目前进度：

* 关于 - pregnancy涉及的、provide parent and family的算不算、teacher development算不算的，等待和shein的交流确认
* 可以和shein交流他那个论文，关于obesity的问卷和score的，我们score的方法也可以从那里学习，但我的想法是对的

目标：

Try to match, 20个flags有哪些topic questions涉及了，涉及了哪些？How many flags are covered by these CDC health questions profile, how we wanna turn that into a score (从每个questions能了解到这个州的policy实施程度如何).

获得的数据：

Flags：[Sex and HIV Education | Guttmacher Institute](https://www.guttmacher.org/state-policy/explore/sex-and-hiv-education)

完整Questionnaire：[Profiles Questionnaires | DASH | CDC](https://www.cdc.gov/healthyyouth/data/profiles/questionnaires.htm)

整理版Questionnaire表格：[Mapping Questions - Google Sheets](https://docs.google.com/spreadsheets/d/1O2KlrNTl_xmzxArHGCchrdgOmCduJKNX0M19tl1KJ0I/edit?gid=1660554617" \l "gid=1660554617)

In the actual data from the survey we get the number of principles/teachers answering yes to each question in each state

计划： 要先了解清楚每个flags的意思

1. 定义和分类问卷问题

确定20个flags：明确这20个flags是什么，并给每个flag一个编号或标签。例如：

Flag 1: 性教育课程覆盖率

Flag 2: 教师性教育培训情况

Flag 3: 学生对性教育课程的满意度

依此类推，直到Flag 20

分类问卷问题：将每个问卷问题与这些flags进行对应。一个问题可以涉及一个或多个flag。建立一个问题与flags的对应表，写excel。例如：

问题1涉及Flag 1和Flag 3

问题2涉及Flag 2

问题3涉及Flag 1、Flag 2和Flag 4

依此类推

1. 统计各个flag的体现情况（分别被提及的次数）

创建统计表：统计每个flag在问卷中的体现情况。可以创建一个表格来记录每个flag涉及的问题数量。例如：

Flag 问题数量

1 10

2 8

3 5

...

20 7

1. 计算每个flag的权重

根据提及次数计算每个flag的权重

计算总提及次数：总提及次数=∑每个flag的提及次数

计算每个flag的权重：

权重i = 提及次数i /总提及次数

1. 定义和计算score

设定评分标准：回答yes得1分，其余为0分

每个问题可以涉及多个flags，因此每个flag的得分需要根据该flag涉及的问题得分加权计算。

* 初始化每个flag的得分：设置每个flag的得分初始值为0。
* 遍历每个问题：如果回答为yes，将该问题的得分（1分）按照flag权重分配给相应的flags。

示例：问题1：回答yes，涉及Flag 1和Flag 3

Flag 1 增加 1 \* 0.10 = 0.10分

Flag 3 增加 1 \* 0.05 = 0.05分

所以也可以把被涉及到且是yes回答的所有问题得分加在一起，最后乘以权重来得到某 个Flag的加权得分。

* 最后把所有flags的得分加在一起

总结

这个流程将帮助你系统化地分析问卷数据，从而得到学校性教育政策实施情况的得分。具体步骤如下：

定义和分类问卷问题与flags的对应关系。

统计各个flag在问卷中的体现情况。

设定每个flag的权重和计算score。

这两部分的分析是连贯的，前期的flags体现情况统计将直接影响后续的得分计算。

问题：

1. 不确定1、2的意思，是否只要涉及性、HIV的全都可以选，那是别的在given sex有的情况下的flag只要涉及就一定包含1或2吗
2. 有很多涉及pregnancy的词，但没有更具体的，那我可以把所有相关的都选上吗，包括Contraception, condom, abstinence.
3. Teacher24a. Individual factors that influence health (e.g., race/ethnicity, sexual orientation, gender identity, socioeconomic status)这算是health的，那他涉及了sexual orientation和sex education吗
4. 问卷里would you like to receive...这种是不是不能算
5. 教师是否收到某些主题的development也可以算体现这些主题的flags吗
6. 涉及STI的，是不是sex education和HIV education的都算