NINE

Researching populations

If we met five men from Mars, and they said they had come to earth for medical treatment because the Martian health service is very bad, we would tend to believe them – they are "real Martians", we haven't met any other Martians, and therefore they seem to represent all Martians. But what would women from Mars say, Martian health professionals, Martians with disabilities, or the whole Martian population? The point of population research is to find out if individual opinions and anecdotes, and small-group characteristics, are evident across large groups of similar people. It relates large numbers of people (C8) to places (C10).[1]

The large-scale national and international studies require large-scale resources, and their methods are explained on websites and critically discussed in academic

publications.[2] Governments have specialist departments for national population research, such as the British *Office for National Statistics*, which provide methodological advice and survey data.[3] Organizations like the OECD and EU research populations regionally. The UN compiles many of these statistics, and the *Population Division* of the *UN Department of Economic and Social Affairs* and the *UN Population Fund* (UNFPA) look specifically at reproductive health, gender equality and population and development strategies.

Most non-specialist researchers will need to use, rather than to do, population studies on a world scale. This chapter therefore concentrates on *understanding* population research, and only outlines how to do *surveys*, and the use of *big data*. Many surveys are designed to be *predictive*, and analysis methods should be considered early (C14.3). But first, it is useful to think about historical population studies, and how *non-digital studies* might still be relevant.

9.1 Non-digital studies

The pervasive use of computers can lead to overlooking non-statistical and non-digital ways to understand populations, which remain relevant. For centuries, historians have studied how certain groups of people inhabit particular places. When linguists studied a particular language with a few people, the findings could apply to the whole population of people who spoke that language. Colonial administrative reports in the 19th–20th century combined basic statistics and descriptive analyses to inform political decision-making, and press reports influenced public opinion.[4] Historical studies can provide valuable data and background information for present-day research, but might also suggest how low-tech methods may still be useful in some circumstances.

An early ***international survey*** was devised by the Scottish anthropologist James Frazer, for his copious collection of magical and mythical beliefs in *The Golden Bough* (1890). He investigated and compared instances of human sacrifice, scapegoats, the burning of humans, the sacrifice of sacred kings, and countless other violent ritual practices from around the world. Yet he had only travelled to Italy and Greece. His method was simple but effective. He sent questionnaires to missionaries and colonial officials in other countries – Japan, China, Africa and the Americas – and got them to contact others and do his research . In modern terms, he *crowdsourced* (C6.4) his *social networks* to do *remote research* and a *snowball survey* of world populations.

Present-day ***opinion polls*** trace their history back to a straw poll carried out by *The Harrisburg Pennsylvanian* in 1824, concerning a presidential election. Then in 1916, the *Literary Digest* carried out a national survey, which predicted Woodrow Wilson's victory. The editors simply sent out millions of postcards and counted those that were returned. In 1936, a similar endeavour failed because there was no accounting for bias in those who received and returned the cards. In the same year, George Gallup carried

out a smaller but more methodologically sound survey, which correctly predicted the success of Franklin D. Roosevelt. Gallup had demonstrated the value of proper sampling, and provided the basis for present-day polls and household surveys.

Another innovative method occurred in Britain from 1937 – ***mass observation*** – the study 'of the masses by the masses for the masses'.[5] Researchers believed that the press was presenting a biased view of public opinion, and organized 500 untrained volunteer observers to keep diaries or reply to open questionnaires. Amateur *crowdsourced* (C6.4) investigators recorded people's discussions and behaviour on the street, in pubs,[6] in factories,[7] during war[8] and at events. Similarly, in 2006, the online *One Day in History* project invited diary entries, and 46,000 blogs were uploaded.[9] The film *Life in a Day* crowdsourced 80,000 clips on *YouTube* (4500 hours), which were edited into a record of micro- and macro-events, in 192 countries, on 24 July 2010. In 2013, the *British Library* (BL) announced that it would capture data from social media sites such as *Facebook* and *Twitter*, including data from pay-sites. *The Listening Project*, by the BL and BBC, is building an archive of short conversations between ordinary, but interesting people.

A population in a particular locality can be studied through *secondary data* about anything relating to that population – goods and services, media, cultural events, church records (C8). Transport networks reveal much about demography. In any location there are also many informal population experts – politicians, police, youth workers, health professionals – who have a lot of information. Interviewing elites and leaders can provide general insights into the populations they head. Many things are already based on population data, and so the population can be studied by studying these things. Adverts are useful because the advertisers will already have done a lot of market research about a local population. Researching the price of a cup of tea in different locations can indicate the wealth gap in a population – in London that would range from tea inside the *Ritz Hotel*, to free tea for homeless people just outside.

9.2 Understanding population studies

Many population studies are local or national, but this chapter only considers types of study that are relevant for world research, and these are generally at three levels:

- *Case studies* are about an aspect of a place or country, and are designed to make comparisons with similar studies elsewhere. ("Women's rights within a northern Nigerian village", to compare with similar studies elsewhere in Islamic Africa.)
- *International studies*, which concern two or more countries, and might be based on comparable case studies or national (often government) data about whole countries. ("School achievement in Europe", to compare European countries.) *Wikipedia* provides useful international lists.[10]
- *Global studies* are a *cross-sectional* investigation of a particular global sector that is not nation-based. ("Lifestyle diseases", to compare the causes globally.)

Population studies start from a broad *purpose* (C4.2) ("To improve health"), but there are many ways to design different studies that have the same purpose.

Demographic studies assess the numbers, types and characteristics of the inhabitants of a particular place. Demographers relate population to factors such as distribution (rural, urban, coastal), structure (age, gender, nationality), groups (religious, ethnic, migrants), status (education level, health, employment), or systems such as economy, education and health (epidemiology). *Dot maps* are a simple method to relate populations to places (C11.1). 'Stocktaking' studies, such as the UNDP *Human Development Index*, assess resources and public services – the availability of water, food, health care, education – and are often created from regular monitoring data ('returns'). National *census* data (based on everyone in a country or place) forms the basis of most national and international demography, but this is often very unreliable, especially in less wealthy countries where groups such as homeless people, immigrants, hill tribes children, and older people are often missing from data. Even in countries such as Britain and the US, census data becomes out of date very quickly, and other methods are being considered.[11] *Population modelling* uses data on births, deaths, migration and life expectancy to predict population size and characteristics.[12] (The 'model' is *built* from algorithms and equations based on relevant variables – life expectancy, technological impacts, desertification, water shortage. The 'simulation' *runs* that model to predict the effects of changes in those variables.) Better computers have greatly improved the modelling of world economic systems.[13] But, even if the models and maths are correct, small data errors, or technological changes, can lead to very inaccurate predictions: the GIGO problem, "Garbage In – Garbage Out". Predictive modelling (C14.3), such as the 'carrying capacity' of the Earth,[14] is very problematic.

Perceptions studies, including *opinion polls* (above), are based on people's stated views not objective facts. They often use *Likert scales* – respondents score questions from 1 (don't agree), to 5 or 7 (strongly agree).[15] There should always be a 'don't know'/'no opinion' option, which is different from the middle score, otherwise respondents are forced to express a numerical opinion that may have no basis. Cultural factors are important. Muslim women may tend to use middle scores, and men the extremes. A scale without a midpoint, of 1–4 or 1–6, may help to overcome this. The *Corruption Perceptions Index* (CPI) from *Transparency International* (TI) explains its methods clearly, and warns that although the 'score' is calculated consistently across years, the 'rank' may change in a way that does *not* imply that countries have "gone up" or "gone down" – the change may simply be because new countries have been added, or previous ones omitted.[16] The objectives and methodologies of studies like this are constantly debated.[17] In a slightly different form, perceptions studies may be used to detect hidden things, such as the level and type of crime that is not reported to the police, which is then compared with similar 'reported crime' data.[18] 'Global' studies need to be viewed with caution. The *Pew Research Center's Global Attitudes Project*[19] ignores most of Africa, Central Asia, Scandinavia and Central and Eastern Europe. Most of its findings show the US favourably.

