Notes:
      1. You are running Small Stata.
      2. Unicode is supported; see help unicode_advice.

```
 1 . do "/Users/dailondolojan/Desktop/Econ 113/H1 Q6.do"

 2 . use "/Users/dailondolojan/desktop/stata_data/smoking_data.dta", clear

 3 .
 4 . *Question 6
 5 .
 6 . *Part A
 7 . * The total number of people included in the data is:
 8 . count
   807

 9 .
10 . * The following table describes how many people are white and non-white:
11 . tab white
```

=1 if white	Freq.	Percent	Cum.
0	**98**	**12.14**	**12.14**
1	**709**	**87.86**	**100.00**
Total	**807**	**100.00**	

```
12 .
13 . *Part B
14 . * The following describes the minimum, maximum, and median annual income:
15 . sum income, detail

                        annual income, $
```

Percentiles		Smallest		
1%	**1500**	**500**		
5%	**3500**	**500**		
10%	**6500**	**500**	Obs	**807**
25%	**12500**	**500**	Sum of Wgt.	**807**
50%	**20000**		Mean	**19304.83**
		Largest	Std. Dev.	**9142.958**
75%	**30000**	**30000**		
90%	**30000**	**30000**	Variance	**8.36e+07**
95%	**30000**	**30000**	Skewness	**-.2386615**
99%	**30000**	**30000**	Kurtosis	**1.826593**

```
16 .
17 . *Part C
18 . * The count of how many people have obtained each level of education and
19 . * most common level of education shown below. This result does make sense
20 . * because the highest frequency is at 12 years which corresponds to graduating
21 . * high school as well as the majority of drop outs corresponding to sophomore
22 . * year of high school. Students could graduate early or drop-out of college
23 . * which corresponds to the values between 13.5 and 15. At 16 years there is a
24 . * minor spike because of individuals graduating on time in a stereotypical
25 . * college setting.
26 . tab educ
```

years of schooling	Freq.	Percent	Cum.
6	**43**	**5.33**	**5.33**
8	**54**	**6.69**	**12.02**
10	**126**	**15.61**	**27.63**
12	**259**	**32.09**	**59.73**
13.5	**128**	**15.86**	**75.59**
15	**32**	**3.97**	**79.55**
16	**86**	**10.66**	**90.21**
18	**79**	**9.79**	**100.00**
Total	**807**	**100.00**	

```
27 .
28 . *Part D
29 . * The education level of 17 year olds in the sample. This result makes sense
30 . * when you consider the fact that some 17 year olds have not completed high
31 . * school or possibly have started school later which will skew the results of
32 . * the data.
33 . tab educ if age == 17
```

years of

schooling	Freq.	Percent	Cum.
10	**17**	**100.00**	**100.00**
Total	**17**	**100.00**	

```
34 .
35 . *Part E
36 . * The mean and standard deviation of ciggarettes smoked per day.
37 . sum cigs
```

Variable	Obs	Mean	Std. Dev.	Min	Max
cigs	**807**	**8.686493**	**13.72152**	**0**	**80**

```
38 .
39 . *Part F
40 . * The mean and standard deviation of cigarettes if a person has more than 12
41 . * years of education.
42 . sum cigs if educ > 12
```

Variable	Obs	Mean	Std. Dev.	Min	Max
cigs	**325**	**7.556923**	**13.99892**	**0**	**80**

```
43 .
   end of do-file

44 . do "/Users/dailondolojan/Desktop/Econ 113/H1 Q7.do"

45 . use "/Users/dailondolojan/desktop/stata_data/smoking_data.dta", clear

46 .
47 . *Question 7
48 .
49 . *Part A
50 . * The correlation between income and cigarettes are given in the table below.
51 . * One reason this data may be true could be that a worker will seek a higher
52 . * income in order to afford a higher intake of ciggarettes.
53 . correlate income cigs
   (obs=807)
```

	income	cigs
income	**1.0000**	
cigs	**0.0532**	**1.0000**

```
54 .
```

```
55 . *Part B
56 . * The correlation between education and cigarettes are given in the table below.
57 . * One reason that supports that the data may be true could be that individuals
58 . * who smoke are more likely to drop out of school or stop their education
59 . * earlier.
60 . correlate educ cigs
(obs=807)
```

