

Data Science Toolbox Portfolio Questions

11 Ethics and Privacy

Daniel Lawson — University of Bristol

Block 11

Portfolio 11

Choose **one question** and write up to **one page** about it. You are free to conduct further experiments to add weight to your results, and any additional material you generate can be submitted as an appendix. See The Assessment Page for advice.

These questions may make reference to the content from the current block.

Question R11.1: For a dataset that you have explored elsewhere, or otherwise, investigate whether you are able to re-identify the data without labels. The portfolio should be about the general ideas for how anonymisation and re-identification work, which are illustrated through case studies (i.e. discussion of specific examples). You are not expected to undertake a practical analysis but any coding you do to support your claims belong in an appendix.

Question R11.2: Briefly explain the goal of differential privacy. Compare and contrast Stochastic Gradient Descent and its differentially private counterpart, Algorithm 1 from Deep Learning with Differential Privacy, relating it to one aspect of the course (spanning lecture notes, coursework you've done, or the portfolio).

Question R11.3: Consider Section 3.2 Measures of Algorithmic Bias of Algorithmic Fairness. With an example (from your own work or from the paper) contrast **Equal opportunity** and one other measure, explaining how they could be used in practice.