Project Brief

Spark Funds, an asset management company wants to make investments in a few companies. The CEO of Spark Funds wants to understand the global trends in investments so that she can take the investment decisions effectively.

Business and Data Understanding

Spark Funds has two minor constraints for investments:

1. It wants to invest between **5 to 15 million USD** per round of investment
2. It wants to invest only in **English-speaking countries** because of the ease of communication with the companies it would invest in

- For your analysis, consider a country to be English speaking only if English is one of the official languages in that country

1. What is the strategy?

Spark Funds wants to invest where most **other investors are investing**. This pattern is often observed among early stage startup investors.

2. Where did we get the data from?

We have taken real investment data from **crunchbase.com**, so the insights you get may be incredibly useful. For this group project, we have divided the data into the following files:

You have to use two main data tables for the entire analysis:

1. Companies.txt
2. Rounds2.csv

3. What is Spark Funds' business objective?

The business objectives and goals of data analysis are pretty straightforward.

1. **Business objective:** The objective is to identify the best sectors, countries, and a suitable investment type for making investments. The overall strategy is to invest where others are investing, implying that the best countries are the ones where most investments are happening.

2. **Goals of data analysis**: Your goals are divided into two sub-goals:
   o **Investment type analysis**: Understanding investments in venture, seed/angel, private equity categories, etc. so Spark Funds can decide which type is best suited for its strategy.
   o **Country analysis**: Understanding which countries have had the most investments in the past. These will be Spark Funds' favourites as well.

Important Note: All your code has to be submitted in one main R file. For every checkpoint, keep writing code in one well-commented R file which you can submit at the end.

Also, you have to prepare a short presentation document to present the results of your analysis to the CEO of Spark Funds. This should briefly describe the important results and recommendations. You should include the plots in this presentation.