3d – A Guide to
Scatter and Line Graphs

Royal Geographical Society with IBG
Advancing geography and geographical learning

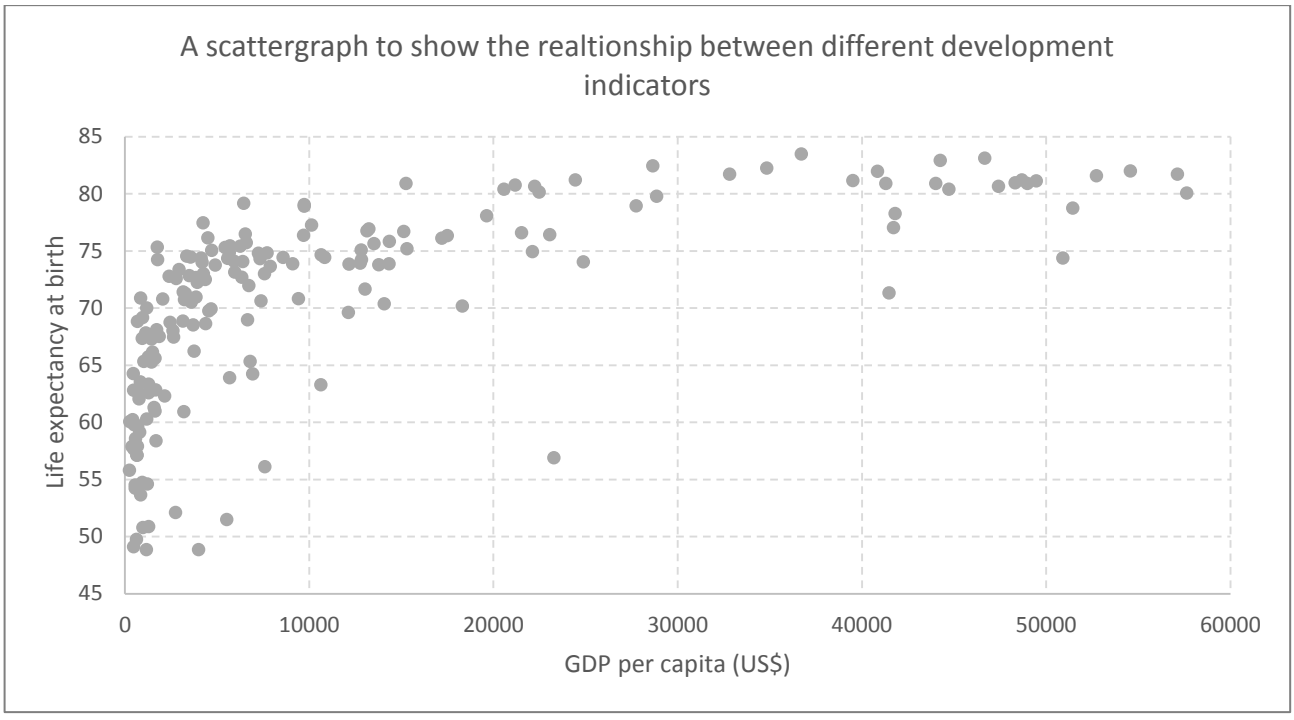
Scatter graphs and line graphs are used to show the potential correlation between two different variables.

Scatter graphs

Scatter graphs can be used when the data from both variables under investigation is continuous. In order to be able to judge whether there is a correlation between the two variables, the researcher should try to include as many points as possible: fewer than thirty points can make it difficult to draw any meaningful conclusions from the data.

Why would we use a scatter graph?

