1. Congratulations! You graduated from UCLA and got a job as the marketing director for a large grocery store chain with an internet site that allows its customers to purchase products on-line. Your supervisor is wondering whether allowing customers to use special store coupons during their on-line purchases will increase the average dollar amount each customer spends on-line.

You decide to randomly sample 200 existing on-line customers. Then you gathered information about their purchasing behavior and demographics. Finally, using a random method, you give 100 of the on-line customers special store coupons for use during their next on-line purchase and you do not give the other 100 any special store coupons. You find that 75 of the 100 or 75% of the customers who were given the special store coupons made a purchase and they purchased an average of $28 worth of groceries. The 100 who did not receive the special store coupons purchased an average of $25 worth of groceries. Approximately 75 of the 100 without the special store coupons made an online purchase.

Data on 200,000 existing online customers reveals that the average dollar amount of groceries purchased is $53. Data on all 2,000,000 online and non-online customers reveals that the average dollar amount of groceries purchased is $41.

Please answer the following questions:
   a. (2 points) The population is:
                   i. $41*2,000,000 or $82,000,000
                  ⓘⓘ 200,000 online customers
                  iii. 2 million customers
                  iv. $28*200,000 or $5,600,000
                  v. 200 customers
   b. (2 points) The sample is:
                  i. 2 million customers
                  ii. 200,000 online customers
                  ⓘ ⓘ ⓘ 200 customers
                  iv. 100 customers given special store coupons
                  v. 100 customers not given special store coupons
   c. (2 points) The parameter is:
                  ⓘ. 53 dollars
                  ii. $53*200,000 or $10,600,000
                  iii. $41*2,000,000 or $82,000,000
                  iv. 28 dollars
   v. $28*200,000 or $5,600,000
   d. (2 points) The statistic of most interest to your supervisor is:
                   i. $28*200 or $5,600
                  ⓘⓘ 28 dollars
                  iii. $53*200 or $10,600
                  iv. 53 dollars
                  v. $41*2,000,000 or $82,000,000
   e. (2 points) This study is:
                  i. an Observational Study with controls
                  ii. an Observational Study that uses a random probability method for sample
                      selection.
                  ⓘ ⓘ ⓘ a Randomized Controlled Experiment
                  iv. a Randomized Experiment without Controls, but it is blind
                  v. a Randomized Experiment without Controls, but it is double-blind

2. Using the graphic below, please answer the following questions:

value of variable 2
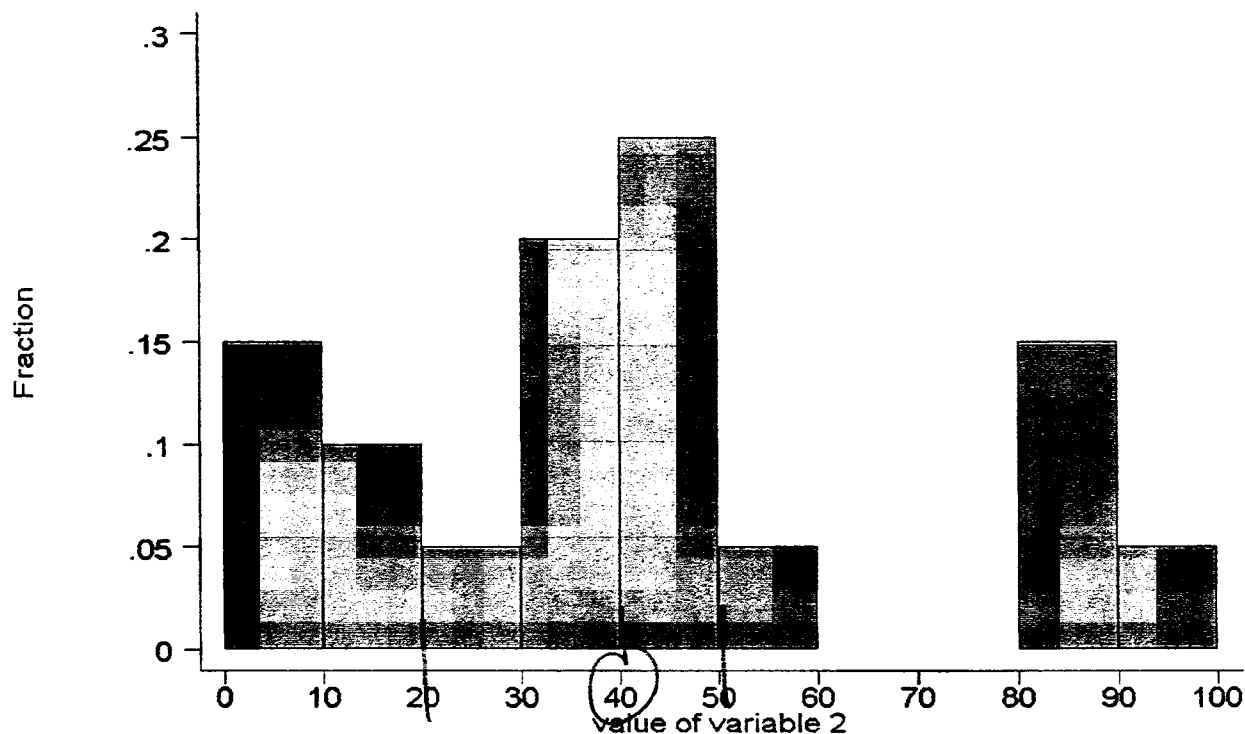
i.   The median for this histogram is (2 points):
      a. Exactly 40
      b. Between 40 and 50
      c. Exactly 50
      d. Between 50 and 60
      e. not calculable because this is <u>NOT</u> a histogram

