2020-21 COURSEWORK

This coursework is issued at the beginning of the term to provide you with ample time to become familiar with the data and to work on various tasks during the free-lab time at the end of each week*.*

This course will use the *Understanding Society* Teaching Dataset as its main teaching dataset. However, the course assessment also requires you to familiarise yourself with other data sources. You are requested to choose one of the four substantive topics listed below.  *In your submission of your coursework your work must be your own,* (i.e. you must do your own, new empirical data analysis for your coursework, and you are not allowed to reproduce analyses done as a group in class).

The coursework is designed to showcase your ability in understanding the relationship between statistical analyses and substantive/sociological theory.

Details of what you are required to do is provided below.

IMPORTANT NOTES:

• You are expected to write a brief commentary or provide an explanation of any methodology used for each task.

• A record of the SPSS syntax used for each question should be included in a separate appendix (it does not count towards the word limit).

• Tables/charts must be numbered. They can appear in the main text or be included as an appendix. O*nly include those that you discuss/refer to*. Including extraneous tables/charts is a sign of a poor assignment and will result in mark deduction. You may paste tables/charts directly from SPSS, or edit them (e.g. in Excel) before including them in your report: this editing choice is up to you.

• Pay particular attention to presentation – marks will be deducted for poor presentation. Unclear language is likely to result in unclear arguments.

Please choose *one* of the Research Questions from the four listed below:

1. Do young people have higher life satisfaction than older people?

2. Are those who work on temporary contracts different from those who have permanent contracts?

3. Which factors predict differences in health outcomes?

4. How many people live without basic household items, describe the characteristics of households without basic household items (Hint: create a scale from some or all of the variables named: c\_cduse1…, c\_cduse2…, etc. ).

*Please note that it is up to you to decide how to measure the key concepts in the research question you choose. It is also up to you to determine which factors may be associated with your dependent variable (explore the relevant literature to come up with sensible factors).*

Follow the instructions below step-by-step to complete your coursework.

1. Introduction and background.

Discuss what previous academic studies have shown with respect to your research question. Include at least three academic references. Discuss any factors that others (in the readings) have found to have an effect on your dependent variable [note: where possible these should be factors that you also examine in your coursework]. For full credit answers must:

• Provide at least three reasonable and relevant references, integrated into the text and described clearly.

• Identify the methods used in reported studies.

Provide background information on the data set *(Understanding Society Teaching Data 2011-12)* and identify the benefits and limitations of using your chosen data source to answer your particular research question(s). Here you may want to provide information on the following:

• Coverage

• Purpose

• Sample size

• Methodology (in particular sampling strategy)

• Questionnaire wording of *some of the key variables relating to your RQs.*

• Benefits

• Limitations

(1) There is useful info on the teaching dataset here:

<https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7549>

(2) Information on the precise questionnaire wording, which sometimes diverge a little from the brief description in the dataset, can be found in the Understanding Society Mainstage Questionnaire <http://doc.ukdataservice.ac.uk/doc/6614/mrdoc/pdf/6614_wave3_questionnaires.pdf>I have provided a copy of this for you on Moodle (under data and data analysis). Please also note the teaching dataset contains fewer variables than the full questionnaire contains.

[20 marks]

2. Select and describe the variables (both dependent and explanatory) that you will use in your analysis.

I. Describe and defend your choice of dependent variable (Y variable). You can elect to have more than one dependent variable if you prefer. Your choice of dependent variable(s) will be informed by the operationalisation you decide on for your research question. Your dependent variable can be an interval-ratio (scale) variable or a categorical variable. Explain why the variable(s) you have selected (or created) are a suitable way of measuring the concepts in your research question. Note: It may be that you will need to recode this variable to do the different analyses described below.

II. Describe and defend your choice of independent variables (X variables). You must include between four and seven explanatory/independent variables in your assignment. These must include at least one interval-ratio and at least one categorical variable. For each of the explanatory variables you have selected explain why you think it may be related to the dependent variable. Ideally, you would link this discussion back to the readings reviewed in part 1.

At least one of your variables (dependent or explanatory) must be recoded. You must explain why you are recoding the variable you are recoding and how you did it. The new variable must be named and have suitable variable labels.

At least one of your variables (dependent or explanatory) must be a scale that you have constructed. You must explain how you have constructed the scale and any choices that you made about which variables to include within the scale, and which to exclude.

Please refer to worksheet 3a for info on how to recode and worksheet 3c, for information on how to compute a scale in SPSS. This can be found under week 3’s teaching materials.

[10 marks]

3. Produce appropriate descriptive statistics for each of your dependent and explanatory variables. These might include, for example: mean, median, minimum, maximum, standard deviation, frequencies or proportions. You will need to decide which measures are appropriate and useful to report for each variable. You should discuss in the text those measures that are most relevant. You may also include any charts that you think are helpful. It is important to only include appropriate statistics. Marks will be lost for irrelevant or inappropriate statistics, tables or charts.

[10 marks]

4. Produce bivariate analyses, showing the relationships between your dependent variable and each of your explanatory variables. (It will be necessary to recode one or more variables to produce these analyses – you should describe all recoding as part of your response to Q3, above). Choose either one or both of the following types of bivariate analysis:

Cross-tabulation: examine how your dependent variable (*recoded if necessary*) varies across categories of any categorical explanatory variable(s). Provide a brief commentary on any similarities or differences. Use a chi-squared test to evaluate the statistical significance of any relationship that appears to exist. Provide a substantive commentary on the association (or lack of association) between the variables, quoting appropriate percentages.

Correlation: explore the relationship between an interval-ratio dependent variable and an intervalratio explanatory variable. Discuss the direction and strength of any relationship that appears to exist.

[15 marks]

5. Produce a series of (at least two) charts. Choose either one or both of the following ways of displaying data:

Histograms show how your interval-ratio dependent variable is distributed across categories of one or more of your explanatory variables (e.g. how hours of work is distributed among males versus among females). Discuss any patterns you see in the histograms and any differences across your different histograms.

Clustered or stacked bar-charts show how your categorical dependent variable is distributed across categories of one or more of your explanatory variables. Discuss any patterns you see in the bar charts and any differences across categories of your explanatory variable(s).

[5 marks]

6. Carry out a Multivariate Analysis. Choose ONE of the following two options:

Multiple linear regression analysis in which you include between four and six explanatory variables. Explain your methodology and write up the substantive implications of your results. You should also identify possible weaknesses in your analysis. You may present a series of nested models (up to a maximum of three) if you wish.

*OR*

Carry out a logistic regression analysis (recode your response variable if need be) in which you include between four and six explanatory variables. Explain your methodology and write up the substantive implications of your results. You should also identify possible weaknesses in your analysis. You may present a series of nested models (up to a maximum of three) if you wish.

[25 marks]

7. What would you conclude on the basis of the results of the analyses you have carried out in this project? How do your findings relate to the existing body of academic work that you discussed in section 1?

Which are your most interesting findings? What, if anything, might you have done differently? How could you develop this analysis in the future?

[15 marks]

Include your syntax in full as an appendix. Include information to clarify to which part of your analysis each part of the syntax refers.