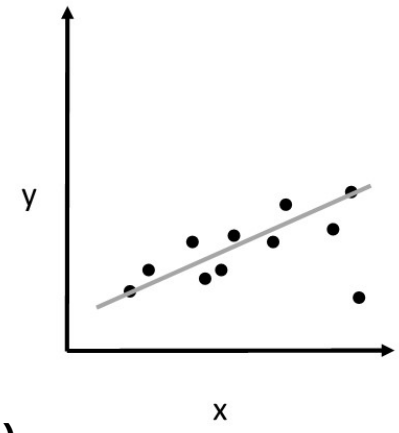
Researchers tend to use scatter graphs when they are faced with very large sets of data. Scatter graphs allow researchers to identify anomalies in the data more easily as well as the overall trend and relationship between the variables. The strength of any relationship can also be visualised.



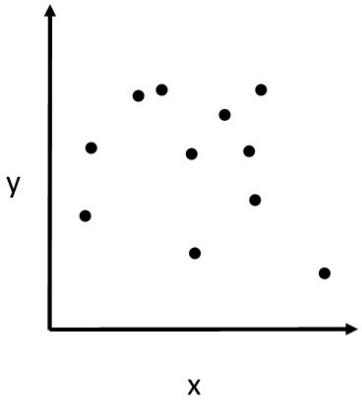
A line of best fit can also be drawn on the graph if the researcher suspects that there is a correlation. This line does not necessarily have to go through any points at all, but should fall in the middle of the points and aim to have an equal number of points above and below the line. An outlying point may be interpreted as an anomaly. The angle and direction of the line of best fit can tell the researcher what type of correlation the data has.

Nuffield Foundation

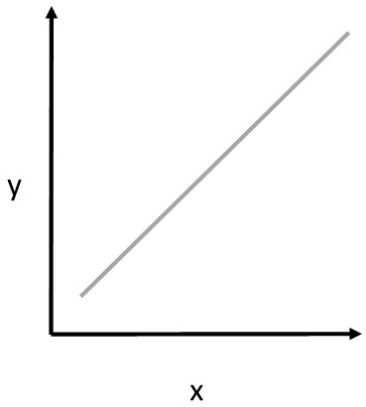
This project was funded by the Nuffield Foundation, but the views expressed are those of the authors and not necessarily those of the Foundation.



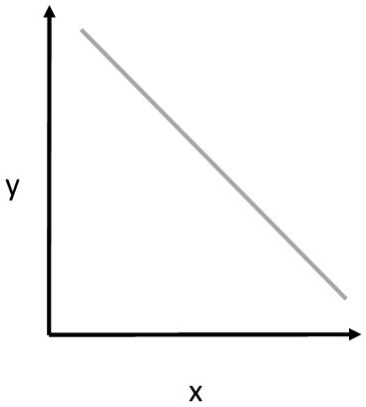
a)



b)



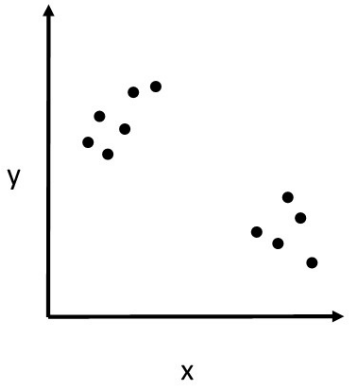
c)



d)

a) Drawing a line of best fit. The outlying point may be thought of as an anomaly.
b) No correlation. It is not possible to draw a line of best fit by eye.
c) Strong positive correlation
d) Strong negative correlation

The nature of the data used in a scatter graph may mean that a line of best fit is not possible. Instead, clusters of data may be identifiable and rings may be drawn on the graph to show how one idea is true for one scale of data and another is true at the other end of the scale.

