INVESTMENT ANALYSIS – SPARK FUNDS ASSIGNMENT

SUBMISSION

Name:

Sachin Karthik Krishna

Email ID:

sachin.karthik.7598@gmail.com

Investment Analysis for Spark Funds

Spark Funds, an asset management company, is currently looking for investment opportunities in English Speaking countries and it is looking to invest anywhere in between $5 - $15 Million per round of investment.

As a data analyst, I was tasked to identify the

- Most suitable type of funding round for Spark Funds, based on the financial constraint they have set for each round of funding

- Top 3 English speaking countries where maximum funding has happened for the most suitable type of funding round

- Most suitable type of sectors for funding based on the number of funds raised by each of these sectors in the top 3 English speaking countries

CRISP-DM Framework

I have used the CRISP-DM Framework to approach this problem and dial down on a viable solution set.

Business objective: Identify the most widely invested sectors in Top 3 English Speaking countries

Data Resources:
→ Companies dataset which contains basic info all the companies which have gotten funding between 1960 and 2015
→ Round2 dataset which contains information about the funding type, amount raised in USD, funded date, etc.
→ Mapping dataset which contains information about the different sub sectors mapped to it's main sector

Analysis Process – Number of unique companies

➢ The analysis process started out with loading the companies and the round 2 dataset.

➢ The companies dataset and the round 2 dataset have a unique identifier known as a permalink. Using that, the number of unique companies were identified in both the datasets. As it turns out, the number of companies in both the datasets are equal. The number of unique companies both the datasets contain are 66,368.

```
1  #calculating the number of unique identifiers for companies in the round2 dateset
2  len(round2.company_permalink.value_counts())
```

66368

```
1  #calculating the number of unique identifiers for companies in the companies dateset
2  len(companies.permalink.value_counts())
```

66368

➢ After verifying the number of companies in both the datasets, the companies and the round2 datasets were merged to form the master frame

Analysis Process – Missing value treatment

- The number of missing values per column in the master frame dataset showed that

  → 23% of the values were missing in the **founded_at** column and 10% of values in the **country_code** column

  → 73% of the values were missing in the **funding_round_code** and 17% of values in the **raised_amount_usd** columns

  → 13%, 12%, 12% of the values were missing in the **state code, city and region** columns, respectively.

  → 8% of the values in the **home page** and 5% of the values in the **category list** column too were missing.

- The exact reasons for the missing values are unknown but one assumption that can be made for the missing values in **raised_amount_usd** column is that amount went undisclosed or the deal didn't cut through but the data got recorded as a false positive.

- The missing values only in the **raised_amount_usd** column and the **country_code** column was dropped since only the values in those 2 columns were essential for the upcoming analysis.

- The missing values in the other columns were left untouched since there wasn't a necessity to treat them for this particular analysis.

Analysis Process – Most suitable funding type & Top 9 Countries

- The different types of funding were grouped and the values of mean and median were extracted. A box plot was also plotted to see the spread of the different funding types.

- The following insights were gained from the mean and median values of different funding types

  → Funding type **venture's mean** values in **$11.74 million** and **median** value was **$5 million**. Both the values fall within the bracket of Spark funds' financial constraint

   → **Seed and Angel** Funding types have a mean and media value in the order of **Hundred thousands**

   → **Private equity funding type** has a **mean** value of **$733 million** and a **median value** of **$200 million**

- Therefore the most suitable type of funding for Spark Funds to invest in is **Venture** since it comfortably **falls in range of its financial constraint.**

- The amount of funds raised by each country was calculated consecutively and the top 9 countries were put into a separated dataset known as Top_9.

Analysis Process- Sector Mapping

- The mapping dataset was modified in a way that the primary sector was beside the main sector category and the other unnecessary rows were removed to make the dataset concise.

- Consecutively the primary sector was extracted in the master frame from the category list column.

- The mapping dataset was merged with the master frame to create master_frame_map.

- Three dataset, D1, D2, and D3 were created which pertained to the **top 3 English Speaking countries – USA, India, Great Britain**.