Market research,[20] which assesses markets, and ***marketing research***,[21] which is about selling things (goods and services), deploy a wide range of population methods. *Market segmentation* analysis uses the demographics – age, sex, income, interests, personality – of a population in relation to the things they are likely to buy. *Market information* concerns supply and demand, competitors, new products, processes and materials, and international laws, regulations and cultural factors. *Market trends* are a specific aspect, measuring the rise and fall of prices, potential customers and interest. Methods such as SWOT analysis (Figure 12.2) assess the potential of companies to respond to the market, and risk analysis (C14.3.1) assesses the degree to which a company can safely invest in something. Advertising companies will use data from market and marketing studies to plan their campaigns. Psychologists study how people make *choices*, using techniques such as *choice modelling*. *Evaluation methods* (C6.2) are used to measure the success of strategies and the effectiveness of market/marketing research. Social media are increasingly studied. *Netography*[22] uses online methods to discover the interests of specific groups of potential customers, though studying blogs, chat rooms and specialist interest websites. *Big data* methods use the mass of data on social network sites (C9.4),[23] and other media such as newspapers, resulting in data such as that on *Google Trends*. *Social marketing* adapts commercial methods to address social issues such as HIV/AIDS and environmental responsibility.[24]

But is simply counting the numbers in a population always going to provide the information we really need? We may use the standard demographics – number of people, age, sex, nationality – but are other characteristics about the people in a population more important, like fat?

thinking zone: how would a butcher research the world population?

carrying capacity

Studies of the Earth's 'carrying capacity' reflect the Malthusian assumption that the problem is population increase. So, most population studies are based on counting people. But is that unit of analysis always best?

human meat

The average global body weight is 62 kg, in North America 80.7 kg, but in Asia 57.7 kg.

'If every country in the world had the same level of fatness that we see in the USA, in weight terms that would be like an extra billion people of global average body mass.' Average *Body Mass Index* (BMI) in the USA is 28.7; in Japan, 22. So living in a wealthy country doesn't need to make you fat.

methods

Country data on BMI and height distribution was used to calculate 'average adult body mass'. 'Average body mass', and population size, was used to calculate 'total [human] biomass'.

car gases

One of the most important determinates of average body mass index is motor vehicle gas consumption per capita. In Arab countries, people eat a lot and they move very little because they drive everywhere." And car use also contributes to greenhouse gases.

cow gases

Methane from ruminants and manure also creates greenhouse gases. More meat means more methane.

therefore

"When people think about environmental sustainability, they immediately focus on population. Actually – it's not how many mouths there are to feed, it's how much flesh there is on the planet." ...and how that relates to car use, and livestock farts.

so

What alternatives to the familiar units of analysis might provide new perspectives on world problems?

[See References for further information.[25]]

9.3 Doing small-scale surveys

Surveys use *questionnaires*, or observation or other *schedules*, to collect data that describes a specific *population* – the inhabitants of a particular place, or a specific group of 'items' (magazines, official documents, websites). For human populations, *demographics* ("age", "sex") and/or *characteristics* ("car use") may be noted, perhaps to test hypotheses ("that women are safer drivers than men"). But there is no point to do a survey if: usable data is already available somewhere else, it is not likely that the results will be accurate, or if the people that the study needs to get data about are very rare within a population – "transgender heads of state".

Surveys are usually statistical, and use software such as *Excel* or *SPSS*. But surveys can also be small-scale, simple, use paper and pencil methods, and be done by calling at households in villages or informal settlements.[26] A population can be human or non-human – "Homeless women in Moscow" – "Gay websites in Russian". Surveys of non-human things will use a schedule in much the same way as a questionnaire, to 'interrogate' the things being researched. But defining a population is not easy. Research based on a population of 'prisoners' may not represent all 'criminals' – only the stupid ones who get caught.[27]

A simple survey design entails deciding:

- whether the research should create *descriptive statistics*, suggest a *causal link* and/or *generalize* from a small sample to a large population?
- the *questions* – will they be 'closed' (yes-no answers), or 'open' (narrative answers), and how will the answers be analysed?

- the *variables* – what are the relevant things that might vary (age, gender, opinion, income), which variables are *independent* (age, gender) and which are *dependent* (opinion, income), and how are these included in the questions?
- the *instrument* – how will the questionnaire or schedule be designed, in relevant languages, and how will it be *piloted* (C4.1)?
- the *population* and *sample* – who or what will be studied, when, where and why? And can they be accessed, and will they respond?
- who will *carry out* the survey – do they need training?
- how will data be *analysed* and *presented* – what are the technology and IT requirements – how will this affect the whole design?

Online surveys can be done through short email questionnaires,[28] or by using online survey tools listed on *Capterra*. Free online basic tools like *Survey Monkey*, software such as the *iSurvey* app for iPads, or reduced cost systems for NGOs such as *EasyGoingSurvey*, all explain the relevant technical methods. *Crowdsourced* (C6.4) *networks* can be used, but likely bias must be considered. Not all those surveyed will respond, and a low or biased *response rate* can make a study invalid.

The general advantages and weaknesses of surveys, and specific considerations for world studies, are explained on the website. For individual researchers or small organizations it is often better to employ professional survey companies. But it is still useful to understand survey methodologies, because companies will usually only take responsibility for the technical aspects of a survey, not for factors like the overall study design, clarity of questions, or the hypotheses or theory underpinning the questions.

From a broad ***purpose*** ("To improve health"), specific *aims* help to focus the study ("To assess the availability of doctors"). The aims are then expressed as *research questions* which the study can answer clearly, by measuring things – "What is the ratio of doctors to population?" But *definitions* are important. What is a 'doctor' – does it include volunteers working for international agencies, and traditional healers? Does the 'population' only include citizens, or also migrant workers, cross-border nomads and aid workers? There may be further significant considerations. Finding out the basic doctor–population ratio will reveal nothing about the distribution of doctors – are they evenly spread across urban and rural areas? *Categories* must be mutually exclusive and unambiguously defined, so that it is clear what counts in which category – in health research 'overweight' and 'obesity' are separate categories, defined by weight, although in general conversation 'obesity' could be included within 'overweight'. Relevant *units of analysis* will be identified, which also needs careful thought. For example children are normally counted like family possessions – like TVs, cars, radios. Using 'household', 'family' or 'child' as the unit will produce different results. The 'number of families' with 5+ children might be 1.5%, but the 'number of children' living in 5+ families might be 5%. 'Child-centred statistics' are just one example of the need for careful consideration of minorities in study design.[29]

It is usually not possible to research a whole population, and so some form of ***sample*** will be used. A *sample* could comprise a few individuals who will be interviewed

in-depth, or thousands who will need to be accessed by an online survey. A similar approach could be applied to other forms of non-human 'population' – a few texts for in-depth documentary analysis, or a *sample* of texts from numerous websites for keyword-sentiment analysis. A *sampling frame* lists all the units in a population from which a sample will be selected – all "African bishops", all "religious websites in China".

Non-probability samples cannot usually be generalized, and the findings are only indicative of what may happen within a population. But statistical methods may still be used, even with a single in-depth interview, for example as an aspect of content analysis (C13.2). Counting words or phrases could be used to identify likely patterns within that interview, such as gendered language, and then to detect differences from a predicted pattern. As the scale and cost of world research is usually very large, the non-probability approaches to selecting samples are often very pragmatic:

- *Opportunistic* selection arises from chance or luck, and is sometimes not considered to create a true sample. But any questioning, however informal and brief, can still be systematic and reflect the purpose and main questions of the research. Opportunities may arise at public meetings, on public transport, or by chatting with people in international settings such as airports.
- *Purposive* samples use units that are chosen for a reason, and the reason should be explained. Respondents may have relevant responsibilities or qualities – politicians, performers, prize winners – or comprise small homogeneous groups – indigenous minorities, refugees, senior managers. A relevant and accessible research site may be used – trade fairs, conferences, sports meetings.
- *Chain-referral*, *network* or *snowball* sampling is arguably more appropriate than probability sampling when studying hidden populations, or exploring social networks or processes. This method is central to investigative research. Investigators "follow up leads", but the aim is to establish specific facts not to provide generalizable findings.
- *Convenience* samples are common in experimental studies. University students might be rewarded for participation in role play by extra credits or small fees. But to what degree can student role play provide indications of how actual people behave in real-life situations?
- *Crowdsourced* samples (C6.4) provide a new dynamic. Although the 'crowd' is self-selecting, it often comprises 'the right people in the right place at the right time'.

These samples are not statistically representative of a population, and cannot usually be generalized to large populations. Representative *probability samples* are necessary for statistical generalization (below C9.3.3).

