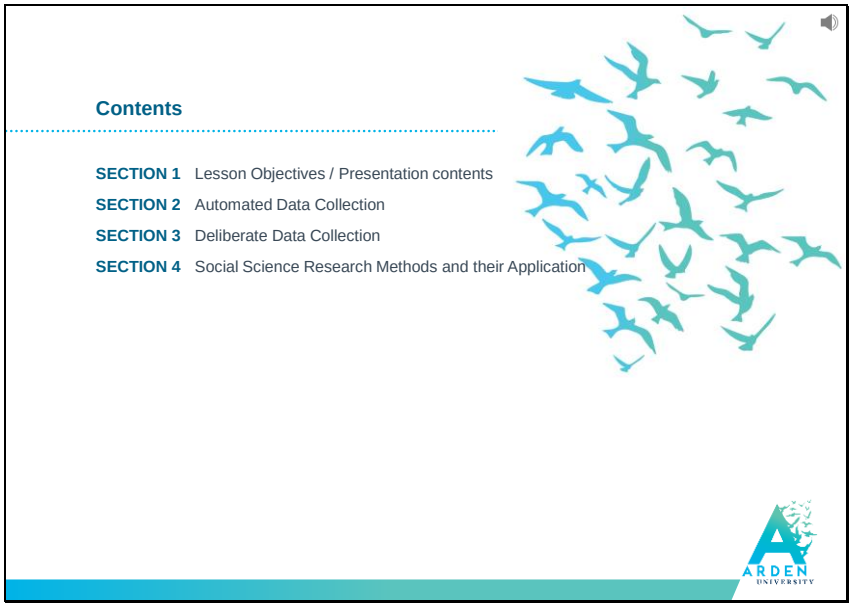
Slide 1

Welcome to lesson four the data design module. In this lesson we'll be exploring designed data collection, particularly surveys and interviews.

Slide 2

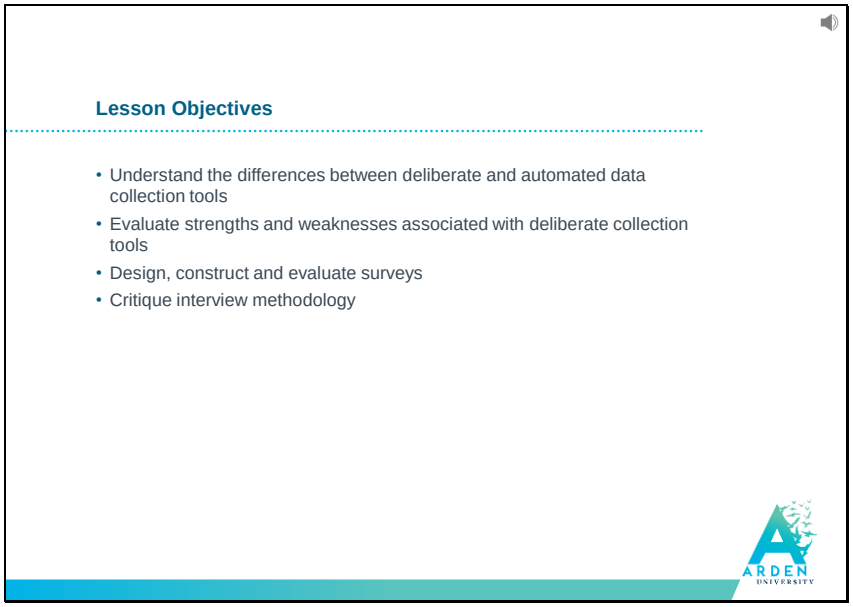


In this particular presentation we're going to explore the lesson objectives and the content of this presentation. We'll have a look at what mean by automated data collection to include the positives and negatives. We'll also explore deliberate data collection methods, again considering their strengths and weaknesses. Finally, we'll explore social research methods and their application in a little more detail.

Slide 3

So firstly let's explore the objectives of this lesson.

