Meeting 3

Present: Conor, Laura, Lewis, Mike

Progress:

Lewis - added Jupyter notebook that calculates correlations between features in the Literacy and Numeracy dataset. He also found a dataset containing information on political gender balance.

Conor has found a method of creating a world map with a colour bar to represent certain features (e.g. female proportion of labour force) - on Github, folder 'World Map Plot'.

Laura found a dataset describing proportion of religions in countries by year.

Discussed

Come up with a plan of how to proceed - from the email, we need to have formulated the key questions, identify the specific datasets needed, and be exploring the data.

Most of the papers found use regression models - might be too simply but good for prediction.

Main question we are asking - How is gender equality relating to education affected by the following features around the world:
- Politics - Proportion of women in parliament
- Religion - Most adopted religion in each country (Note: may need to consider ethics)
- Wealth - GDP is obvious, but 'Happy Planet Index' might give more interesting results

By next week we will have all the datasets we are using in one folder.

What Machine Learning models can we use?
- Regression
- Time Series Modelling (possible predictive techniques?)
- Clustering (k-means, t-SNE, etc.)