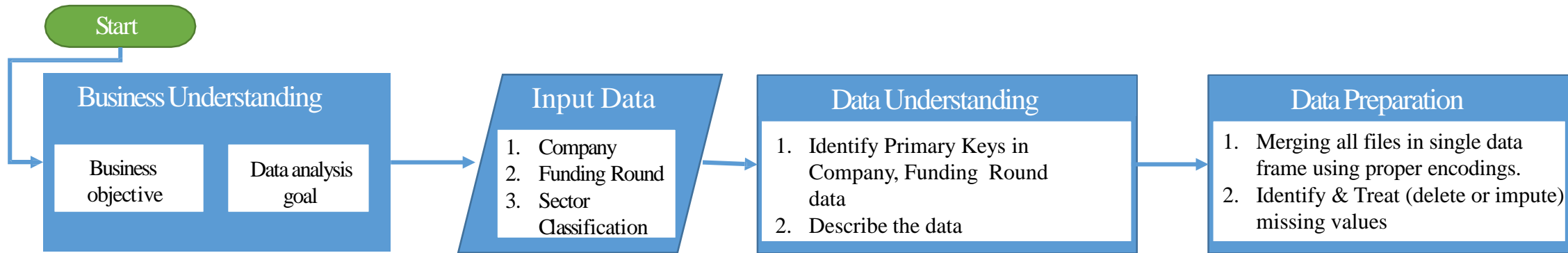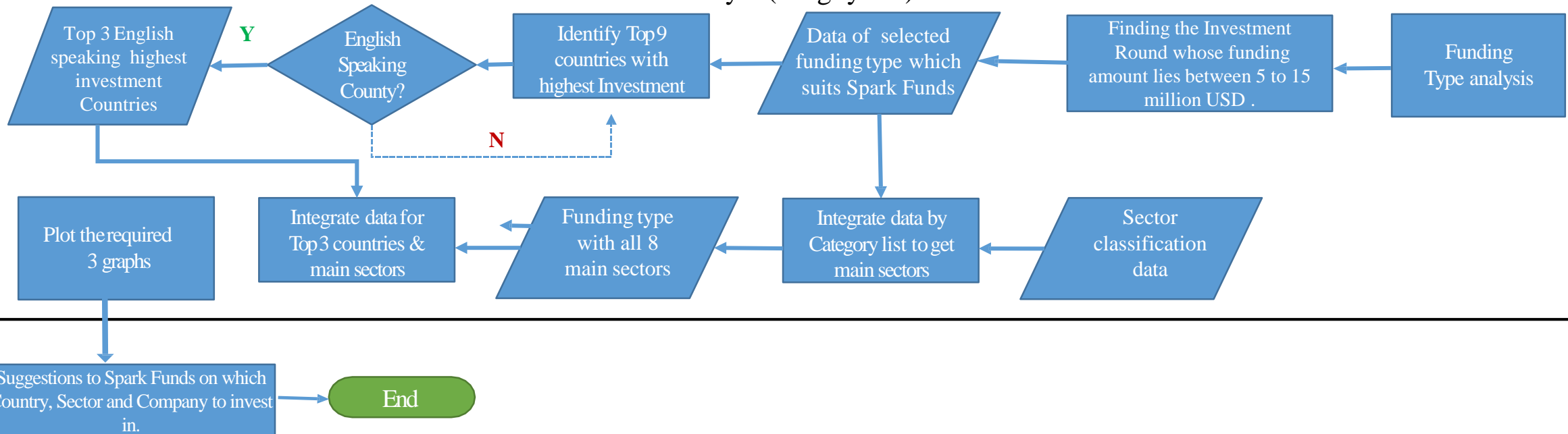# INVESTMENT ASSIGNMENT

SUBMISSION

Name: Sakshi Shrivastava

Abstract

- **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 Objective:** The objective is to identify the best sectors, countries, and a suitable investment type for Spark Funds to make investments.

- Two minor constraints of Spark Funds are:
  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.

- **Goals of data analysis:**
  - **Investment type analysis:** Suggest the best suited investment type from different investment rounds(like angel, seed, private equity, venture, etc.) present in the data.
  - **Country analysis:** Identifying the countries which have been the most heavily invested in the past.
  - **Sector analysis:** Understanding the distribution of investments across the eight main sectors and identify the top sectors to invest in.

