression model in question (8) reflect the causal
> relationship between a mother's smoking habits and her infant's
birthweight? Explain. (Hint: What is contained in ui?)
==================================================================*/

The regression model in question 8 reflects the causal negative
relationship between a mother's smoking habits and an infant's
birthweight. The regression suggests that mother's who smoked during
pregnancy gave birth to infants with a lower birthweight than mother's
who did not smoke during pregnancy.

We reject our hypothesis that average birthweight of infants born to
mothers who smoked vs non-smokers is the same.

However, despite our findings that there is a causal relationship between
smoking and low infant birthweight, the model fails to reflect all of the
other factors that might contribute to low birthweight. Such as the
environment the mother lives in, household income, diet, exercise.
All of these factors plus more can impact an infants birthweight.
(ui is the error term, it contains all other factors besides smoking
which impacts and infant's birthweight).*/

/*==================================================================
==================================================================*/
.
. cap log close _all
```