Applied Data Science   **Wed 11th March 2020**

Meeting 1

Present: Conor, Sarah, Mike, Lewis, Laura

Discussed:

- Admin & Organisation
    - Github repository easier for version control & sharing files.
    - Coding in Python
    - Report in LaTex
    - Plan
        - Week 1-2 – data digging, filling holes, find what we want to tackle
        - Week 3 - we won't get much/anything done
        - Easter – bulk of the work (stay in touch)
        - Week 4-5 – write up, make pictures pretty
- Want to start with a broad theme – need to choose which aspects and datasets we'll look at.
- Datasets (Unicef) – can also get data from other sources, will group under these broad headings
    - Education – attendance rates, literacy
    - Violence & Social Norms – Violence, FGM, Child marriage, Child labour, Social expectations etc
    - Health – sanitation, disease
    - Financial & Urbanistion
- Could be worth choosing a theme and then everyone looks at whether areas of that theme are viable.
- Violence is more likely to give a better picture but could be interesting to tie in with education.
- Focus on education
    - Primary school rates vs secondary school rates
    - Literacy rates
    - Completion rate vs attendance rate
- Could be worth clustering countries and investigating further

Actions:

- Conor - Github repository
- Laura – set up Overleaf
- Start with Education data to investigate – have a dig around, see if you can find anything
    - Drop of between primary to secondary - Laura
    - Attendance - Mike
    - Literacy & Numeracy - Lewis
    - Transition to Work, what kind of work - Conor
    - NEET (not in education, employment or training) - Sarah