Slide 4



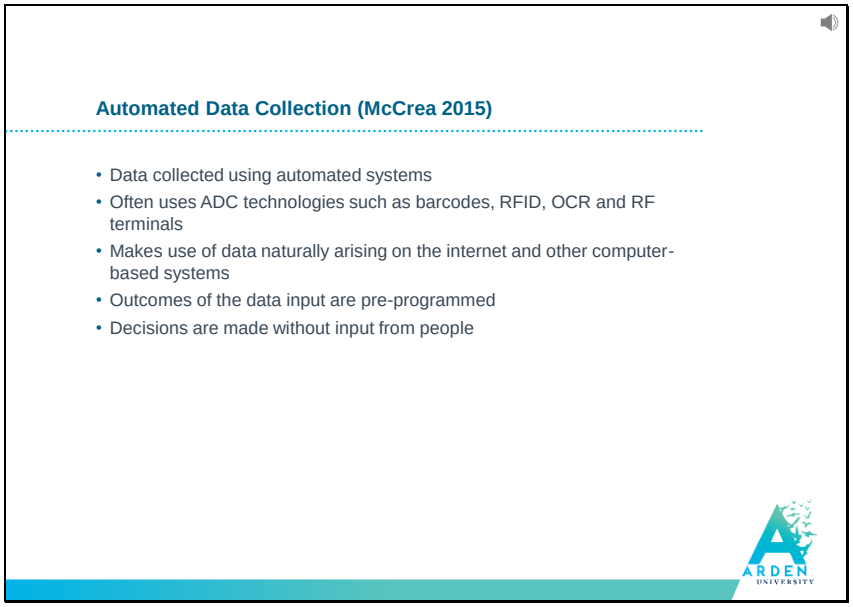
The first objective is to understand the difference between deliberate and automated data collection tools and we'll have a look at some of the different tools that are available and we'll look at evaluating their relative merits.

We'll then look in more detail at the deliberate data collection tools that we're exploring in this lesson, namely surveys and interviews. Again, we'll have a look at their relative merits. We'll explore how to design, construct and evaluate surveys through a series of activities, and then we'll do the same through the interview methodology.

Slide 5

So in order to fully understand these deliberate data collection tools that we're looking at, let's have a look at what we mean by automated data collection initially. This will be explored further in a later lesson as well.

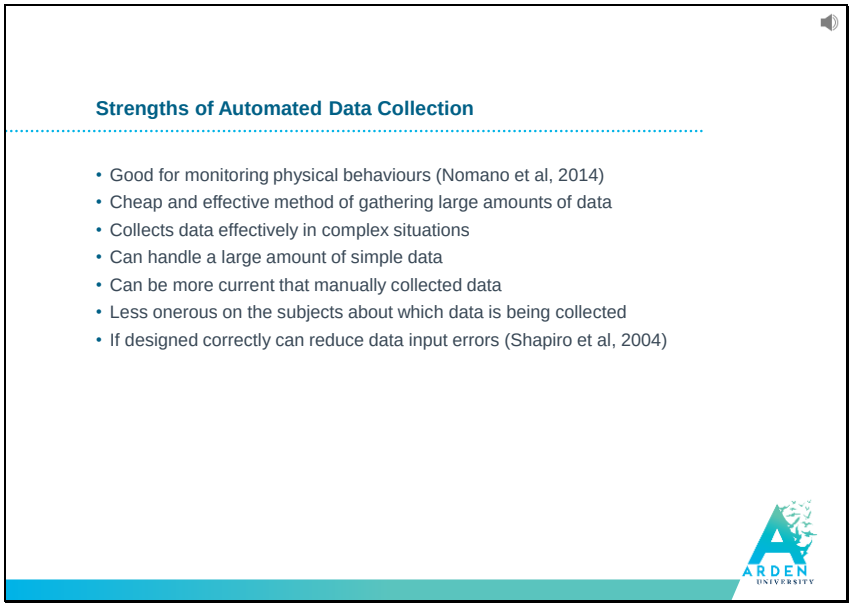
Slide 6



Automated data collection tools is defined by Macree 2015 as any kind of system that automatically collects data. Now in the world of operations and sales these quite often use ADC technology such as barcodes, which can be easily read and therefore data gathered about a particular item. They also might make use of radio frequency identification tools which can transmit to a receiver and allow us to collect data that way. These are quite often used say to track the movements of animals. We can also have a look at optical character recognition tools. Really, these are not as sort of broadly employed as they may well be. We could also consider things like radio frequency type tools that are based around the same kind of things as the radio frequency identification tools and help us to pick up data in a sort of automated and remote way. Now automated data collection also makes use of data that naturally arises through systems like the internet and other computer based systems.

