

INVESTMENT ASSIGNMENT SUBMISSION

Name: Sumit Arora

Spark Funds Global Investment Analysis

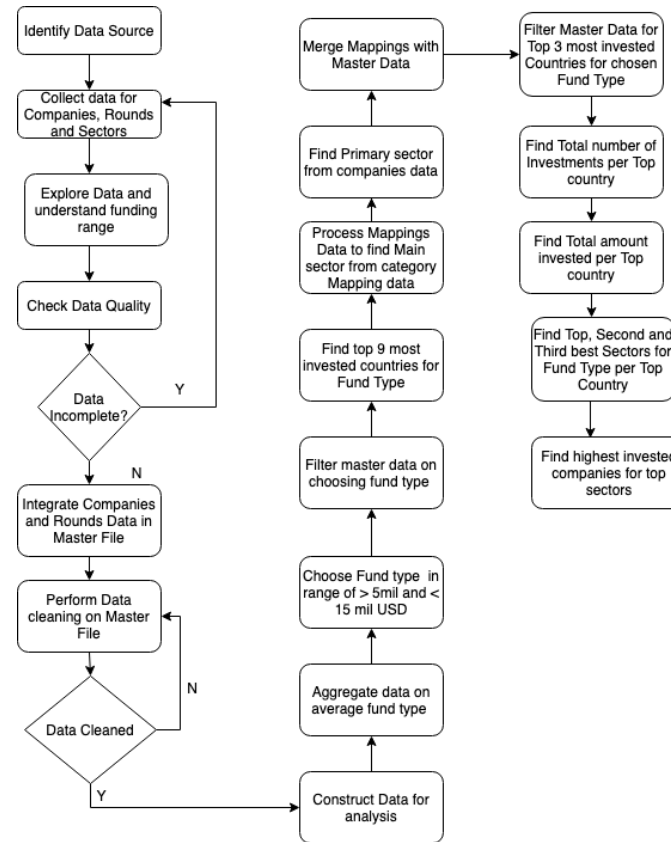
Spark Funds needs to make investments in few companies globally.

The objective of this case study is to identify best sectors, countries and a suitable investment type for making investments.

Spark funds has two minor constraints for investments:

- The range of per round investment is between 5 to 15 million USD
- It wants to invest only in English-speaking countries where English is one of the official languages in that country. This decision is based on the ease of communication with the companies it would invest in.

Analysis Methodology



Investment Type Analysis Assumptions

Data Integration

1. Permalink is used as unique key column for each company.
2. The master_frame is created with Rounds2 and Companies data by lowercasing permalink unique key column to avoid case sensitivity mismatch issues

Data Cleanup

The following data was removed from master_frame comprising of Rounds2 and Companies file

1. funding_round_code column had 72% of missing values
2. founded_at column had 18% of data missing
3. homepage_url was removed since the data was not utilised during analysis.
4. raised_amount_usd has 17% of values missing that resulting in removing 17% of data from master_frame since this column is used during analysis
5. country_code had 6% of missing values that resulted in removing rows.

Overall there's 20% of data reduction from original data which is enough to carry out analysis.

Country Analysis Assumption

1. Master Data is filtered with chosen Venture funding type to reduce data
2. Top nine countries which have received the highest total funding (across ALL sectors for the chosen investment type)
3. Top9 data frame is created with the top nine countries (based on the total investment amount each country has received)

