**Practical Data Science Team Project One**

**Meeting One Notes**

**Summary:**

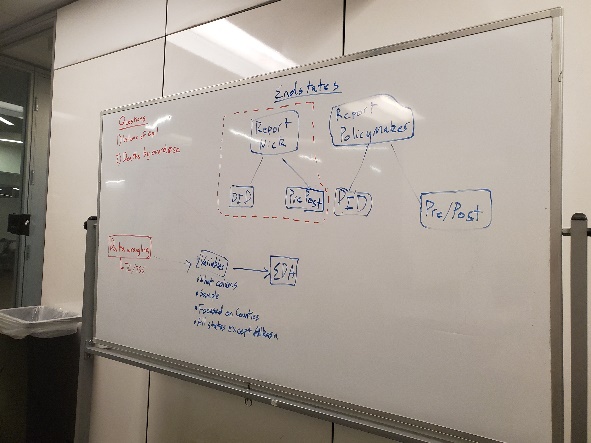
This project has two required reports listed below.

Report One: Dr Eubank

Report Two: Policymaker but they will include the same information

* They will focus on policy changes in Florida, Texas, and Washington
* Examining the effects of the policy changes on the following questions
  + The volume of opioids prescribed
  + Drug overdose deaths
* Using difference-in-difference and pre-post analysis
* Note: We are also short at least one source to complete the analysis

**Whiteboard Sketch:**



**Due Outs:**

* By **Sunday at 5:00pm** (submit via Slack and the documents folder on Github) answers to the following questions.
  + What variables do I need in this data (links below). I think the top one tells you almost everything but I am not for sure. Probably best to check all of them.
    - <https://github.com/wpinvestigative/arcos-api/blob/master/data/data_dictionary.csv>
    - <https://github.com/wpinvestigative/arcos-api/tree/master/data>
  + What sample (years, counties, etc) need to be converted
  + What should a single row look like
  + What do I need to do? What do I need my input datasets to look like to get to this final analysis dataset?
  + What intermediate datasets, what source datasets do I need? What variables do they need?
  + Then we can have a team meeting to delegate these out.
* **Everyone also needs to make their own branch in Github**

**Way Ahead:**

I will take what I get on Sunday night and bring it to Nick for office hours Monday. If you know someone who has a good idea of how to chunk data that will be EXTREMELY helpful.

**Pending Questions:**

* General Github Strategy (code review, etc)
* How to best organize/chunk the huge dataset