

2.2: Scatterplots

When we have bivariate data, the first thing we should always do is draw a graph of this data, to get some feeling about what the data is showing us and what statistical methods it makes sense to try to use. The way to do this is as follows

[def:scatterplot] Given bivariate quantitative data, we make the **scatterplot** of this data as follows: Draw an x - and a y -axis, and label them with descriptions of the independent and dependent variables, respectively. Then, for each individual in the dataset, put a dot on the graph at location (x, y) , if x is the value of that individual's independent variable and y the value of its dependent variable.

After making a scatterplot, we usually describe it qualitatively in three respects:

[def:scattershape] If the cloud of data points in a scatterplot generally lies near some curve, we say that the scatterplot has [approximately] that **shape**.

A common shape we tend to find in scatterplots is that it is **linear**

If there is no visible shape, we say the scatterplot is **amorphous**, or **has no clear shape**.

[def:scatterstrength] When a scatterplot has some visible shape – so that we do not describe it as amorphous – how close the cloud of data points is to that curve is called the **strength** of that association. In this context, a **strong** [linear, *e.g.*,] association means that the dots are close to the named curve [line, *e.g.*,], while a **weak** association means that the points do not lie particularly close to any of the named curves [line, *e.g.*,].

[def:scatterdirection] In case a scatterplot has a fairly strong linear association, the **direction** of the association described whether the line is increasing or decreasing. We say the association is **positive** if the line is increasing and **negative** if it is decreasing.

[Note that the words *positive* and *negative* here can be thought of as describing the *slope* of the line which we are saying is the underlying relationship in the scatterplot.]

This page titled [2.2: Scatterplots](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Jonathan A. Poritz](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.