Specific ***questionnaire questions*** derive from the *research questions* ("Is access to doctors adequate?"), to investigate relevant details ("What prevents you from getting to the doctor?"). With thought, smart questions can be devised which are immediately adaptable to different populations and criteria, and automatically provide statistical findings:

> "Do you think you are above or below average {intelligence / physical attractiveness / leadership / sexiness}, within your {college, workplace, country, the world}?"

Logically, a population should be a 50–50 split, above and below 'average'. Usually, more people think they are above average, and so the data from the smart question *automatically* evidences the theories of human over-optimism.[30] From most population studies, *statistics* – numbers resulting from counting or measuring things – can be produced and analysed in many different ways. Numerous books explain survey methodologies in-depth,[31] and this section only outlines basic *descriptive, inferential and causal* statistics (Figure 9.1). Follow up the references for further explanations in the book *Statistics for Research*.[32]

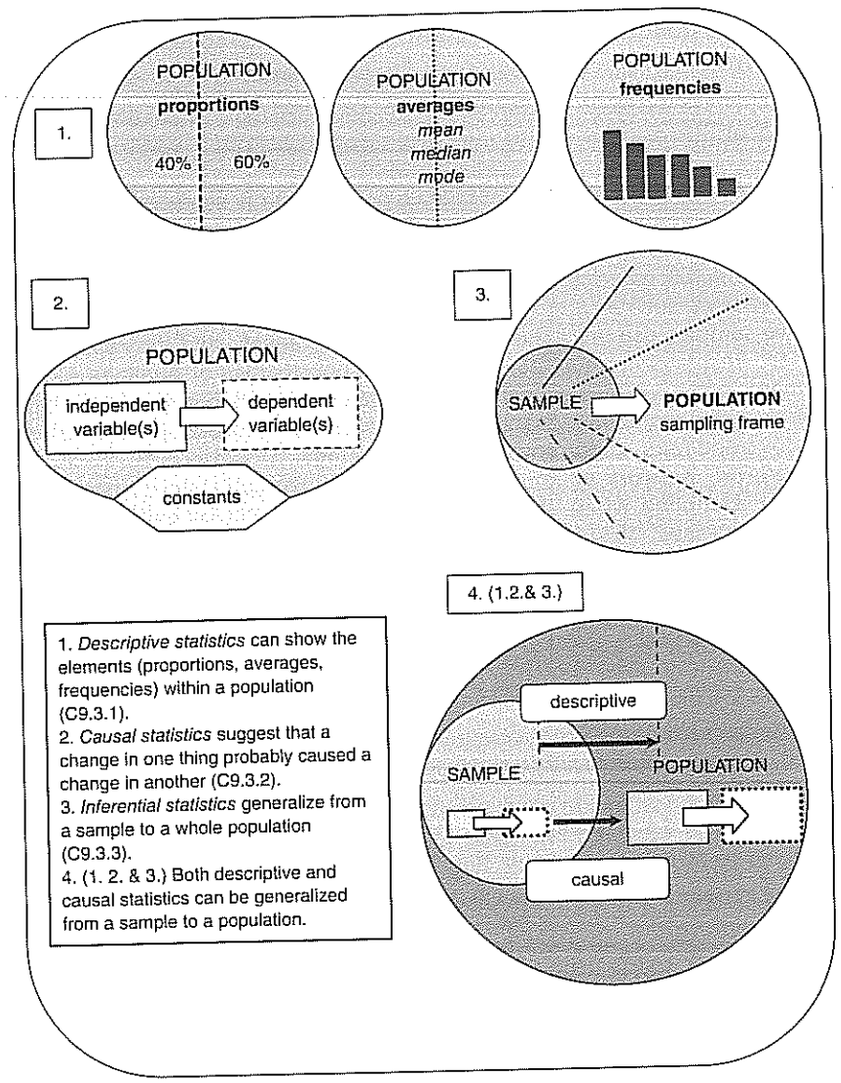


Figure 9.1 Descriptive, inferential and causal statistics

9.3.1 Descriptive statistics

From a quantitative survey, the basic numbers – ***raw data*** – are usually presented as ***descriptive statistics*** – easily understandable tables, or simple visualizations such as graphs and pie-charts. Data might be *disaggregated* to show numbers in relation to particular sub-groups within populations (male/female, rural/urban), or *aggregated* to combine different (perhaps uncertain) data sources, in the hope that, if some data is weak, it will not bias the overall findings because sources are so numerous – *newspaper circulation + libraries + computer owners + books published + self-definition = literacy levels*. Indexes (C14.4) are often aggregated. But aggregation embodies a lot of (usually unacknowledged) qualitative judgement – how does '1 newspaper' equate with '1 library'?

Every form of statistical presentation and analysis hopefully increases the understanding of something, but this also loses information. Percentages lose information about actual numbers, and even basic numbers lose information about who said what. So the choice of how to present and analyse numbers should closely reflect the *purpose* of a study (C4.2). A significant misunderstanding, even when experts explain research, concerns the difference between a population statistic and a personal statistic. If "1 in 5 people die within five years" of contracting a particular disease, that does not imply that every individual person with that disease has a 1 in 5 likelihood of dying. The likelihood of dying for each person is 100% or zero.

Raw data might be presented as simple numbers (26 women, 43 men), or then ***standardized*** to make comparisons easier when the numbers of those surveyed differ. It might show (Figure 9.1):

- the *elements* of population in the form of *proportions* (percentages – 60%).
- *averages* – the *mean* (the total from all scores divided by the number of scores), the *median* (the central point when scores are ordered high to low), or *mode* (the most frequent score). But it is important to consider whether these forms of presentation make sense. A *mean* of 'type of government' does not seem sensible; the *mean* of 'numbers of MPs' might be interesting.
- *frequencies* (1 doctor per 100 patients, 10 times each year).

Frequency tables[33] (Figure 9.2) show how often the same things occur. Data is first *listed* (or entered) for each case (Table A). This can then be converted into a frequency table, *ranked* in the order of highest to lowest frequency, and the frequencies could also be converted into percentages or proportions, which would create a *relative frequency table* (Table B). The data could be *disaggregated* into 'large countries' and 'small countries', and the *cross-tabulations* shown in a *bivariate table*, i.e. it is quickly possible to see that only two 'small countries' are 'theocracies', or that most 'large

countries' are 'parliamentary republics' (Table C). But, unless the 'type' is unambiguous ('age', 'nationality') this is not always straightforward and, again, definition is important. The delineation between 'small' and 'large' countries will be arbitrary, and Britain could be classified as a 'theocracy'.

Type of government (Table A – *list*)	
Case (countries)	**Type of government**
1	Single-party state
2	Absolute monarchy
3	Parliamentary republic
4	Parliamentary republic
5	Theocracy
6	Parliamentary republic
7	Single-party state
8	Absolute monarchy
9	Single-party state
10	Parliamentary republic
etc....to 50	

Type of government (Table B – *frequency*)		
Type of government	**Frequency**	**%**
Parliamentary republic	20	40
Single-party state	15	30
Absolute monarchy	10	20
Theocracy	5	10
Total	50	100

Type of government (Table C – *bivariate frequency*)			
Type of government	**Small countries**	**Large countries**	**Total**
Parliamentary republic	8	12	20
Single-party state	10	5	15
Absolute monarchy	8	2	10
Theocracy	2	3	5
Total	28	22	50

Figure 9.2 Frequency tables

9.3.2 Causal statistics

Population studies often try to suggest ***cause and effect***, that if one thing changes ("a new poverty reduction policy"), something related has changed ("poverty reduced") (C14.2).

Qualitative analysis will argue causation in terms of the existence or absence of variables and the probable nature of cause and effect – "*How* does a scientific education influence the problem-solving skills of politicians?" Statistical analysis assesses *if* a specific causal relationship is evident – "*Does* a scientific education influence the problem-solving skills of politicians?" Statistical analysis assesses the degree of a causal relationship in various ways:

- How the *value* of variables may be relevant – do more educational qualifications equate with greater income?
- The comparative *strength* of the causal relationship – how much qualifications affect income?
- The *completeness* of the causal relationship – what is the gap between the average science graduate's income, and the lowest and highest science graduate's income – big (less complete) or small (more complete)?

But to investigate whether a causal link exists requires a plausible theory/hypothesis that variables are related. It seems plausible that 'education qualifications' and 'income' are related in many countries, but it is not plausible that the number of robots in Japan is related to the number of *Imams* in Iran, yet.

