ECON 491

Week 12 Assignment: HFCs

Group 4

December 9, 2022

When working with measures of mental health such as with the Public Health Questionnaire (PHQ), the question responses are mostly related to each other. For example, someone responding to the question, “How often have you been bothered by little interest or pleasure in doing things in the past two weeks” as “Nearly every day” would likely not respond to the question of “Feeling down, depressed, or hopeless?” as “Not at all”.

Because of the potential issue with consistency, we implement a high frequency check to measure the total sum of differences between question responses. We expect that for those answering the questions truthfully, the total sum of differences would be relatively small because the question responses are correlated. We can set an arbitrary cutoff point (relative to the number of questions) to flag seemingly “inconsistent” responses for further scrutiny.

Admittedly, the check is coded very heavy-handedly as a proof of concept. A more proper check might only test the difference between two variable responses that we believe should be highly correlated, rather than all of the responses together.

We also ran checks to see if there are any duplicates (one student submitted multiple entries, or a student used someone else’s id number) or missing entries (a student’s id number is not used in any of the entries), to see if there are any blank responses for questions, and to see if the duration is either high (they might have left it running for a long time) or too low (they didn’t read the questions and just answered randomly).