2) The histogram of a set of discrete numerical data is shown below.

![Histogram of x]

(a) Find the mode and the median.

\[ \text{median} \Rightarrow 55 \times 0.5 = 27.5 \rightarrow 28^{th} \rightarrow \frac{3}{2} \text{ (median)} \]

mode \Rightarrow \text{(most frequent data point)} \Rightarrow \frac{3}{2} \text{ (mode)}

(b) Find the sum of the data points? \( \left( \sum_{i=1}^{n} x_i \right) \)

What is the sum of the squares of deviations?

\[ \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 \]

\[ \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x} \times n = \frac{\sum_{i=1}^{n} x_i}{n} = 3.364 \times 55 = 189.02 \]

\[ \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 = s^2 \times (n-1) = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1} \times n-1 \]

\[ = (1.532)^2 \times (55-1) = 126.74 \]
c) The 100\textsuperscript{th} percentile is 4.5. What is \( p \)?
(give the answer in rational form)

\[ n \times p = 55 \]

43\textsuperscript{nd} \to 4

44\textsuperscript{th} \to 5

43 \times p = 55 \implies 55 \times p = 43 \implies p = \frac{43}{55}

4) What is the 1\textsuperscript{st} quartile (\( Q_1 \))?

\[ 1\text{QR} = Q_3 - Q_1 \]

\[ Q_3 \implies n \times 0.75 = 55 \times 0.75 = 41.25 \to 42 = 43.75 \text{rd} \]

\[ Q_1 \implies n \times 0.25 = 55 \times 0.25 = 13.75 \to 14\text{th} = 2 \]

\[ 1\text{QR} = 4 - 2 = 2 \]

a) What is the relative frequency of 2?

\[ \text{frequency of 2} = 12 \]

\[ \text{relative frequency} = \frac{12}{55} \]
1) 20 data points, some sums and the correlation coefficient are given below for a data set.

<table>
<thead>
<tr>
<th>x</th>
<th>20</th>
<th>55</th>
<th>44</th>
<th>35</th>
<th>52</th>
<th>63</th>
<th>39</th>
<th>69</th>
<th>75</th>
<th>95</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>33</td>
<td>64</td>
<td>62</td>
<td>54</td>
<td>55</td>
<td>78</td>
<td>64</td>
<td>82</td>
<td>85</td>
<td>90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>x</th>
<th>86</th>
<th>72</th>
<th>28</th>
<th>88</th>
<th>84</th>
<th>39</th>
<th>18</th>
<th>25</th>
<th>65</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>88</td>
<td>75</td>
<td>38</td>
<td>95</td>
<td>81</td>
<td>52</td>
<td>42</td>
<td>39</td>
<td>76</td>
<td>75</td>
</tr>
</tbody>
</table>

\[
\sum_{i=1}^{20} x_i = 1123, \quad \sum_{i=1}^{20} y_i = 1328, \quad \sum_{i=1}^{20} x_i^2 = 74829, \quad \sum_{i=1}^{20} y_i^2 = 97,478, \quad \sum_{i=1}^{20} x_i y_i = 83078, \quad r = 0.951.
\]

a) What is the value of \( s_{xx} \)?

\[
s_{xx} = \frac{\sum x_i^2 - (\sum x_i)^2}{n} = \frac{74,829 - 63,056}{11,773} = 11,773
\]

b) What does \( r = 0.951 \) tell you? Explain.

There is a strong positive correlation between the \( x \) and \( y \) variable.

c) Find the mean and standard deviation for the \( x_i \) and the mean for the \( y_i \).

\[
\overline{x} = \frac{1123}{20} = 56.15 \quad \overline{y} = \frac{1328}{20} = 66.4
\]

\[
s = \sqrt{\frac{1123 - 56.15^2}{20}} = 7.423
\]

1) What is the value of \( \bar{A} \) (the sum of squares of \( y_i \))?

\[
r = 0.951 = \frac{\sum x_i y_i - (\sum x_i \sum y_i)}{n}
\]

\[
\sqrt{\frac{\sum x_i^2 - (\sum x_i)^2}{n}} \sqrt{\frac{\sum y_i^2 - (\sum y_i)^2}{n}} \sqrt{\overline{A} - (\sum y_i)^2}
\]

\[
= \sqrt{\frac{74,829 - (1123)^2}{20}} \sqrt{\overline{A} - (66.4)^2} = A = 881,180
\]
e) Obtain the regression line for these data using the method of least squares:

\[ \hat{b}_1 = \frac{S_{xy}}{S_{xx}} = 0.72 \]

\[ \hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x} = 66.4 - 0.72 \times 56.18 = 25.97 \]

\[ \hat{y} = 25.97 + 0.72x \]