1. (2 pts) A sociologist was interested in whether people’s attitudes about raising taxes to support the building of a new school were related to whether they had children in school. One hundred voters with children in the ninth grade were asked whether they supported the idea of higher taxes for the purposes of building a new high school. Five years later they were asked the same question. Is this a cross-sectional or longitudinal study? Explain.

2. (5 pts) In order to determine what voting aged Americans think of President G.W. Bush, a recent Gallup poll surveyed 1005 adult Americans and found that 90% of those surveyed approve of the way President G.W. Bush is handling his job.

(a) Describe the sample. Is the size of the sample given, if so what is it?

(b) Describe the parameter of interest. Is its value known, if so what is it?

(c) Describe the statistic of interest. Is its value known, if so what is it?

3. (4 points) (True or False)

(a) _____ The median is to the right of the mean in data that is skewed to the right.

(b) _____ The standard deviation of data closely bunched around its mean is smaller than the standard deviation of data that is much more dispersed.

(c) _____ The mean for a finite collection of data with an odd number of data is always one of the data values.

(d) _____ The median is not a resistant measure because it is influenced by extreme values.
4. The following represents time in minutes it took a group of 24 students to complete an English test.

30 36 37 37 38 39 40 41 42 43 44 45
46 49 52 53 54 56 61 68 77 79 83 99

(a) (3 pts) Construct a stem and leaf plot for this data using stems 3, 4, 5, 6, 7, 8, 9.

(b) (2 pts) Find the percentile rank of the time of 77 minutes.

(c) (2 pts) Find the third quartile $Q_3$ for the data.

(d) (4 pts) Given that $\sum x = 1249$ and $\sum x^2 = 72041$ for the data, find the mean and sample standard deviation.

(e) (2 pts) Does the data appear to be skewed. If so, in which direction? Explain.
5. Consider the following data of 28 numbers.

\[
\begin{align*}
30 & \quad 32 & \quad 38 & \quad 40 & \quad 51 & \quad 62 & \quad 68 & \quad 72 & \quad 74 & \quad 75 & \quad 75 & \quad 75 & \quad 76 & \quad 77 \\
80 & \quad 85 & \quad 88 & \quad 89 & \quad 90 & \quad 91 & \quad 92 & \quad 94 & \quad 99 & \quad 100 & \quad 110 & \quad 111 & \quad 115 & \quad 129
\end{align*}
\]

(a) (2 pts) Find the median of the data.

(b) (6 pts) Given that \( Q_1 = 70 \), and \( Q_3 = 93 \) find the IQR and construct both a boxplot and a modified boxplot for the data.

(c) (4 pts) Construct a relative frequency histogram for the data where the first class has limits 30–49, be sure to list the class boundaries.
6. A farmer is interested in the relationship between the amount of fertilizer in hundreds of pounds (x) and the number of soybeans produced in bushels per acre (y). The farmer obtained the following data.

<table>
<thead>
<tr>
<th>Amount of Fertilizer (x)</th>
<th>1.0</th>
<th>2.5</th>
<th>3.0</th>
<th>3.0</th>
<th>3.4</th>
<th>4.0</th>
<th>4.5</th>
<th>4.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bushels per Acre (y)</td>
<td>25</td>
<td>32</td>
<td>35</td>
<td>32</td>
<td>35</td>
<td>39</td>
<td>41</td>
<td>42</td>
</tr>
</tbody>
</table>

For this data: $\sum x = 26.2$, $\sum x^2 = 96.1$, $\sum y = 281$, $\sum y^2 = 10089$, $\sum xy = 967.1$.

(a) (4 pts) Find the equation of the least squares regression line.

(b) (2 pts) Use the regression line to predict how many pounds of fertilizer should be used to produce 37 bushels of soybeans per acre.

(c) (1 pt) From just looking at the data, do you expect a positive or negative correlation coefficient? Explain.

(d) (3 pts) Find the correlation coefficient for this data and explain what it means concerning the goodness of the linear fit.