INVESTMENT CASE STUDY

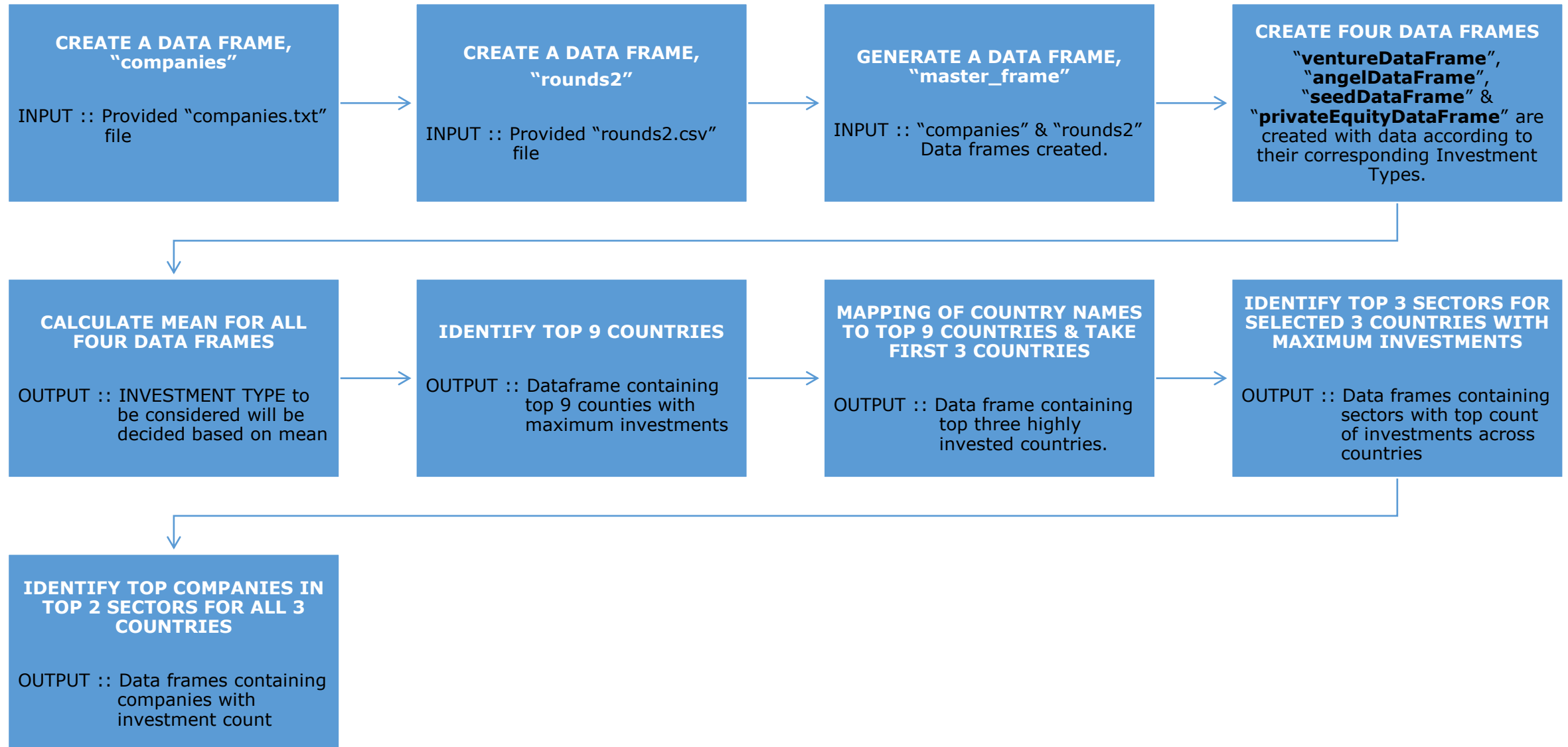# SUBMISSION

Group Name:

1.   Subhanshu Rathi
2.   Pavan ML
3.   Ravi Shekhar Rai
4.   Srikanth K

Abstract

Case study is to provide insight about investment strategy to Spark Funds, an asset management company by taking various factors into consideration which are as mentioned below.