So a lot of data can be gathered very, very rapidly and stored and then analysed as required. Now quite often the outcomes of the data input are pre-programmed, so if we think in terms of say a barcode, all the data that is stored in relation to the outcome of scanning a barcode for example is pre-programmed and requires some kind of entry. But once that's done then that will remain and every time that particular barcode is scanned the same information will be recorded. Then decisions can also be automated from this. So without requiring input from people. For example, in a stock control and warehousing environment we could make decisions on the reorder of products without actually having to get input from people because stock control levels can be monitored automatically and then inventory can be updated.

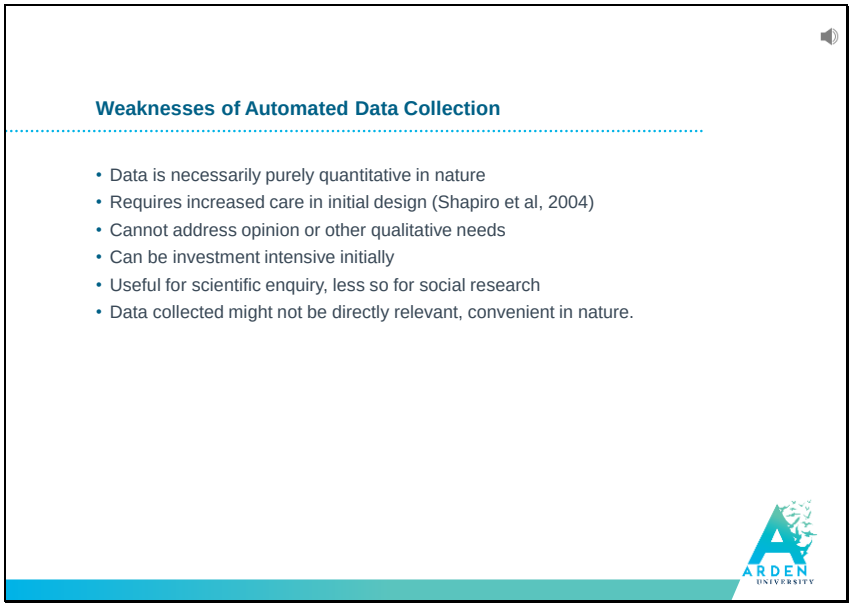
Slide 7



Now automated data collection does have some significant strengths and Nomano used it to monitor physical behaviours. So again that's akin to say the tracking of animal positions.

It can be a very cheap and effective method of gathering very large amounts of data. We can, for example, on the internet ask a system to simply record every instance of something and it will then trawl and find that for us. It can also help us to collect data in quite complex situations which might be very difficult for an individual human researcher to keep a track of. They're also used for handling a large amount of very simple data and can give us very quick outputs. One of the beauties of automated data collection is it can be more current than manually collected data. That sampling period essentially can be ongoing and we can dip into it at any point and get very, very up-to-date results. It can also be less onerous on the subjects about which data is being collected. So for example, monitoring people's movements and locations as part of a study could be very easily done with a GPS receiver, rather than asking them where they've been. This also helps us to reduce the sort of data input errors that may occur through other data collection methods.

