Exam 1

For problems 1-10, circle the letter next to the response that best answers the question or completes the sentence. You do not have to show any work or write any explanations here. 2 pts each

1. Which sampling method is most likely to give a representative sample
   a) systematic sampling
   b) convenience sampling
   c) stratified sampling
   d) simple random sampling

2. The number of classes in a frequency distribution depends on the size of the data set.
   In general, the …
   a) larger the data set, the larger the number of classes
   b) larger the data set, the smaller the number of classes
   c) number of classes should be equal to the number of values in the data set divided by 6
   d) smaller the data set, the larger the number of classes

3. A student is told that his score on an exam was in the 20th percentile. This means that
   a) about 20 students in the class scored better than him
   b) about 20 students in the class scored worse than him
   c) about 20% of the students in the class scored better than him
   d) about 20% of the students in the class scored worse than him

4. Suppose you have a data set given in two versions, one is the original ungrouped data, and the other is a set where the data values have been grouped. If you calculate the mean and standard deviation for the two versions of the data sets, you would expect to get
   a) exactly the same mean and standard variation no matter which set you used for your calculations
   b) exactly the same mean but a slightly different standard variation
   c) slightly different mean but exactly the same standard variation
   d) slightly different mean and a slightly different standard variation

5. The z-score of a data value tells us
   a) what spread the data have
   b) the percentile rank of that data
   c) how outlier-sensitive the data is
   d) how many standard deviations that value is from its population mean
6. The distribution shown in the graph to the right is...
   a) skewed right
   b) skewed left
   c) bell-shaped
   d) uniform

7. We would expect that the mean is …
   a) less than the median
   b) more than the median
   c) the same as the median
   d) the same as the mode

8. A rough estimation of the standard deviation of the data values in this distribution could be
   a) 2  b) 12  c) 42  d) 92

Using the dotplot graph on the right, answer the following questions:  

9. How big is the sample size?

10. Find the mode of the distribution.

11. The following graph shows the ethnicity distribution at a specific college.

   (a) The college has about 12,000 students. How many of the students identify as Hispanic?

   (b) This is an example of a
       i) bar graph
       ii) histogram
       iv) polygon
       v) line graph
12. An insurance company would like to determine the proportion of all medical doctors who have been involved in one or more malpractice lawsuits. The company selects 500 doctors at random from a professional directory and determines the number in the sample who have been involved in a malpractice lawsuit.

Match each statement with correct term regarding this study:

(a) _____ is all medical doctors listed in the professional directory.
(b) _____ is the number of medical doctors who have been involved in one or more malpractice suits.
(c) _____ is the proportion of medical doctors who have been involved in one or more malpractice suits in the population.
(d) _____ is the 500 doctors selected at random from the professional directory.
(e) _____ is the proportion of medical doctors who have been involved in one or more malpractice suits in the sample.

13. The following table gives the frequency distribution of ice-cream flavors sold.

<table>
<thead>
<tr>
<th>Ice-cream flavor</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla</td>
<td>12</td>
</tr>
<tr>
<td>Chocolate</td>
<td>9</td>
</tr>
<tr>
<td>Strawberry</td>
<td>4</td>
</tr>
</tbody>
</table>

a) How big is the sample size?

b) Which type of variable are we studying in the above sample?
   a) qualitative   b) quantitative   c) discrete   c) continuous

c) Find the percentage of chocolate ice-creams sold.

d) Which of the following would it make sense to find for this data set:
   (a) mean   (b) median   (c) mode   (d) variance
15. The following sample data give the number of cars owned by some families in a specific neighborhood. 
*Please show all work for problems below! Use formulas, not just calculator.*

a) Find the sample size.  
2 pts

b) Find the median.  
3 pts

<table>
<thead>
<tr>
<th>Number of Cars</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

c) Find the mean. Include correct notation and units of measurements. Round to one decimal.  
4 pts

d) Find the standard deviation. Include correct notation and units of measurements. Round final answer to two decimals.  
8 pts
16. IQ scores are bell-shaped and have a mean of 100 and a standard deviation of 15. Approximately what percentage of all people have an IQ between 70 and 130? Please show work or explain how your find your answer.

17. Suppose you have the following stem-and-leaf display:

```
4 | 2 2
5 | 3 9
6 | 1 3 8
7 | 0
```

a) What is the value of the biggest data point in this data set?

b) How many observations are in this data set?

18. Use the given boxplot to answer the following questions.

```
0  5  10  15  20  25  30
Percent score
```

a) What percentage of the data values lie below 14?

b) Approximate the median.

c) How many outliers are there?
19. Draw arrows to match each histogram with the corresponding box-and-whisker plot: 

4 pts

20. Circle the histogram above that has the largest standard deviation. 

2 pts

21. What is one of the main reasons we usually prefer taking a sample rather than a census? 

2 pts