ii.   The inter-quartile range for this histogram is (3 points):
      a. 30
      b. 40
      c. 50
      d. 60
      e. not calculable because this is <u>NOT</u> a histogram

3. You are on the verge of investing some of your hard-earned money in the stock market and you cannot decide between two funds, let's call them A and B. Your investment adviser, I'll call him The Oracle, gives you some information on their performance (as measured by percentage return) over the last 200 trading days:

Fund A:

Variable	Obs	Mean	Std. Dev.	Min	Max
dailyret	200	4.000000	4.00000	-12.00	20.00

Fund B:

Variable	Obs	Mean	Std. Dev.	Min	Max
dailyret	200	10.000000	12.0000	-22.00	52.00

Assume they are both very close to <u>normally distributed</u>, so for our purposes they should be considered normal variables. Please answer the following questions:

A. The Oracle advises you that while there are periods of time when you may get very low (near zero) returns on your investment, you should avoid negative returns at all costs. Keeping his advice and the information above (i.e. mean, standard deviation, normally distributed) in mind, which Fund will have the larger proportion of negative (less than zero) percentage returns? (circle one, 2 points)

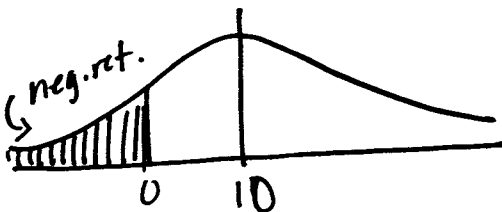   FUND A   (FUND B)

Show the numerical calculations that helped you to arrive at your conclusion. (8 points)

Fund A
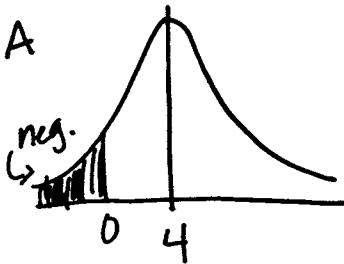
$$z = \frac{0-4}{4} = -1.0 \quad \text{area is } .1587$$
$$\text{or } 15.87\%$$

Fund B

$$z = \frac{0-10}{12} = -.83 \quad \text{area is}$$
$$.2033$$
$$\text{or } 20.33\%$$

20.33 > 15.87 so FUND B has more neg.

(using information from the problem, but not the solution for part A)

B.   Please give the five number summary for variable dailyret for Fund B and show your work for full credit (10 points). Please list the values from lowest to highest. If you do not have enough information for a complete answer, please write "not enough information", give as much of the five number summary as you can from the information given above, and tell us what you would need to give a complete answer.