We Received 3 Data Files to perform this analysis and suggest the best verticals for Spark Funds to Invest :

1. **companies.csv** : This file consists of the data of companies like their name, location of registration, company permalink, date of establishment, sectors in which the company operates etc.

2. **rounds2.csv** : This file consisted of the details of rounds in which the companies registered and the funding details they received in that round.

3. **mapping.csv** : This file consists of a mapping list of 600+ categories to 8 main sectors.
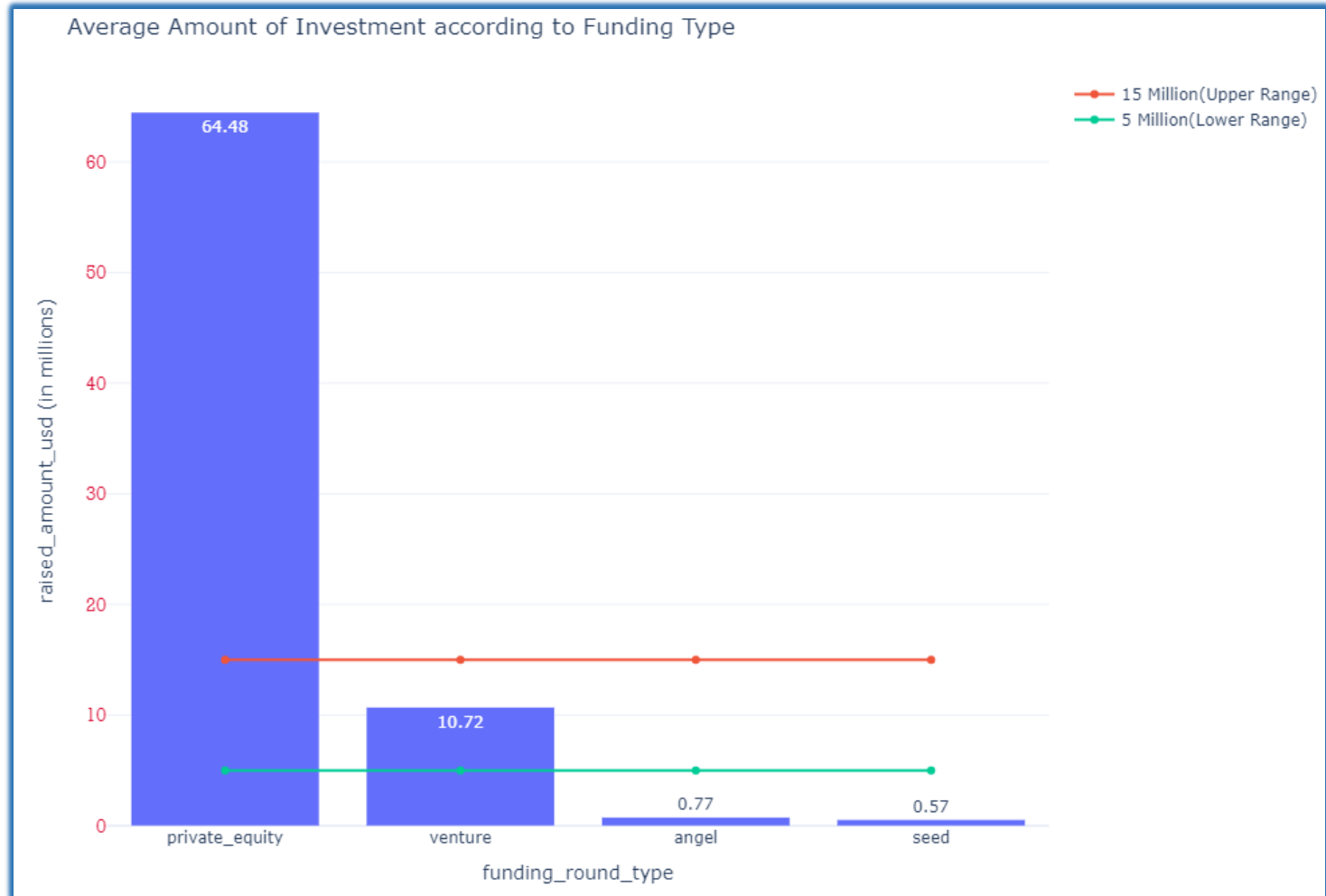
During initial analysis we came across a few data quality issues, which are as follows :

*   Both the files rounds2.csv and companies.csv contains lot of Unicode characters. These characters were then decoded into ascii format for processing

*   Missing values were replaced by 0 for raised_funding_amount variable as there can be companies which did not received any funding in the round they registered for. Also, Missing values for the columns with 'character' values are replaced by blanks ('') as blanks are also used in mapping file.