- These three datasets also were filtered to contain values of only venture type funding round and amount raised was in between $5 million and $15 million.

- After this, the top 3 sectors were identified in each country. The top 3 sectors were according to the number of investments made in each sector. The top 3 sectors in each country are

    → **Others and Social, Finance, Analytics, Advertising** took the **top 2 spots** in **all the three countries**.

    → **Cleantech/ semiconductor** took the **3rd place** in the **USA and Great Britain**

    → **News, Search and Messaging** took the **3rd spot in India**

- The total amount of funding raised by all the sectors were calculated for each country.
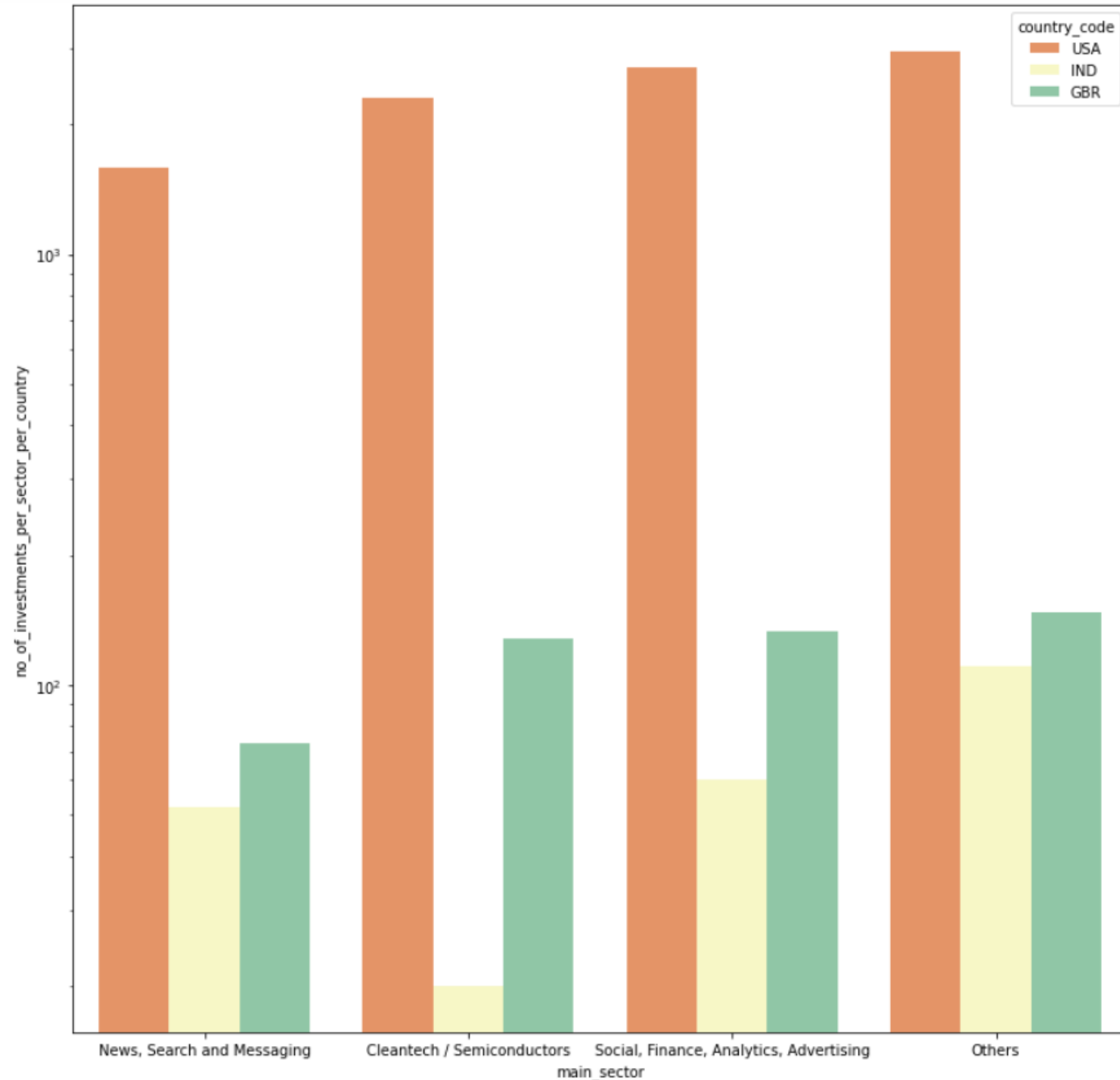
Total Amount and Average amount invested in seed, angel, venture and private equity
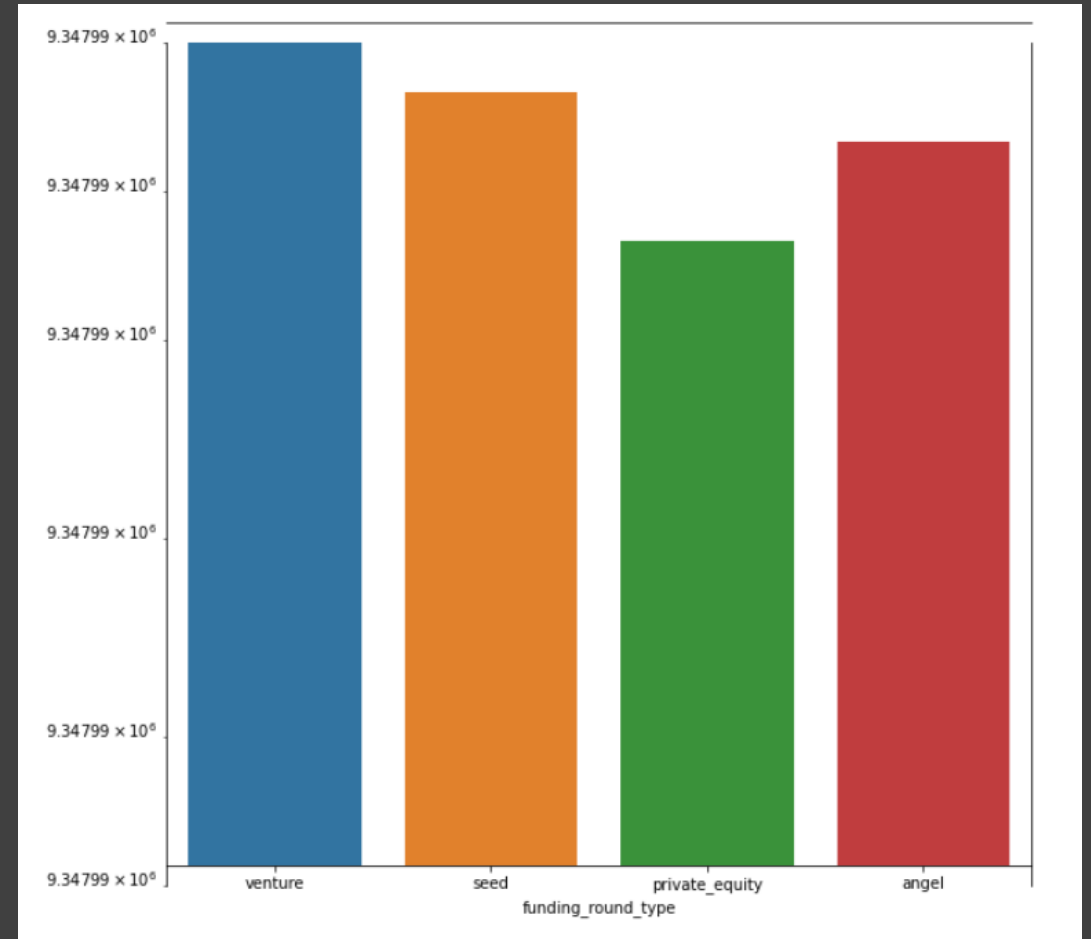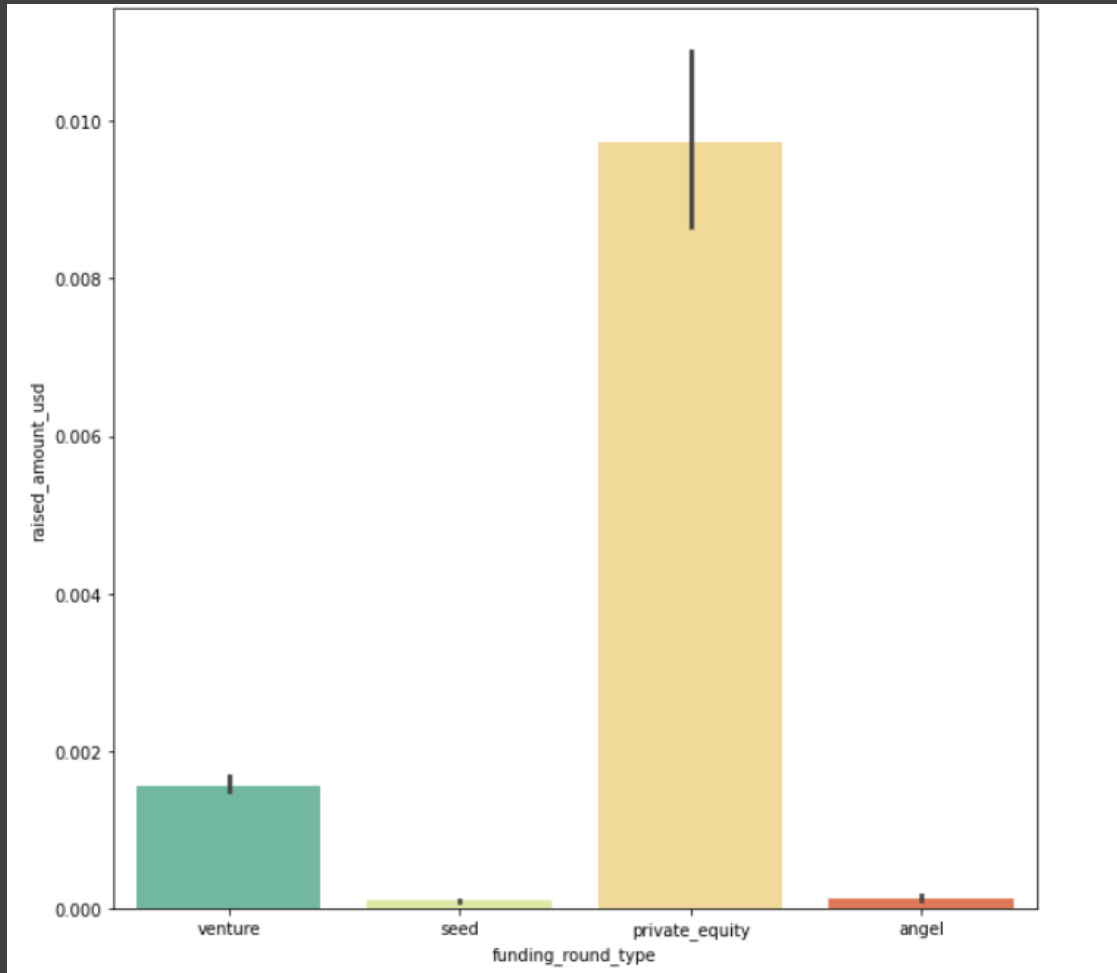