	educ	cigs
educ	**1.0000**	
cigs	**-0.0487**	**1.0000**

```
61 .
62 . *Part C
63 . * High income, low education seem to be the likely group to have the highest
64 . * rate of smoking. According to the data, there is a slighty positive
65 . * correlation between income and ciggarettes thus meaning higher income
66 . * individuals smoke more ciggarettes. Also, there is a slightly negative
67 . * correlation between education and ciggarettes thus meaning lower education
68 . * individuals smoke more ciggarettes. This can be seen by the correlation table
69 . * shown below
70 . correlate income educ cigs
(obs=807)
```

	income	educ	cigs
income	**1.0000**		
educ	**0.3344**	**1.0000**	
cigs	**0.0532**	**-0.0487**	**1.0000**

```
71 .
end of do-file

72 . do "/Users/dailondolojan/Desktop/Econ 113/H1 Q7.do"

73 . use "/Users/dailondolojan/desktop/stata_data/smoking_data.dta", clear

74 .
75 . *Question 7
76 .
77 . *Part A
78 . * The correlation between income and cigarettes are given in the table below.
79 . * One reason this data may be true could be that a worker will seek a higher
80 . * income in order to afford a higher intake of ciggarettes.
81 . correlate income cigs
(obs=807)
```

```
             |   income     cigs
-------------+----------------------
      income |   1.0000
        cigs |   0.0532   1.0000
```

```
. 
. *Part B
. * The correlation between education and cigarettes are given in the table below.
. * One reason that supports that the data may be true could be that individuals
. * who smoke are more likely to drop out of school or stop their education
. * earlier.
. correlate educ cigs
(obs=807)
```

```
             |     educ     cigs
-------------+----------------------
        educ |   1.0000
        cigs |  -0.0487   1.0000
```

```
. 
. *Part C
. * High income, low education seem to be the likely group to have the highest
. * rate of smoking. According to the data, there is a slighty positive
. * correlation between income and ciggarettes thus meaning higher income
. * individuals smoke more ciggarettes. Also, there is a slightly negative
. * correlation between education and ciggarettes thus meaning lower education
. * individuals smoke more ciggarettes. This can be seen by the correlation table
. * shown below
. correlate income educ cigs
(obs=807)
```

```
             |   income     educ     cigs
-------------+---------------------------
      income |   1.0000
        educ |   0.3344   1.0000
        cigs |   0.0532  -0.0487   1.0000
```

```
. 
end of do-file

. do "/Users/dailondolojan/Desktop/Econ 113/H1 Q7.do"

. use "/Users/dailondolojan/desktop/stata_data/smoking_data.dta", clear

. 
```

```
103 . *Question 7
104 .
105 . *Part A
106 . * The correlation between income and cigarettes are given in the table below.
107 . * One reason this data may be true could be that a worker will seek a higher
108 . * income in order to afford a higher intake of ciggarettes.
109 . correlate income cigs
(obs=807)

             |   income     cigs
-------------+----------------------
      income |   1.0000
        cigs |   0.0532   1.0000


110 .
111 . *Part B
112 . * The correlation between education and cigarettes are given in the table below.
113 . * One reason that supports that the data may be true could be that individuals
114 . * who smoke are more likely to drop out of school or stop their education
115 . * earlier.
116 . correlate educ cigs
(obs=807)

             |     educ     cigs
-------------+----------------------
        educ |   1.0000
        cigs |  -0.0487   1.0000


117 .
118 . *Part C
119 . * High income, low education seem to be the likely group to have the highest
120 . * rate of smoking. According to the data, there is a slighty positive
121 . * correlation between income and ciggarettes thus meaning higher income
122 . * individuals smoke more ciggarettes. Also, there is a slightly negative
123 . * correlation between education and ciggarettes thus meaning lower education
124 . * individuals smoke more ciggarettes. This can be seen by the correlation table
125 . * shown below
126 . correlate income educ cigs
(obs=807)

             |   income     educ     cigs
-------------+---------------------------
      income |   1.0000
        educ |   0.3344   1.0000
        cigs |   0.0532  -0.0487   1.0000
```

```
. 
end of do-file

. 
```