Things that could *not* normally change, are termed ***constants***. If the things could change (vary), they are called ***variables***. Both can be human or non-human. The analysis of variables reflects three main forms:[34]

- *Univariate* (using one variable) – 'religion'.
- *Bivariate* (using two variables)[35] – 'religion' *and* 'attendance at religious meetings'.
- *Multivariate* (using many variables) – 'religion' and 'attendance at religious meetings' and 'place of birth' and 'age' and 'coercion'.

Univariate statistics only show a simple relationship between basic numbers – higher/lower – but these can still be *ranked* into *indexes* or *indicators* (C14.4).

Bivariate and *multivariate* statistics can be ranked, but might then provide the basis for an ***association/correlation***. Some *variables* may also seem ***causally*** related to others – if an *independent variable* changes, the *dependent variable* will change. But influence can be in one direction, or more. It might seem obvious that 'education level' would be *independent*, coming before the resultant *dependent* 'income', but direction also needs theorizing. If the type of 'education level' being studied were 'PhD', that 'education level' might depend on *previous* 'income', because doing a PhD is usually expensive and often happens after or during a period of employment. But when there appears to be a relationship between variables, there may not be *causation*. The number of young smokers may decrease following an anti-smoking TV campaign,

but that decrease may have been caused by another factor – a *confounder* – such as a famous rock-star dying of lung cancer. *Probability* theory provides ways to express the strength of a belief that a causal link exists, or could exist – the risk that "the death of the president will affect the economy". But there is never 100% certainty.

Significance tests contribute to validity (C7.6) by assessing how likely a result is due to chance (but not the degree of the causal relationship). Tests set up null hypotheses – that there is no difference between two variables, which the tests confirm or not. In statistical terms 'significant' implies that there is little likelihood that a result occurred by chance – usually less than 5%. But these 'significance levels' have arisen from tradition not objective statistical fact.[36] The word significance can have political, social or statistical meaning, denoting importance. To learn that a prime minister believes that genetically modified (GM) foods are harmful is a politically significant finding; for supermarket CEOs to decide to stop selling GM foods is socially significant; to show that 66% of a population does not want GM foods may be statistically significant. Statistical tests also do not assess the validity of data. If the data was collected carelessly, that will probably not be detected or corrected by tests.

In all quantitative analysis there will be unexplained factors, including data collection errors and the effects of other unknown or omitted independent variables. It is incorrect to assume that these unexplained factors are all known and can therefore be named, because they have not been measured and tested. In an attitude survey about "war", retrospectively naming all the unknown factors, "fear" would not be correct unless fear had been measured.

9.3.3 Inferential statistics

Studies may involve the whole *population* of a particular place – a *census* of all *units of analysis* in a *population*. But often a smaller manageable ***sample*** of *units* will *represent* the *population* that is being studied (C7.4). It is then possible to *infer* ('*extrapolate*', '*generalize*') from the sample data, to describe the whole population.

To create inferential statistics, ***probability*** ('representative', 'random') ***samples*** will be used to create data that can be generalized to the whole population.[37]

- *Simple random* samples comprise *units of analysis* that each have an equal likelihood of being chosen – numbers-out-of-a-hat". (But the selection of one unit must not influence the selection of any others.)
- *Systematic samples* may select every *n*th unit from a list – every 10th item on *eBay*. (But the list must not be organized in a way that may bias the study.)
- *Stratified samples* divide the population into groups ('strata'), which have clear differences, such as geographical districts with different language groups, and then take random or systematic samples from each group. (But the groups must not have different characteristics that may bias the relevant data.)
- *Quota sampling* is often used for commercial surveys, especially opinion polls. The quotas reflect the main demographics of the whole population (age, sex, etc.), but the actual respondents are chosen opportunistically not randomly, e.g. by an interviewer in a railway station. (But there is no control over who interviewers actually select.)
- *Cluster sampling* – uses relevant geographical areas, and selects a random sample from these – "cities with an airport". (But the choice of clusters should reflect the aims of the research, e.g. airport noise.)

These *representative* samples can usually be generalized statistically to larger populations, but there must be a theoretical basis for the generalization, for example explaining why the whole *population* is culturally similar to the *sample*.

Whatever the scale, type or complexity of a survey, a central problem for world population studies is the difficulty of accommodating ***cross-cultural differences***,[38] and this problem is most evident in opinion/perceptions research. Initially, methodologists tried to create, refine and pilot types of question that were 'culture-proof', but this is difficult, not least because cultures constantly change. There were also attempts 'to measure directly response category incomparability and to correct for it'[39] – tests for cultural difference were put into questionnaires, and the results were then used to weight the responses to the main questions. But this approach can make the questionnaire long, irritating and irrational to the interviewee. The problem becomes more complex in globalized, cosmopolitan communities – 'nation' is no longer a simple determinate of 'culture'. Big data researchers take this discussion a stage further and argue that, with massive data sets, cultural differences become irrelevant.

9.4 Big data

Online, big data methods provide many new opportunities to study populations (C6.5). *Social media analysis* looks specifically for patterns in people's online interaction – web use, mobility and transport, public service demand. *Crowdsourced* data (C6.4) may help to predict and analyse disasters (typhoons, earthquakes, industrial disasters), or events such as epidemics or pollution. Many methods have been adapted for *market research*[40] and other commercial endeavours,[41] under the headings 'webnography',[42] big data,[43] or 'predictive analytics'.[44] The approach is increasingly an aspect of government planning, even in less developed countries.[45] Sites like the *COSMOS* and *GDELT* platforms are making commercially captured data more widely available for academic and civil society researchers, together with tools for analysing and visualizing data, such as IBM's *Many Eyes*.

Big data research uses algorithms to process large amounts of data very quickly, to assess:

- *events* – infectious disease outbreaks, disasters, emergencies, environmental change, political violence. These can be predicted by comparing new patterns with norms, or spikes in keywords in tweets ("wind" "crowd" "earthquake")

- *behaviours* – how people *respond* to events. This uses 'sentiment' – like/dislike – and 'moods', such as 'tension' which may predict events like riots.
- *style* – a *totality* of *behaviours* – "Do international visitors differ from local people?"
- *trends* – significant patterns and changes in *direction* within populations – "Which types of clothes are being bought, in which countries?"[46]
- *networks* – the significant *links* and *influences*, and how are they used – "How are infectious diseases spreading around the world?"
- *interactions* – modelling *influences* and *effects* from some form of online communication or contact – "Which countries have internet dating links?"
- *the future* – forecasting, modelling, predicting events, economic scenarios, social trends (C14.3) – "Might climate change increase violence and conflict?"

It is possible to link big data with standard data, for example the geotags on tweets with standard government data about socially deprived areas. Natural data might also provide a type of 'control group' which can be compared with similar social media data – interview data with tweets, census demographics with website user profiles. Big data also creates the possibility for creating ongoing real-time statistical data – commuter flow, pollution, births-and-deaths.

Many questions are raised about the soundness of big data research, and these are discussed further on the website. In general, big data seems likely to augment rather than replace standard social science methods.

main ideas

When researching ***populations*** consider: whether the data already exists in some form, digital *and* non-digital methods, unambiguous definitions and cross-cultural aspects.

Research ***design*** is likely to be at the level of:

- *case studies* to compare with similar things in other countries.
- *international studies* to compare more than one country or region.
- *global studies* which concern aspects of world society that are not nation-based.

Examples of ***relevant studies*** will include:

- *demographic studies* – how people relate to places.
- *perception* (opinion) *studies* – which assess what people think about things, not objective facts.
- *market research* which assesses markets or investigates how to sell things.

Small-scale surveys may utilize:

- *univariate, bivariate, or multivariate* data.
- *descriptive statistics* – numbers which may *standardize* data, and show *proportions*, *'averages'* and *frequencies*.
- *causal statistics* which hypothesize/theorize causation, and then try to evidence the cause and effect.
- *inferential statistics* which generalize from sample data to a whole population.

Big data studies may use data from:

- social media networks.
- crowdsourced evidence.
- "exhaust data" from other forms of mass data collection.

key reading

Argyrous, G. (2011) *Statistics for Research*. London: Sage.

Blair, J. et al. (2014) *Designing Surveys*. London: Sage

Davidov, E. et al. (2011) *Cross-Cultural Analysis: Methods and Applications*. London: Sage.

