

Exploring the Social World

Quantitative Block Handbook (Weeks 7-12)

1. Module Programme

Lectures:

Mondays 4 - 5pm

Brodie Tower, Lecture Room 106

Lectures will introduce and explain the fundamentals of quantitative methods, with the opportunity to apply the method introduced in the labs later in the week.

PC Practical sessions:

Thursday 4 - 6 pm

Week 7: PC Teaching Centre (Room 217), Central Teaching Laboratory

Week 8-12: Room 202a, Harrison Hughes Building

The computer practical sessions, will give you the chance to use and apply quantitative methods to real-world data. These are primarily self-directed sessions, but with support on hand if you get stuck. Support and training in SPSS will also be provided through these sessions. Weekly sessions will be driven by empirical research questions.

Week	Topic	Staff
7	Data Visualisation <i>Computer Lab Practical</i>	GF
8	Multiple Linear Regression <i>Computer Lab Practical</i>	GF
9	Prediction and Linear Regression Model Evaluation <i>Computer Lab Practical</i>	GF
10	Dummy Variables <i>Computer Lab Practical</i>	ZY
11	Logistic Regression <i>Computer Lab Practical</i>	ZY
12	Summary and Final Q&A <i>Computer Lab Practical</i>	ZY

2. Software

For quantitative training sessions, ensure you have installed and/or have access to SPSS and the various datasets used in the module. You can decide to conduct your analysis/assignment in Excel, R or Python if you prefer.

Accessing SPSS

You have three ways of accessing this software. We strongly recommend Option 1.

1. SPSS can be downloaded onto your own computer for academic use under 'Data Analysis' from <https://ulp.onthehub.com/>. You will need to login using your University username and password.
2. SPSS can be accessed online from any PC with a good internet connection via <https://www.liverpool.ac.uk/csd/apps-anywhere/>
3. SPSS is pre-installed on all University PCs and can be accessed via the new University's Remote Teaching Centre Service: <https://www.liverpool.ac.uk/csd/pc-centres/remote/>

SPSS Installation Instructions Guides are available on the module website on Canvas under Modules.

3. Data

Example datasets could be accessed through Canvas. These include:

- 2021 UK Census Data
- 2021 Annulation Population Survey

Note: The Annual Population Survey requires the completion of a form prior to its usage, as it is licensed).

4. Assessment:

4.1. Data analysis and exploration

Deadline: 2pm Tuesday 9th January

The assignment consists of writing a research report using one of the regression techniques learned during the module. The report should have the following sections: an introduction, a brief literature review, methods and data, results and/or discussion, and discussion or/and conclusion. The Section 2 below in this document provides detail on the required content for each of these sections. The literature review should be kept very brief, one or two paragraphs. The core of the assignment should be the results/discussion sections.

The basic idea is to put in practice the methods learned during the quantitative component of the module. You are required to apply a linear or logistic regression model to the data provided for the module. To this end, you will need to identify a relevant research question to be addressed. Based on the chosen question, you will need to identify a dependent or outcome variable which you want

to explain, and at least three relevant independent variables that can use to explain your chosen outcome variable. The selection of variables should be informed by the literature and empirical evidence.

Once variables have been chosen, you will need to describe the data and type of regression to be used for the analysis. You need to explain any transformation done to the original data source, such as reclassifying variables, or changing variables from continuous to nominal scales. You also need to briefly describe the data use: source of data, year of data collection, indicate the number of records used, state if you are using individual records or geographical units, explain if you are selecting a sample, and any relevant details. You also need to identify type of regression to be used and why. You need to describe the model(s) to be estimated. It is good practice to write down the model equation to be estimated. Each of the model components should be clearly identified and described.

You need to provide two types of analyses. First, you need to provide a descriptive analysis of the data. Here you could use tables and/or plots reporting relevant descriptive statistics, such as the mean, median and standard deviation; variable distributions using histograms; and relationships between variables using correlation matrices or scatter plots. Second, you need to present an estimated regression model or models and the interpretation of the estimated coefficients. You need a careful and critical analysis of the regression estimates. You should think that you intend to use your regression models to advice your boss who is expecting to make some decisions based on the information you will provide. As part of this process, you need to discuss the model assessment results for the overall model and regression coefficients. Remember to substantiate your arguments using relevant literature and evidence, and present results clearly in tables and graphs rather than copying and pasting the results from SPSS.

