CS7DS3 Main Assignment

To be submitted by **5 p.m.** on **Monday 28th April, 2025**. The assignment is worth 70% of the module mark. This assignment consists of a report summarising your analysis of the data set described below. Submit solutions through blackboard. Submission of all code used in your analysis is not required, however please provide an outline description of the methods and packages that you have used.

If you have any questions about the assignment, email me: arwhite@tcd.ie.

Assignment

I would like you to analyse the `USvotes` dataset. This dataset is available to download from blackboard. It has been adapted from several data sources. The data set contains voting records for US counties across several states in the last presidential election. Information regarding depression levels and racial status of counties is also available. An accompanying data description file on blackboard gives more details. For further details, see here: https://freerangestats.info/blog/2024/12/23/depression-and-vote and here: https://freerangestats.info/blog/2025/01/03/depression-and-vote-again.

Using the methods and ideas developed in this module, I would like you to build a statistical model that addresses the following question: "Is there any evidence that voting for Trump in the last election was associated with increased rates of depression?" You can model voter proportion per county directly, on the log-odds scale, or more crudely by modelling whether or not the county voted for the Republican party in a majority (i.e., $> 50\%$).

Using simple language and clear graphics, identify which data features are associated with an increase in votes for the Republican party, and whether or not a positive association with depression can be identified once additional aspects of the data have also been appropriately modelled. I am especially interested in the effect that different states might have on the data result.

Report Structure

Your report of this analysis should consist of no more than eight A4 pages, not including a title page. Your report should have the following structure:

- An objective statement summarising your main findings;

- A central figure illustrating your main findings for the objectives outlined for your chosen option;

- An outline of the approach taken to address your project objectives.

The report will be marked subject to the following criteria. Creativity will be rewarded:

- Appropriate use of methods [40%]

- Quality of figures [40%]

- Clarity of writing and exposition, overall report structure [20%]