Dr Florian Reiche
F.Reiche@warwick.ac.uk

WARWICK
THE UNIVERSITY OF WARWICK

DEPARTMENT OF POLITICS
& INTERNATIONAL STUDIES

1 | The Case Study

1.1 Codebook

Variable	Label	Year
ward	Ward name	n/a
inner	Binary classification of whether the ward is inner or outer London (based on ONS classification)	n/a
area	Land area (km^2)	n/a
population	Population	2015
children	Number of people aged 0-15	2015
adults	Number of people aged 16-64	2015
elderly	Number of people aged 65+	2015
age	Mean age of population	2013
education	Percentage of population with level 4 qualifications and above	2011
crime	Crimes committed per 1,000 residents	2015
crime_bin	Crime Dummy (0: Low Crime, 1:High Crime)	2015
employed	Number of people aged 16-64 in employment	2015
benefits	Percentage of population claiming work-related benefits	2011
migration	Net rate of worker-aged migration	2012
income	Median household income (GBP)	2013
houseprice	Median house price (GBP)	2014
cars	Average number of cars per household	2011
turnout	Turnout at the 2012 mayoral election (%)	2012

Table 1: Codebook for london_exercises Data Set

The data are taken from London Data Store (2013).

1.2 Regression

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: Turnout				
Average Age	0.740***		0.389***		0.324***
	(0.063)		(0.067)		(0.072)
Household Income		0.0004***	0.0003***		0.0003***
		(0.00002)	(0.00003)		(0.00003)
Crime Level (High)				−1.902***	−1.010**
				(0.455)	(0.413)
Constant	7.555***	19.551***	8.689***	34.760***	10.905***
	(2.284)	(0.986)	(2.107)	(0.256)	(2.286)
Observations	625	625	625	625	625
R^2	0.180	0.267	0.305	0.027	0.311
Adjusted R^2	0.179	0.266	0.303	0.026	0.308
Note:					*$p<0.1$; **$p<0.05$; ***$p<0.01$

Table 2: Regression Models

1.3 Two-Sample Test

```
Two Sample t-test

data:  turnout by crime_bin
t = CCC , df = 623, p-value = 1.669e-05
alternative hypothesis: true difference in means between group
Low Crime and group High Crime is greater than 0
95 percent confidence interval:
 1.15259    Inf
sample estimates:
 mean in group Low Crime mean in group High Crime
            AAA                     BBB
```

2 | Questions

2.1 About Regression (Section 1.2)

1. Hypotheses
    a. Formulate the alternative hypotheses underpinning Models 1, 2, and 4.
2. Significance
    a. Which coefficients in Models 1-6 are statistically significant at a 95% confidence level? What does this mean?
    b. What is the t-value for the slope coefficient in Model 2?
    c. How many degrees of freedom does Model 4 have? Why?
    d. How many degrees of freedom does Model 5 have? Why?
3. The Coefficients
    a. What does the intercept in Model 2 mean?
    b. Interpret the slope coefficient of Model 1.
    c. Interpret the intercept in Model 4.
    d. Interpret the slope coefficient in Model 4.
    e. Interpret the slope coefficient for `Household Income` in Model 3.
    f. Which models can explain a turnout of less than 15%?
    g. How would you find out if we need to explain a turnout of less than 15%?
    h. Why is the slope coefficient in Model 2 much smaller than in Model 1?
    i. Why is the intercept in Model 4 so much larger than in Models 1-3?
4. Why have I asked you questions about significance first, and then about substantive interpretation of the coefficients? (no, I didn't just feel like it)
5. The Sample Regression Function
    a. Specify the sample regression function (SRF) for Model 5.
6. Model Fit
    a. Which Model has the best overall model fit?
    b. Interpret the model fit measure for Model 3.
    c. Describe the role of $\bar{Y}$ in the coefficient of determination.
7. Model Specification
    a. Which assumption of the CLM would you likely violate if you estimated the following:
       `model6 <- lm(turnout ~ income + houseprice + age, data=london)?`

2.2 About the Two-Sample Test (Section 1.3)

1. What is the correct numerical value for AAA?
2. What is the correct numerical value for BBB?
3. What is the correct numerical value for CCC?

References

London Data Store. (2013). Ward Profiles and Atlas. available online at `https://data.london.gov.uk/dataset/ward-profiles-and-atlas`.