INVESTMENT ASSIGNMENT

SUBMISSION

Name: Supriya Raman

SPARK FUNDS HAVE A PROBLEM STATEMENT !!!

Spark Funds 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 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'

Following are some of the considerations –

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

3. Spark Funds wants to understand the trend and invest where most other investors are investing.
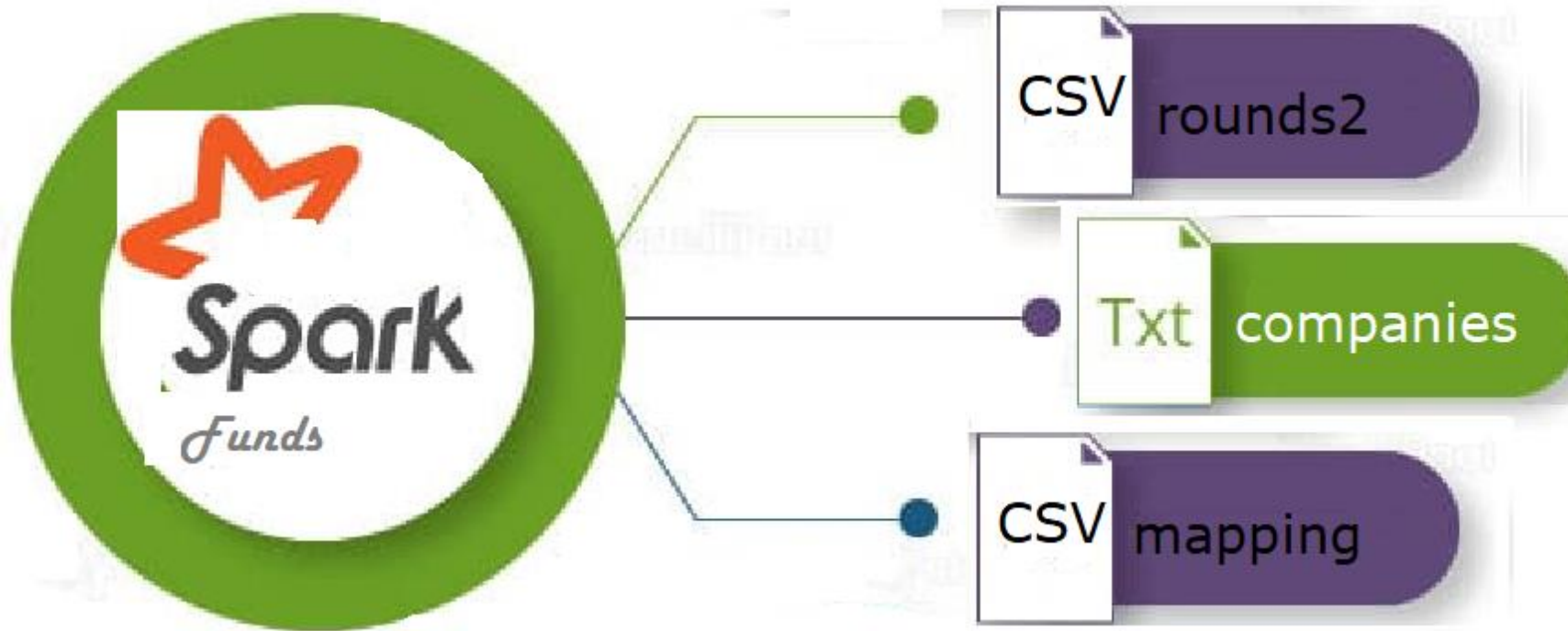
LET'S SET THE GOAL

Business Objective has been divided into following sub-goals:-

- **Investment type analysis:** Compare the typical investment amounts in each of the below funding types to choose the type one which best suites for Spark Funds' strategy-

    venture

    seed

    angel

    private equity

- **Country analysis:** Identify the top three English-speaking countries which have been the most heavily invested in the past.

