Instructions. Complete each of the following eight questions. Please show all appropriate work in your solutions in order to obtain maximum credit. You may use a calculator.

1. (6 pts) Categorize the following data according to level: nominal, ordinal, interval or ratio.
   (a) Ans: ___________ The length of time it takes for someone to run a marathon.
   (b) Ans: ___________ The time of day the marathon starts.
   (c) Ans: ___________ The name of the city in which the marathon is run.

2. (6 pts) To compute the average amount of medical insurance its patients had in 2006, a hospital considered taking the following types of samples. For each case classify the sample as simple random, stratified, systematic, cluster, convenient.
   (a) Ans: ___________ The hospital numbered its complete list of patients from 2006 and used a random number generator to select 100 patients from the list from which to collect the data.
   (b) Ans: ___________ The hospital decided to collect the data from the first 50 patients admitted on July 4, 2006.
   (c) Ans: ___________ The hospital randomly chose a patient, and collected data from that patient and every 200th patient admitted thereafter.

3. (8 pts) Consider the collection of 30 data
   
   | 2 | 3 | 4 | 4 | 5 | 6 | 8 | 9 | 11 | 12 | 14 | 15 | 18 | 19 | 21 |
   | 23 | 23 | 25 | 27 | 34 | 40 | 45 | 51 | 56 | 58 | 62 | 73 | 83 | 91 | 98 |

   Complete the following table given that the first class has limits 1 - 20.

<table>
<thead>
<tr>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
<th>Midpoint</th>
<th>Frequency</th>
<th>Cumulative Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>.5</td>
<td>20.5</td>
<td>10.5</td>
<td>14</td>
<td>14</td>
<td>14/30 ≈ .467</td>
</tr>
<tr>
<td>21</td>
<td>40</td>
<td>20.5</td>
<td>40.5</td>
<td>30.5</td>
<td>7</td>
<td>21</td>
<td>7/20 ≈ .35</td>
</tr>
<tr>
<td>41</td>
<td>60</td>
<td>40.5</td>
<td>60.5</td>
<td>50.5</td>
<td>4</td>
<td>25</td>
<td>4/30 ≈ .133</td>
</tr>
<tr>
<td>61</td>
<td>80</td>
<td>60.5</td>
<td>80.5</td>
<td>70.5</td>
<td>2</td>
<td>27</td>
<td>2/20 ≈ .107</td>
</tr>
<tr>
<td>81</td>
<td>100</td>
<td>80.5</td>
<td>100.5</td>
<td>90.5</td>
<td>3</td>
<td>30</td>
<td>3/20 ≈ .15</td>
</tr>
</tbody>
</table>
4. Consider the following table

<table>
<thead>
<tr>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
<th>Midpoint</th>
<th>Frequency</th>
<th>Cumulative Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>.5</td>
<td>15.5</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>.04</td>
</tr>
<tr>
<td>16</td>
<td>30</td>
<td>15.5</td>
<td>30.5</td>
<td>23</td>
<td>5</td>
<td>7</td>
<td>.10</td>
</tr>
<tr>
<td>31</td>
<td>45</td>
<td>30.5</td>
<td>45.5</td>
<td>38</td>
<td>24</td>
<td>31</td>
<td>.48</td>
</tr>
<tr>
<td>46</td>
<td>60</td>
<td>45.5</td>
<td>60.5</td>
<td>53</td>
<td>13</td>
<td>44</td>
<td>.26</td>
</tr>
<tr>
<td>61</td>
<td>75</td>
<td>60.5</td>
<td>75.5</td>
<td>68</td>
<td>6</td>
<td>50</td>
<td>.12</td>
</tr>
</tbody>
</table>

(a) (5 pts) Construct a relative frequency histogram using the information given in the table above.

(b) (5 pts) Construct an ogive using the information given in the table above.
5. Consider the following sample consisting of 20 numbers.

\[
\begin{array}{cccccccccccccccc}
20 & 23 & 24 & 24 & 25 & 26 & 28 & 29 & 31 & 32 \\
33 & 34 & 35 & 37 & 41 & 46 & 52 & 53 & 58 & 98
\end{array}
\]

(a) (6 pts) Given that \( \sum x = 749 \) and \( \sum x^2 = 34109 \) find the mean, variance and standard deviation for this sample.

\[
\bar{x} = \frac{749}{20} = 37.45
\]

\[
\text{Variance: } s^2 = \frac{34109 - \frac{749^2}{20}}{19} = 318.892
\]

\[
\text{St. Dev: } s = \sqrt{318.892} \approx 17.858
\]

(b) (8 pts) Find the \( Q_1, Q_2, Q_3 \) and the IQR for the data and then construct a box-and-whisker plot for the data.

\[
Q_2 = 32.5
\]

\[
Q_1 = 25.5
\]

\[
Q_3 = 43.5
\]

6. (8 pts)

(a) Ans: \( 17 \) Given a collection of data with lowest number 4 and highest number 100. What class width should be chosen if 6 classes are desired?

(b) Ans: \( 33 \text{rd} \) Given a collection of ordered data with 65 numbers, in what position is the median?

(c) Ans: \( 75^{th} \) What percentile is the third quartile \( Q_3 ? \)

(d) Ans: \( 37\% \) In a set of data, approximately what percentage of the data lie at or above the 63rd percentile?
7. (4 pts) Estimate the mean depth of ground water using the grouped data in the following table.

<table>
<thead>
<tr>
<th>Distance from ground to water level (ft), $x$</th>
<th>15 - 19</th>
<th>20 - 24</th>
<th>25 - 29</th>
<th>30 - 34</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of wells, $f$</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

\[
\bar{x} = \frac{(17)(3) + (22)(5) + (27)(4) + (32)(4)}{20} = 25.25
\]

8. (4 pts) A population is known to have a mean of 50 and a standard deviation of 6.

(a) Use Chebyshev’s theorem to find an interval that contains at least 75% of the data.

\[
\mu \pm 2\sigma \quad \Rightarrow \quad 38 \text{ to } 62
\]

(b) At least what portion of data is contained in the interval from 26 to 74?

\[
26 = \mu - 4\sigma \\
74 = \mu + 4\sigma
\]

\[
\therefore \text{ at least } 1 - \frac{1}{4^2} = \frac{15}{16} \text{ of data} \\
\text{(or at least } 93.75\% \text{ of data})
\]