

The Hackathon! Assignment – deadline May 9th

For this assignment, I want you to work with a large, open dataset (a minimum of 1,000 data points) of your choosing – it may be an open dataset in your area. But if you can't find one you could use a dataset from somewhere such as the following:

- The Google dataset search toolbox:
<https://toolbox.google.com/datasetsearch>
- The Tidy Tuesday datasets:
<https://github.com/rfordatascience/tidytuesday>
- The gapminder datasets:
<https://www.gapminder.org/data/>
- The Kaggle datasets:
<https://www.kaggle.com/datasets>

The dataset doesn't have to be psychological in nature. But whatever dataset you choose, I want to see evidence of data wrangling and tidying (if appropriate), visualisation, and modelling - with a summary of what meaning you have extracted from the data (and any caveats about the interpretation that you think are worth raising).