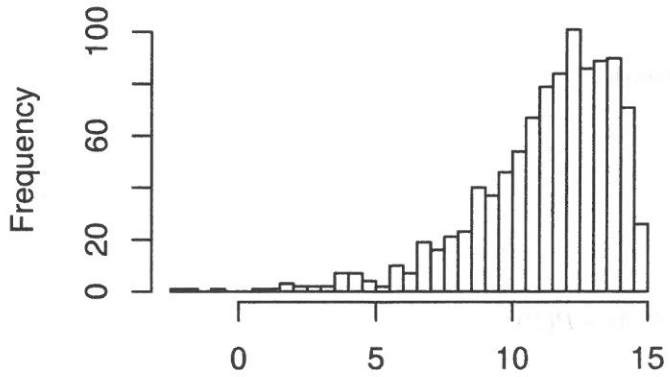
For each problem in this section, please circle one of the following answers. No justification is required.

1. [2 points]




The histogram shown above represents data that are

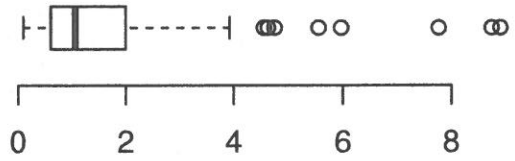
a) skewed right

b) skewed left

c) not skewed

d) not enough information provided

2. [2 points]




The box plot shown above represents data that are

a) skewed right

b) skewed left

c) not skewed

d) not enough information provided

[Could be a bimodal distrⁿ. See histogram and boxplot of old faithful waiting times on Prelim Review Slide #22.]

3. [2 points]

Which of the following can be used effectively to display two numerical variables?

a) color-coded box plots

b) a histogram

c) stacked bar plots

d) none of the above [Scatter plot is the best option].

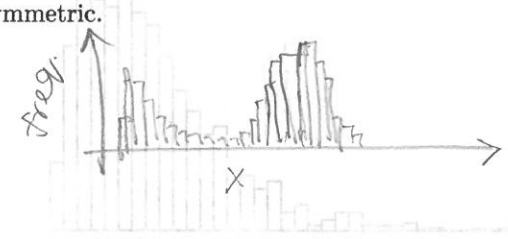
Option (a) 'Color-coded box plots' is also acceptable.

**True or False?** For each of the following questions, please answer either *True* or *False*. While justification is not required, it is encouraged and may allow in some cases for partial credit to be awarded.

4. [2 points]

If a distribution is bimodal, then it must be symmetric.

FALSE




5. [2 points]

If events $A$ and $B$ are disjoint, then $P(A \cup B) = P(A) + P(B)$.

TRUE

6. [2 points]

If $A$ and $B$ are independent, then $P(A|B) = \frac{P(A)}{P(B)}$.

FALSE

$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$.

(use independence)

7. [2 points]

The events $A$ and $A^C$ are disjoint.

TRUE

8. [2 points]

(Recall that a deck of playing cards has 26 red cards and 26 black cards.) Suppose you shuffle a deck of cards thoroughly and then select the top card. Let $X$ be 1 if the card is red and 0 if the card is black. The random variable $X$ is Bernoulli(0.5).

TRUE

9. [2 points]

You shuffle a deck of cards thoroughly and then select the top 5 cards. Let $X$ be the number of red cards. The random variable $X$ is Binomial(5, 0.5).

FALSE

$P(\text{Card 1 is red}) = \frac{26}{52} = \frac{1}{2}$.

$P(\text{Card 2 is red}) = \frac{51 \times 26}{52 \times 51} = \frac{1}{2}$.

$P(\text{Card 2 red} \mid \text{Card 1 red}) = \frac{25}{51} \neq P(\text{Card 2 red})$.

So trials are not independent.

**Please answer the following question.**

10.

a) [5 points] Suppose you want to know the average length of leaves grown in a forest. In each case below, give the name of the type of sampling described:

i. You take a leaf from the lowest branch of some trees close to your lab.

