Online library of Quality, Service Improvement and Redesign tools

Scatter diagram (correlation)
Scatter diagram (correlation)

What is it?

It is important to test a theory before changes are made to ensure you avoid quick fix solutions to problems that you think you know the cause of (but don’t actually). A scatter diagram can help you identify the true strength of the relationship between the cause and effect of two variables and factors. This can help to ensure that you are focusing improvement efforts on the true cause of a problem.

When to use it

You can use a scatter diagram to determine the significance and strength of a relationship between two variables before making changes in practice. If the factors are related, evaluate the relationship by visually interpreting the width and tightness of the scatter.

How to use it

1. Using a data collection sheet, collect 50 or more samples of paired data (data related to both variables). You must have two measurements for each observation point or item – for example, if your theory is that there is a relationship between the age of patients and the length of their recovery time, you would collect sets of paired data relating to 50 or more patients ie the age of each patient and the length of their recovery time.

2. Draw lines on the diagram representing the suspected causes on the x-axis (horizontal) and the suspected effect on the y-axis (vertical). Determine the measures and increments and label each line. In the example above, we suspect that patient age has an impact on recovery time, so patient age would be shown on the x-axis and recovery time on the y-axis.

3. Plot the results on the scatter diagram.

4. Interpret the results by visual interpretation, considering the width and tightness of the scatter.
**Example**

You want to test the theory of a possible relationship between the age of a patient and recovery time following an operation.

1. Collect the data

**Figure 1: Data table**

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age</th>
<th>Recovery time (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>40</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>55</td>
<td>6</td>
</tr>
<tr>
<td>E</td>
<td>63</td>
<td>7</td>
</tr>
</tbody>
</table>

2. Produce the diagram

3. Plot the results

**Figure 2: Plot results on diagram**

Plot all of the paired measurements onto the scatter diagram. If data values are repeated and fall on the same point, draw a circle around that point as many times as it is repeated.
4. Interpret the results

The scatter diagram cannot prove a cause and effect relationship but it does suggest the strength of a relationship between two variables. The stronger the relationship suggested, the greater the likelihood that a change in one of the variables will affect a change in the other.

The five diagrams below and overleaf show the various patterns that scatter diagrams can have, together with explanations of how you would interpret each pattern.

**Figure 3: A positive correlation**

Recovery times are likely to increase as patient age increases.

**Figure 4: A possible positive correlation**

Other variables in addition to patient age may affect recovery time.

**Figure 5: No correlation**

There is no demonstrated connection between patient age and recovery time.

**Figure 6: A possible negative correlation**

Other variables in addition to patient age may affect recovery time.
**Figure 7: A negative correlation**

This diagram suggests that recovery time is likely to decrease as patient age increases.

---

**What next?**

If the scatter diagram shows an effect then you can confidently continue with your improvement effort. If the results show no relationship, you could test another theory. This way you have avoided jumping to conclusions and making changes that may make matters worse.

A cautionary note: a scatter diagram shows patterns in data and can help you to indicate the existence of a relationship, but the diagram may not confirm for sure that there is a cause and effect relationship between the two variables as there may be other factors that affect the variables tested.

**Additional resources**


---

Written by the ACT Academy for their Quality, Service Improvement and Redesign suite of programmes.

Contact: nhsi.act@nhs.net