

## Group work and presentations

As a group you will:

1. Use the provided data (Social Contact Data set -- details below) to generate an appropriate hypothesis, based on the data and the study that collected the data
2. Design and conduct an appropriate statistical analysis to test your hypothesis
3. Prepare and present your work as a 15-minute presentation
  - a. You should cover the aims, background/context, methods, results of your work, and discuss the implications as well as limitations of your approach. It is up to the group to decide who presents (single or multiple presenters is fine)
4. Include a slide on contributions to the work (who did what)
5. Answer questions (as a group) for up to 5min on your work

## SOCIAL CONTACT SURVEY DATA

This is data collected for a study on social contact rates conducted in 2009, seeking to measuring the number of people encountered by participants during a single day.

Researchers conducted a cross-sectional survey of households and individuals within Great Britain (GB), asking for self-reported information regarding social encounters made during a specified day.

There were two recruitment arms to the study: a postal survey using a paper-based questionnaire sent to households in Great Britain (England, Wales and Scotland), and a web-based survey using an on-line questionnaire which was open for anyone to participate.

The postal survey was distributed to randomly selected households within GB from the post office address list database, with a total of 140,000 posted during 2009. Each survey had a randomized day of the week for which participants were asked to report their social contact during that day.

The study website hosting the on-line survey was further promoted ad hoc via university press releases, social networking sites and other media outlets (local radio, local and national newspapers). Participants for the on-line survey were asked to report their social contacts for "yesterday".

Basic demographic data of participants were collected, including age and gender of participants, the number of people in their household and the first part of their home postcode, providing an approximate location.

The fields included in the summary data are:

- id – participant identifier (integer)
- postal – whether the survey was a paper-based postal questionnaire (0 = no, 1 = yes)
- unmatched\_postcode – whether the home postcode provided by the participant could be matched to a known postcode (0 = no, 1 = yes)
- web – whether the survey was completed on-line (0 = no, 1 = yes)

## CHIC601 Group work and presentations

- age – age, in years, of participant
- date – date questionnaire completed (and submitted, for on-line versions)
- day\_of\_week – the day of the week for which the contact information refers (0 = Monday, 1 = Tuesday, ..., 6 = Sunday)
- postcode – imputed home postcode
- sex – sex of participant (0 = female, 1 = male, -1 = unspecified)
- household.size – number of household members, including the participant. (-1 = unspecified)
- occupation – categorical inferred occupation types. based on free-text occupation information provided by participants
- total\_contacts – the total number of social contacts made during the day

Occupations were provided as a free text field – that is, participants were free to describe their occupation in their own words, rather than selecting from a provided list of roles. These were classified by the two researchers independently into the following categories:

- entertain – roles involved in the entertainment industry
- health – roles involved in primary or secondary healthcare
- home – stay-at home parents
- labour – roles involving manual labour, including construction
- mechanic – engineering and mechanical roles, including garage mechanics
- office – general office workers
- preschool – children not yet attending school
- public – roles that are public facing, such as librarian or receptionist
- research – academic and research roles
- retired – no occupation
- schoolchild – children attending primary or secondary schools
- service – roles in public service, such as civil servants, council workers, clerks
- student – undertaking higher education, in universities or colleges
- teaching – roles in primary and secondary schools
- transport – roles involving public transport, aviation, goods delivery, taxi and minibus drivers
- unemployed – no occupation but not of retirement age
- unknown – unable to categorise

Sources of possible additional data:

- <https://geoportal.statistics.gov.uk/> -- holds lookups and data on geographical units (e.g., upper-tier local authorities, postcodes, census information)
- <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0050074>  
POLYMOD – a study on social contacts across 8 European countries, including Great Britain
- <https://webarchive.nationalarchives.gov.uk/ukgwa/20160105160709/http://www.ons.gov.uk/ons/index.htm> -- ONS 2011 Census information

Do not feel you are restricted to using only suggested data sources – it is up to you and your analytical design!

Further information on the study is available in this paper:

<https://royalsocietypublishing.org/doi/full/10.1098/rspb.2013.1037>