Holdsworth, C. et al. (2013) *Population and Society*. London: Sage.

Journal – *Big Data*. Sage.

WEF (2012) *Big Data, Big Impact: New Possibilities for International Development*. Geneva: World Economic Forum.

online resources

To access the resources – search on the name in italics, use the http, or search on the generic term in 'quote marks'.

Survey Monkey – free basic survey tools

Capterra – lists online survey sites

Snap Surveys – attractive presentation of online surveys – www.snapsurveys.com

EasyGoingSurvey – low cost services and tools for NGOs

International Survey – international social, political and economic surveys – www.international-survey.org

iSurvey – app for iPads

Office for National Statistics – methods guidance – www.ons.gov.uk/ons/guide-method/index.html

ESRC *National Centre for Research Methods* (NCRM) – current information on statistical and online methods – www.ncrm.ac.uk

COSMOS – free big data from commercial sites, and tools

GDELT – big data harvested from international media

IBM *Many Eyes* – open source visualization software

MORE ON THE WEBSITE

TEN

Researching places

In any *place*, things that other visitors may find annoying – poor transport, shortages in shops, decaying buildings – often present opportunities for interesting research. The adverts on the road from the airport, or other entry points, are a fascinating source of data. Data collection just entails being prepared, and a small camera with a quick shutter action.

Distinctive locations include *street, urban, rural* and *coastal* settings, and the '*global commons*'. Within these 'human-influenced' environments (Figure 1, p. xxxi) *texts, objects* and *buildings* (including *infrastructure*) provide primary data. There is, of course, overlap – texts might be on objects, objects might be built. The important point is not to miss something significant because it does not fit a preconceived idea about what is relevant.

The approaches of *human geography*,[1] *archaeology*[2] and *ethnography*[3] ('culturology'[4]) (C6.1) provide useful insights into researching 'place'.[5] Data collection may entail *watching* and *listening* to people (C8.3), and probably *mapping*, which is discussed in C11. Data may be *captured* by electronic devices, not least digital cameras, but binoculars and a magnifying glass are also useful. Audio data about places is also valuable.[6] But the first way to understand any place is to read about it (C3) in guidebooks, commercial country-guides, *Wikipedia* and embassy websites. In 1030, al-Bīrūnī provided a neat discussion of the merits of 'hearsay' (speech and documents) versus 'eye-witness' (observation) (Figure 10.1).[7]

> No one will deny that in questions of historic authenticity *hearsay* does not equal *eye-witness*, for in the latter the eye of the observer apprehends the substance of that which is observed, both in the time when and in the place where it exists, whilst hearsay has its peculiar drawbacks.
>
> But for these, it would even be preferable to eye-witness; for the object of eye-witness can only be *actual* momentary existence, whilst hearsay comprehends alike the present, the past, and the future, so as to apply in a certain sense both to that which *is* and to that which is *not* (i.e. which either has ceased to exist or has not yet come into existence).
>
> Abū Rayhān al-Bīrūnī, *India* (circa 1030)

Figure 10.1 Hearsay versus eye-witness

Research about peoples (C8) and places is related, and *following up* observations is often important. To discover national values – the "national psyche" – go beyond current politics. Ask what people learned about their own country at school. Even if they dismiss it as nationalistic propaganda, school learning – good or bad – is a shared experience which unites people in relation to their country. In Belarus, a bison billboard was explained by local students, then by experts, and then through a print from the *British Museum*. It symbolizes the resilience of the Belarusian people.

10.1 The street

A quick walk around a new town or village is a good start for ***urban ethnography.*** Chatting in bars and cafes can provide a wealth of *contextual* information about an unfamiliar place, and clues about how to access hidden groups (Figure 7.1). But

street research is not new, and books such as Mayhew's *London Labour and the London Poor* (1850) provide insights into potential methodologies, and international data for historical comparisons.[8]

Researching *people* is considered fully in Chapter 8, but ***street ethnography***[9] provides ways to understand non-specific street populations – 'crowds',[10] transient groups,[11] gangs[12] and street children.[13] A particular problem for a street ethnographer is that *participatory observation* and *interviews* may be difficult,[14] yet *non-participatory observation* can lead to superficial findings. 'Lurking' is one alternative – the researcher 'hangs out at the periphery of a social setting'.[15] Similarly, *phenomenological* approaches might use a 'go-along' method.[16]

But first, it is sensible to find out about local ***law*** and ***customs*** (C5.2). Although photography in a 'public place' is usually lawful, taking photos of sensitive sites and following people may be unlawful. And what is, or is not, a 'public place' may not always be clear. Many street people – beggars, prostitutes, traders – have "minders", who watch them and can become violent. And these "minders" may be local police officers. Noting political posters is useful research, but can also indicate that making jokes about certain political leaders may be unlawful. International buskers have useful research networks about street law, such as *Buskers Advocates*.

From simple ***street observation***, homeless people may evidence discrimination, and graffiti might indicate gang culture and street style. *Transient peoples* – nomads, farmers, traders – may congregate in commercial hotels, markets, or *caravanserai* (camel hotels). *Political data* comes from posters and murals. Simply taping a cheap video "spy camera" to a bike helmet, or car roof-rack, can record large amounts of street data unobtrusively. Put a camera on someone who uses a wheelchair, and you will find out how wheelchair users view a place, and how the public views them.

Streets are also the site of, or gateway to, informal and formal ***events*** – political gatherings, social activism,[17] riots,[18] and places where people have traditionally met to express opinions like Hyde Park Corner (London), or Hong Lim Park (Singapore). The assessment of formal events and festivals has become a useful aspect of *audience research*[19] – why people attend, how they behave, commercial implications. Research about international *event management* is similar.[20] Balloons and kites are useful for aerial observation and audio recording at outdoor events, as they are quiet and safer than motorized UAVs (C11.2). Satellite data can help to assess the size of a crowd,[21] new software is likely to detect faces in crowds,[22] and predicting when crowd movement indicates danger is increasingly significant.[23] *Streamweaver* coordinates mobile phone data, captured by different people at the same event, which can later be analysed simultaneously on a split screen.

10.2 Urban areas

Street research is an aspect of ***urban research***. This takes a broader view, to include less visible data about *organizations* and *systems* – transport, justice, public

services – and *permanent communities* – housing, town planning, labour. Urban study has a long academic history,[24] including *ethnographic*[25] and *anthropological* approaches.[26] Many university departments specialize in research about their local environs, and are a good place for outsiders to start their studies. Local public administrations have their own innovative methods.[27] *Cities* are increasingly used as a unit of analysis to explain national and international phenomena.[28] New topics include *megacities*,[29] *conflict* in cities[30] and communities with *declining pollutions*[31] explored by the *Shrinking Cities International Research Network*. The *Urban Affairs Annual Reviews* series provides a good record of significant issues.

Urban spaces also provide access to hidden ***underground*** data,[32] which is often ignored. Tourists in London's beautiful Embankment Gardens are not aware they are walking on a large sewer built to stop the 'Big Stink' of 1858. Different places have different underground characteristics – shopping malls (Montreal), metro architecture (Moscow), lost cities (Derikuyu), abandoned mines (Philadelphia),[33] Cold War defences (Europe, US, Beijing),[34] wartime shelters (London),[35] robotic cycle parks (Tokyo), and sewers (Paris). Understanding underground networks is becoming important in the context of probable flooding from sea level rise, as in Paris.

Primary data might relate to gender issues,[36] migrant labour or social exclusion, and work for independent tourist guides such as *Lonely Planet*. There is a wealth of ***secondary data*** for outsider-researchers, in local media archives, official records and local libraries. ***Mixed-methods*** approaches can provide new perspectives on topics not covered fully by formal public administrative studies, such as social capital.[37] A study might combine formal and non-formal data, for example drug adverts to assess local health service problems, and products available for men and women to compare local gender aspects.