- **Sector analysis:** Understanding the distribution of investments across the eight main sectors. Listed below
    - **Automotive & Sports**
    - **Cleantech / Semiconductors**
    - **Entertainment**
    - **Health**
    - **Manufacturing**
    - **News, Search and Messaging**
    - **Others**
    - **Social, Finance, Analytics, Advertising**

DATA FILES

Data collected from real investment data from crunchbase.com

rounds2 - Funding round details

Attributes	Description
company_permalink	Unique ID of company
funding_round_permalink	Unique ID of funding round
funding_round_type	Type of funding – venture, angel, private equity etc.
funding_round_code	Round of venture funding (round A, B etc.)
funded_at	Date of funding
raised_amount_usd	Money raised in funding (USD)

Companies - companies' details

Attribute	Description
Permalink	Unique ID of company
name	Company name
homepage_url	Website URL
category_list	Category/categories to which a company belongs
status	Operational status
country_code	Country Code
state_code	State

mapping - Sectors classification

Attributes	Description
primary sector	category names in the companies table
main sector	eight broad sector names

DATA EXPLORATION AND PREPARATION

Companies Table –

- List of companies present in companies.txt are loaded into dataframe named "df_companies"

- Data is read using "palmos" encoding and Tab as a separator

- Column "permalink" is converted to lower case to enable merge of dataframes.

rounds2 Table –

- This table contains funding rounds details present in rounds2.csv are loaded into dataframe named "df_rounds2"

- Data is read using "palmos" encoding

- Column "company_permalink" is converted to lower case and the column name is renamed to "permalink" – this is done to enable merge of dataframes.

Both the above dataframes are merged to form "master_frame" dataframe to facilitate all analysis.

FUNDING TYPE ANALYSIS

- Sparks Funds want to invest between 5 to 15 million USD per round of investment

- The objective of this analysis is to choose one suitable funding type from among below four funding types based on the most representative investment amount –
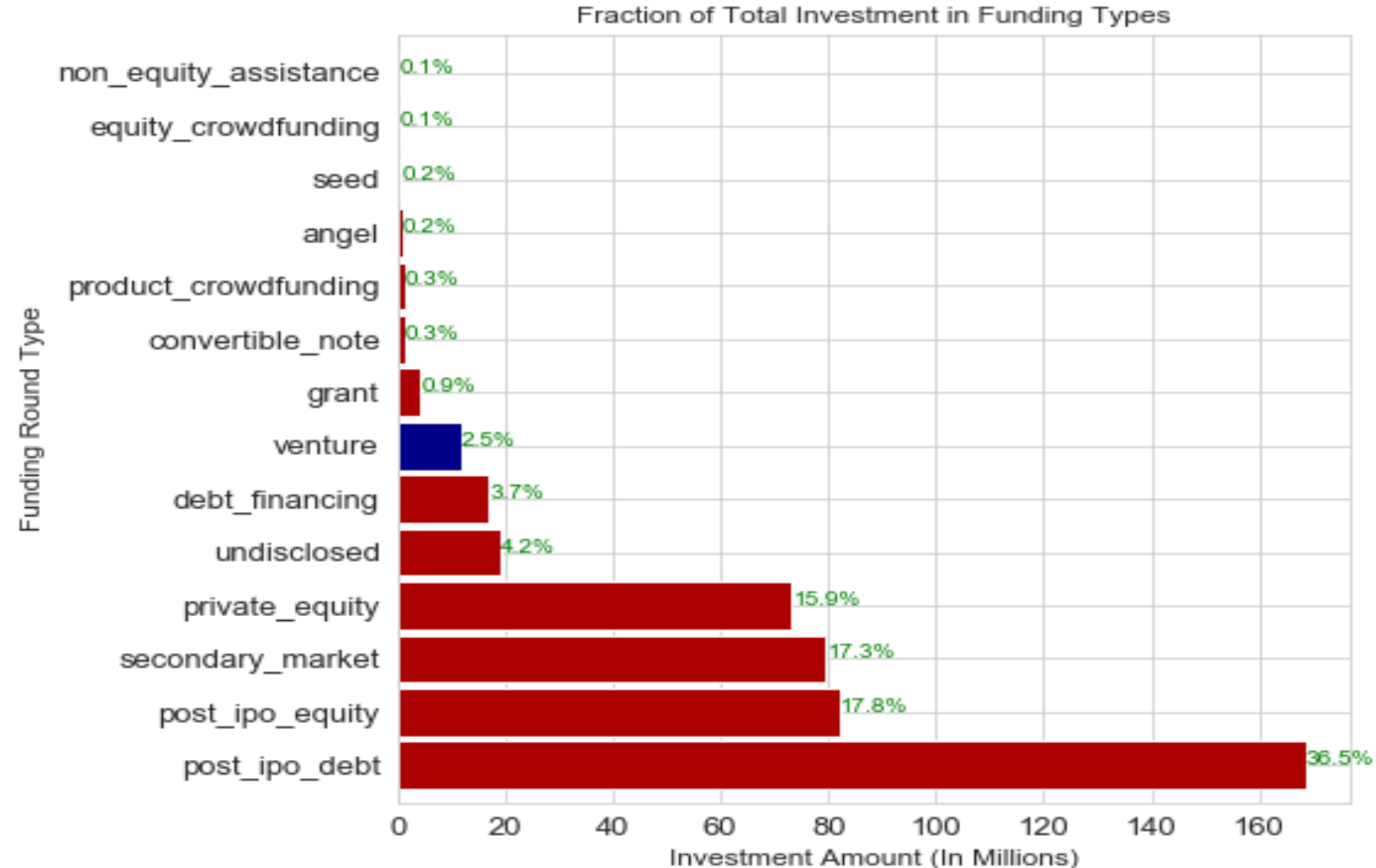
    venture

