

# INVESTMENT ASSIGNMENT SUBMISSION

Name: D. Ajay Chandra

## <Abstract>

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

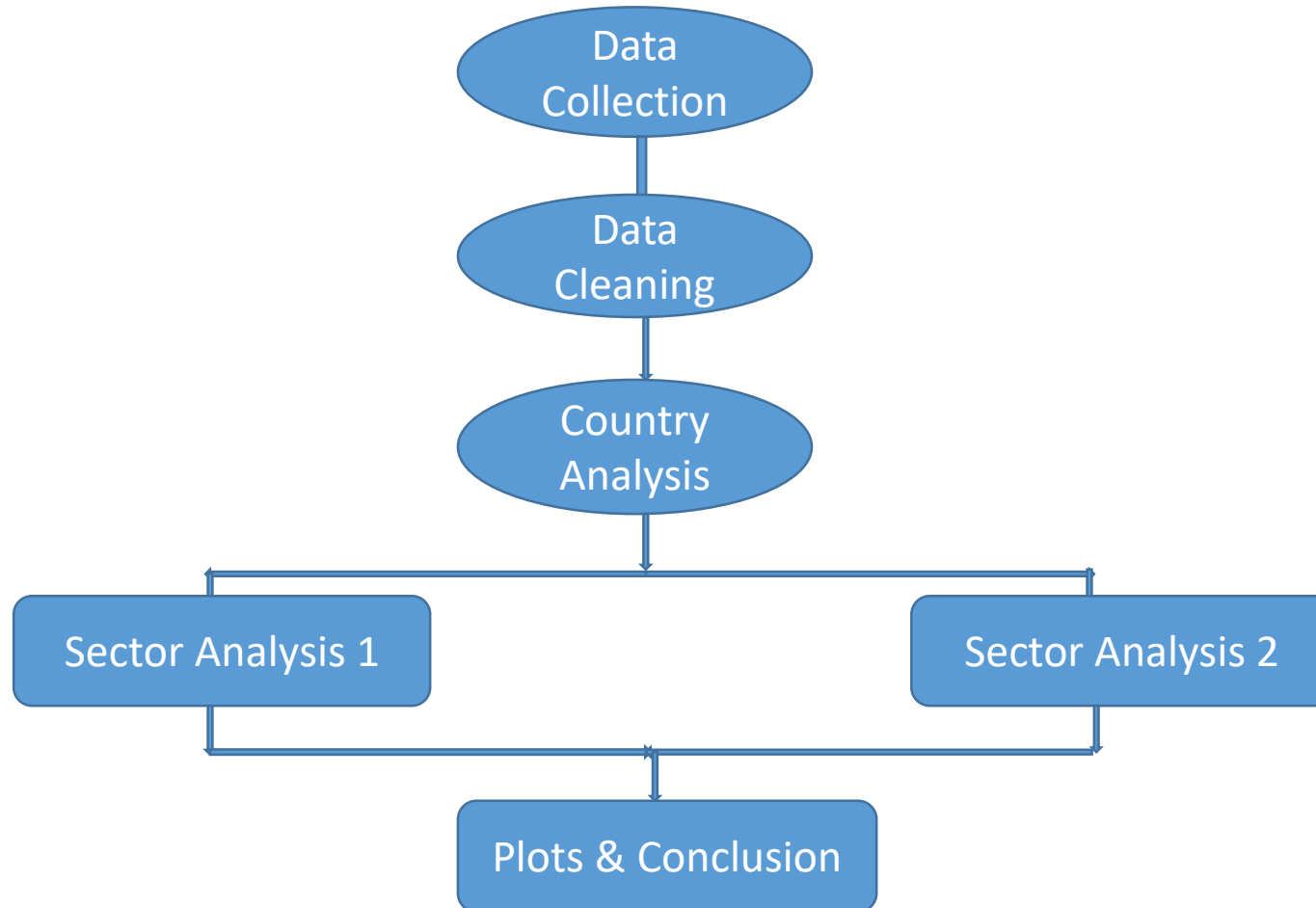
**Spark Funds has two minor constraints for investments:**

- It wants to invest between **5 to 15 million USD** per round of investment
- It wants to invest only in **English-speaking countries** because of the ease of communication with the companies it would invest in
- Spark Funds wants to invest where most **other investors are investing**. This pattern is often observed among early stage startup investors.

**The business objectives and goals of data analysis are pretty straightforward:**

- **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' sectors and countries are the ones 'where most investors are investing'.
- **Goals of data analysis:** Your goals are divided into three sub-goals:
  - **Investment type analysis:** Comparing the typical investment amounts in the venture, seed, angel, private equity etc. so that Spark Funds can choose the type that is best suited for their strategy.
  - **Country analysis:** Identifying the countries which have been the most heavily invested in the past. These will be Spark Funds' favourites as well.
  - **Sector analysis:** Understanding the distribution of investments across the eight main sectors.