*   Many columns which were not effective in terms of analysis were directly dropped to reduce the data width.

*   A Major data Quality Issue was analyzed in the mapping file, the names of categories were misspelled as in many instances 'Na / na' was replaced by '0' . We corrected this anomaly before performing further analysis.

Assumptions made during analysis:

*   Replacing NAs in raised_amount_usd with 0 by assuming that the companies whose data for funding amount is missing raised 0 funding in that round.

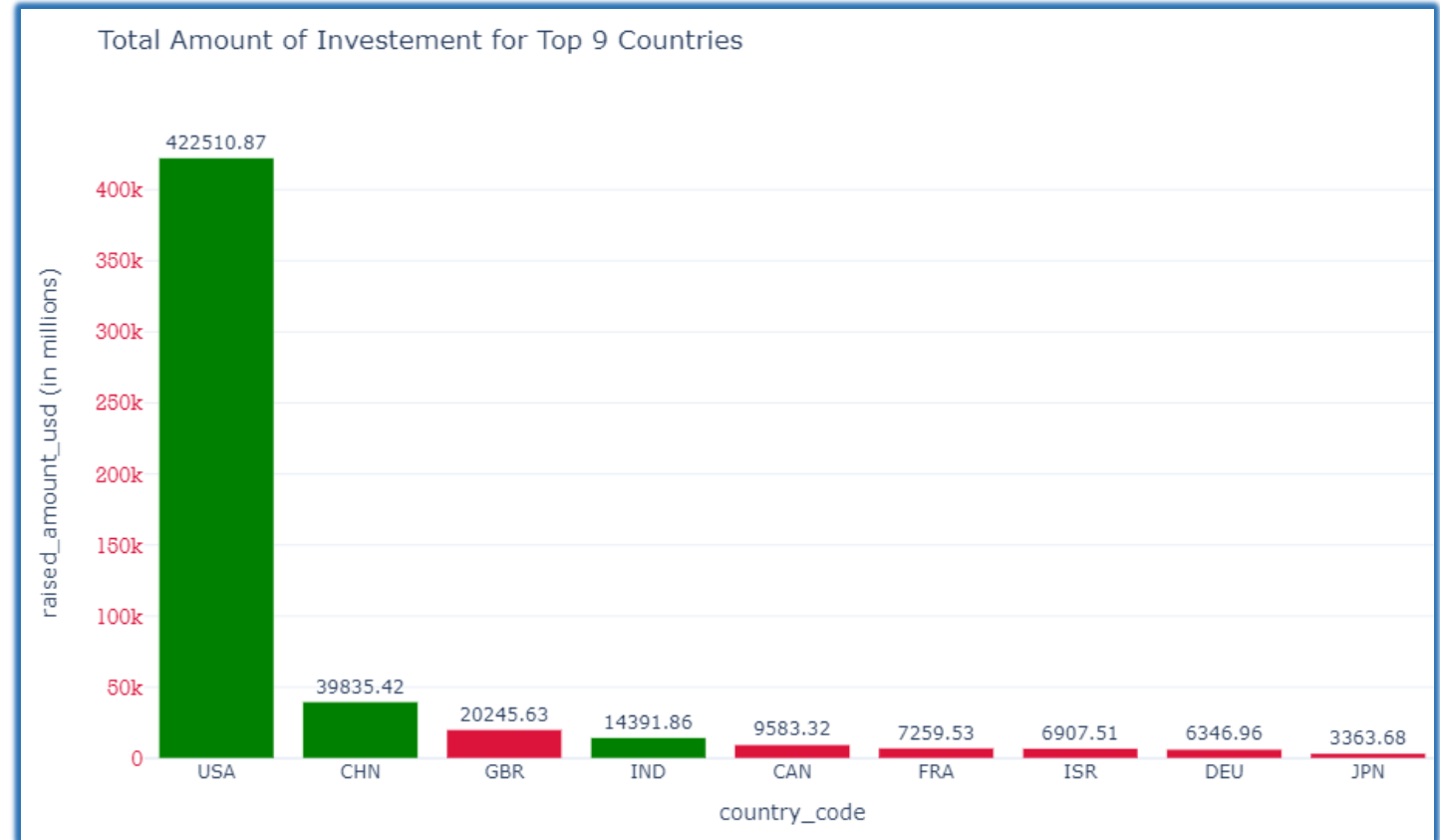*   Replacing NAs in other variables by blanks(") assuming that mapping data also replaces NAs by blanks(").