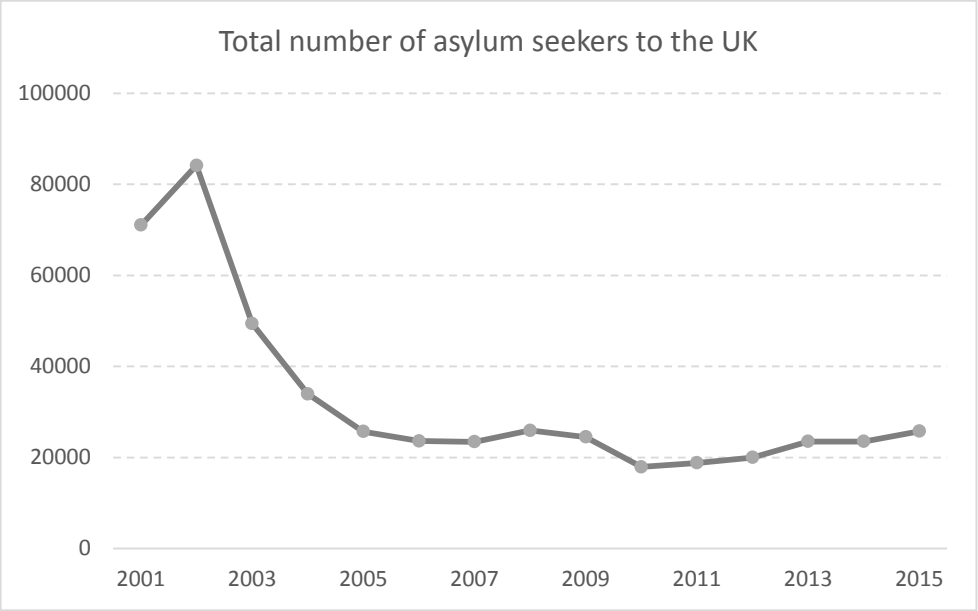


Line graphs

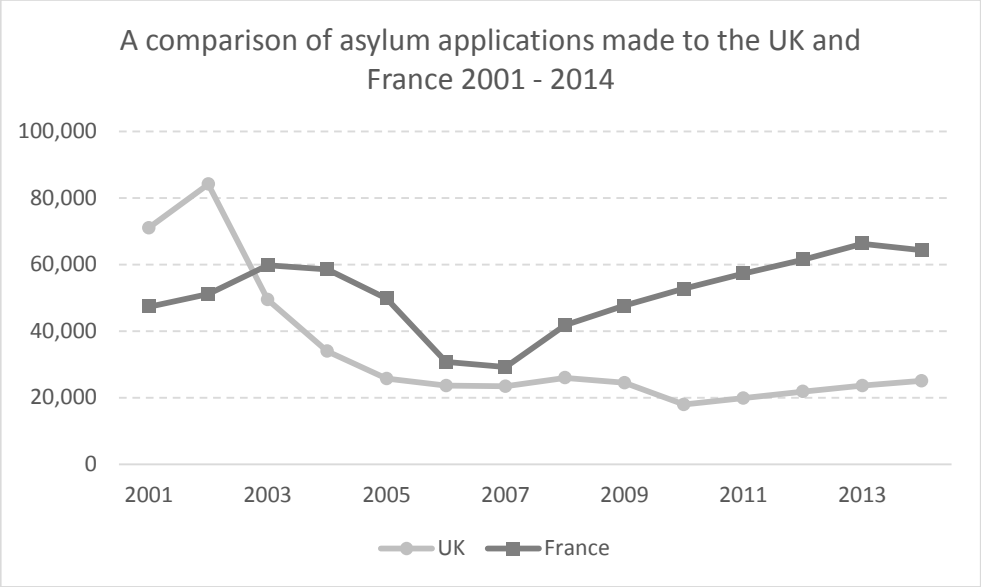
Line graphs are used to show the change in a relationship between two variables over a form of continuous data such as distance, or most often, time. Points are plotted in the same way as one would for a scatter graph but then joined together in sequence.

Why would we use a line graph?

Researchers would use a line graph to show the nature of a relationship, not whether there is a relationship or not between two sets of data. They can be used to instantly see the 'journey' made by a geographical entity over time or distance which can be useful for creating a narrative around a geographical idea.

