Form A

1–4. A 1997 net income report (in millions of dollars) for Fifth Third Bancorp is given below.

<table>
<thead>
<tr>
<th>yrs since 1992</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net Income</td>
<td>172.0</td>
<td>206.2</td>
<td>244.4</td>
<td>287.6</td>
<td>335.0</td>
</tr>
</tbody>
</table>

1. [10 points] Sketch a scatterplot of this data below. Clearly label the axes with tickmarks and values to give a sense of the scale of the data. Also label the axes with appropriate variable names and units of measurement.

![Scatterplot](image)

2. [15 points] Find a formula for the linear function that agrees exactly with the data for the years 1993 and 1995. From this model, estimate net income for the bank in the year 2002.

The data for years 1993 and 1995 give the points (1, 206.2) and (3, 287.6). The slope of the line between these points is \((287.6 - 206.2)/(3 - 1) = 40.7\). Since the linear function has the form \(I(t) = mt + b\), we now have \(I(t) = 40.7t + b\). Using the first of the two points, we also then have \(206.2 = 40.7\cdot1 + b\), so \(b = 165.5\). The complete model formula is \(I(t) = 40.7t + 165.5\). This gives as an estimate for net income in 2002 \(I(10) = 40.7\cdot10 + 165.5 = 572.5\) million.

3. [10 points] Find the linear regression model. From this model, estimate net income for the bank in the year 2002.

The calculator gives the regression model as \(I(t) = 40.74t + 167.56\). This provides the following estimate: \(I(10) = 40.74\cdot10 + 167.56 = 574.96\) million.
4. [15 points] Discuss the reliability of the estimates you calculated in #2 and #3.

Both estimates are quite close, but the one in #3 is likely more accurate because it is generated from the regression model which uses all of the original data; the function in #2 uses only two data points. Still, both estimates suffer from a high degree of extrapolation: Fifth Third’s net income levels might not maintain the same linear trends from the early 1990’s all the way into the year 2002.

5. [15 points] According to Bankrate.com, the best current (February 4, 2002) 2-year CD is offering 4.16% as a nominal annual rate, compounded monthly. If you were to purchase $5,000 worth in this investment, how much would you earn at maturity in two years?

Use the compound interest formula $A = P(1 + r/m)^{mt}$; so $A = 5000(1 + 0.0416/12)^{12 \cdot 2} = $5433.01 is the amount of money in the account in two years. Hence, the interest earned amounts to $433.01.

6–8. According to the annual Hawaii Youth Risk Behavior Survey, the following percentages of 12th graders in Hawaii admitted smoking marijuana at least once within the last 30 days.

<table>
<thead>
<tr>
<th>year</th>
<th>1987</th>
<th>1989</th>
<th>1991</th>
<th>1993</th>
<th>1995</th>
</tr>
</thead>
<tbody>
<tr>
<td>percent</td>
<td>17.2</td>
<td>13.9</td>
<td>14.6</td>
<td>17.9</td>
<td>23.5</td>
</tr>
</tbody>
</table>

6. [10 points] Assign appropriate variable names and indicate units of measurement.

$t =$ years since 1987

$P(t) =$ % of 12th graders admitting to smoking marijuana in the last 30 days

7. [10 points] View a scatterplot of this data on your calculator. What type of function model is appropriate to use here, and why?

The scatterplot is concave up, showing a decrease followed by an increase in values: use a quadratic model.

8. [15 points] Find the regression model of the type you indicated in #7 above and write it below. What does the model predict for the percentage of 12th graders in Hawaii in 2001 who will admit to smoking marijuana at least once within the last 30 days this year? Is this a reasonable estimate?

The regression model is $P(t) = 0.36t^2 - 2.08t + 17.01$. Therefore, the percentage of 12th graders in Hawaii in 2001 who will admit to smoking marijuana at least once within the last 30 days in 2001 is $P(14) = 59.2\%$. This value is unreasonably high, a poor estimate due to gross extrapolation.
1–6. U.S. News & World Report (June 4, 1989) reported the following data on the number of deaths in the U.S. each year caused by lightning strikes.