- Plot shows the average funding round amount for the 4 funding rounds venture, angel, private_equity and seed.

- Plot clearly shows that average funding amount for **Venture** round falls in the investment range of spark funds i.e. 5-15 M USD range.

- The green and red line shows the lower and upper range of investment for Spark Funds' i.e. 5 Million USD and 15 Million USD respectively.

Average Amount of Investment according to Funding Type

Country Type Analysis

- Steps to find top 9 countries for funding type **Venture**:
  1) Filtering the data where 'country_code' variable is not (") blank.
  2) Grouping the data by 'country_code' and finding the total funding raised by companies in each country.
  3) Sorting the countries based on total raised amount of each country.
  4) Selecting the top 9 countries based on sorting.
- USA receives the highest investment amount (approx. 422k Million USD)
- The top three English-speaking countries in the data frame top9 are(these were checked manually from pdf):
  1.USA
  2.GBR
  3.IND
- Neglecting China as English is not the official language in China.

Country	Total Investment Amount Received in Million USD
USA	422510.872
CHN	39835.418
GBR	20245.629
IND	14391.859
CAN	9583.323
FRA	7259.532
ISR	6907.514
DEU	6346.960
JPN	3363.676

- The above plot shows the top 9 countries based on the total amount of investments of funding type **Venture**.
- **USA, GBR and IND** are marked in green color as they are top 3 English speaking countries and the other countries are marked in red as English is not their primary language.

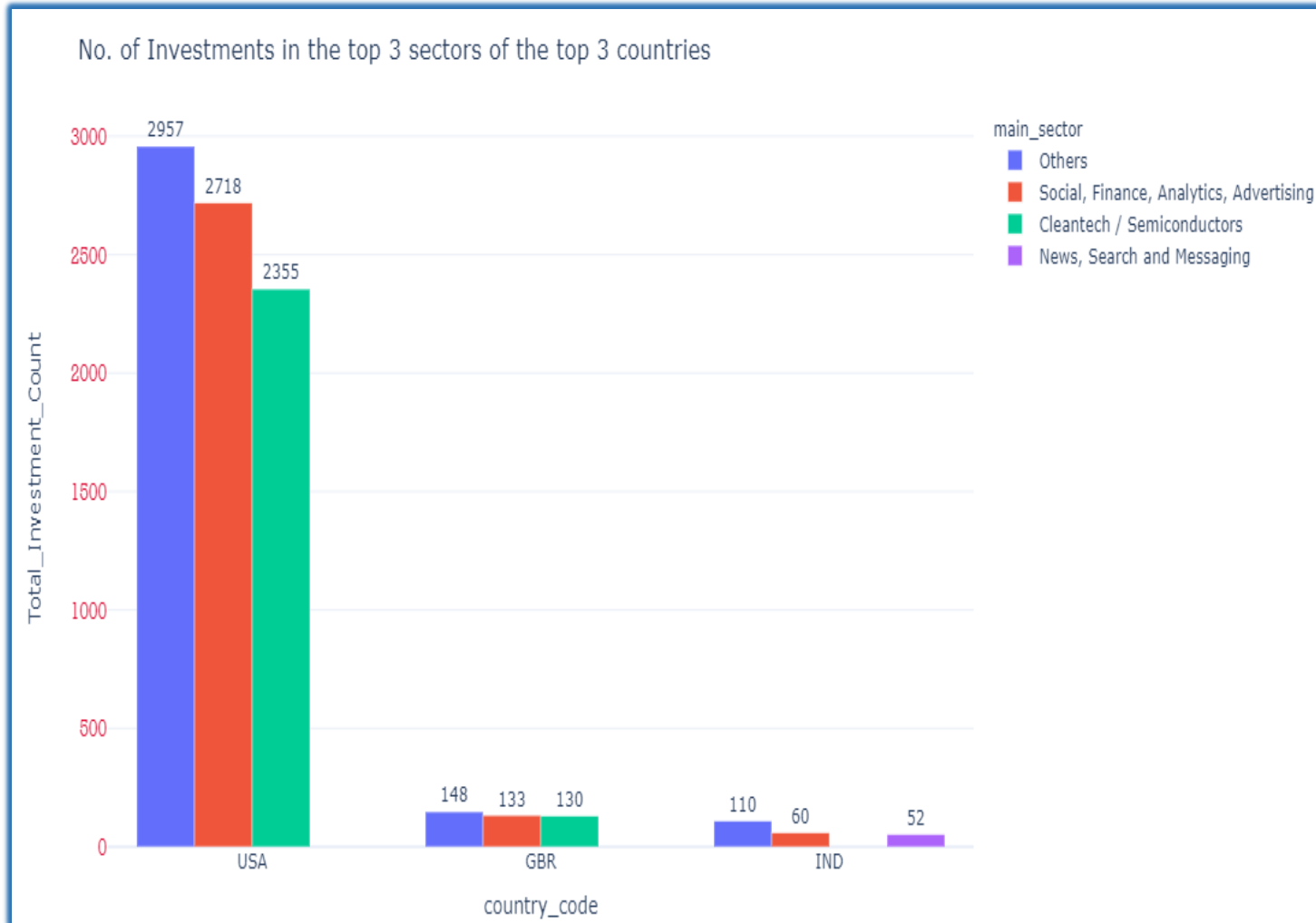Total Amount of Investement for Top 9 Countries