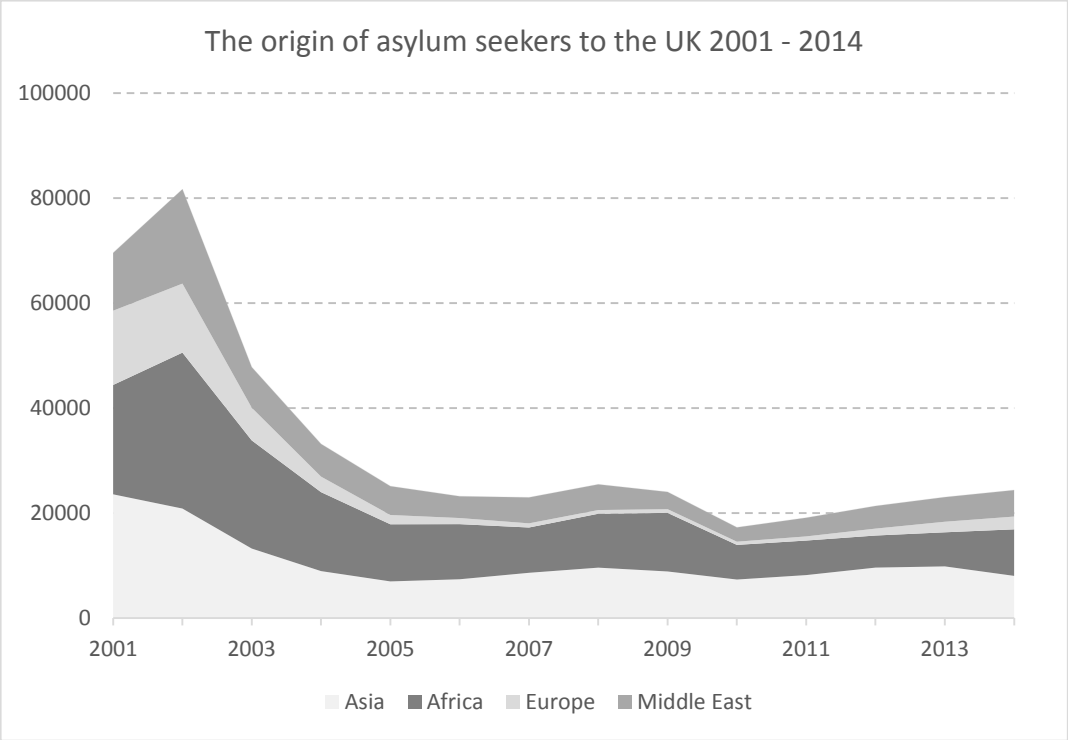


More than one line can appear on a graph, along with a key, to show different sets of data.



Compound line graphs

For some data there may be a total value, and further component values, that can be shown on the same graph. For example, the total number of asylum seeker applications a country receives may change over time but the places from which those asylum seekers come will vary in different years.



Using Logarithmic Scales

Sometimes the spread of the data is so large between the minimum and maximum values that producing a graph with a meaningful scale on its axes can be challenging. In these cases, a logarithmic scale might be helpful. A logarithmic scale is one where values along an axis are compressed to make the whole spread of data fit into one graph (known as a log graph). Usually this compression occurs by powers of ten, and instead of the axis showing units going up one at a time in a linear fashion (0, 1, 2, 3, 4 etc), the units go up in non-equal parts (1, 10, 100, 1000, 10000 etc). If both sets of data have a large spread, both sets of axes can be converted into a log scale (and the resulting graph is known as a log-log plot).

Drawing a graph that uses a logarithmic scale can be done using online downloads and packages designed to work with very large data sets. If you wish to draw a logarithmic graph by hand, log graph paper can be found online and printed off. Plotting the points on a logarithmic scale is done in exactly the same way as when using a normal scaled axis: the researcher needs to simply take a little more time as it can be easy to misplace a point using these scales.