Sector Analysis Assumptions

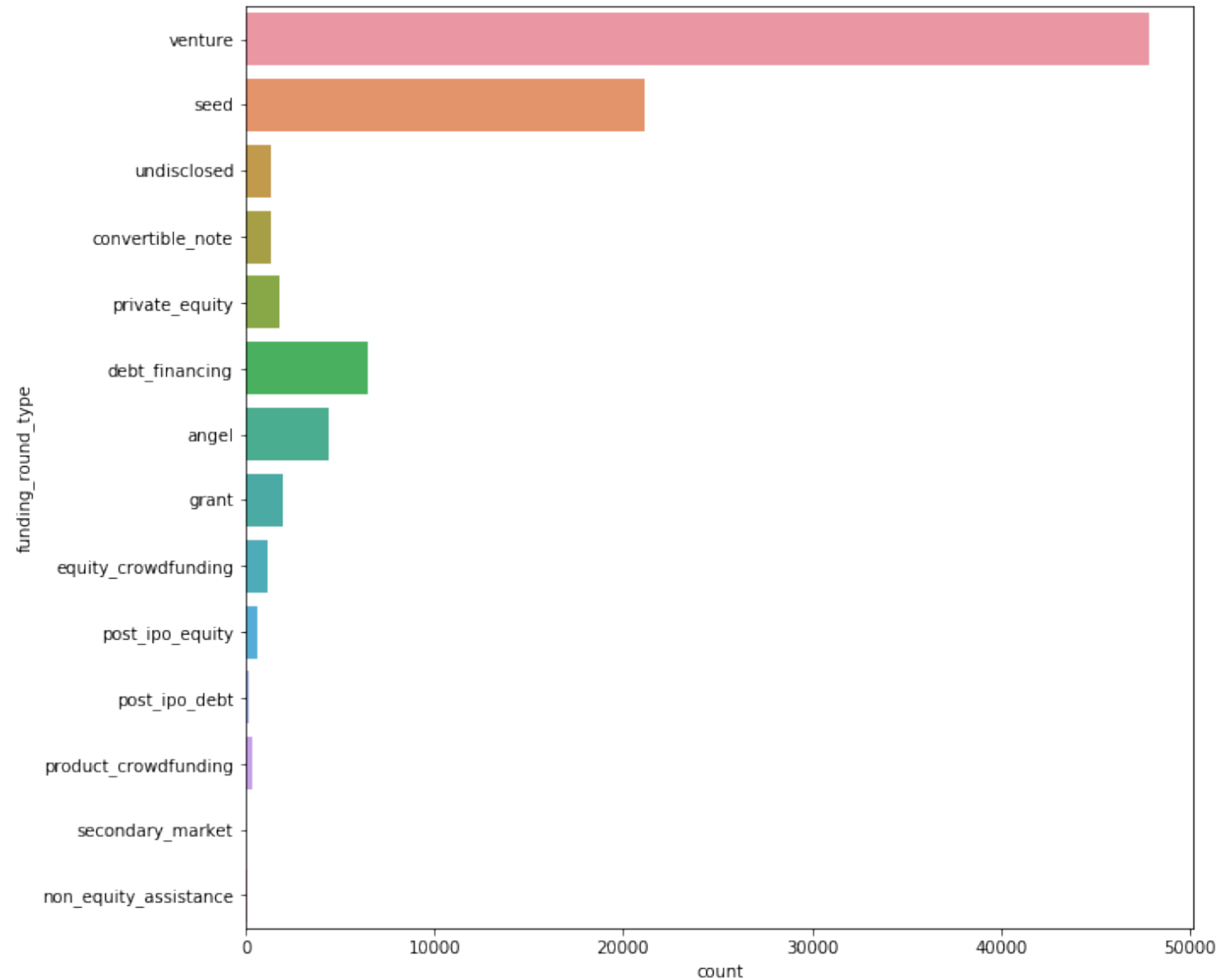
Mapping Data

1. Any missing Category List data has been removed
2. Also created a new column `main_sector` that was derived by Mapping data
3. Created a new column `primary_sector` based out of `category_list` and also lowercased for data joining.