    angel

    seed

    private equity

- "master_frame dataframe data was grouped by "funding_round_type" field.

- Mean of each funding type was calculated ignoring the null values.

- "Venture" was identified as the most suitable investment type.

FUNDING TYPE ANALYSIS - RESULT

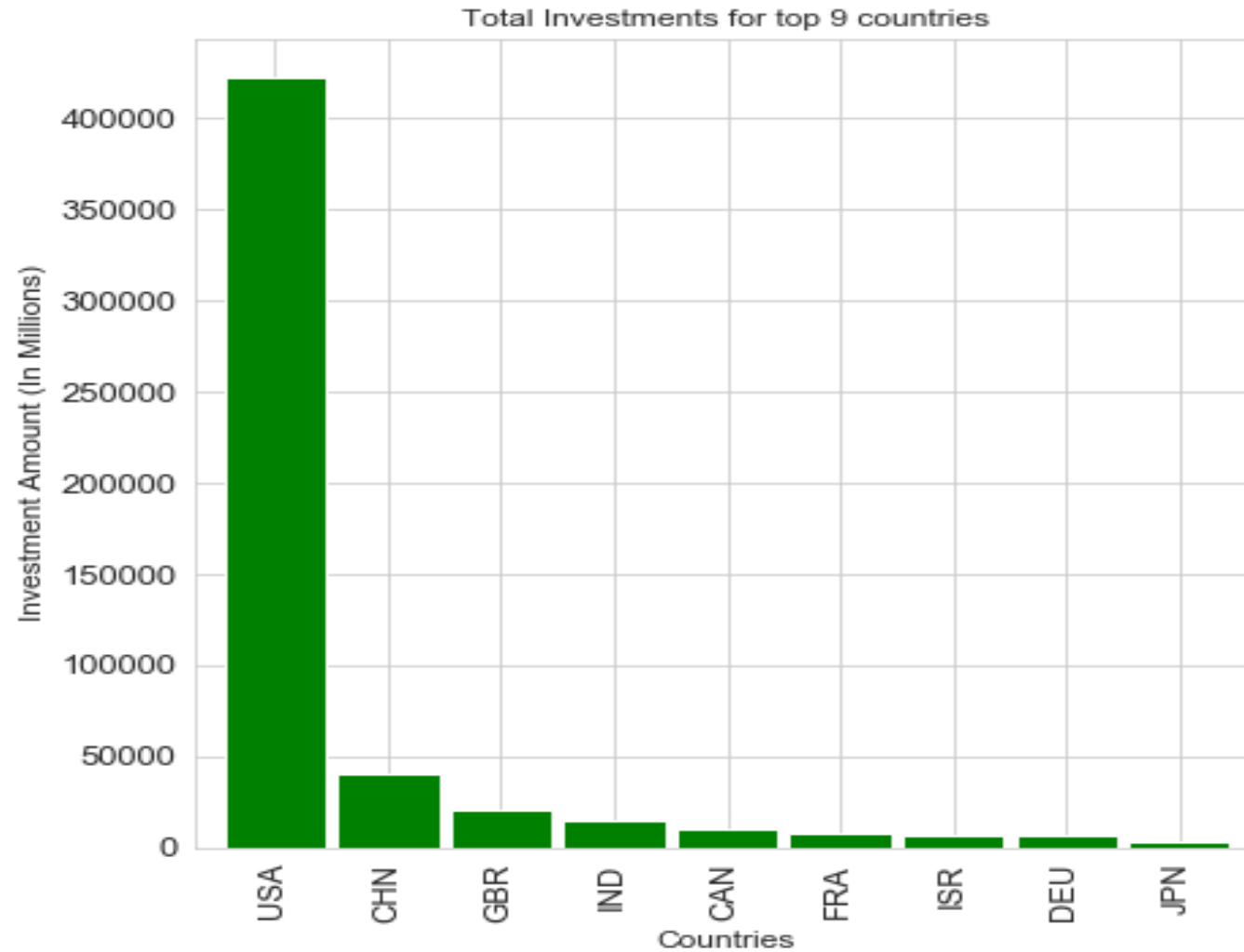Fraction of Total Investment in Funding Types