MIN	Q_1	M	Q_3	MAX
-22	1.96	10.0	18.04	52.00

1) median = mean for normal variables

2) for Q_1

25%

$z = -.67$
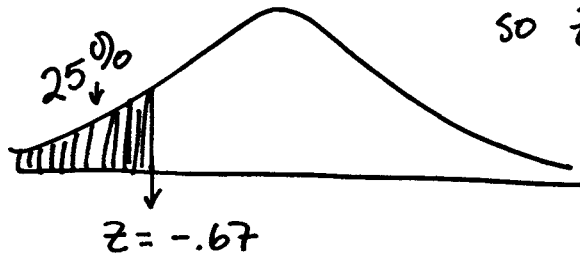
so $z = \dfrac{X - M}{\sigma}$ so $-.67 = \dfrac{X - 10}{12}$
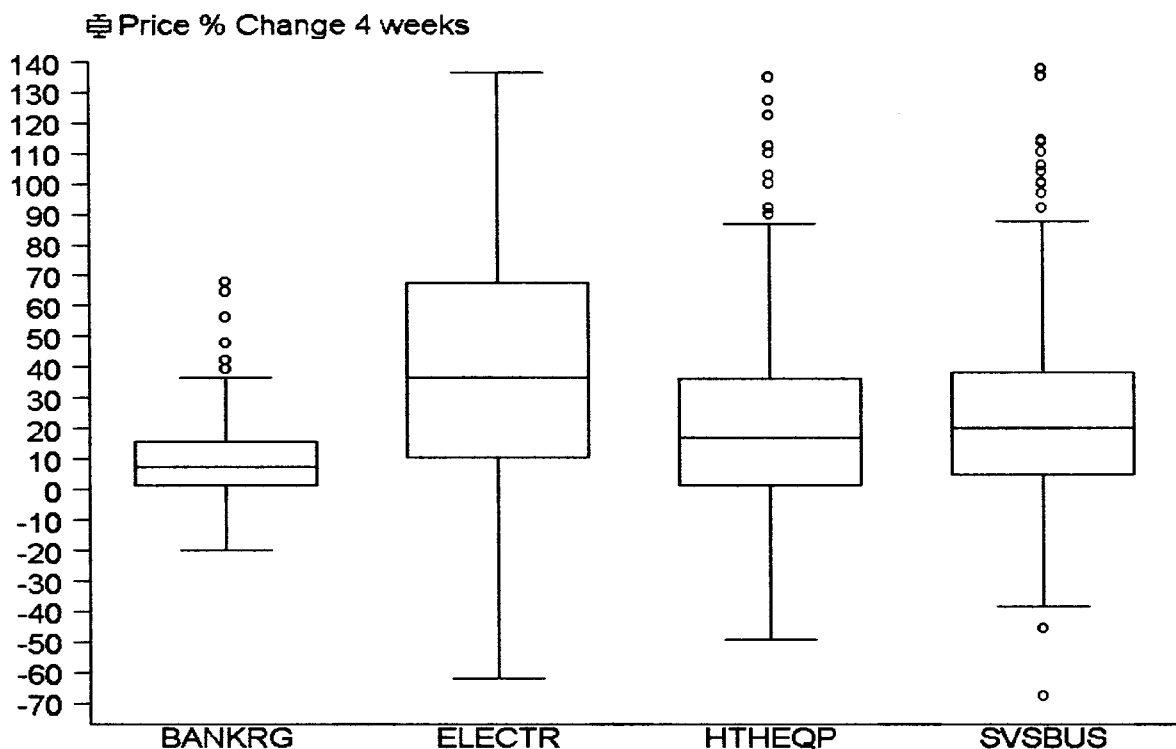
Solve for X

1.96

3) for Q_3

75%

$z = +.67$
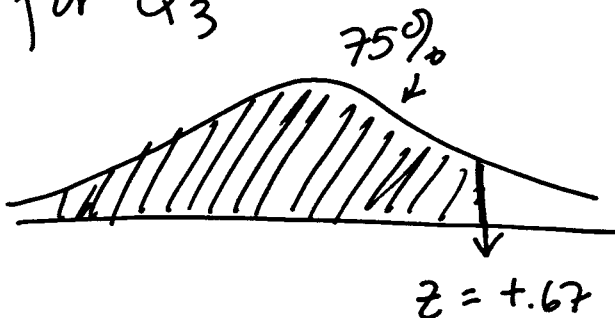
$+.67 = \dfrac{X - 10}{12}$

Solve for

$X = 18.04$

4. The box plots below summarize the percentage change in stock price for 4 industries (each industry has many companies) for January 2001.   Please answer the following questions on the basis of this graphic.

⊟ Price % Change 4 weeks

(a)   For the Banking industry (BANKRG), the 75th percentile for percentage change in stock price is (3 points):

      A.   Over 40%
      B.   A little less than 40%
      C.   Between 20% and 30%
      (D)   slightly less than 20%
      E.   not calculable from this graphic

(b)   Is the distribution of percentage change in stock price nearly symmetrical for any these industries? (circle one, 2 points)

    (YES)   NO   NEED MORE INFORMATION

(c)   Please tell how you arrived at your answer in part (b) above, use calculations where appropriate (5 points)

ELECTR

Median is nearly centered between 75th + 25th

Min + Max are nearly equal distance from Q₁ and Q₃

No outliers to cause skewness