

EDA Summative Assessment

Isaac Beight-Welland

March 2023

1 Question 1.1

- a) *How would you characterise the design (cross-section, repeated cross-section, panel, time-series...) of the “European Values Study”?*

The European Values Study is a repeated cross-sectional survey as a different samples are used between the different time periods: the sample from the first year is different from the next.

- b) *Is the source of the data experimental or observational?*

I would argue that the European Values Study as a survey is observational as it does not discriminate between a treatment and control group. Instead observations are made by comparing associations between answers of different respondents, this makes it more difficult to discern a causation.

- c) *Life satisfaction and Employment are self-reported variables. What does “self-reported” mean? Explain why self-reported variables might create issues in terms of the accuracy and comparability of the data.*

Self-reported variables ask participants to measure themselves, this can create issues with data validity as participants may not be honest, or their experiences may impact the severity of their response despite having similar circumstances to others.

- d) *How would you interpret a value higher than 90 for the variable percentile? Briefly explain why the variable percentile is better suited than the absolute level of income (Monthly household income) to compare individuals at different points of the income distribution and across countries.*

A value higher than 90 means that the absolute value of the variable is in the top 10% of the total sample. In terms of monthly household income, a value of 90 would indicate that they have an individual income greater than 90% of the country population's incomes.

- e) *The variable Full employment is constructed from the variable Employment with the following formula:*

`=IF(OR(R2="Full time", R2="Unemployed"), IF(R2="Full time",1,0),"")`

Where R is the column for Employment, Explain how the formula works. What is the value returned for an individual who only works part time?

The formula consists of a nested if statement. First it checks whether the variable is "Full time" or "Unemployed", if so, the values 1 and 0 are applied respectively. If the value is neither of the two (as indicated by the OR operator), then the value remains blank "".