- Following table shows the Top 3 sectors and their investment count as well as the invested amount in top 3 English speaking countries. Best Company is the company with highest amount invested in that particular sector.

Country Name	Top 3 sectors	Count of Investments	Amount of Investment(in corresponding sector)	Best Company (in that sector & has Highest Amount)
USA	Others	2957	26376.009 M USD	virtustream
	Social, Finance, Analytics, Advertising	2718	23837.382 M USD	shotspotter
	Cleantech / Semiconductors	2355	21678.441 M USD	biodesix
GBR (United Kingdom)	Others	148	1298.623 M USD	electric-cloud
	Social, Finance, Analytics, Advertising	133	1089.407 M USD	celltick-technologies
	Cleantech / Semiconductors	130	1163 .988 M USD	eusa-pharma
IND	Others	110	1013.410 M USD	firstcry-com
	Social, Finance, Analytics, Advertising	60	550.549 M USD	manthan-systems
	News, Search and Messaging	52	433.835 M USD	gupshup-technology-india-pvt- ltd

- Plot shows the number of investments in the top 3 sectors of the top 3 countries in Funding Type Venture.

- Each country shows 3 main sectors which are the top sectors based on the investment count in that country.

- Clearly, **Others**(blue color) category is the top sector in all the 3 countries.

- **Social, Finance ,Analytics and Advertising**(red color) is the second most funded main sector in all the 3 countries.

- **Cleantech / Semiconducters**(green color) is the 3rd main sector in USA and GBR while **News, Search and Messaging**(purple color) is the 3rd most funded sector in IND.

No. of Investments in the top 3 sectors of the top 3 countries

main_sector
- ■ Others
- ■ Social, Finance, Analytics, Advertising
- ■ Cleantech / Semiconductors
- ■ News, Search and Messaging

- Based on analysis(keeping all the objectives, constraints in mind) **Venture Round** is the best suited funding round for Spark Funds.
- Under Venture Funding type, Spark Funds should invest in the following countries, sectors and companies:

Country to Invest	Top 3 Sectors to Invest	Top Companies to Invest
USA	Others	virtustream
	Social, Finance ,Analytics and Advertising	shotspotter
	Cleantech / Semiconducters	biodesix
GBR(United Kingdom)	Others	electric-cloud
	Social, Finance ,Analytics and Advertising	celltick-technologies
	Cleantech / Semiconducters	eusa-pharma
India	Others	firstcry-com
	Social, Finance ,Analytics and Advertising	manthan-systems
	News, Search and Messaging	gupshup-technology-India-pvt-ltd