4.2. Required Report Structure

1. Introduction

- Aim.
- Why is the topic relevant or worth being investigated?
- Knowledge gap & Research questions.
- Brief structure of the article.

2. Literature review

- What do we already know about this subject?
- Rationale for including certain predictor variables in the model.
- What knowledge gap remains that this article will address? (includes ‘not studied before in this area / year’). *Note: there is no expectation on totally original research. The focus is on a clean, sensible, data analysis situated in existing ideas.*

3. Methodology

- A *brief* introduction to the dataset being analysed (who collected it? When? How many responses? etc.).
- A description of any transformation made to the original data i.e. Turn a continuous variable of income into intervals or reducing the number of age groups from 11 to 3.
- A *brief* description and justification of the statistical techniques in the subsequent analysis.

Then, either:

4. *Results and Discussion*

- Results and interpretation, including relevance to existing literature.
- Selective illustrations (graphs and tables) to make your findings as clear as possible.

5. *Conclusion*

- Summary of main findings
- Limitations of study (self-critique)
- Implications of the findings of the study for others

Or:

4. *Results*

- Brief description of results.
- Selective illustrations (graphs and tables).

5. *Discussion*

- Interpretation of results, including relevance to existing literature.
- Selective illustrations (graphs and tables).

6. *Conclusion*

- Summary of main findings.
- Limitations of study (self-critique).
- Implications of the findings of the study for others.

4.3. Weekly Tasks

Below we have broken down the assessment into a series of discrete tasks along a timeline to help guide you all.

Week	Tasks
7	<ul style="list-style-type: none">• State/write your research question(s).• Select a dataset: choose from those supplied in class, or one you have sourced yourself.• Identify a dependent/outcome variable of interest.
8	<ul style="list-style-type: none">• Select 3 theoretically relevant independent variables.• Briefly explain / justify the theoretical importance between the dependent variable and independent variables with reference to relevant literature / evidence.
9	<ul style="list-style-type: none">• Transform your data as needed e.g. Creating age group variable from a continuous age variable; or dummy variables.• Briefly explain and justify the specific way you have adopted to transform your data.
10	<ul style="list-style-type: none">• Present descriptive statistics for each independent variable and the dependent variable: counts, percentages, a centrality measure, a spread measure, histograms or any relevant statistic.

	<ul style="list-style-type: none"> • Report the observed association between the dependent and independent variables: correlation plus a graphic or tabular visualisation. • Briefly describe and critically discuss the results.
11	<ul style="list-style-type: none"> • Select a regression strategy and explain why a linear or logit model is appropriate. • Perform one or a series of regression models, including different combinations of your chosen independent variables to explain and/or predict your dependent variable.
12	<ul style="list-style-type: none"> • Assess the overall model fit: is it good? Why? • Assess individual coefficients: are they reliable? • Offer an interpretation of what the model shows.
13	Final Submission

4.4. Additional Tips

1. Make sure you use the type of correlation and regression that matches the types of dependent and independent variables chosen for your analyses i.e. categorical or continuous.
2. Structure your write-up as much like a journal article as possible (see suggested structure above), subject to the following caveats:
 - Explicitly address all of the points above in your write up.
 - There is no need for an extensive literature review – the focus of this assignment is on demonstrating mastery of data analysis exploration, regression, visualisation and interpretation – but a good/very good answer would include a few citations to relevant ideas/theories at relevant points, based upon a *quick* google.
 - Because the emphasis of the assignment is on mastery of technique, more extensive (but still not exhaustive!) coverage of data preparation and model diagnostics than that found in a conventional academic article would be appropriate.
3. The guideline word count for this assignment is up to c. 2000 words, excluding bibliography.
4. Marking criteria: standard school ‘report’ marking criteria apply, but with less emphasis on evidence of reading, and a strong emphasis on the use of regression models, critical interpretation of results and presentation standards.

5. Feedback

Verbal feedback on understandings of all module material will be provided during lectures and computer session labs if requested. Formative feedback will be provided within three working weeks of the intermediate assessment submission deadline (beginning of week 10).