

**Unit 1 - Summarizing Data**  
**Week #2 - Practice Problems**

**Due: Monday September 22, 2008**

1. **Before you begin:** *This exercise gives you practice with the calculation and interpretation of some numerical summaries. See lecture notes pp 36-47 and/or the text pp 8-25 of your text. Part "c" gives you some additional practice working with and understanding a box and whisker plot. For this, see lecture notes pp 25 and/or the text pp 30-33.*

- a. The following are behavioral ratings as measured by the Zang Anxiety Scale (ZAS) for 26 persons with a diagnosis of panic disorder:

53	51	46	45	40	35
59	51	45	60	35	
45	38	53	43	31	
36	40	41	41	38	
69	41	46	38	36	

Compute the mean, median, mode, range, variance, and standard deviation, and the 25th and 75th percentiles.

- b. The following are behavioral ratings as measured by the Zang Anxiety Scale (ZAS) for 21 healthy controls:

26	26	25	25	25
28	26	26	25	
34	30	31	28	
26	34	25	25	
25	28	25	25	

Compute the mean, median, mode, range, variance, and standard deviation, and the 25th and 75th percentiles.

- c. Construct Box and Whisker plots using the data from parts "a" and "b". In one or two sentences, compare the two groups.

2. **Before you begin:** This exercise provides additional practice with the frequency and relative frequency table. See lecture notes p9 and text pp 25-27. From there, you gain practice working with grouped data and, in particular, the challenge of estimating the values of numerical summaries when you don't have the individual values. See Lecture notes page 33 "weighted mean". In applying these ideas, use the midpoint of each age interval. For a bit of extra help in completing this exercise, visit <http://www.duncanwil.co.uk/average4.html>:

The following table shows the age distribution of cases of a certain disease reported during a year in a particular state.

Age	Number of Cases
5-14	5
15-24	10
25-34	20
35-44	22
45-54	13
55-64	5
TOTAL	75

- 2a. Construct a frequency table with columns for class endpoints, class midpoint, frequency, relative frequency, cumulative frequency, and cumulative relative frequency.
- 2b. Construct a cumulative relative frequency plot of the data. Use this plot to estimate the 10th, 25th, 50th, and 75th percentiles.
- 2c. Compute the mean, median, variance, and standard deviation .

3.

***Before you begin:*** This exercise makes sure that you understand the specifics of a box and whisker plot (apart from clicking on a link!) For review visit <http://stattrek.com/AP-Statistics-1/Boxplot.aspx>

For women undergoing in vitro fertilization, various therapies are used to stimulate the ovaries. In one study comparing the effectiveness of a new hormone therapy on three groups of women with different types of fertility problems, an outcome of interest is the number of oocytes that 'ripened'. Some summary statistics on the number of ripened oocytes per woman for each of the three groups are reported below.

Statistic	Group		
	1	2	3
n	38	19	21
mean	13.6	6.4	8.2
median	8	8	7
P25	5	4	5
P75	11	11	12
minimum	5	1	4
maximum	40	13	14

3a. Compute box and whisker plots for the three groups.

3b. In your opinion, which statistics are best for comparing these three groups? Why?