ST 350 
LECTURE WORKSHEET #5  
Reiland

Name________________

BOXPLOTS

Who: students in ST 350-002  
What: amount spent on textbooks for current semester  
Why: many reasons are possible, for example:
    i) change financial aid formula to allow more for textbooks  
    ii) bookstore mgr wants to track textbook price changes

1. The data are shown below (note that the data are NOT ordered!).

   370  320  390  280  390  55  350  400  475  420  450  350  305  325

   ordered data:
       55  280  305  320  325  350  350  370  390  390  400  420  450  475

2. Calculate median: 14 numbers (even number of numbers)
   position of overall median: between 7th (350) and 8th (370) ordered numbers
   median = \( \frac{350 + 370}{2} = 360 \)

3. Calculate \( Q_1 \):
   \( Q_1 \) = median of “lower half” (smallest 7 numbers) = 4th observation from “small” end = 320

4. Calculate \( Q_3 \):
   \( Q_3 \) = median of “upper half” (largest 7 numbers) = 11th observation from “small” end = 400

5. Label the number line below and draw a box with ends at the quartiles calculated above:
6. Show the location of the median with a "+" in the box. (In the boxplot above the location of the median is indicated by the vertical line inside the box).

**Lines from ends of box; outliers**

7. a. $IQR = Q_3 - Q_1 = 400 - 320 = 80$
   b. $1.5(IQR) = 1.5(80) = 120$
   c. $Q_1 - 1.5(IQR) = 320 - 120 = 200$
      $Q_3 + 1.5(IQR) = 400 + 120 = 520$

mark these boundaries on the number line above (these boundaries are just for construction and are NOT part of the boxplot display)

d. From each end of the above box, draw lines to the most extreme observations that are within 1.5(IQR) of the end of the box.

From the low end of the box draw the line to the smallest value that is greater than $Q_1 - 1.5(IQR) = 200$, so the line is drawn to 280.

From the high end of the box draw the line to the largest value that is less than $Q_3 + 1.5(IQR) = 520$, so the line is drawn to 475. Since 475 is the largest data value there are no outliers at the high end.

e. Indicate by a "*" each observation that is more than 1.5(IQR) from the ends of the box. These observations are outliers. The data value 55 is an outlier at the low end since it is less than $Q_1 - 1.5(IQR) = 200$. 