## Project: CE Idea Interest over Time Analysis

**GOALS:**

P1: Refine our understanding of how much incubatee's idea preferences change over time

P2: Understand the main drivers behind incubatee's preferences changing

P3: Draw conclusions/recommendations for the different departments based on the data

Process and Summary:

* Ask self: What is a good numerical version of this question?
* Clean data → Try to dust off python/excel skills → find this is slower than I’d hope! :) → Pull what I can with some AI assistance & do visual sweeps
* Got partway through P1. If these skills are needed in what I end up doing, very happy to spend some time refreshing my skills (used to get into flow with this!). Think it would take a few days?
* Answers below to two questions from P1 (i.e., “moved negative→pos” & “did not found program offerings”) P1
  + I think the thought that people update from neg to pos may have come from the 2021 round.

**SCOPE:**

* Take 3 hours on Thurs to make as much progress as possible on the above goals, using
* All idea data from weekly surveys 2020-H1'2025 [Idea Interest Over Time Data for Elizabeth](https://docs.google.com/spreadsheets/d/1WYidF5Zat8W789Ig2b_hW3J8UjwbTlV9LfhW7gsF5fo/edit?gid=0#gid=0)
  + *Note that there's a chance some context in the open-text is sensitive (ie. a person mentions how they feel about another incubatee), although I've tried to clean it. Please use discretion and sensitivity with this data*
* It might be useful to get some additional "soft" context from me to help you draw conclusions/scope the work further. Can discuss in our chat

*I imagine thorough completion of these goals will take more than 3 hours - we can continue the work in projects over the coming weeks if you find it interesting :) My hope is that you can at least 80/20 P1 in 3 hours, but I could be off-base depending on how messy the data is.*

**Further breakdown of goals/specific data we want to know:**

P1: Refine our understanding of how much incubatee's idea preferences change over time

* What is the average “trajectory” an incubatee goes through in terms of idea openness
  + Do they become more or less “open” during the program
  + ***If they start negative on an idea, how likely is it they become positive?***

Data covered: 2021 - H1’25 (does not include H1’25 pre-program data)

Specific question: If participants provide a negative score (1-3) in week 1, how likely is it they become positive (5-7) on that question *at any week?*

*→* Depending on interests, question could be asked differently: E.g., If participants provide a negative score (1-2) in week 1 or 2, how likely is it they resolve positively (6-7) on it in week 4 or 5?

**Instances of participants moving neg → pos**

H1’25

* Adnaan - CFME - Week 1:1 → Week 5:6
* Amy - CFME - Week 1:1 → Week 2:6

H2’24

* Uttej - PLA - Week 1:3 → Week 2:5

H1’24 -N/A

H2’23

* Habiba - Week 1: 3 → ORS → Week 5:5

H1’23

* Miri - Tobacco Taxation - Week 1: -2→ Week 2:1
* Nicoll - KC 1 - Week 1: -1 → Week 4:1
* Victoria - Tobacco Taxation - Week 1: -2 → Week 3: 1

2022

* Sarah - Tobacco Taxation - Week 1: -1 → Week 2: 1 → W3 0

2021

* + Andres - Fortification feed - Week 1: -3 → Week 2: 0, W3 2, W4 1, W5 0
  + Devon - Fortification feed - Week 1: -2 → Week 3:1 → W4 0
  + Isaac - Fortification feed - Week 1: -2 → Week 3: 0, W4 0, W5 3 AND Shrimp Welfare - Week 1: -3 → Week 3: 2, W4 3, W5 3If they start positive, how likely is it they become negative?
* For each idea that has been launched, what was the founders’ journey with that idea?
  + What was their relative rating of it over the program?
* How often do incubatees found ideas that are offered in the program vs not?

Data covered: 2020 - H1’25

Slightly expanded question scope because grabbed it automatically. Includes names of incubatees who were listed as not founding at all.

* + Founded idea not in survey
    - Tom → Lead Acid Batteries
    - Micaella → Lead Acid Batteries
    - Laura → CBT for Crime Reduction
    - Henning → CBT for Crime Reduction
    - Tuberculosis → Habiba
    - Keyur - Syphilis testing
    - Nils - Syphilis testing
  + Founded own Idea:
    - Haindavi → India Welfare Policy
    - Uttej - India Welfare Policy
    - Nicoll - Nicol’s idea
    - Joined other organization
  + Joined other organisation
    - Adnaan → Joined ARMoR
    - Lizzie → Joined Ansh
    - Stuart → Almost launched CFME, Joined Animal Advocacy Africa
    - Jessica → Almost launched CFME, Joined SWP
    - Amy → Joined LEEP
  + Not listed in founding results

H2’24

* + - Oliver

H1’24

* + - Jonas

H2’23

* + - Jaipal
    - Samuel

H1’23

* + - Ben
    - Moritz
    - Miri
    - Jonas
    - William
    - J.T.

2022

* + - James

2021

* Following note on excel to exclude everyone not listed in founding results

2020

* Following note on excel to exclude everyone not listed in founding results
* How convergent are people on cause area? How many truly “agnostic” founders have there been? (how often do people end up only preferring ideas in one cause area)

P2: Understand the main drivers behind incubatee's preferences changing

* Using the open-text, pull out sentiment analysis (using AI, likely) to understand the primary reasons people have become more or less excited about ideas
  + Idea-by-idea, especially for ideas that have not been founded
  + Across-cohorts or causes (ie. policy ideas)
* Are there notable “popularity” effects? Ie. ideas that start more or less popular, middling people updating toward that average more than is normal

P3: Draw conclusions/recommendations for the different departments based on the data

* Steve and I communicate that peoples’ preferences change over time; is this actually true?? If not, what might be likely consequences (pos and neg) of communicating the data transparently
* How might this/should this influence either the ideas research recommends (low likelihood this should), or more likely, how we communicate the ideas/supplement the reports pre-or-during program?
* How should recruitment update based on this data, both in
  + Getting a more accurate understanding of people’s likely idea preferences pre-program
  + How to outreach based on the data