# <Problem solving methodology>



## <Analysis>

### **Data Collection:**

1. Reading companies and rounds data CSV files into data frames using pandas, finding unique companies in companies & rounds data files and also the primary key.
2. Merging the two data frames into one master data frame with the help of the primary key available.

### **Data Cleaning:**

1. Identifying the missing values in the merge data frame i.e. (Master Frame) both in columns and rows and dropping the rows with null values.
2. Checking the percentage of the null values in both columns and rows and cleaning the data by removing the rows with null values greater than 5 and then removing the rows with even one null value.
3. Before making the analysis need to make sure that the master frame is completely filled with values and there are no null values.

## <Analysis>

### **Country Analysis:**

1. Finding the most representative value of each fund type invested across the countries.
2. Identifying the top nine countries who have received most number of funds through the investments.
3. Identifying the top three countries out of the nine where they are among the top 3 English speaking countries in the world.

### **Sector Analysis 1:**

1. Extracting the primary sector of each category list from the category list column.
2. Use the mapping file 'mapping.csv' to map each primary sector to one of the eight main sectors.

## <Analysis>

### **Sector Analysis 2:**

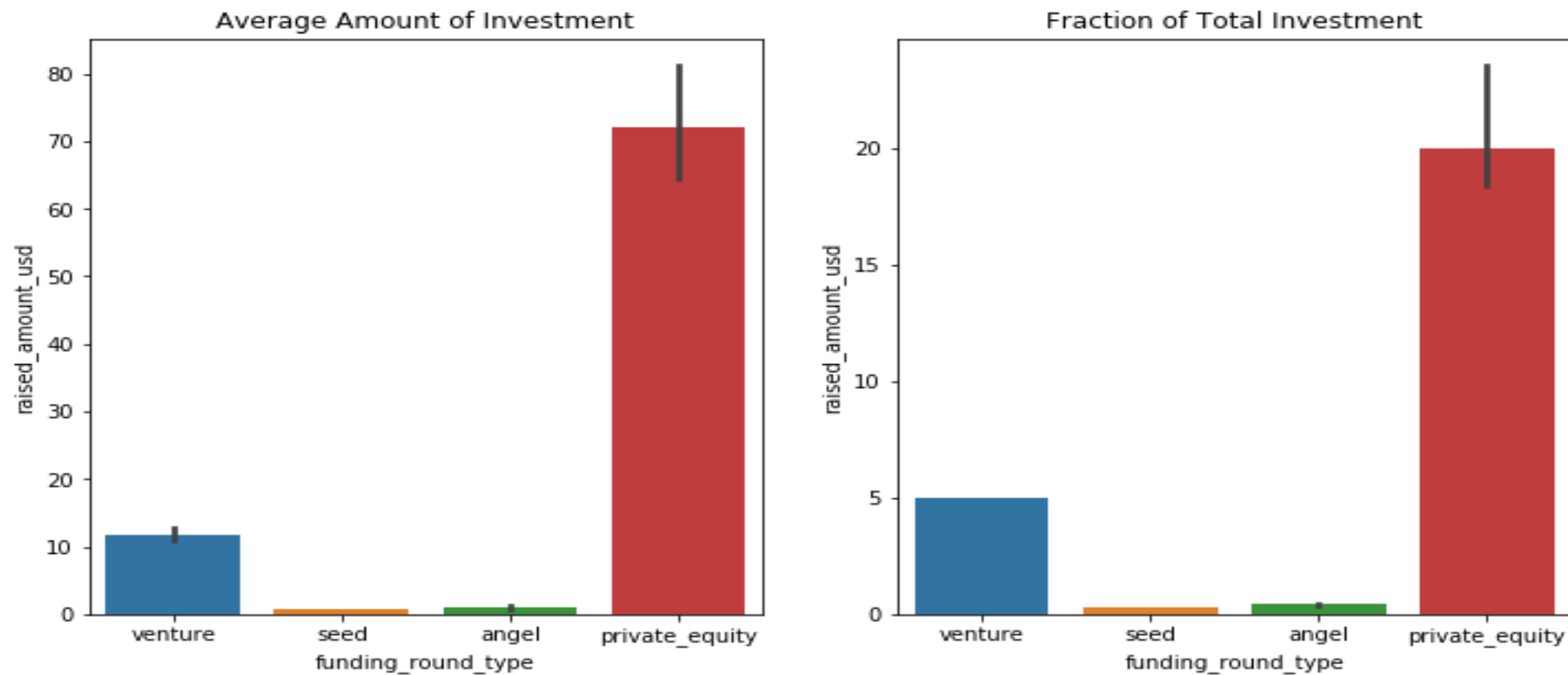
Created three separate data frames D1, D2 and D3 for each of the three countries containing the observations of funding type FT falling within the 5-15 million USD range.