<table>
<thead>
<tr>
<th>years since 1940</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
<th>49</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of deaths</td>
<td>340</td>
<td>268</td>
<td>219</td>
<td>181</td>
<td>129</td>
<td>149</td>
<td>122</td>
<td>91</td>
<td>74</td>
<td>74</td>
<td>70</td>
</tr>
</tbody>
</table>

1. [10 points] Consider first only the data from 1940 to 1960. Sketch a scatterplot of these five data points below. Clearly label the axes with tickmarks and values to give a sense of the scale of the data. Also label the axes with appropriate variable names and units of measurement.

![Scatterplot of data](image)

2. [15 points] Find a formula for the linear function that agrees exactly with the data for the years 1950 and 1955. From this model, estimate the number of lightning-related deaths for the year 1958.

The data for years 1950 and 1955 give the points (10, 219) and (15, 181). The slope of the line between these points is \( \frac{219 - 181}{10 - 15} = -7.6 \). Since the linear function has the form \( D(t) = mt + b \), we now have \( D(t) = -7.6t + b \). Using the first of the two points, we also then have \( 219 = -7.6 \cdot 10 + b \), so \( b = 295 \). The complete model formula is \( D(t) = -7.6t + 295 \). This gives as an estimate for the number of lightning-related deaths in 1958: \( D(18) = 158.2 \), or 158 deaths.

3. [10 points] Find the linear regression model based on all five data points. From this model, estimate the number of lightning-related deaths for the year 1958.

The calculator gives the regression model as \( D(t) = -10.18t + 329.2 \). This provides the following estimate: \( D(18) = 145.96 \), or 146 deaths.

4. [15 points] Discuss the reliability of the estimates you calculated in #2 and...
Both estimates are quite close, but the one in #3 is likely to be more accurate because it is generated from the regression model which uses all of the original data; the function in #2 uses only two data points.

5. [10 points] Now view a scatterplot of all the data (1940-1989) given in the table. What type of function model is appropriate to model all this data, and why?

The shape of the scatterplot (steady decline with end behavior tending to 0) indicates that an exponential function will provide the best model, but other models are also possible.

6. [10 points] Find the regression model of the type you indicated in #5 above and write it below. What does this model predict for the number of lightning-related deaths for the year 2002? Is this a reasonable estimate?

The calculator gives \( D(t) = 306.34(0.94794)^t \) as the regression model. It provides the estimate \( D(62) = 40.6 \) in 2002, we estimate 41 lightning-related deaths. The estimate is not very reliable, since there is a fair amount of extrapolation error involved.

7. [15 points] What type of function is described by the input/output table below, and how can you tell? You may not answer by graphing the points or using a regression formula; work from the table exclusively.

<table>
<thead>
<tr>
<th>x</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>8</td>
<td>7</td>
<td>11</td>
<td>20</td>
<td>34</td>
<td>53</td>
<td>77</td>
<td>106</td>
<td>140</td>
<td>179</td>
</tr>
</tbody>
</table>

Inputs are equally spaced, so we can look at first and second differences:

\[
\begin{array}{ccccccccccc}
8 & 7 & 11 & 20 & 34 & 53 & 77 & 106 & 140 & 179 \\
1st: & -1 & 4 & 9 & 14 & 19 & 24 & 29 & 34 & 39 \\
2nd: & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 \\
\end{array}
\]

Constant second differences mean that the function is quadratic.

8. [15 points] A young girl receives a $6000 inheritance which her parents will invest towards her college education. They choose an investment that earns interest at 6% compounded annually. How long will it take for the investment to double in value?

From the compound interest formula \( A = P(1 + \frac{r}{m})^{mt} \), we get

\[
12000 = 6000(1 + 0.06/1)^{1t} \\
12000 = 6000(1.06)^t
\]
\[2 = 1.06^t\]
\[\ln 2 = t \ln 1.06\]
\[t = \frac{\ln 2}{\ln 1.06}\]
\[t = 11.89\]

It will take nearly 12 years for the investment to double in size.