COUNTRY ANALYSIS

- **The analysis was done to identify the top three English-speaking countries which have received the highest total funding (across ALL sectors for the "venture" investment type).**

- Filter was applied on "master_frame" dataframe to extract required data as following:-
    - ✓ country code is not null
    - ✓ Raised amount USD (Investment amount) is not null
    - ✓ Funding type is venture
    - ✓ Columns "funding_round_type","raised_amount_usd" and "country_code" were extracted for analysis.

- "top9" dataframe was created with sum of funding amount grouped by Country code.

- From the top9 countries listed in the dataframe, top three English speaking countries were selected.

1	Top English speaking country	USA	United States of America
2	Second English speaking country	GBR	United Kingdom of Great Britain and Northern Ireland
3	Third English speaking country	IND	India

Total Investments for top 9 countries

SECTOR ANALYSIS

- **The analysis was done to identify the sectors in which Spark Funds would be recommended to invest in.**

- "primary sector" is extracted using business rule as first string before the vertical bar in "category_list" column and was stored in "master_frame.

- Sector classification file mapping.csv was read using "palmos" encoding and was loaded into dataframe named "df_mapping".

- "category_list" column in df_mapping is found with spelling mistakes. Data was corrected to replace 0 with substring text as 'na'

- Each "primary_sector" value in master_frame was mapped to "main sector" field value in "df_mapping" dataframe. Extracted new data was added in dataframe "final_list".

- Three dataframes D1, D2 and D3 were extracted from "final_list" for three top English speaking countries with funding type as "venture" only and investment amount in range of 5 to 15 millions (both inclusive)