The three data frames contain:

- All the columns of the master frame along with the primary sector and the main sector
- The total number (or count) of investments for each main sector in a separate column
- The total amount invested in each main sector in a separate column
- Using the three data frames, calculated the total number and amount of investments in each main sector.

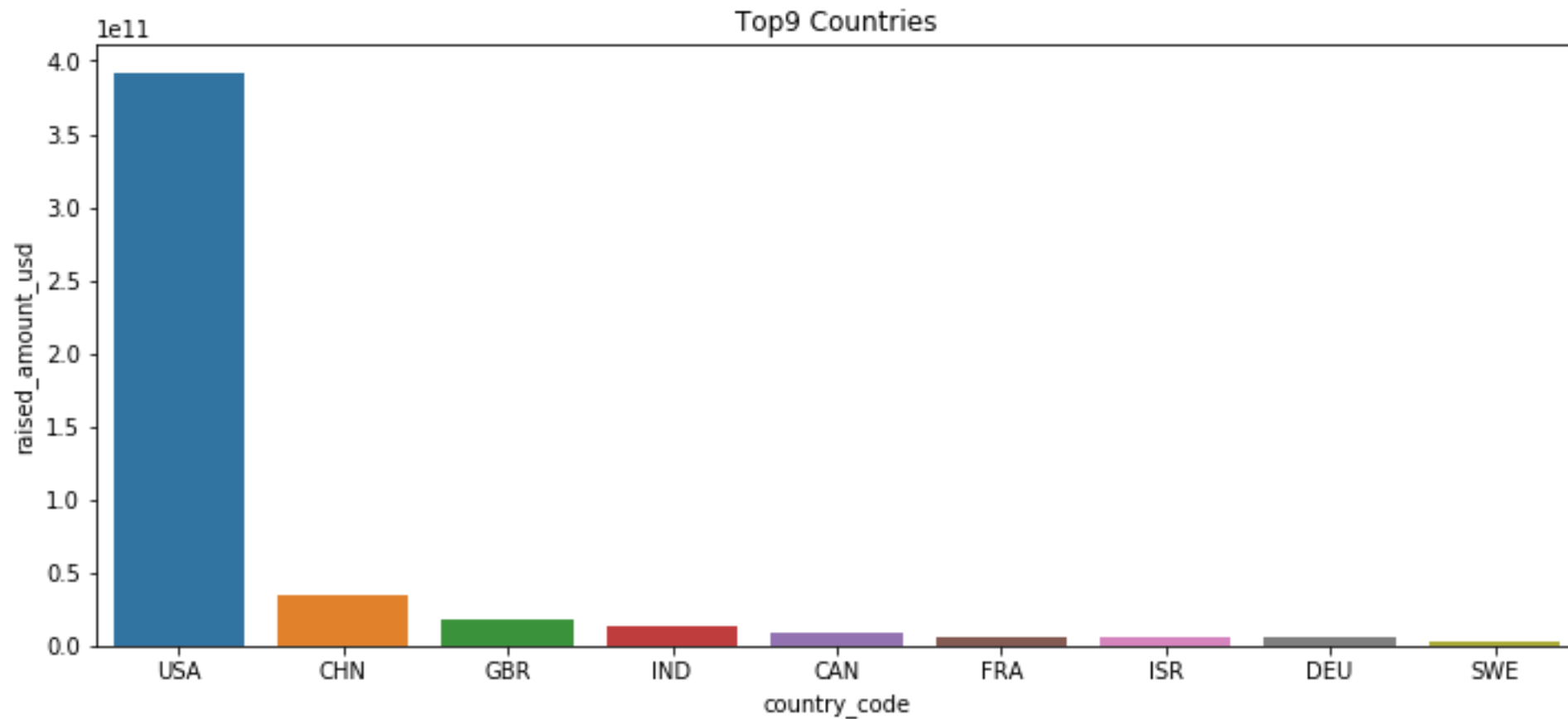
## <Results>

Plot 1: A plot showing the fraction of total investments (globally) in venture, seed, and private equity, and the average amount of investment in each funding type. So we can conclude that investing in **Venture** funding type (FT) is suitable for Spark Funds as it wants to invest between **5 to 15 million USD** per round of investment.