10.3 Rural areas

Rural research considers different ***terrains*** – deserts, steppes, wetlands, coastal, mountains, forests. Many rural places host ***transient populations*** – nomads, traders, hunters – who are often united by common world narratives and cosmologies. Rural research studies remote communities such as hill tribes,[38] but this is sometimes politically motivated.[39] Other rural topics include agricultural techniques,[40] land use,[41] land reform[42] and gender aspects.[43] Demographers have developed specific methods for rural areas,[44] and health[45] and food security.[46] Crowdsourced approaches to rural research (C6.4) address aspects such as *land cover*.[47] Semi-rural, ***peri-urban*** areas – "slums", "shanty towns", "informal settlements" – fall between urban and rural delineations, and present new topics such as urban agriculture and forestry,[48] water and waste disposal.[49] The traditional methods of *anthropology* and *ethnography* have generated a wealth of historical data about rural peoples, and there are recent perspectives.[50]

Historically, ***botanical artists***, who accompanied international explorers and mariners, meticulously recorded plants that might be useful, or saleable, in Europe. They depicted roots, blossom, flowers, seeds, pollinators and different stages of growth in a single *scientific drawing*, while maintaining the aesthetic value. *Kew Gardens* (London) maintains the tradition, a library, archives and a seed bank, which includes samples collected by Darwin. Artists now record pertinent concerns such as 'plants in peril' from environmental change. Landscape artists, including photographers, also created useful records of rural terrains, especially in the European colonies. But much of this work was idealized propaganda, showing how nature has been tamed, to encourage immigration and investment.

Practical methodologies come from ***development studies***,[51] and are often specific to particular communities.[52] *Participatory* approaches adapt standard methods, including using stones as counters and assessing responses to drama presentations about topics such as HIV/AIDS,[53] and try to involve all sectors of the local population including children.[54] *Rapid Rural Appraisal* (RRA)[55] uses informal interviews with villagers, and discussions with key people such as village chiefs and local officials. *Transect walks* with small groups use a 'sample' of land and crops – perhaps from the top of a mountain down to a river valley – which is discussed on site. The key research questions are simple but comprehensive and inclusive:

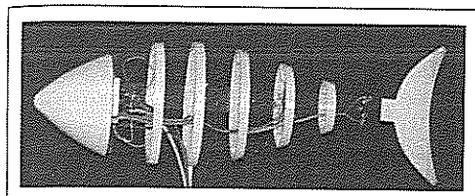
- What has changed over recent years, and what may change in the future?
- What are the problems?
- How do people think these problems can be avoided or solved?

The *World Bank* provides useful research tools, and case studies.[56] But development studies methods must be carefully piloted to ensure people all understand the processes including the images (Figure 15.3).

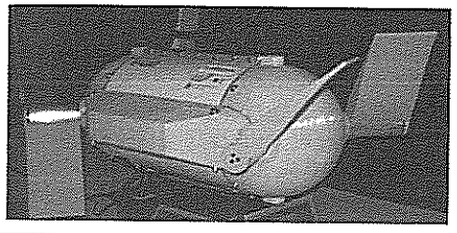
10.4 Coastal regions

Coastal peoples live on natural borders, yet perceive their territory as borderless, and 'the sea as our bank'.[57] They inhabit both nation state territory, and the global commons (C10.5). This raises questions about identity, worldviews, belief systems, environmental impacts and cultural systems. Not least of the unique problems is how to measure a coastline. A coast is never straight, and so as the ruler (unit of measurement) gets smaller, the measurement gets longer, because it measures into smaller bays, inlets, rock formations, erosion and eventually microscopic fractals.[58] Disasters – tsunamis, typhoons, pollution, radiation – are prompting innovative research. *Oxfam* researchers found that the 2005 Asian tsunami had killed four times more women than men, because misinterpreted Muslim traditions had prevented girls learning to swim.[59]

The coast is the intersection of diverse areas of ***multidisciplinary*** study – oceanography, wetlands, agriculture, meteorology, archaeology, tourism, engineering, geology, marine biology, ecotoxicology, navigation, customs and excise, security. But studies have often focused on fishing[60] rather than whole coastal populations. Work on *coastal management* assesses government policies and commercial development,[61] and the likely impacts of climate change.[62] Ecosystems are unique,[63] which demands innovative approaches. Much relevant data is 'invisible' because it is 'time-hidden' (too slow to be perceived), 'scale hidden' (too big or small), 'distant' (influenced by large ecosystems or ocean dynamics), 'dynamic' (constantly moving), or 'submerged' (underwater or underground).[64] Underwater robots and imaging devices are providing access to this 'invisible data' (Figure 10.2).[65] *Ramsar* provides ongoing discussion about wetland regions.



The *iTUNA* swims like a fish to record fish behaviour. The spine is a smart electro-responsive material. The case is 3D printed. (CSIC, Madrid)



This "turtle submarine" swims in deep water with cameras, sensors and tools to recover archaeological and other samples. (ARROWS, TUT/CfB Tallin, Estonia)

Figure 10.2 Underwater robots

Photos: Author's own

A significant example of coastal research concerns the *Chisso Company* mercury poisoning at Minamata Bay (Japan). Victims and fishing families became the first environmental activists in the world, and the *Soshisha Minamata Disease Museum* has remarkable data about their research. From 1925, Minamata research was complex and innovative. Local nurses observed strange health problems, and local people noticed there were too many rats, and worked out that this was because there were too few cats. Doctors tried to study the mysterious environmental health problems, before ecotoxicology existed.[66] They mapped affected humans, cats and fish. Poisoned fish could be caught easily by the cats and children, because they floated or swam slowly, and so cats died and children became ill. This, and infrastructure such as railways that took fish to distant markets were mapped. Workers in the

factory drew diagrams of the processes, from memory, showing that an initial filtering system had been removed as production increased. Lawyers could therefore argue that the company knew the discharge was hazardous, and brought negligence cases to court without environmental legislation. Noting the electric pylons is the key to understanding why the *Chisso* factory started, and continues to exist (C12.2). The region has cheap hydro-electric power. In 2013, the agreement of the UN *Minamata Convention on Mercury* marked the significance of the disaster, and research continues about caring for disabled, and often alienated, survivors.

10.5 Global commons

From the 1960s, arguments about 'the tragedy of the commons'[67] inspired research about the world's shared resources, which informed research about international security, law and regulation.[68] The UNEP/DELC works to consolidate the subsequent disparate treaties. Places of interest include:

- Outer space (UN *Committee on the Peaceful Uses of Outer Space*; *Outer Space Treaty*; *Moon Treaty*) – communications satellites,[69] space commerce,[70] security.[71]
- Atmosphere (*Vienna Convention for the Protection of the Ozone Layer*; *Convention on Long-Range Transboundary Air Pollution*; UN *Framework Convention on Climate Change*, IPCC) – agriculture and fisheries,[72] health,[73] livelihoods,[74] the dynamics of 'environmental refugees',[75] human impacts on the biosphere.[76]
- Sea (*Convention on the Law of the Sea*; *Global Ocean Commission*) – livelihoods and safety of sailors and fishing communities,[77] marine pollution affecting coastal populations,[78] marine resources and potential for related conflict,[79] transport of passengers and goods – ship safety,[80] economic impacts on port workers and communities,[81] modern piracy,[82] environmental change – rising sea levels,[83] loss of island territories, freshwater cycles.[84]
- Polar regions (*Antarctic Treaty Systems*) – exploration,[85] melting ice,[86] human impacts.[87]

There are convincing arguments that *cyberspace* also represents part of the global commons, and research now concerns freedom of expression, access to the digital commons, censorship, privacy and surveillance.[88] The ICRC is investigating how exploiting internet-based vulnerabilities – hospitals, navigation, disaster warning systems – through 'cyber warfare', might amount to a war crime.[89]

The ***human dimensions***[90] reflect how science relates to social research,[91] and that global resources represent *global public goods*.[92] A fundamental area of research is the perceptions and rights of indigenous peoples about the commons.[93] New technologies are making relevant low-tech data collection methods more feasible. These are often a basis for crowdsourced research – mapping polluters,[94] measuring greenhouse gases with mobile sensors[95] and tracking oil spills.[96]

10.6 Texts

Textual analysis can be an aspect of documentary analysis (C13.2). But in public places, texts are usually short sets of symbols – words and images[97] – which communicate messages through media such as inscriptions, posters and memorial plaques. ***Visual research*** methods cover a wide range of topics,[98] and are discussed in the journal *Visual Studies*. *Graphic design* is a significant aspect.[99] Whether visual data is *quantitative* (can be measured) or *qualitative* (can be interpreted) depends on how data is collected and recorded, and that relates to how it will be analysed.[100]

Data about how a country promotes its national identity is readily available on official images on *coins, banknotes and stamps*. Collecting this data does not always entail seeing the actual objects, they can be viewed on collectors' sites and *eBay*. In public places, state posters and murals provide a way to "read" and interpret political messages.[101] Of course, texts and images are often altered – manipulated, airbrushed, selectively framed, defaced – for political or journalistic reasons. Researching alterations is sometimes more interesting than the original text, and often reflects geopolitics.

