Welcome to the Specter Data Science Challenge! 👻

**Introduction to Data**

Specter offers three data feed products:

1. Company Signals
2. Talent Signals
3. Strategic Intelligence Signals

Today you’ll work with a [redacted sample company dataset](https://drive.google.com/file/d/16BYjfjVmibBxs-qaBvOSMr0ddoZ5gK1l/view?usp=sharing). We suggest you explore the dataset before diving into the challenge so you can understand what each data points mean, and value each of these later when you’re building the Ranking. For additional context you can segment the data points into different categories:

* Informative / Firmographic Data
* Funding Data
* Website / Social Media / App URLs
* Web Data
* Team Data
* Social Data
* App Data

**Task #1: Ranking Algorithm**

Imagine you are an early stage venture capital investor primarily interested in companies that have raised less than $20m in Total Funding and are showing High Growth 🚀

You are interested in all Geographies and in all Verticals / Industries / Categories.

Your goal is to build a ranking algorithm that scores companies from 1 to X, with 1 being the highest priority for investment. The ranking should prioritize companies that fit the investment criteria outlined above. You should use as many data points and growth variables, as well as historical data, as possible. Keep in mind that there may be missing values in the dataset.

**Task #2: Taxonomy Model**

In the provided dataset, there is a blank column called 'New Industry'. Your goal is to build a taxonomy model that can populate the 'New Industry' column with 'Finance', 'Health', or 'Finance; Health', depending on the following datapoints 'Description', 'Industry', 'Category Groups', and 'Tags'.

Good Luck! 🍀