STAT 2507 Assignment # 1 Fall 2011

Total = 20 marks

Last Name Solutions, First

Student # Your LAB Section

Due date: **If your lab is on a Friday afternoon, give your assignment to your TA, at the end of your lab, on Friday September 30. If your lab is on a Monday, Tuesday, Wednesday, or Thursday, hand in you assignment to me no later than 3:00 pm. on Friday September 30.**

IMPORTANT NOTES:

1. **You MUST write your lab section in the space provided above (A1, A2, A3, etc.)**
2. **You MUST work by yourself NOT in groups.**

Click on 'Editor' and then on 'Enable Commands' to enable the 'MTB' prompt. Type the following Minitab commands to generate 200 numbers to be saved in a vector called c2:

```
MTB > set c1
DATA> 1:200
DATA> end
MTB > random 200 c11;
SUBC> normal 1400 100.
MTB > let c2=1000*log(c1)+c11        (c2 contains the data)
```

Think of these 200 numbers in c2 as the prices of a sample of 200 used cars, in dollars.

Question 1. Draw a stem-and-leaf plot of these 200 prices (you can do this by typing 'stem c2') and then use your plot to answer the following questions:

[0.5] (a) The maximum price is $ 6800 . } ±200

[0.5] (b) The minimum price is $ 1200 .

[1] (c) The median price is $ 5900 . ±100

[1] (d) 36% of the prices are less than or equal to $ 5600 .

[1] (e) What is the shape of the distribution of the summer incomes? Answer: skewed left

[2.5] **Question 2.** Use the 'describe' command to answer parts (a), (b), and (c) of Question 1, (type describe c2). (±150) (±150) (±50)

Answer: (a) 6829.2 , (b) 1282.7 , (c) 5995.2 .

(d) The average price and the standard deviation of the price of a used car are, respectively, $ 5717.2 and $ 952.6 .
(±150) (±50)

Question 3.

[1] (a) What proportion of the prices are within 2 standard deviations of the mean price (i.e., what proportion of the prices fall in the interval $\bar{x} \pm 2s$)? Answer: 189/200. (189±3)/200

[0.5] (b) Answer part (a) using Tchebysheff's Theorem. At least 150/200 = 75%.

[0.5] (c) Answer part (a) using the Empirical Rule. Approximately 190/200 = 95%

Question 4. Now make the following transformation of the used car prices (that are in c2): (new price) = 1.6× (price) + 200. This can be done in Minitab by using the 'let' command:

MTB > let c10=1.6*c2 + 200

[4] (±0.8)

(a) The z-scores corresponding to the two smallest values in c10 are -4.65 and -3.80 (±0.5) respectively. Are these two vlaues outliers? yes.
Why ? |Z-score| > 3.

(b) Use a Boxplot of the new prices in the vector c10 to answer the following questions (click on 'graph' and then on 'Boxplot' and click 'ok'):

[1] (i) The median is approximately equal to $ 9800 (±200)

[1] (ii) The interquartile range is approximately equal to 1800 (±200) (±300)

[1] (iii) The range of the data (new prices) is approximately equal to $ 11200-2200 = 9000

[0.5] (c) Obtain a histogram of the new prices in c10 and comment on the shape of the distribution of these prices skewed left

ALSO, do the following questions:

[2] **A.** The height of a particular tree is known to have a mound-shaped distribution with a mean 180 centimeters and a variance of 100 centimeters. Approximately, what proportion of such trees have a height between 170 and 200 centimeters? (Use the empirical rule.)

(a) 85% (b) 81.5% (c) 85.5% (d) 90%

[2] **B.** For a sample of size 4, it is known that $x_1 - \bar{x} = 10$, $x_2 - \bar{x} = -8$, and $x_3 - \bar{x} = -6$. The sample variance is

(a) 16 (b) 66.67 (c) 72 (d) Impossible to compute