Data needs to permit an analytical "reading" of texts, beyond the obvious meaning, to address questions such as:

- What is the image?
- Where and when was it created?
- Who created it, and how?
- Who is the image for, and why was it created?
- What obvious 'message' does it send?
- What are the hidden messages?
- What has changed over time, and why?
- What are the power implications – gender, wealth, age, status, nationalistic?
- What languages are used, and why?

Heraldry (Figure 2.5), *logos* and *insignias* indicate the values and identity of states and organizations. The emblem of Kazakhstan is based around the *shangyrak*, the central skyward roof window of the yurt. This not only supports the whole yurt frame, but also acts as a sundial, and is a form of Zenith Telescope to observe the movement of stars.

10.7 Objects

Objects (artefacts)[102] are central to research by archaeologists, anthropologists and museum curators, who have their own specialist methods, including photography.[103] In the past, objects from other countries were usually *acquired*, which often amounted

to stealing. New technologies are creating more ethical ways to find, record and research objects, and to transfer data (C7). This includes 3D scanning and photographic replication, or 3D printing. Many interesting objects can now be found by searching on museum, gallery or auction sites such as *eBay*.

Archaeologists and ***curators*** usually deal with objects related to people who are long dead.[104] In the past, this has often entailed guesswork about their meaning and function, which often needs to be reassessed. Objects that were not easily categorized were sometimes termed 'votive', 'ritual', 'symbolic', 'religious' or 'sacred' without much evidence. Their use may have been more prosaic, as maps, calendars, or memory aids. We now realize that some "sacred" objects were probably children's toys,[105] but this did not fit the male-dominated world of archaeology. In contrast, ***anthropologists*** may be studying objects that have contemporary significance for living people. Investigating these may mean talking directly to people for whom the objects have great meaning, which should be done in a sensitive way. Researchers study how and why buildings are created and used, but also the influences between the construction of buildings and objects.

Forensic anthropology and ***archaeology*** links traditional ethnographic methods with police investigation techniques, to study things like objects in mass graves, and this also requires a combination of science and sensitivity.[106] Systematic collection and annotation is crucial, preferably by neutral parties. Reputable police forces have robust systems for ensuring the integrity of forensic evidence.[107] Spent munitions in a war zone can evidence the use of illegal weapons such as cluster bombs and chemical weapons, and many munitions carry precise details about technical specifications, country of origin, dates of manufacture and batch numbers, which can evidence systematic abuse of power. Robert Fisk provides an example:

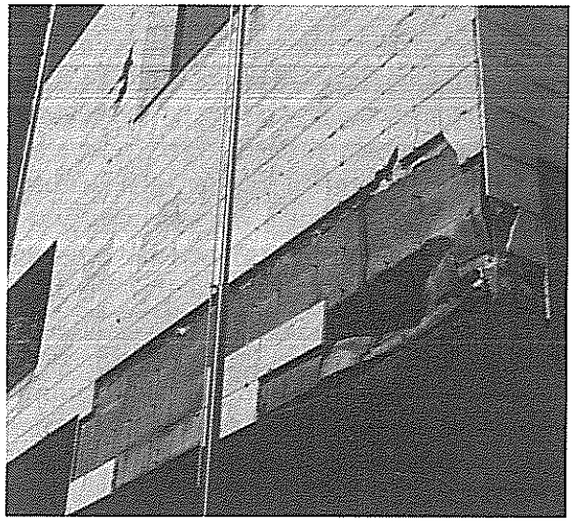
> Pieces of the US manufacturer's [Lockheed] Hellfire air-to-ground missile lie in the local civil defence headquarters in Bethlehem...The missile engine, fuel pipe and shreds of the wiring system have been sorted into plastic bags by ambulance drivers and paramedics, alongside shrapnel from dozens of US-made fuses for shells fired by Israeli tanks into the Christian village of Beit Jalla.[108]

Investigators need to keep, and photograph or document, contextual evidence to verify location and, if possible, timing. A bullet still embedded in something – a brick or human tissue – is better evidence than a clean bullet. Photographing an object together with a newspaper, clearly showing a date, provides evidence that the photo was taken after the paper was published. A public clock, or the sun and shadows, might verify time. Plants and trees can indicate the season.

New methods also provide new ways to ***reinterpret*** old objects. Improved dating can revolutionize world history.[109] The first sound recording was on a 'phonautograph' in 1857, but the machine could only record graphically and not play back. In 2008, computer-graphic software was able to convert those visual patterns back into sound.[110]

10.8 Built environment

Research about ***buildings***, uninhabited ***built objects*** and ***infrastructure*** systems ranges from historic stone buildings[111] and memorials, to electric pylons. *Google Street View* is useful for remote close-up viewing, when there are precise details about the location. Research is often associated with urban history,[112] or restoration and conservation.[113] Damage and alterations are worth investigating. The damage above the entrance to a government building in Sarajevo has not been repaired, as an ongoing reminder of the 1992–1996 siege (Figure 10.3). Similarly, World War II bomb damage to the *Tate Britain* gallery and *V&A* museum, London, has been left as a memorial.



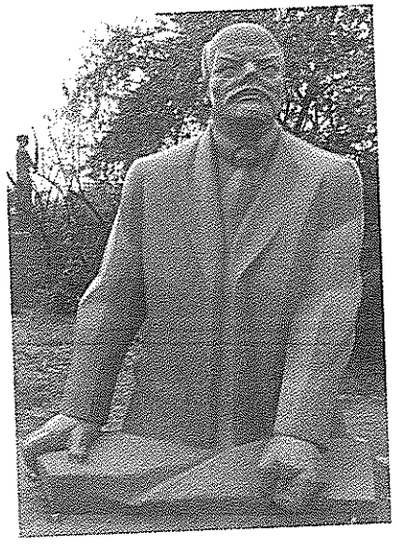
Entrance to a government building
(Capajeвo ¬ Sarajevo, 2007)

Figure 10.3 The symbolism of war damage
Photo: Author's own

Public buildings, such as courts, public baths, churches and museums, are often open to researchers. From observation and interviews, museum researchers might analyse *visitor behaviour*.[114] Older buildings have often changed their use dramatically, this is not always obvious, but can indicate local political values. A former KGB building in Budapest is now a museum, one in Almaty is now a bland-fronted college, and the vast KGB building in Minsk includes a colonnade and panopticon and is still closed for business.

Built objects are usually uninhabited structures – statues, monuments, memorials, gravestones – and data often needs to be collected in a way that permits *textual analysis* (above). Gender researchers might observe war memorials to compare involvement of women in conflict. A British memorial emphasizes the anonymity of British women soldiers by showing empty uniforms of non-combat ranks such as nursing. But in Kazakhstan, women war heroes are presented as named fighters. Kazakh nomadic tradition considers women as strong and equal to men. Other sources can validate visual analysis. The 2012 WEF *Global Gender Gap Index* ranked Kazakhstan highly at 32; France was down at 45.

After *regime change*, new rulers are faced with the conundrum of what to do with the symbolic monuments of their predecessors. To keep them suggests that previous ideas are still valid, but to destroy them equates with political vandalism. Protestant iconoclasts in Reformation England destroyed and defaced – but also punished, tried, tortured and even hanged – Catholic monuments. In 2003, the US military stage-managed images of Iraqi people seeming to pull down a statue of Saddam Hussein.[115] Ideological monuments may have artistic merit, and preserving a history of the former regime can be important to legitimize the new regime. Budapest solved this conundrum by creating a *Memento Park*, which preserves communist era statues (Figure 10.4).



A truncated Lenin
Memento Park, Budapest

Figure 10.4 Communist monument
Photo: Author's own

Infrastructure connects places and spaces – it provides the physical networks and vectors, which transfer people, things, knowledge and ideas along prescribed areas. Methodologies are undeveloped, but *systems* and *network analysis* can provide a basis (C12). Infrastructure *hubs* can provide access to otherwise unobtainable data about a place – sewage treatment plants contain data about the degree of drug use in a locality.[116] Where do things that are stored in 'the cloud' actually exist, and what are the environmental implications?[117] The absence of data can also be significant – the empty highways of Pyongyang reveal a lot about life in North Korea.

