

Probability, Statistics and Modelling II

Tutorial 1 – Introduction to R

Our running example for the next five weeks will be from the ‘Policing the pandemic’ research project which was carried out by members of the department during the first wave of the pandemic (Chris Pósch, Julia Yesberg, Arabella Kyprianides, Ben Bradford, Zoe Hobson, Jonathan Jackson (LSE), Reka Solymosi (Manchester)).

‘Policing the pandemic’ was a multi-wave panel study tracking the experiences, attitudes, and behaviours of 1,200 people recruited on the platform Prolific Academic for six consecutive waves—300 living in London and 100 living in each of Edinburgh, Newcastle, Cardiff, Leeds, Liverpool, Manchester, Birmingham, Sheffield and Glasgow.

We will analyse a selection of variables from the first wave data, which were collected on 21 April 2020. To give some context to the data: the first lockdown in the UK started on 23 March, the Queen gave her address to the nation on 5 April, and Prime Minister Boris Johnson was hospitalised between 6-12 April. 21 April was close to the first peak of the pandemic with 8,237 deaths related to Covid-19 that week.

<i>Constructs</i>	<i>Variable</i>	<i>Item wording</i>	<i>Response options</i>
Duration it took for someone to finish the survey	duration	N/A	seconds
Confidence in the handling of the COVID-19 crisis	pm	Prime minister	No confidence at all – A lot of confidence
	nhs	National Health Service	
	pol	Police	
	js	Justice System	
	gov	Government	
Coronavirus status	cov	Have you had Covid-19 (coronavirus)?	Yes, diagnosed and recovered ... Prefer not to say
Coronavirus attitudes	covknow	How would you rate your knowledge level on Covid-19?	Bad-Excellent
	covconc	How concerned are you about getting Covid-19?	Not concerned at all – Very concerned
	covidexp	How long do you expect it will be until the coronavirus outbreak is over and things are back to normal in the UK?	Less than 1 month - Never
Gender	gender	What is your gender?	Male/Female/Non-binary
	male	Binary variables created from gender	Male/Not
	female		Female/Not
	nonbin		Non-binary/Not
Age	age	Which of these age bands do you fall into?	16-24 – 65+

	age1	Binary variables created from age	16-24/Not
	age2		25-44/Not
	age3		45-64/Not
	age4		65+/Not
Area	area	Which city's metropolitan area do you live in?	Birmingham ... None of these
Ethnicity	ethnic	Please select the option which best describes your ethnic group:	Recoded to: Asian ... White
	asian	Binary variables created from ethnic	Asian/Not
	black		Black/Not
	mixed		Mixed/Not
	ethnico		Other ethnicity/Not
	white		White/Not
Key worker	keywork	Are you currently fulfilling any of the government's identified 'key worker' roles (listed below)?	Recoded to: Key worker/not

Table 1 Variables in the dataset

Please carry out the tasks and answer the questions below.

1. Create a folder for the seminars/tutorials and set up a working directory using RStudio.
2. Install the packages needed for today's class. Browse the dataset by looking at the top and bottom lines of the data as well as the whole dataset. What is your first impression?
3. Look at the descriptive statistics of all variables in the dataset. How do these statistics correspond to the description of each of the variables in Table 1?
4. Now consider only the questions regarding the respondents' confidence in the various institutions' handling of the pandemic. Based on the descriptive statistics, which institutions are trusted most/least? Why could this be the case?
5. So far, we have only used descriptive statistics. Using crosstabs and the associated chi-square tests, we will carry out some inferential statistics. What is the null hypothesis for the Chi-square test? Is there an association between age/ethnicity and public confidence in the government? How would you interpret the results?
6. Some scholars would argue that public confidence in the handling of the pandemic might be associated with a general trust in institutions. To examine this, run correlation analysis with the public confidence variables. What is the null hypothesis? How would you interpret the results of these correlations? What does this tell you about public confidence during the first peak of the pandemic? What would you tell the scholars regarding institutional trust?

If you want to read more about the ‘Policing the pandemic’ project, visit any of the following pages:

<https://blogs.lse.ac.uk/politicsandpolicy/lockdown-social-norms/>

<https://blogs.lse.ac.uk/covid19/2020/05/01/what-makes-britons-trust-police-to-enforce-the-lockdown-fairly/>

<https://blogs.lse.ac.uk/politicsandpolicy/lockdown-compliance-may/>

These videos can also give you some further insight into the findings from the project:

https://www.youtube.com/watch?v=eGnoS7QmBmM&ab_channel=UCLJDI

https://www.youtube.com/watch?v=Q1zFwh6BAH0&ab_channel=UCLJDI

https://www.youtube.com/watch?v=1b2RiymmWnI&t=184s&ab_channel=UCLJDI

https://www.youtube.com/watch?v=8_awxIMhVL0&ab_channel=UCLJDI

https://www.youtube.com/watch?v=IUVVzHChy3c&ab_channel=UCLJDI