## <Results>

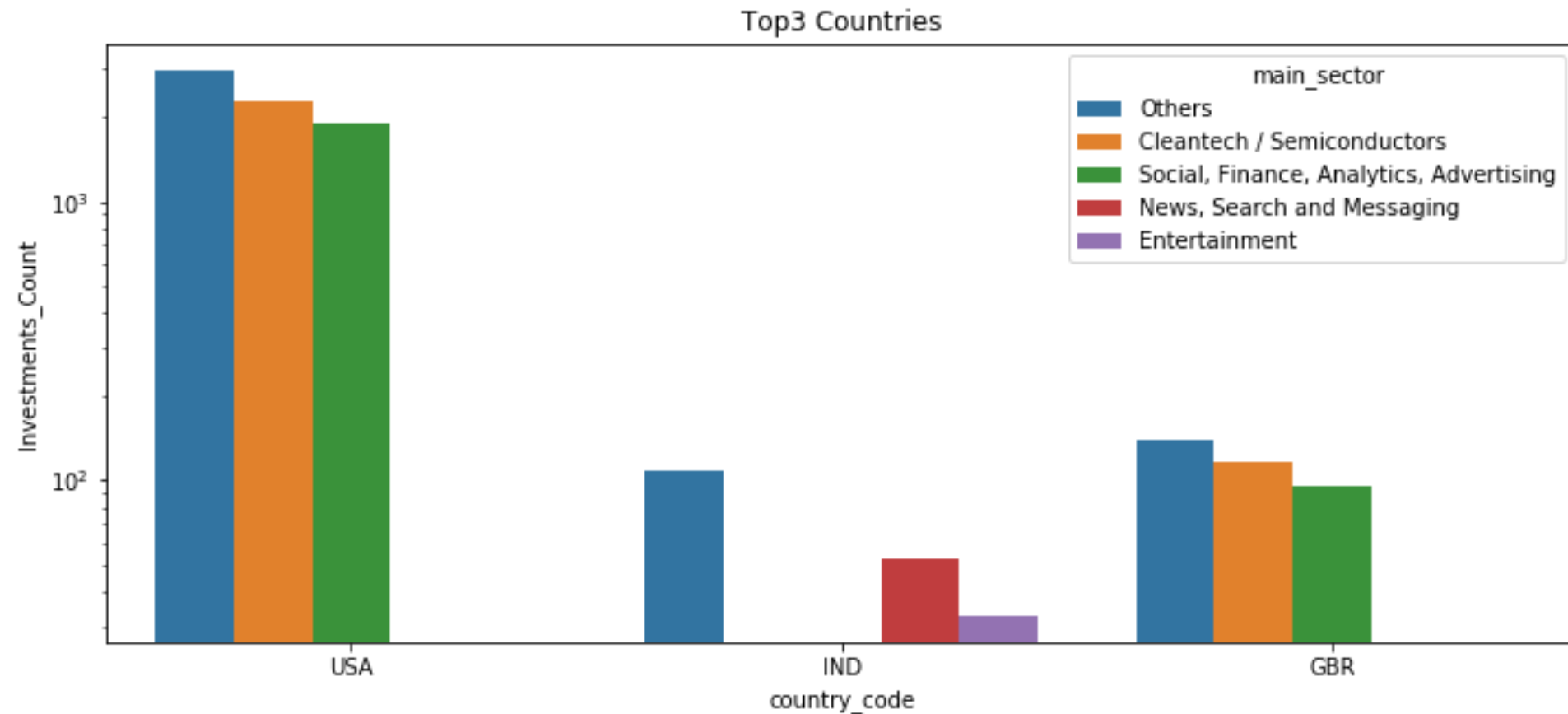
**Plot 2:** A plot showing the top 9 countries against the total amount of investments of **Venture** funding type FT.





## <Results>

**Plot 3:** A plot showing the number of investments in the top 3 main sectors of the top 3 countries in one chart (for the chosen investment type FT).



## <Conclusions>

Based on the assumptions made while cleaning the data, analysis of the data as per the constraints, business objective and goals of the company. I believe the company can invest in the following funding type & companies:

1. The company can invest in **Venture** funding type (FT) as it wants to invest between **5 to 15 million USD** per round of investment.
2. Among the top 9 countries against the total amount of investments of **Venture** funding type the company can invest in **USA, IND** and **GBR** as it wants to invest only in **English-speaking countries**.
3. Amongst the top 3 English speaking countries the company can invest in **USA** mostly on Others, Cleantech / Semiconductors & Social, Finance, Analytics, Advertising sectors as it is where most other investors are investing
4. In USA among the sectors mentioned above the company can invest in **Virtustream** and **Biodesix** companies as they are the companies which received highest investments.