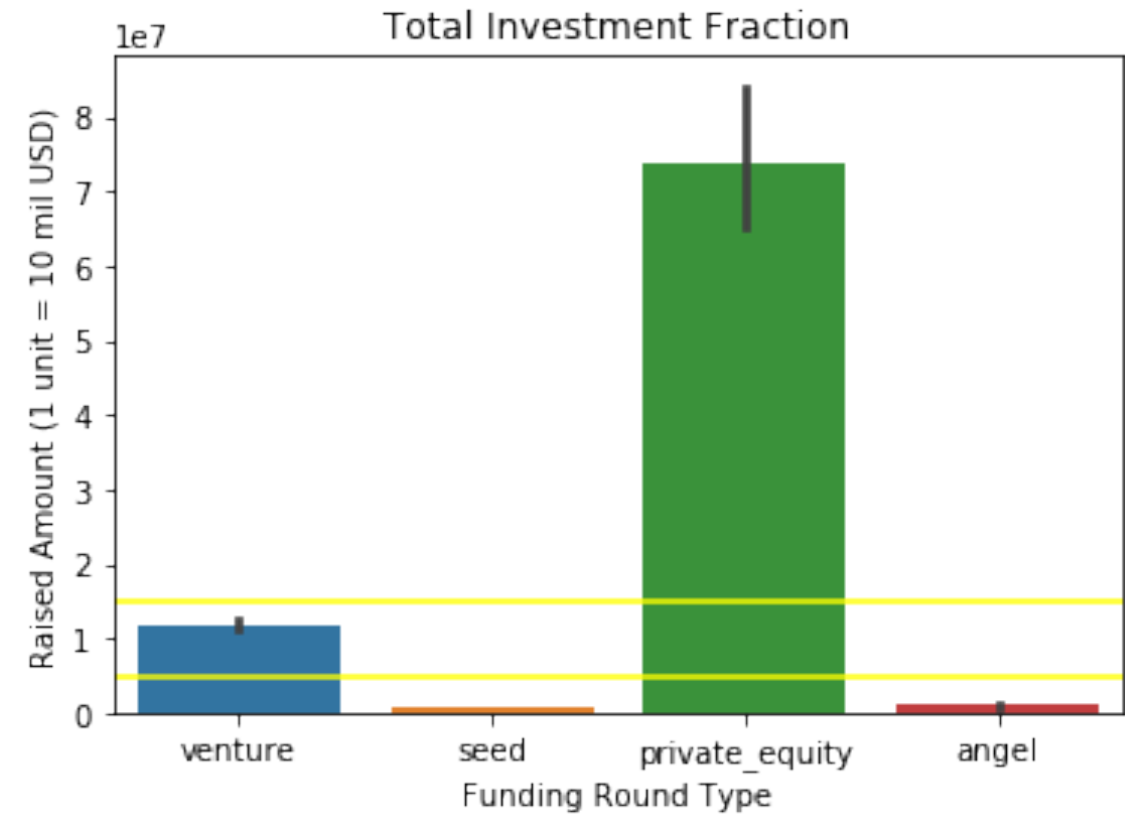
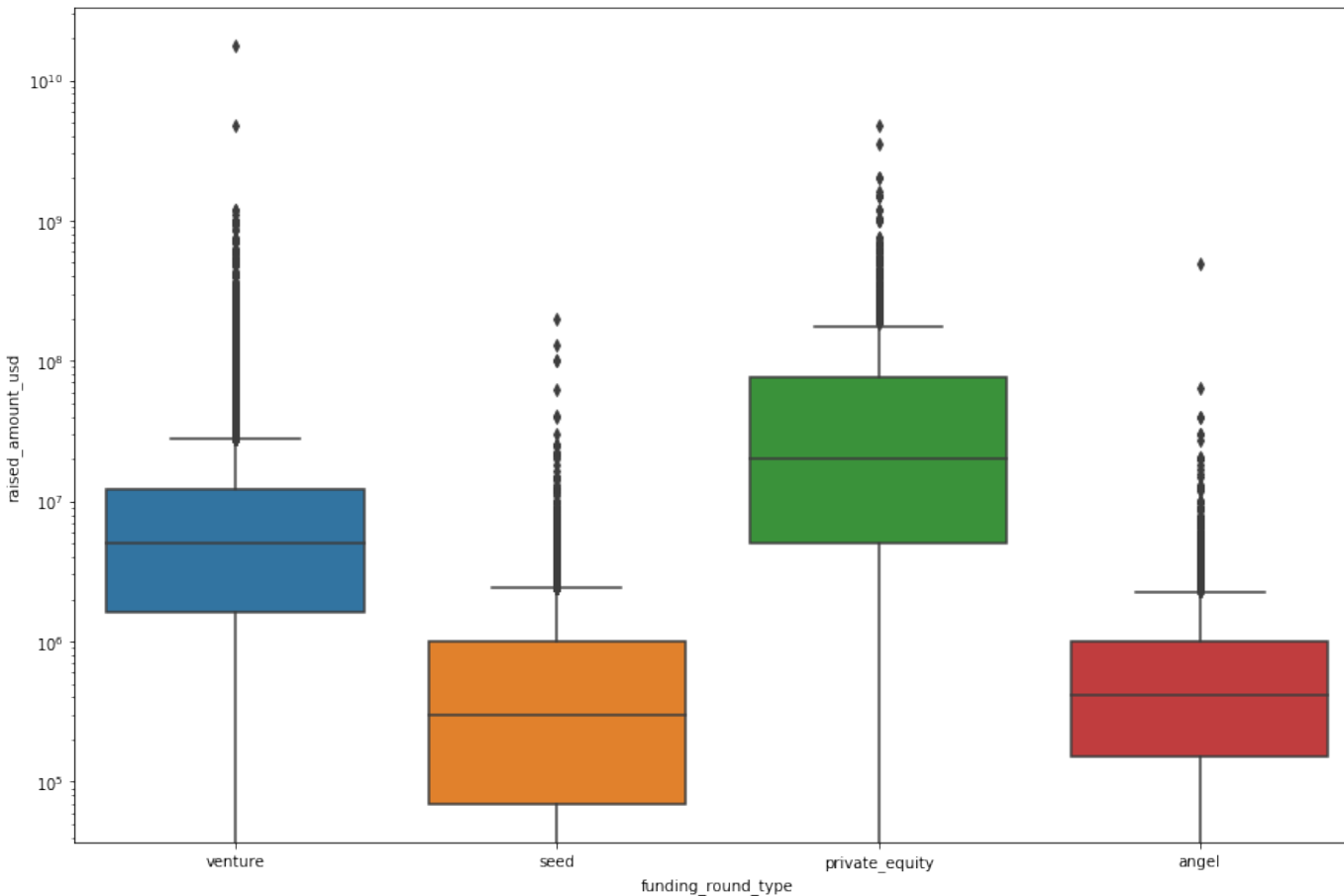
Company Data

