

Examples from newspapers

Assignment 4 contained the following question.

Find an article in a newspaper or magazine published in 2012 that gives the results of a survey, observational study or experiment. Cut out the article, which should be no bigger than half of an A4 page, and staple it to the top half of an A4 page. Write down the name of the newspaper or magazine, the page where the article appeared, and the date of publication.

In the second half of the page, comment briefly on the article. For example, does it describe a survey, an observational study or an experiment? If a survey, how was the sample chosen? If an observational study, what conditions were compared? If an experiment, what were the treatments and the experimental units? What was measured?

Do you think this investigation was conducted well?

When we think about data we have to think about its

collection What question were the data-collectors trying to answer? How did they collect the data? What did they measure or ask?

summary and display See the lectures on Exploratory Data Analysis, and also Mini-tab Practicals 3 and 7.

analysis This includes estimation and hypothesis testing. More advanced techniques are covered in later Statistics modules.

interpretation We have to explain our results in a non-technical way without losing precision.

The assignment question was about the *collection* phase. These notes contain a number of examples submitted in the assignment, to illustrate important points about data collection.

An experiment

Source *Metro*, 6 February 2012.

Question of interest Can a filter reduce harmful emissions from London buses?

Treatments $\left\{ \begin{array}{l} \text{Nothing} \\ \text{New filter} \end{array} \right.$

Did they really include engines without the new filter in the trial, or did they assume that the figures for emissions without the new filter were already known?

Experimental units Buses (or maybe engines?)

How many?

What was measured? Emissions of nitrous oxide and particulate matter PM10.

Conclusion The new filter reduced PM10 emissions by 77% and nitrous oxide emissions by 88%.

An experiment

Source *PubMed*, **62(3)**, 2003, page 208.

This is hardly a 'newspaper or magazine', and certainly was not published in 2012.

Question of interest Can diet improve the health of people with rheumatoid arthritis?

Treatments $\left\{ \begin{array}{l} \text{Mediterranean diet} \\ \text{Ordinary Western diet} \end{array} \right.$

Experimental units 51 patients with rheumatoid arthritis, who were randomly allocated to the diets.

What was measured? Disease activity index, and other things.

Conclusion The Mediterranean diet reduced inflammation.

An observational study

Source *Daily Mail*, 6 February 2012.

Question of interest Does green tea keep old people more agile?

Observational units Nearly 14,000 Japanese people aged 65 or more, followed for three years.

What was recorded? Quantity of green tea drunk; some measure of ‘functional disability’.

How often did they ask about the tea-drinking?

Analysis 13% of those drinking less than one cup of green tea per day became functionally disabled; 7% of those drinking at least 5 cups of green tea per day became functionally disabled.

An observational study

Source *Guardian*, 6 February 2012, page 11;
also *Daily Mail*, 6 February 2012, page 25,
also *Metro*, 6 February 2012, page 3.

Question of interest Does diabetes in a mother-to-be increase the risk of birth defects in her baby?

Observational units 410,149 pregnancies in the North of England in 1996–2008.

What was recorded? Diabetes status of the mother; presence of any birth defect in the baby.

Analysis The risk of a birth defect is

$$\begin{cases} 1 \text{ in } 13 & \text{if the mother is diabetic} \\ 1 \text{ in } 50 & \text{otherwise.} \end{cases}$$

Two newspapers reported these risks as 7% and 2% respectively. Which way of reporting the risks do you think is more likely to be properly understood by the general public?

A survey

Source *Daily Mirror*, 6 February 2012, page 24.

Questions of interest What books do children under 10 read or have read to them?
How often?

Sample “2,000 parents and children under 10”.

The Early Learning Centre conducted the poll: how did they choose these 2000 people?

What exactly was asked? The newspaper did not report this exactly, but it seems that there was a list of questions including something like the following:

- “What is your children’s favourite book?”
- “How often do you/they read it per year?”
- “How often do you read to your children each week?”
- “Do your children need to be bribed to read books?”

Headline Result “*The Very Hungry Caterpillar* ... is read around nine times a year on average.”

A survey

Source *Evening Standard*, 6 February 2012, page 41.

Question of interest Do people think that bankers are overpaid?

Sample *Evening Standard* readers, self-selected.

How many?