- Sum and count of investment grouped by eight main sectors was extracted and added as separate columns in D1, D2 and D3.
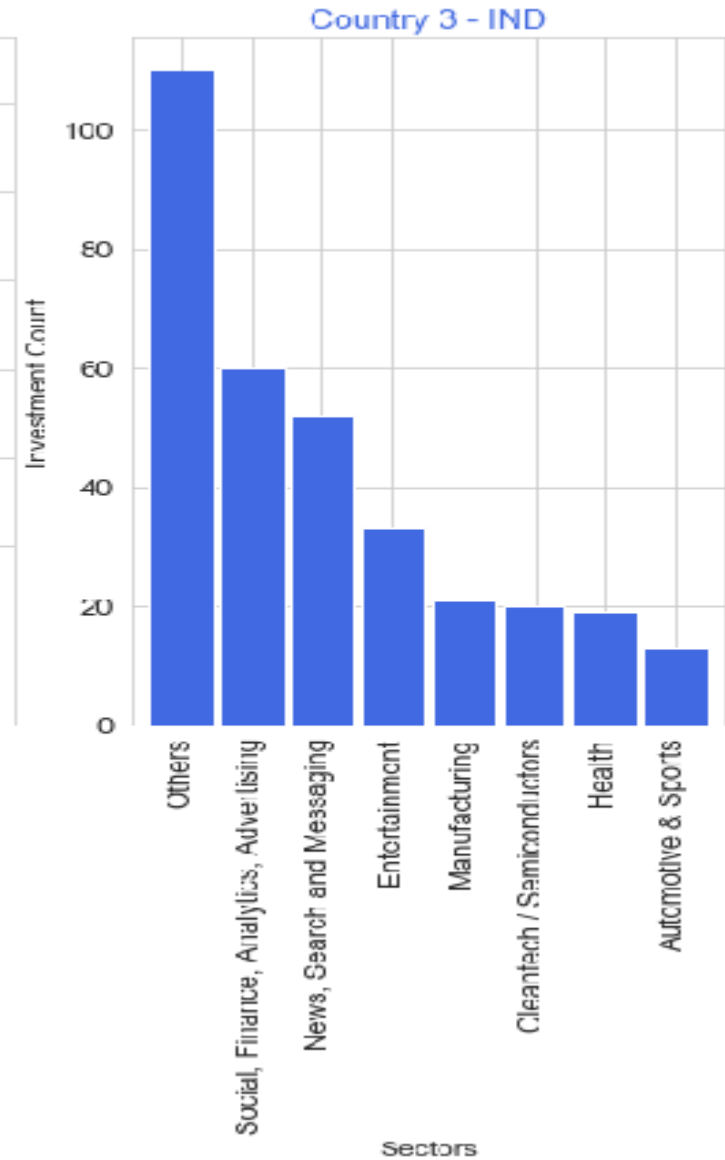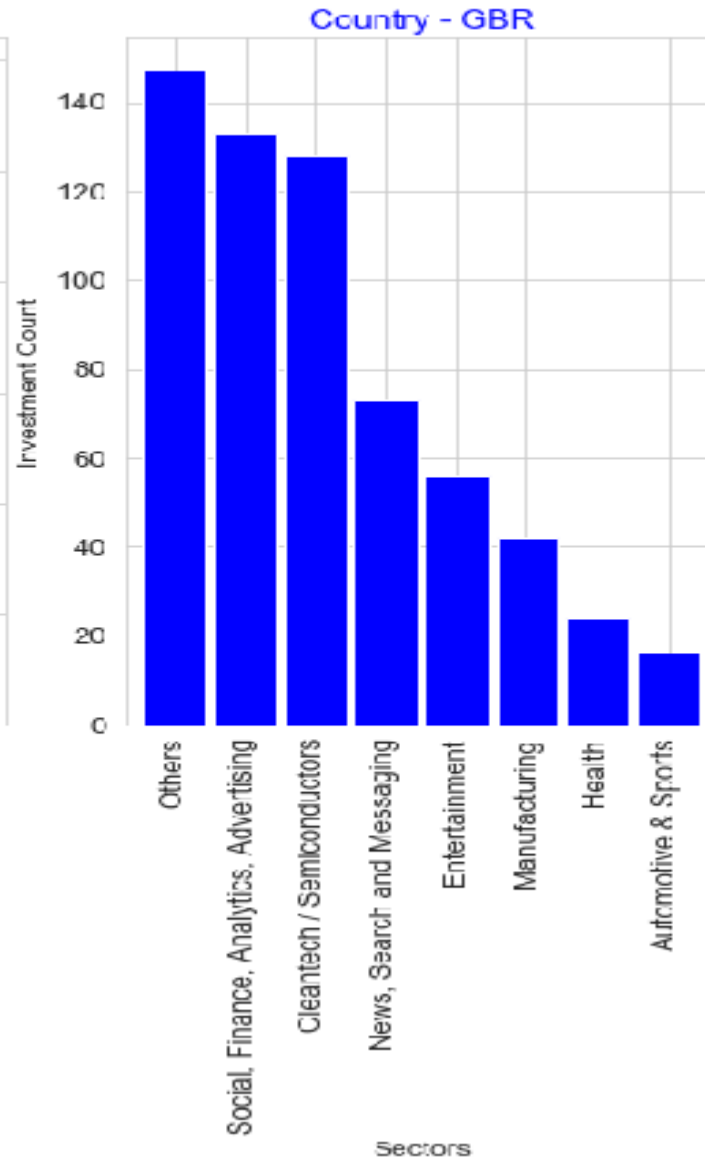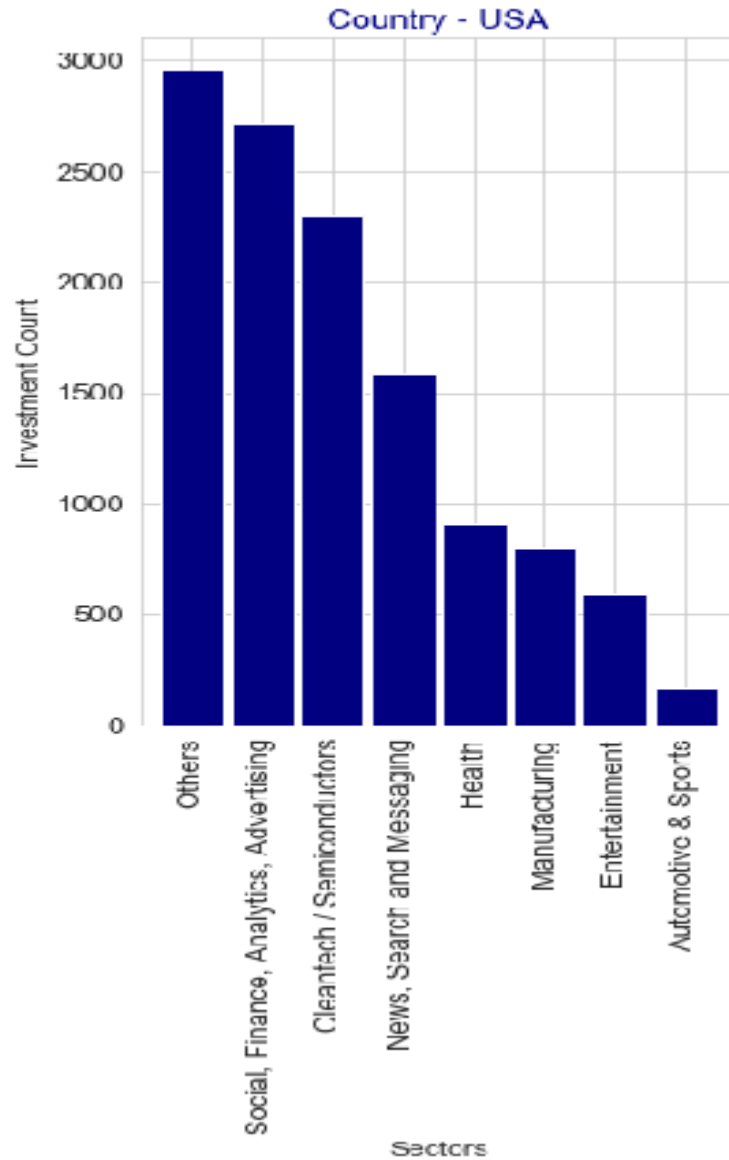
SECTOR WISE INVESTMENT ANALYSIS- RESULTS

- **The detailed analysis helped to list the observations as below -**

Sl.no	Questions	Country 1 - USA	Country Two	C3
	Sector-wise Investment Analysis			
1	Total number of Investments (count)	12150	628	330
2	Total amount of investment (USD)	108531347515.00	5436843539.00	2976543602.00
3	Top Sector name (no. of investment-wise)	Others	Others	Others
4	Second Sector name (no. of investment-wise)	Social, Finance, Analytics, Advertising	Social, Finance, Analytics, Advertising	Social, Finance, Analytics, Advertising
5	Third Sector name (no. of investment-wise)	Cleantech / Semiconductors	Cleantech / Semiconductors	News, Search and Messaging
6	Number of investments in top sector (3)	2950	147	110
7	Number of investments in second sector (4)	2714	133	60
8	Number of investments in third sector (5)	2300	128	52
9	For point 3 (top sector count-wise), which company received the highest investment?	Virtustream	Electric Cloud	FirstCry.com
10	For point 4 (second best sector count-wise), which company received the highest investment?	SST Inc. (Formerly ShotSpotter)	Celltick Technologies	Manthan Systems

ANALYSIS SUMMARY

Funding Type = Venture