Application 1

Practice

1. Using data set campus.dta, write code that confirms that the average enrollment of universities in the data set is 16,076.35
2. Using data set ACS\_major.csv, write code that shows the distribution of education by gender.

One measure of health of newborn babies is birthweight. Low birthweight is an indicator of poor health. Another measure of health is APGAR – a quick test taken at 1 and 5 minutes to measure health. The one-minute score is a 10-point score to determine how well the baby tolerated the birthing process. The 5-minute score measures how well the baby is doing outside the mother’s womb. The data set we have is a random sample of births in 1993 in Pennsylvania. It can be found in the following data set:

 birth\_data.xlsx

Step 1: Load the data set into STATA

Step 2: Sum stats

1. What is the average weight of newborn babies in grams?
2. How many individuals are in the data set?
3. What fraction of the sample of mothers smoked while pregnant?
4. Among all the respondents, what is the average number of cigarettes smoked?
5. Write code to recreate the following table:

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Obs | Mean | Std. Dev.  |
| Mother age | 13,725 | 27.58 | 5.70 |
| Mother years of education | 13,725 | 13.10 | 2.24 |
| Father age | 13,725 | 29.91 | 6.41 |
| Father years of education | 13,725 | 13.15 | 2.28 |

1. What is the average one-minute APGAR score?
2. What fraction of the sample had a one-minute APGAR score of 7 or less?
3. The national average of birthweight in grams is 3371. Is your sample different from the national average? Explain.
4. Is the birthweight different between boy and girl babies? Explain.
5. It is argued that one-minute APGAR scores are no different from five-minute APGAR scores. What can you say about this argument? Explain.