What exactly was asked? “Are investment bankers overpaid?”

Analysis 81% of respondents said “yes”; 19% said “no”.

Discussion It is absolutely clear what was asked and what the outcome was, but has anything useful really been learnt? It seems plausible that readers who feel strongly that bankers are overpaid would be more likely to take the trouble to send a text to say so.

An observational study

Source *Metro*, 3 February 2012, page 7.

Question of interest Does the risk of dying in hospital depend on the day of admission?

Observational units All admissions to English hospitals in 2009–2010.

Does this mean 12 months or 24?

What was recorded? Day of the week when the person was admitted to hospital; whether or not the person died within 30 days of admission.

Analysis The risk of death within 30 days if admitted on a Sunday is 16% higher than if admitted on a Wednesday.

An observational study

Source *Metro*, 10 January 2012, page 5.

Question of interest What method is best to help giving up smoking?

Observational units 787 adult smokers trying to give up, followed for five years.

What was recorded? The method used (such as counselling, nicotine patches, a mixture of these, cold turkey); success in giving up without relapsing.

Analysis In the long term, nicotine patches are no more effective than going cold turkey.

An experiment

Source *Daily Mail*, 2 February 2012.

Question of interest Can silver chemicals kill cancer cells as effectively as standard chemotherapy?

Treatments { silver compounds
standard chemotherapy drug

Or did they use “historical controls”?

Experimental units Cancer cells extracted from humans with breast or colon cancer.
How many?

What was measured? The number of cells killed.

Conclusion The silver compounds are as effective as chemotherapy.

An experiment

Source *The Economist*, 7 January 2012.

Question of interest Are 17th and 18th century violins from Cremona (made by people like Stradivari) better than modern imitations?

Treatments { Three 18th century violins from Cremona
Three modern violins “made to Cremonese patterns”

Experimental units 21 expert violinists each playing 2 one-minute sessions (this gives 42 experimental units in total).

What was measured? With sight and smell severely reduced, each player played one old and one new violin, and rated them.

Conclusion Two of the modern violins beat all of the old ones.

A census (or maybe not)

Source *New Scientist*, 21 January 2012, page 5.

Question of interest Are more US pedestrians being killed by vehicles while wearing headphones?

Population All deaths in the USA.

What is recorded? Identity of person; cause of death; date of death.

Analysis The number “has almost trebled from 16 in 2004 to 47 in 2011.”

Discussion The student who found this article pointed out that it would have been more useful to show the annual figures, to see if there was a consistent upward trend. I located the original article at

Headphone use and pedestrian injury and death in the United States: 2004–2011

This makes it clear that the information did not come from the census of all causes of death at all. Several electronic databases were searched. Data was extracted about incidents between January 2004 and June 2011 in which a pedestrian using headphones had been hit by a vehicle, whether or not they died. The number of such incidents retrieved was as follows.

2004–2005	2006–2007	2008–2009	2010–2011	Total
16	19	34	47	116

Of these 116 incidents, 81 resulted in the pedestrian’s death, and it is not stated how those deaths were spread over the years. Note also that it is not obligatory to report incidents that do not result in death, so there may have been under-reporting.

This is a case where finding and reading the original research article shows that the journalist mis-reported the findings in several important respects. Unfortunately, this is not uncommon.

A survey

Source *Daily Mail*, 4 February 2012, page 30.

Questions of interest In the UK, how many people gamble? Who? How often? What type of gambling?

Sample The 2010 British Gambling Prevalence Survey appears to have collected data from a large number of adults, but there is no information about how many or about how they were chosen.

What exactly was asked? Many questions, include something like “Have you gambled during the past week?” and “What did you gamble on?”

Headline result “The number of ‘grey gamblers’—over-75s—who have taken up betting has increased” since 1999.

Is this just because the total number of over-75s has increased? Later in the article it is stated that the proportion has increased. Does the writer understand the difference between the number increasing and the proportion increasing?

Examples which I do not regard as data collection

- Adverts
- A collection of opinions about students and alcoholism
- A collection of so-called experts stating “facts” about vitamins with no supporting evidence.
- Various stories which simply contained a few numbers
- An article comparing Victorian entertainment with today’s, with *no* information about where the data came from
- Details of goals in a single football match
- Numerical announcements by individual companies, such as their sales for the past year.