

More on Stata

- Stem and leaf plot

We will use a built-in Stata dataset called “auto” to illustrate. Load this dataset by typing:

‘use <http://www.stata-press.com/data/r13/auto>’ on your Stata command line.

- ❖ Browse the dataset to see the nature and variables on it by typing
 - browse

Create a stem-and-leaf plot for the variable mpg by typing the following command: *stem mpg*

Comment: The above plot is a five-line plot with *lines()* equal to 5 (five lines per interval of 10) and *width()* equal to 2 (two leaves per stem).

- We can split the stems into two with the following command: *stem mpg, lines(2)*
Remark: *stem mpg, lines(2)* and *stem mpg, width(5)* produce the same plot .
- No split of the stems, use *stem mpg, lines(1)*.

stem price, lines(1) digits(3) Stem-and-leaf plot for price (Price)

Create a stem-and-leaf plot for another variable in the dataset called price (the values for this variable are in thousands) by typing: *stem price*

- *round()* option: We can use the *round()* command to round numbers. For example, *round(100)* rounds the prices to the hundreds place.

Type ‘*stem price, round(100)*’ on the command line. What do you see?

Comment: the variable ‘price’ has four or five digits. Observe that stem presented the display in terms of units of 100, so a car that cost \$3,291 was treated for display purposes as \$3,300.

Or we can specify each price to be rounded to the tens place:

- type ‘*stem price, round(10)*’ on the command line
- browse the dataset with ‘*browse if price < 4000*’. Use the output on the screen to explain leaf of your stem and leaf plot.

we can use the command ***prune*** to remove any stems without leaves:

- Type: ‘*stem price, round(10) prune*’ on the command line. What happen?
- You can do cosmetic works to make your plot look good as you want. E.g:
stem price, lines(1) digits(3)
produces one stem and 3 digits for each value in the dataset.

Menu

Statistics > Summaries, tables, and tests > Distributional plots and tests > Stem-and-leaf display