CONVENIENCE SAMPLING

ii. You get a simple random sample of trees; then, on each of these trees, you select a simple random sample of branches; then, on each branch, you select a simple random sample of leaves.

CLUSTER SAMPLING
(More precisely, MULTISTAGE SAMPLING)

iii. You walk through the forest with the park ranger who points out which leaves seem to be the most typical/representative of the overall population of leaves in this forest.

JUDGEMENT SAMPLING

iv. For each species of tree, you get a simple random sample of leaves.

STRATIFIED SAMPLING

v. You walk through the forest gathering leaves until you have 50 from each species of tree.

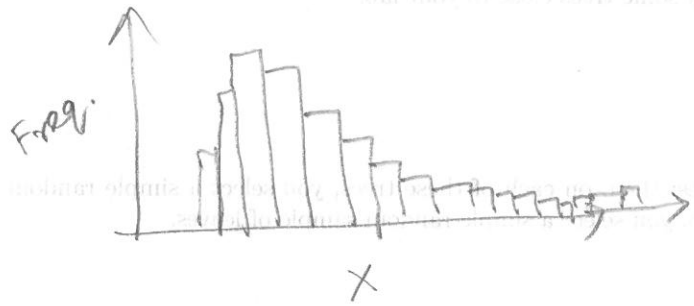
QUOTA SAMPLING

b) [2 points] Describe why it would be difficult to get a simple random sample of leaves in the forest.

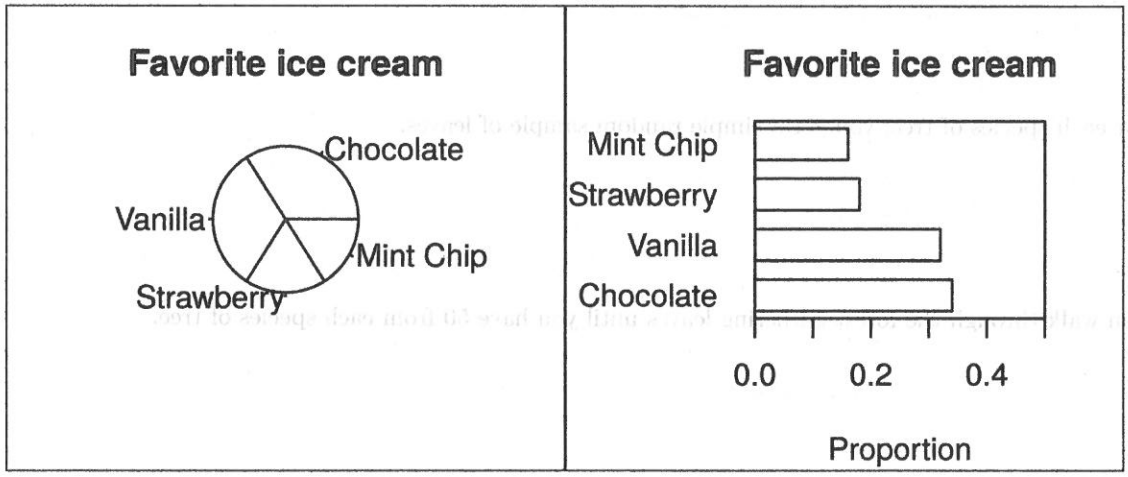
MAKING A LIST OF ALL LEAVES IN THE FOREST, AND
CREATING A SAMPLING FRAME
WILL BE VERY DIFFICULT, IF NOT IMPOSSIBLE.

**Please answer the following questions.** An answer without justification will not receive full credit.

11. [2 points] Draw a histogram in which the mean is higher than the median.




12. [3 points] Here are two ways to display the same data. Which is the more effective graphic? The more justification you give, the better.




The bar chart is more effective than pie chart.
It shows, e.g.,
(i) Proportion of chocolate a little higher than proportion of vanilla;
(ii) Proportion of strawberry a little higher than proportion of mint chip;
(iii) Proportions of MC, S, V and C are approximately 0·15, 0·17, 0·32, 0·33.
None of these information is clear in the pie chart.