1. Category list contains multiple categories and as per the requirement, first value of category is mapped to `primary_sector` column.
2. `primary_sector` is lowercased to avoid mismatching while joining.
3. Companies data is merged with Mapping data
4. D1, D2 and D3 data-frames are created with master data having budget range 5-15 million USD for Top 3 investment countries
5. Country wide analysis is performed on D1, D2 and D3 data frames to find out
 - Total number of investments
 - Total amount of investment (USD)
 - Top sector
 - Highest investment companies in Top investment countries

Current Investment Types

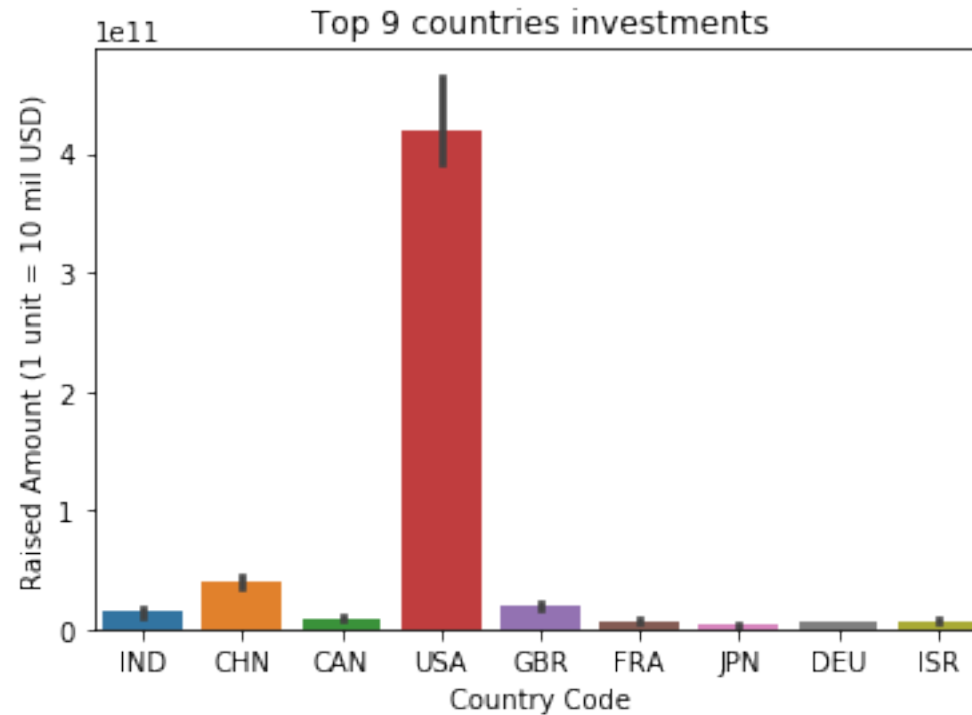


Raised funding amount (USD) for chosen 4 Funding Types

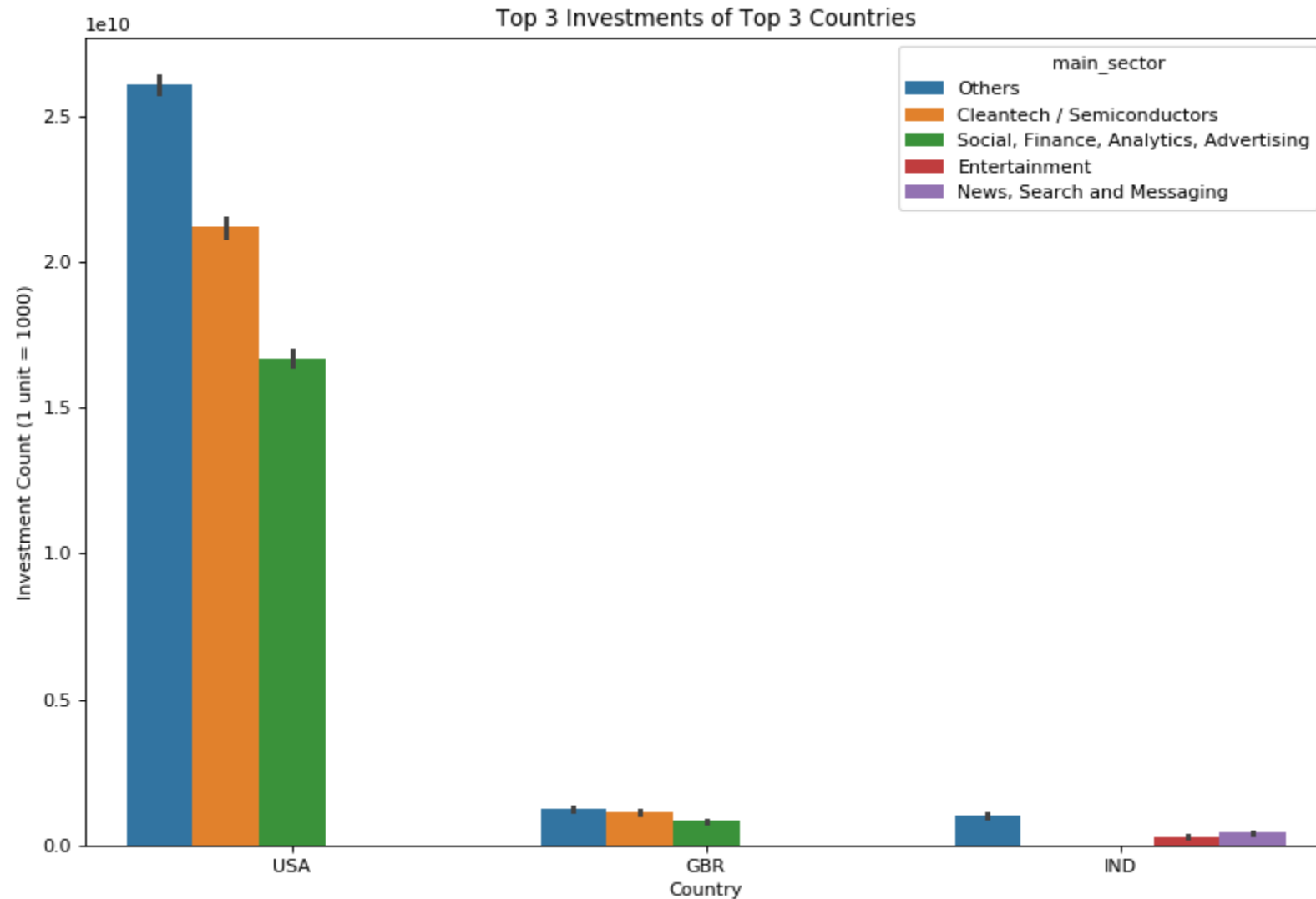


Venture Fund Type has maximum investments within 5-15 million range

Top 9 countries against the total amount of investments



Number of investments in the top 3 sectors of the top 3 countries



Conclusions

The analysis shows the range of budget to be invested 5-15 mil is best suited for Venture Fund Type.

As per the chosen funding type Venture, USA is the most preferred country for the investment.

The top sectors of investment are Others, CleanTech Semiconductors that are common sectors of investments for other top countries GBR and INDIA.

The Top countries to be invested for **1st top category** are

Company	Category	Country
1. virtustream	Others	USA
2. electric-cloud	Others	GBR
3. firstcry-com	Others	IND

The Top countries to be invested for **2nd top category** are

Company	Category	Country
4. biodesix	CleanTech/ Semiconductors	USA
5. eusa-pharma	CleanTech Semiconductors	GBR
6. gupshup-technology-india-pvt-ltd IND	News, Search and Messaging	IND

The investment budget could be divided among the above Top companies for safer investment strategy.