Total amount of funds invested in the highly funded 9 countries
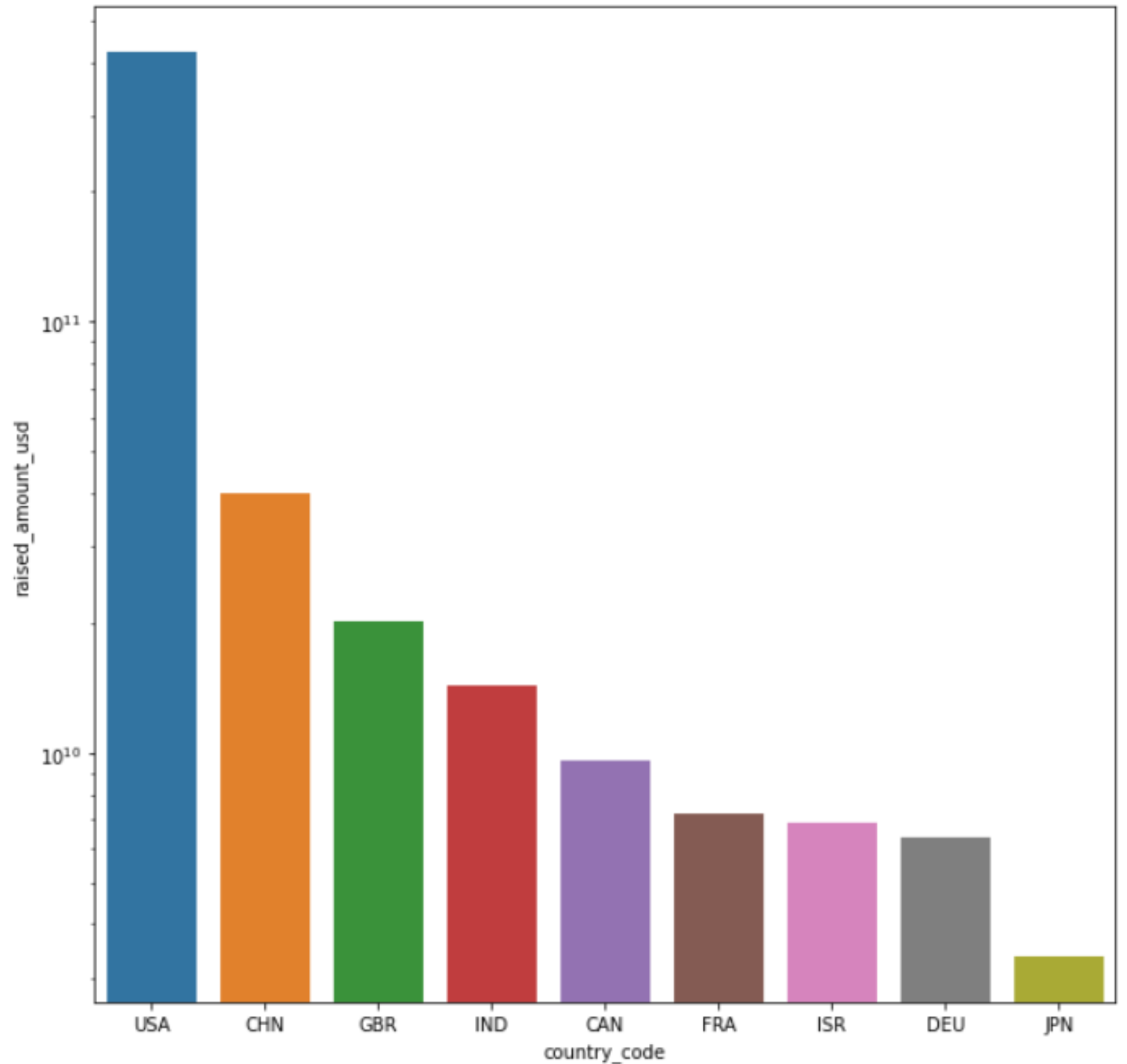
No of investments made in the top 4 sectors in the top 3 English speaking countries(USA, India, Great Britain)

Conclusions

➢ Since Spark funds is looking to invest in the most widely invested country and in one of the most widely invested sectors, it should place its money in the United States.

➢ The reasons for choosing United States over the other 2 countries are

→Number of investments raised by USA is almost 20 times as much as Great Britain and 36 times as much as India

→Ease of communication since English is the de facto language in USA whereas it is only the official language in India, among many other languages.

➢ Spark Funds could invest its money in either Others or Social, Finance, Analytics, Advertising sectors since the number of investments raised by both the sectors in the United States in relatively similar.