Slide 8



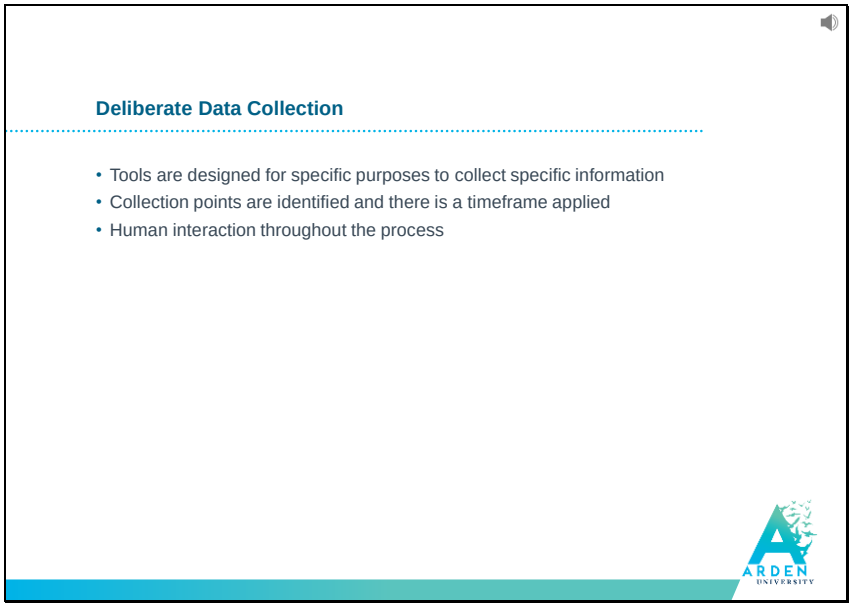
Now there are some significant weaknesses and it's worth considering that data collection collected this way is necessarily purely quantitative in nature. It doesn't really allow us to collect much in the way of qualitative and longer answer type things. It does require increased care in the initial design process so we need to be very clear about what our outcomes are, what it is that we're looking for and how we're going to recognise that. Because if we pre-programme it incorrectly to start with then all the outputs we're going to get are not going to meet our needs. We can't really use this to address opinion or other qualitative needs. Also it can be investment intensive initially, so setting up these systems can be very, very costly.

Now, this can be very useful for scientific inquiry or as we've seen before in the business context, things like, gathering large amounts of data about GPS locations as mobile phone companies do. But somewhat less so for social research, that still requires human input. So, a very good example of this would be the way that Google collects data for Google maps. A lot of that can be automated, so people's locations or mobile phone data usage and things like that, can be very rapidly and automatically collected. That's actually used very effectively in say live traffic updates, which Google maps use to show the users where traffic may be moving slower than usual and that's gleaned from GPS locations on devices. However, the real rich information that you find on the service, such as, the different, say, restaurants that are along the way or the quality of different tourist attractions is asked for from users and it's asked for using a deliberate data collection approach.

Slide 9

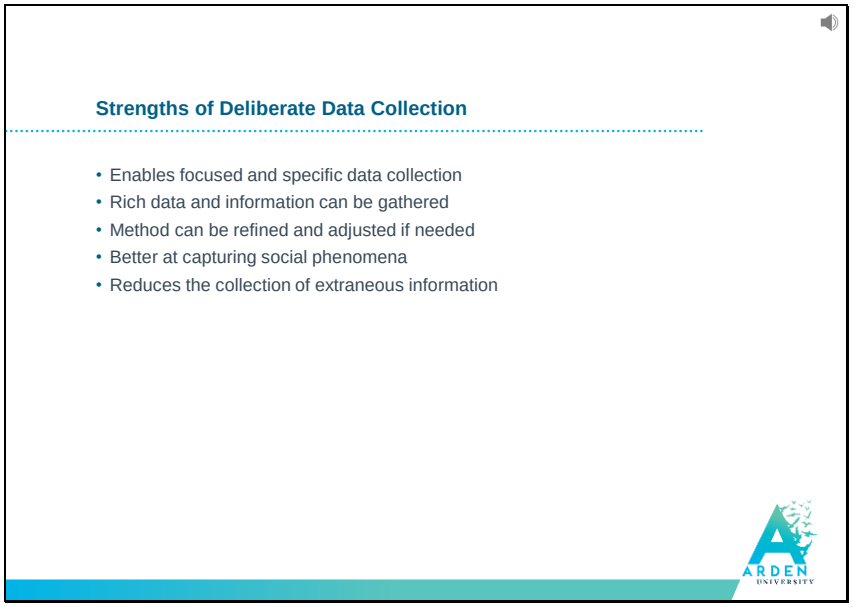
So moving on from these automated data collection tools, let's explore some of the deliberate data collection tools.

