Probability, Statistics and Modelling II

Tutorial 3 – Linear regression, multiple linear regression, dummy variables

We will continue with the 'Policing the pandemic' dataset from the last two weeks.

Constructs	Variable	Item wording	Response options
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. The government's behavioural science unit is interested in what shaped the expectations of Britons during the first peak of the pandemic and they asked you to investigate. First, they want you to revisit the model you fitted last week, with concern regarding catching the virus as the explanatory variable. Remind yourself of the results of this model.

2. The behavioural science unit wants you to give them a sense regarding the uncertainty surrounding your estimates. To answer them, derive the confidence intervals and interpret the results. How did you calculate these?

3. In addition to concern regarding the virus, the behavioural science unit are curious whether confidence in the government and the NHS were associated with people's expectations regarding the length of the pandemic. Fit this model and answer the following questions:

   a) How would you interpret each of the coefficients in the model?
   b) What do the confidence intervals tell you and how do they correspond to the results?
   c) To what extent are the results different/similar compared to the earlier model? Why might this be the case?
   d) What would you tell the behavioural science unit (i.e. your substantive interpretation)?

4. The behavioural science unit is grateful for your help. Now, they ask you to consider two additional variables that are important to them: key workers (i.e. 'essential workers' for the functioning of the society and the state) and people in different age groups. They expect that key workers must have had a better sense of how long the pandemic might last. They also expect that the youngest people might have expected a shorter pandemic. Add these variables to the model, and answer the following questions:

   a) Have the results changed? How would you explain the changes or the lack thereof?
   b) Does changing the reference category matter?
   c) What would you tell the behavioural science unit regarding their hypotheses?