$(46, 66, 73, 83, 95)$, $\bar{x} = 73.1$.

1. (25 pts) $(7, 12, 17, 20, 24)$, $\bar{x} = 16.5$.

   (a) The hypotheses should be

   > H_0: The patients who took Vytorin had the same cancer rate as the control group.

   > H_1: The patients who took Vytorin had a higher cancer rate than the control group.

   (b) The dummy pill is called a *placebo* and the group that was administered the dummy pill is the *control group*.

   (c) The placebo eliminates response bias. The patients do not know how they are expected to respond because they do not know whether they are receiving the treatment.

   (d) The link was not statistically significant. The article states that "the number of cancer cases did not differ significantly." Just because there were more cases does not mean that there were *significantly* more cases.

   (e) This is an experiment. The researchers administered prescribed doses of Vytorin and they administered the placebo. In other words, they manipulated the explanatory variable.

   (f) One explanatory variable is whether the patient received Vytorin. One response variable is whether the patient had cancer. Remember, a variable must be something that can be observed individually for each member of the sample. The "risk of cancer" cannot be observed for individuals. The number of patients with cancer or the percentage of patients with cancer is a statistic, not a variable.

2. (15 pts) $(4, 9, 12, 14, 15)$, $\bar{x} = 10.8$.

   (a) The direction of extreme is to the right. In the null-hypothesis picture, you can see that there are very few values at the right end of the distribution. The reverse is true of the alternative-hypothesis picture.

   (b) One reasonable decision rule is "Reject H_0 if the observed SAT Math score is 700 or greater." You could use 600 or even 800 instead of 700. Using 500 is going a bit too far to the left.

   (c) For the rule stated in part (b), in the null-hypothesis picture (Downtown High School), there are two X's in the rejection region (≥ 700) out of 12, so $\alpha = \frac{2}{12}$. In the alternative-hypothesis picture, there are 5 X's in the acceptance region (≤ 600) out of 12, so $\beta = \frac{5}{12}$.

(d) To compute the p-value of 500, use the null-hypothesis and count the X's to the right (direction of extreme) starting at 500. There are 6 X's, so the p-value is $\frac{6}{12}$.

3. (25 pts) $(0, 14, 18, 19, 25)$, $\bar{x} = 16.0$.

   (a) Selection bias. The people with cell phones only are less likely to be in the sample.

   (b) The "Wilder effect" is an example of response bias. Some people are embarrassed to say publicly that they would not vote for a black man.

   (c) If they intentionally selected a certain number of cell-phone users and a certain number of landline users, then it would be a stratified sample. The two strata would be the cell-phone users and the landline users.

   (d) The 48% figure is a statistic. It is based on a sample, not the population.

   (e) The variable of primary interest is whether the person will vote for Obama or McCain.

   (f) The bar graph:
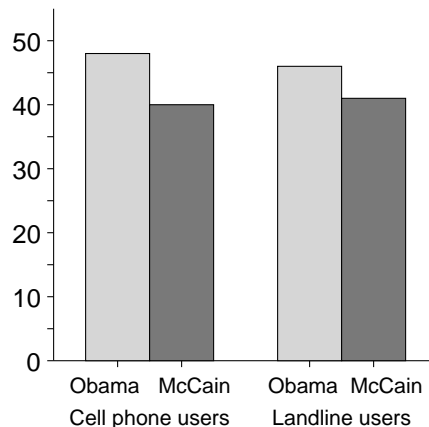
4. (25 pts) $(13, 20, 22, 25, 25)$, $\bar{x} = 21.7$.

   (a) The procedure: Number the halls 1 through 5. Then randomly select 2 integers from 1 to 5. You could use `randInt(1,5)`. Then include all the students from the two selected halls in the sample.

   (b) With a seed of 34, the two random numbers are 2 and 1. So the halls are Paul Hall and Ball Hall.

   (c) Paul Hall has 30 students and Ball Hall has 20 students, so the sample has 50 students.

   (d) Yes, all students are equally likely to be selected. Any given student is selected if and only if his hall is selected. Each hall has a 2/5 chance of being selected, so each student has a 2/5 chance of being selected.

(e) To select a simple random sample of 80 students, number all the students from 1 to 200. Then select 80 random numbers from 1 to 200, perhaps using `randInt(1,200)`. The selected students would constitute the sample.

(f) With a seed of 17, we get the numbers 135, 111, 167, 196, 135 (reject), and 89. Be sure to use `randInt(1,200)`, not `randInt(1,80)`. We are selecting from a group of 200 students, not a group of 80 students.

5. (10 pts) $(0, 8, 9, 10, 10)$, $\overline{x} = 8.2$.

All three displays will show essentially the same shape.

The frequency plot (easiest):
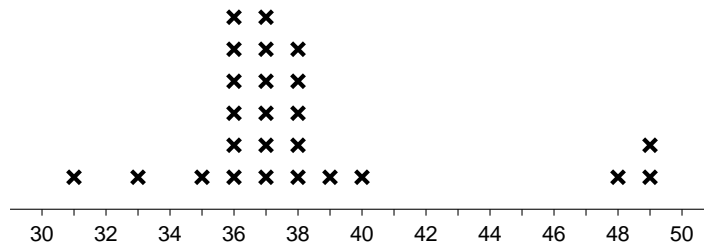
The stem-and-leaf display:

Stem	Leaf
31	0
32	
33	4
34	
35	6
36	4 6 6 0 5 5
37	6 6 5 1 4 5
38	1 0 6 6 2
39	5
40	2
41	
42	
43	
44	
45	
46	
47	
48	5
49	2 3

The histogram (hardest), starting at 30000 with a class width of 2000 (other choices are possible):

3