

**Solution: HW #1**

Title: Foundations of Econometrics  
Course: Econ 367

Fall/2015  
Instructor: Dr. I-Ming Chiu

**Part I By Hand**

Q1. 6 (Exercise 1.1, pp. 9)

**Answer:** Simple random sampling or stratified sampling method.

Q2. 30 (Exercise 1.3, pp. 31) [exclude part (d)]

**Answer:** (see 367\_SolHW#1.xls)

(a) Mean (640.5) > median (582.5), the distribution is skewed to the right.

(b) Mean becomes smaller (610.5), but median remains the same; it is resistant.

(c) 20% trimmed mean = 591.17

Q3. 38 (Exercise 1.3, pp. 32)

**Answer:**

There are 7 Ss and 3 Fs.

a) 7/10

b)  $\bar{x} = 7/10$ , they are the same.

c) Suppose there is additional y Ss among these 15.

$(y + 7) / 25 = 0.8$ ,  $y = 13$ .

Q4. 42 (Exercise 1.4, pp. 41~42)

**Answer:** (see 367\_SolHW#1.xls)

(b)  $s^2 = 0.48$ ,  $s = 0.69$

(d) sample variance (standard deviation) is not affected by rescaling the data.

Q5. 46 & 50 (exercise 1.4, pp. 42)

Answer: (see 367\_SolHW#1.xls)

46 (a)

$$\sum_{i=1}^{17} x_i = 2.75 + 2.62 + \dots = 56.8, \bar{x} = \frac{56.8}{17} = 3.3412; \sum_{i=1}^{17} x_i^2 = 2.75^2 + 2.62^2 + \dots = 197.804$$

$$(b) S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2 = 197.804 - 17 \cdot 3.3412^2 = 8.0252$$

$$\text{Var}(x) \text{ or } s^2 = 8.0252/16 = 0.5016; \text{Sd}(x) \text{ or } s = 0.7082$$

50 (a)  $Q1 = 2.74$ ,  $Q3 = 3.88$ ; (b)  $IQR = 1.14$ ; (c) Not affected; it's resistant; (d) to 2.74; (e)  $IQR = 1.19$

Q6. 52 (Exercise 1.4, pp. 43)

**Answer:** (see 367\_SolHW#1.xls)

- (a) Fourth or  $IQR = 33$
- (b) No outliers
- (c) You can use Stata to plot it by yourself
- (d) No more than 32 ( $= 424 - 392$ )

Q7. 58 (Exercise 1.4, pp. 44)

**Answer:**

- (a) GVC is higher at 2:00 PM; the corresponding  $Q1$ , median, and  $Q3$  are higher than the same measures at other time points.
- (b) The distributions are mostly symmetric at all five time points.
- (c) There are outliers and extreme outliers when GVC is measured at 6:00 PM.

## Part II Use Stata

Q8. (a) What is the dimension of the “auto” data?

- (b) What kind of data file that “auto” belongs to? (Use method B and C, respectively, from Handout #1, pp. 3)
- (c) Differentiate the variable type using “mpg” and “foreign” variable. (Use method A from Handout #1, pp. 3)
- (d) Find the Five Number Summary and  $IQR$  for “mpg” variable, meanwhile; find a way to decide its skewness without using the histogram plot.
- (e) Present the four diagrams (Stem-and-Leaf, Dotplot, Histogram, and Boxplot) for “mpg” variable.
- (f) Present the comparative (i.e., side-by-side; see textbook, pp. 40) Boxplot of “mpg” based on the “foreign” variable. What can you find from this plot?

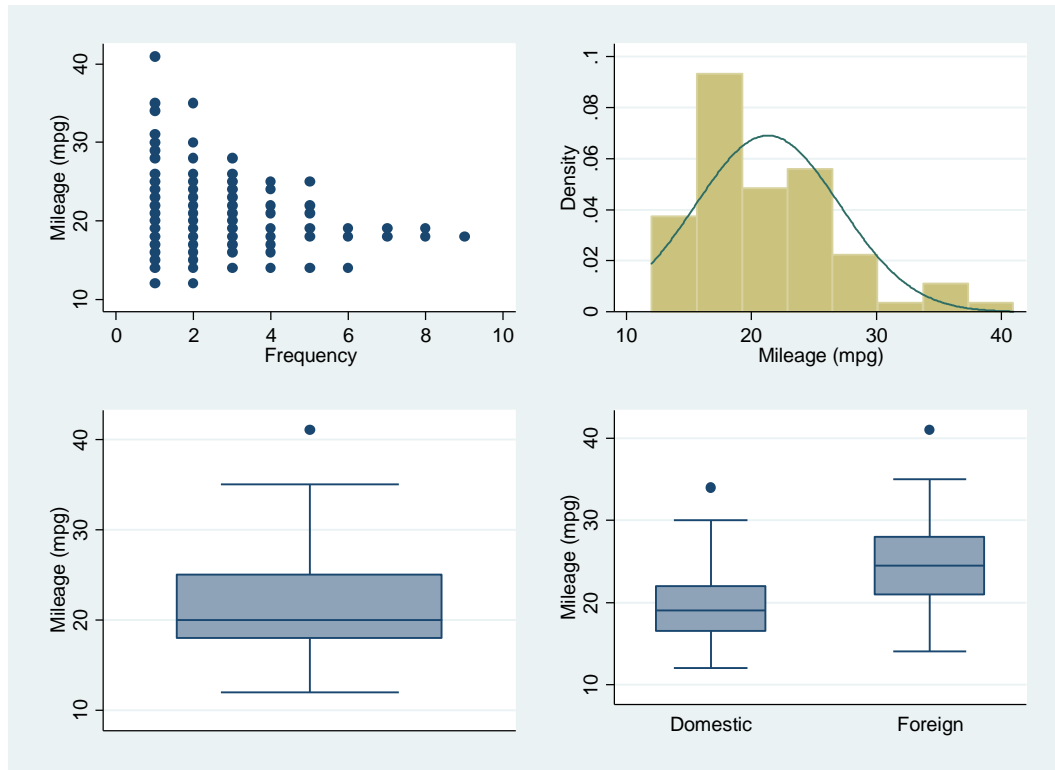
Answer:

- (a) 74 (cases) \* 12 (variables); (b) Cross-section & observational
- (c) “mpg” is a quantitative and foreign is categorical variable
- (d)  $\min = 12$ ,  $Q1 = 18$ , median = 20,  $Q3 = 25$ , max = 41
- (e) Stem-and Leaf

Stem-and-leaf plot for mpg (Mileage (mpg))

```
1t | 22
1f | 44444455
1s | 66667777
1. | 888888888999999999
2* | 00011111
2t | 22222333
2f | 4444555555
2s | 666
2. | 8889
3* | 001
3t |
3f | 455
3s |
3. |
4* | 1
```

(e) & (f)



The right corner comparative boxplots indicates “Foreign” cars are more efficient using all five-number summary.

Stata Do file for your reference

--

```
sysuse auto, clear
```

```
log using 367_SolHW01.smcl
```

```
describe
```

```
summarize mpg, detail
```

```
*for stem-and-leaf diagram, copy as "picture" and paste to your word file
```

```
stem mpg
```

```
dotplot mpg, name(panela) nodraw
```

```
histogram mpg, normal name(panelb) nodraw
```

```
graph box mpg, name(panelc) nodraw
```

```
graph box mpg, over(foreign) name(paneld) nodraw
```

```
graph combine panela panelb panelc paneld, cols(2)
```

```
log close
```