Slide 10



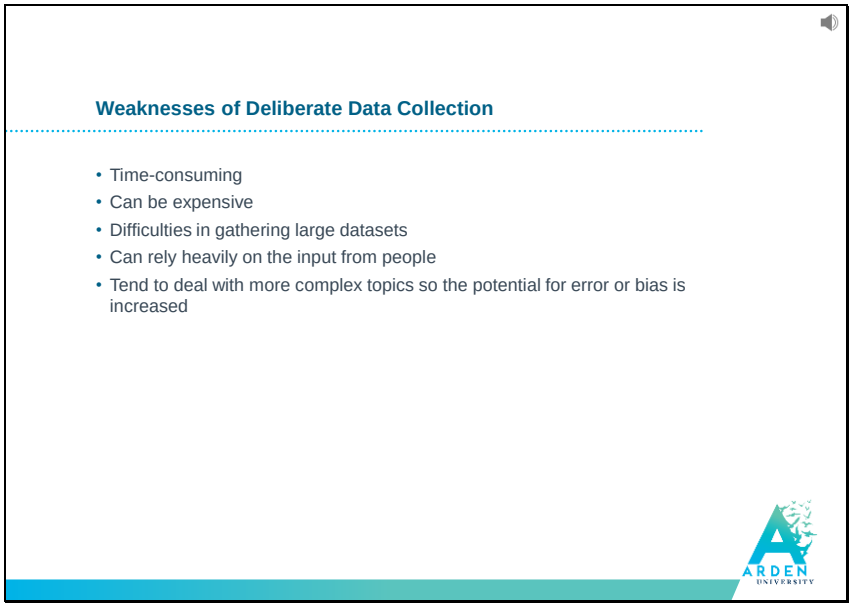
Now deliberate data collection, we have tools that are designed for a specific purpose to collect very specific information, normally about a given phenomenon. Now whilst automated data collection does collect data on very specific things, it tends to be a sort of broader net rather than something very, very focused. We identify collection points, so participants for a study for example and we tend to apply a timeframe to our study so that we can fit it in with what we're doing. It's human interaction throughout the whole process. So people are involved and the researchers are involved throughout the whole of the research.

Slide 11



Now there are certain strengths associated with this and it allows us to focus on very specific data, the very specific things that we want to know. We can gather very rich data and rich information. So, we can ask those longer answer questions and we can adjust and refine our method if we need to, and it can be less costly to do that. This operation tends to be better at capturing social phenomenon, so the interactions of people where we require perhaps a subjective observation. It can help us to reduce the collection of extraneous information, things that we don't necessarily need to know in a given context, but it also allows us to consider perhaps some of the relationships that we do need and we can build those in as we go.

Slide 12



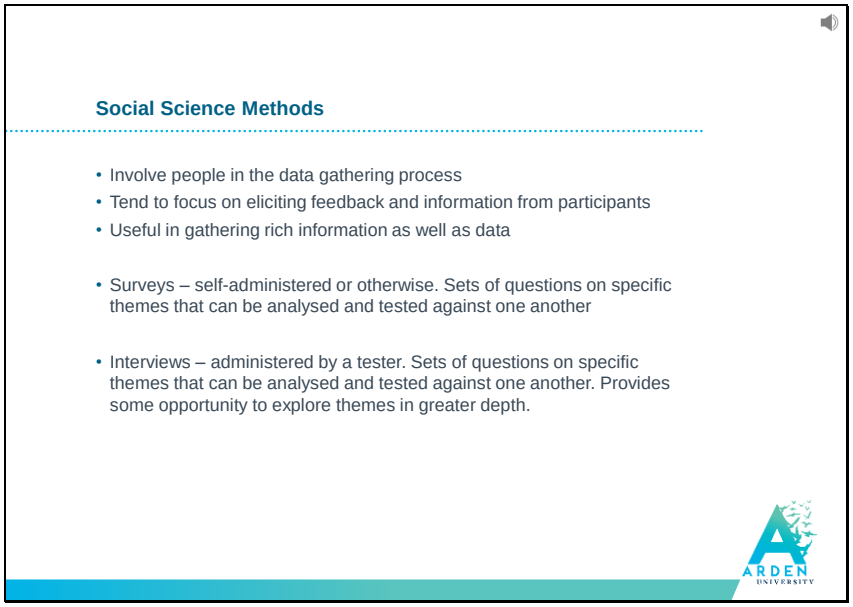
Now on the down side a lot of deliberate data collection is very time consuming. We need people to be involved in the process and that can then increase the amount of time they take. It can also be quite expensive, so applying these methods can ... again, because they take people's time that can increase the expense. There's also often significant difficulties in gathering large datasets because of the approach.

