



## Visual Displays- Two Variables

---

### Before you Watch

Visual displays if done well can more effectively or succinctly communicate key information than text.

Visual displays in Statistics are referred to as graphs, and just as the type of graph, or visual display, we use to represent a *single variable* is determined by the **variable** type, so too the variable types of the **pair** of variables dictates the visual display that is most appropriate when we wish to display pairs of variables.

It is very important to clearly and unambiguously name the chart, and label what is being plotted (graphed or displayed) on each of the vertical and horizontal axes.

Graphs summarise data and in doing so provide information. Patterns in the graphs of two variables show information about the **relationship or association** between the two variables.

The appropriate graphs used to display each combination of variable types are:

Both numerical                               => **scatterplot**

Both categorical                              => **clustered bar chart**

1 numerical and 1 categorical => **histograms vertically aligned (or side by side boxplots)**

### The Video Content

There are times when you'll measure two variables upon each individual item under investigation. How you report upon pairs of variables depends upon their types.

You may have two categorical variables, or two numerical variables, or one of each type.

Is there a relationship between suffering from tonsillitis and snoring? A study into this measured two categorical variables across 50 children, where one variable is tonsillitis, with possible responses 'yes' or 'no', and the other variable is Snore, with possible responses 'yes' or 'no'.

So for each child, we have information on whether or not they snore and whether or not they have suffered tonsillitis. Here are the data for the first ten children in the study. Each row represents a child.

As the two variables are categorical you can consider cross-tabulations. There are 17 children who responded NO to Tonsillitis and NO to Snoring. 3 children responded NO to Tonsillitis and YES to Snoring. Clustered bar charts help to visually assess if the response to tonsillitis, yes or no, is associated with snoring or not.

Of all those who responded as not having suffered tonsillitis, almost all also indicated they don't snore. Of all those who responded as having suffered tonsillitis, did almost all indicate they don't snore too? NO, the responses are equally shared between snoring and not snoring.

So, it depends upon the response to tonsillitis as to how people responded about snoring. This indicates there is a relationship or association between tonsillitis and snoring. Distribution of snoring responses differs depending on the response to tonsillitis.

What about when we have two numerical variables, for example the amount of a catalyst used, in grams, and reaction time, measured across multiple experiments. Is the amount of the catalyst used (which is measured in grams) associated with the reaction time?

The experiment was repeated and also with varying amounts of catalyst to produce the following results from 21 experiments. We have 21 pairs of values, for the two variables 'catalyst amount' and 'reaction time'.

Can you see a relationship between the two variables? What if we plot the data? We may use a scatterplot to do so. We plot each pair of values by a single point on this graph. Repeating for all points we get our final graph and we can see that generally as the catalyst amount increases the reaction time decreases.

What about when we have one numerical variable and one categorical variable? Perhaps an engineer wants to test the efficiency (the numerical variable) of his



newly-designed mechanical device against the current device (the categorical variable would identify the device tested as either the current one or new one).

Or, perhaps a chemist or chemical engineer has designed a new epoxy resin and wishes to compare the pull off force (that would be the numerical variable) of two groups of epoxy resin (a categorical variable represents whether resin is the current one or newly designed one).

Consider the situation of a randomised controlled trial, where a sample of people from the population with a disease are randomly allocated to one of two groups. The first group receives a new medication, whilst the other group uses an existing treatment.

After a period of time, all individuals across both groups are tested and a measurement recorded indicating percentage improvement (that's the numerical variable). The categorical variable is 'Group' with responses 'Treatment' if receiving the new medication, or 'Control' if only the existing treatment is received. We could consider vertically-aligned histograms.

We compare these using the 3Ss - Shape, Centre and Spread They are both approximately bell-shaped, or normal. The group corresponding to the New Medication has a higher mean percentage improvement than the group undergoing the current treatment, and we can consider the means and standard deviations.

The methods shown have been exploratory or descriptive analyses, we can also perform formal tests and use statistical models for each situation, as described in further videos in this series.

## Now What?

Now that you have learnt about visual displays for two variables you should consider the methods of testing for associations (relationships) between two variables in the [Hypothesis test](#) videos (Comparing two groups: t tests and [Chi-squared test](#)) and the [Correlation](#) and [Regression](#) videos .



## ***But, when am I going to use this?***

Across all fields of study we measure characteristics, or random variables. We often are interested in more than one measure, and also whether the two measures (variables) are related. For example, is a chemical reaction time associated with the amount of a catalyst used, or is response time, or even cognitive load associated with age, or sex? Visually displaying data arising from pairs of variables well enables the identification of patterns and effective communication of key information. This may be used in organisational reports, presentations to key stakeholders, reporting upon research and investigations or any other similar forum or endeavour. Say less, show more!

## ***Other Links***

- Learn about the importance of labels at [http://www.gcse.com/science/AQA\\_ISA\\_graphs.htm](http://www.gcse.com/science/AQA_ISA_graphs.htm)

