

Direct data sources

Direct data sources are sources that provide the data gatherer with original data.

Questionnaires

Questionnaires are often used to collect data from individuals. They can be hard copy or completed online. Questionnaires can make it easier to analyse information because all respondents are asked the same questions.

Interviews

Interviews allow you to collect more data from people as responses can be given in greater depth.

Observation

Data gatherers observe what is happening during an activity, then record and analyse the resulting data.

Data logging

This is the use of sensors to produce data that can be gathered and interpreted.

Uses of direct data

An example of use of direct data could be planning a new high speed train line

The government are investigating the feasibility of introducing a new high speed rail link between the capital and a major industrial city. Before they decide whether to proceed or not, they will need to collect some direct data.

This direct data will include:

- the time it takes to get from the capital to the other city using the existing rail line
- the number of trains and passengers who use the existing rail line
- how many passengers would use the new system
- what people who live on or near the existing route think about the effect it would have on their environment.

Here are some examples of how the data could be collected.

The time it takes to get from the capital to the other city using the existing rail line:

This information can be collected from existing train timetables, however this method would not be using a direct data source. Original data could be collected by actually travelling on trains periodically and timing the journeys, but this might not be practical given the time it would take.

The number of trains and frequency on the existing line:

The suggested method to be used is a data logger. A sensor is placed on the rail line. This sensor is attached to a roadside data logger. As trains pass over the sensor, their speed, time of day, number of carriages and frequency are logged. The advantage of a data logger is that it gathers physical data automatically.

The number of passengers:

The method could be to use infra-red sensors fitted around each door on the train to count the number of passengers getting on and off the train at each station. From these it can be calculated how many passengers are on the train at any point along its route. The data is fed back to a microprocessor.

How many passengers would use the new system:

This could make use of questionnaires: passengers on the existing route and airline passengers in the capital are asked to complete the questionnaires.

The advantage of questionnaires is that they can be collected and analysed reasonably quickly. The disadvantage is that only a proportion are returned, making the sample size quite small.

What local residents think:

Face-to face interviews would be best. The advantage of interviews is that they may gather some unexpected data and obtain personal attitudes that a simple questionnaire would not. However, it takes time for many interviews to take place.

Indirect data sources

Indirect data sources are third-party sources that the data gatherer can obtain data from.

Electoral register

This is a list of adults who are entitled to vote in a local or national election. An edited version of the register can be purchased and used for any purpose.

Businesses collecting personal information

Businesses sell the information that they collect from their customers. For example when someone purchases something online they are often asked to tick a box authorising the business to share this with other organisations. Customers often provide personal information that has a commercial value. Businesses use this information to create mailing lists that can be purchased by any other organisation/individual to send emails or even brochures through the post.

Uses of indirect data

Apart from elections and other government purposes, the electoral register can only be used to select individuals for jury service or by credit reference agencies. These agencies are allowed to buy the full register to help them check the names and addresses of people applying for credit. They are also allowed to use the register to carry out identity checks in an attempt to deal with money laundering.

Businesses which collect personal information often use it to create mailing lists that they then sell to other organisations, which are then able to send emails or even brochures through the post.

Any organisation that provides data or information to the general public for use by them can be said to be an indirect source.

Another scenario could be studying pollution in rivers. Direct data sources could be used, of course; questionnaires could be handed out to local landowners and residents in houses near to the river, asking about the effects on them of the pollution, and they could also be interviewed. Computers with sensors could be used to collect data from the river. However, indirect data sources could also be used; documents may have been published by government departments showing pollution data for the area.

Advantages of direct data sources

- Only as much or as little data is gathered as needed
- Exactly where the data came from, and therefore how reliable it is, is known
- There may be an opportunity to sell the data for other purposes
- Gathering data directly addresses specific issues, as the data gatherer controls the methods of collecting the data to fit their needs.

Disadvantages of direct data sources

- Data gathering may be expensive as other companies may have to be hired to get it.
- It may involve having to buy equipment such as data loggers and computers
- It may not be possible to gather original data due to the time of year e.g. winter snowfall data may be required but it is now the middle of summer
- Compared to indirect data sources, using direct data sources may be very expensive in preparing and carrying out the gathering of data. Costs can be incurred in, for example, producing the paper for questionnaires, or the equipment for an experiment
- It takes longer to gather data than to acquire data from an indirect data source
- By the time the project is complete the data may be out-of-date
- The sample size may be small.

Advantages of indirect data sources

- Indirect data sources may allow a larger set of data to be examined using less time and money than direct data collection would require
- The use of indirect data sources allows data to be gathered from subjects (e.g. people) to which the data gatherer does not have physical access
- A larger sample size can be used. Direct data gathering can have limitations due to the availability of the people being interviewed, but by using indirect data sources, the size of the sample can be increased giving rise to greater confidence in the findings
- Using indirect data sources can be done at a relatively low cost, although this varies.
- Quite often the data can be in an easily accessible location such as the internet whereas for direct data sources, travelling expenses and time taken to collect data can be great
- Information can be of a higher quality. Data collected indirectly has already been collated and grouped into meaningful categories and, for example, poorly-written responses to questionnaires or interview transcripts do not have to be read through to create the data source.

Disadvantages of indirect data sources

- The various purposes for which data was collected originally may be quite different to the purpose of the current research and unnecessary data may need to be filtered out
- There may be no data available – the data required has simply never been recorded
- There may be sampling bias – data from only one section of the community (whether it is based on educational level, level of income etc.) may have been collected but what is required is data from a representative cross-section of the community.