Where an automated tool could collect a very large dataset about a very simple topic in a very short period of time, doing the same sort of things, using a deliberate data collection tool, can take longer and also give us a smaller dataset. They tend to rely very heavily on the input from people, so getting subjects can be very difficult. Because of the complexity of the topics that we can deal with, there tends to be more potential for error or bias within the process. Now in an automated tool we can pretty much work to eliminate that before the tool even gets applied. But it can be very difficult to achieve that with some of the deliberate data collection tools that are used.

Slide 13

So how does this impact on some of the social science research methods that we use?

Slide 14



Well our social science methods tend to involve people in the data gathering process. We tend to be observing behaviours and we want to elicit feedback and information from participants. So whilst automated tools have some value in achieving that, they can't necessarily achieve everything, especially not the gathering of rich information as well as simple data.

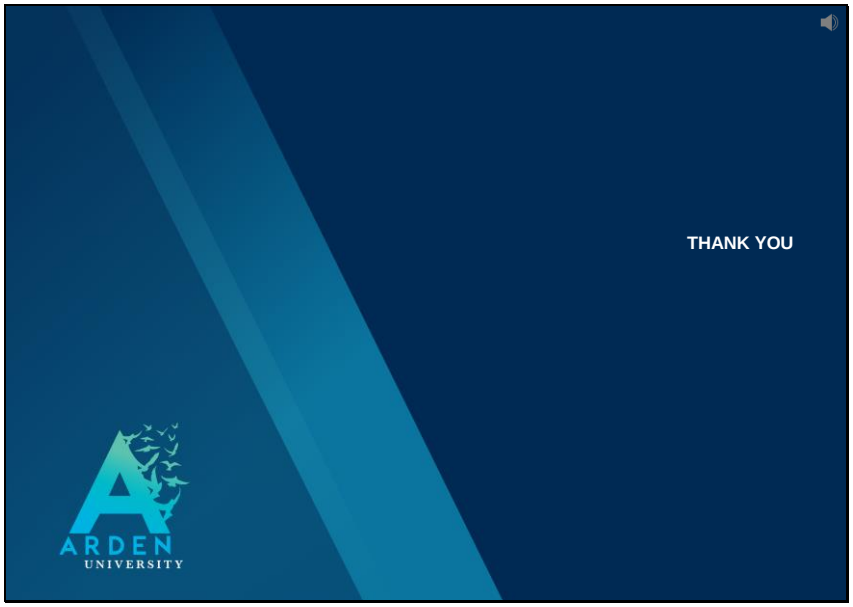
So we can apply two different approaches. Surveys, either self administered or otherwise, which are built up on sets of questions on very specific themes, that we can analyse and test against one another, tend to be used quite heavily as a social science method because we can think about the items that we want to include and we can plan it very, very deliberately and we can apply it in practice. Interviews are another approach that follow along the same kind of lines and again, administered by a tester, we have sets of questions on specific themes and again we can analyse and test them. It also gives us an opportunity to explore themes in greater depth, so where things start to emerge from questions, we can then go a little bit further than perhaps a self administered questionnaire might. So it's these two approaches that we're going to explore in further detail within this lesson.

Slide 15

References

- McCrea, B. 2015. Warehouse/DC management: ADC's improved efficiency is at hand, *Logistics Management*. April 2015.
- Nomano, F., Browning, L., Nakagawa, S., Griffith, S., Russell, A. 2014. Validation of an automated data collection method for quantifying social networks in collective behaviours. *Behavioural Ecology and Sociobiology*, 68, 8, 1379-1391.
- Shapiro, J., Bessette, N., Baumlin, J., Ragin, D., Richardson, L. 2004. Automating research data collection, *Academic Emergency Medicine*, 11, 11, 1223-1228

Slide 16



So take your time to work through everything in the lesson and in particular ensure that you engage with the activities and report back to your peers, and gather and provide feedback. Thank you.