Historically, important ***vectors*** were based on natural phenomena – currents and prevailing winds, pathways and passes, routes between oases and wells. And these might be useful when studying things like cross-border pollution, or trafficking networks. On land, early navigation may have entailed using 'ley lines' – 'straight tracks' between landmarks such as hills and trees, and later between buildings such as burial mounds and spires.[118] Cathedral spires (or tall cedar trees by desert monasteries) helped religious travellers to navigate. But this navigation was based on knowledge, and knowledge was based on power. To know that the main entrance to a Christian church faced west turned cathedrals into compasses.

Local buildings and infrastructure may also be used for local research and ***surveillance***.[119] Multi-purpose buildings often contained additional sophisticated research technologies. Towers seem an architectural human universal across nations and throughout history, yet few do anything uniquely useful, except to provide a way for the powerful to observe the people. With negotiation, researchers may be able to utilize some of these facilities for their own research, especially for mapping (C11).

10.9 Environment

'Places' exist in the context of a physical ***environment***. And *texts*, *objects* and *buildings* may all exist, change, or have disappeared because of environmental factors. Although environmental research is normally the domain of natural scientists, social science increasingly explores the *human dimensions* through activist groups, citizen science,[120] crowdsourced research (C6.4) and environmental victims.[121]

Satellite data evidences long-term global change, such as the shrinking of the Aral Sea, and environmental impacts on the global commons. Cheap UAVs are providing ways for aerial environmental monitoring by non-specialists, which may be better than satellite images. Quadcopters, kites and balloons[122] can collect data cheaply and quickly, as for the BP oil spill in Mexico Bay[123] (C11.1). Numerous mobile and handheld sensors make data collection by non-specialists increasingly easy. With many devices, samples can be analysed instantly on-site, and data sent

through the internet, which avoids the problem of posting or carrying hazardous substances (C7.7).

Small ***robots*** with cameras or sensors can reach inaccessible terrains. *Rosphere*[124] is a robot ball that will roll along the ground collecting data about the land and plants. Low level *flying robots* such as *GimBall* can hold a course yet negotiate hazards, as in forests. *Unattended ground sensors* can be camouflaged as rocks and use solar power. It seems probable that soon *transient electronic devices* will self-destruct or biodegrade. Thousands of these could be dropped in the sea to monitor pollution, and then dissolve.

An interest in electronic devices should not rule out innovative unobtrusive *low-tech data collection*, including 'human sensors'. Identical white T-shirts and shoes, worn in polluted and clean cities, can collect comparative data about human exposure for analysis in a laboratory back home. Shoes can pick up sample soil, especially if the soles are smeared with grease or *Vaseline*. Keeping samples of water from washing hair is similar – hair is an excellent pollution-collection device. The data might evidence toxic releases from factories or military sites. Data must be dated and tagged, and sealed in separate plastic bags to avoid cross-contamination. Independent photos or videos of the collection can help to verify that data is genuine. Officials – police or emigration – are unlikely to suspect a grubby T-shirt, or smelly trainers, as evidence of subversive evidence gathering. And, if challenged, the objects could be explained as the basis for an art work, or even used to create one.

10.10 Data capture

Former data capture techniques can be relevant to present-day research, and inform the use of modern digital devices. An English engraver, E.W. Lane (1801–1876), provided ethnographic descriptions and meticulous ***scientific drawings*** of daily life in Egypt. His depiction of a street in Cairo shows different trades, goods, inscriptions, ages, sexes, classes, transport, animals, architectural design, female modesty, construction techniques, clothing, and even the sunny weather. But he omits irrelevant information. With electronic devices, the first consideration must be, as for scientific drawing, how to "select the signal from the noise".

The website explains and lists numerous devices that can help with data capture. Cheap small devices are available to collect data unobtrusively, especially in sensitive places – "spy cameras", video and audio recorders, and sensors. They need testing before use, to use the controls quickly and efficiently, and to aim and stabilize shots. But, note the warning at the end of C7 and C8. While you are observing and recording other people, other people may be using similar devices, and buildings, to observe and record *you*.

Sometimes researchers visit a particular place because recently something has happened, an *event* or *incident* – revolution, disaster, environmental change. The

purpose is to assess something that is different from the norm, and so a wide range of non-standard data capture may be relevant. Researchers working with political prisoners might ask them about smells and sounds, to identify or verify the place of imprisonment. Good audio recordings in war zones can help experts to identify types of munitions and aerial weapons. Smell could provide initial evidence of unusual chemicals, or disposal on a large scale by burning. In his study of Hiroshima after the US use of an A-bomb, Lifton provides a rare example of 'olfactory imagery', describing 'a terrible smell, like broiling sardines'.[125] Cheap sensors can now capture environmental data easily, and odour recorders – "electronic noses" – are capturing usable data about smell[126] to understand 'smellscapes'.[127] Increasingly, places can be "seen without looking", and non-visual data can provide a wealth of insights, if we remember to note it.

thinking zone: how would a blind person "see" the data?

What did wars, crucifixions and Chinese "death by a thousand cuts" sound like? How did the Catholic burnings of Protestants smell? There are seemingly no documentary accounts, yet this data could easily have been noted.

"seeing"

- If sighted, observe a research site with eyes closed. Are some forms of non-visual data interesting and relevant?
- Ask blind researchers to observe a research site, or listen to recordings. What do they notice that others do not?

recording

Use an audio recorder during general observation in public places, and note the counter number when visual or audio data seems interesting. Check these audio data 'hotspots' later.

validating

Check the audio tracks on secondary video data, e.g. from mobile phones in war zones, or YouTube clips. Does the sound match the video data, and claims about where and when it was taken – clock bells, *muezzin* calls, seasonal birds? If a video appears to be taken from a vehicle (helicopter, bike) or specific place (cafe, coast), does the background sound match? Is a contextual sound missing – engine noise, sea, weather?

future history

- What sounds may soon be lost – dialects, glaciers, technologies, backgammon in cafes, local radio?
- How would odour data improve historical evidence?

[See References for further information.[128]]

main ideas

When researching ***places*** consider methods from *human geography, archaeology, anthropology* and *ethnography – reading, watching, talking, recording*. Places can be categorized in terms of:

- **street –** "invisible" people, graffiti, posters, adverts, street markets and shops, informal and formal events.
- **urban –** organizations, public services, permanent populations.
- **rural –** different terrains and transient populations.
- **coastal –** coastal management and communities, "invisible" factors hidden by time/scale/distance or because they are dynamic or submerged.
- **global commons –** shared resources, human dimensions, satellite data.
- **environment –** human dimensions, aerial observation, sensors.

Data about places is likely to come from visual research on:

- **texts** – banknotes, stamps, logos, insignia, political posters, murals.
- **objects –** archaeological and anthropological evidence and museums, reinterpreting old objects with new technologies.
- **buildings –** dwellings, public buildings, industrial and commercial centres – and uninhabited **built objects** – monuments, 'observatories', infrastructure.

key reading

Andranovich, G.D. and Riposa, G. (eds) (1993) *Doing Urban Research*. New York: Sage.

Flowerdew, R. and David, M. (2005) *Methods in Human Geography: A Guide for Students Doing a Research Project*. Harlow: Prentice-Hall.

Kusenbach, M. (2003) 'Street phenomenology: the go-along as ethnographic research tool', *Ethnography*, 4 (3): 455–485.

Sinton, D. and Lund, J. (eds) (2007) *Understanding Place*. Redlands, CA: ESRI Press.

Twumasi, P.A. (2001) *Social Research in Rural Communities*. Accra: Ghana University Press.

online resources

To access the resources – search on the name in italics, use the http, or search on the generic term in 'quote marks'.

Street ethnography – http://blogs.ubc.ca/qualresearch/2012/11/12/street-ethnography/

Buskers Advocates – information for street users – www.buskersadvocates.org

Digital Humanitarian Network (DHN) – interface between formal humanitarian organizations and global networks of tech-savvy digital volunteers

Wiki Sensor – environmental monitoring software platform – http://wikisensor.com

CITI-SENSE – facilitates citizens' observatories – www.citi-sense.eu

PublicLab – development of cheap citizen science devices – http://publiclab.org/blog

Tactical Technology Collective – info activists

Crytek CryENGINE – converts maps and scientific drawings into 3D visualizations

MORE ON THE WEBSITE