1. Investments are to be done in to an "Investment Type" whose average investments are in between 5 million USD – 15 million USD

2. Investments are to be done only in english speaking countries

3. Investments are to be done in a sector where maximum count of investments are done

Problem Solving Methodology – Flow Chart

CREATE A DATA FRAME, "companies"

INPUT :: Provided "companies.txt" file

CREATE A DATA FRAME, "rounds2"

INPUT :: Provided "rounds2.csv" file

GENERATE A DATA FRAME, "master_frame"

INPUT :: "companies" & "rounds2" Data frames created.

CREATE FOUR DATA FRAMES
"**ventureDataFrame**", "**angelDataFrame**", "**seedDataFrame**" & "**privateEquityDataFrame**" are created with data according to their corresponding Investment Types.

CALCULATE MEAN FOR ALL FOUR DATA FRAMES

OUTPUT :: INVESTMENT TYPE to be considered will be decided based on mean

IDENTIFY TOP 9 COUNTRIES

OUTPUT :: Dataframe containing top 9 counties with maximum investments

MAPPING OF COUNTRY NAMES TO TOP 9 COUNTRIES & TAKE FIRST 3 COUNTRIES

OUTPUT :: Data frame containing top three highly invested countries.

IDENTIFY TOP 3 SECTORS FOR SELECTED 3 COUNTRIES WITH MAXIMUM INVESTMENTS

OUTPUT :: Data frames containing sectors with top count of investments across countries

IDENTIFY TOP COMPANIES IN TOP 2 SECTORS FOR ALL 3 COUNTRIES

OUTPUT :: Data frames containing companies with investment count

Data Cleaning

S.No	Decision	Explanation
1	Convert "company_permalink" column values of "rounds2" data frame to lower case	Repetition of entities exist in the table. Same entities differ with each other because of case sensitivity.
2	Convert "permalink" column values of "companies" data frame to lower case	For comparing the difference between dataframes, "companies" & "rounds2" we have converted "permalink" entities to lower case which acts as PRIMARY key.
3	Replace "NA" values of column, "raised_amount_usd" with numeric value 0	Instead of dropping the entries where investment value is not declared, we have taken them as 0 investment
4	Replace "NA" values of column, "country_code" with "UNKNOWN" string.	Instead of removing the entries where country information is not provided, we have considered as "UNKNOWN". This will make us to know how much investments have happened in countries which are unknown to us.
5	Replace "NA" values of column, "category_list" with empty string.	Entries having category_list value as empty will be getting main_category as "Blanks". This mapping will be done using provided mapping.csv file
6	Replace "0" with "na" for entries under "category_list" column of mapping.csv file	We have identified a date entry error. "0" has been entered where "na" should be entered. Ex: "A0lytics", "Alter0tive Medicine"

Analysis – Investment Type

Investment Type Identification is first step we have considered in this case study.

❑ We have calculated the average amount of investment happened in four sectors, "Venture", "Angel", "Seed" & "Private Equity" and arrived at the below figures.

Investment Type	Average Amount Invested (USD)
Ventures	1,06,34,054
Angel	7,64,564
Seed	5,56,606
Private Equity	6,21,11,788

❑ From the above table we have taken "**Venture**" investment type as a go forward abiding to the condition that company needs to invest 5-15 million USD per round.

Analysis – Country

Second analysis factor for us in this case study is Country

❑ For Investment type, "Venture" we have calculated the investment amounts grouped by country wise and picked the top 9 countries as mentioned below

Country	Investment Amount (USD)
United States of America (USA)	422510842796
China (CHN)	39835418773
UNKNOWN	25506108414
United Kingdom (GBR)	202445627416
India (IND)	14391858718
Canada (CAN)	9583332317
France (FRA)	7259536732
Israel (ISR)	6907514579
Germany (DEU)	6346959822

❑ We have not considered CHINA as it is not a english speaking country.

❑ We have ignored UNKOWN countries as this is taken during analysis for knowing how much investment has happened in countries which we don't have information.

❑ So we have considered the top three countries as "**United States of America (USA)**", "**United Kingdom (GBR)**" and "**India (IND)**".
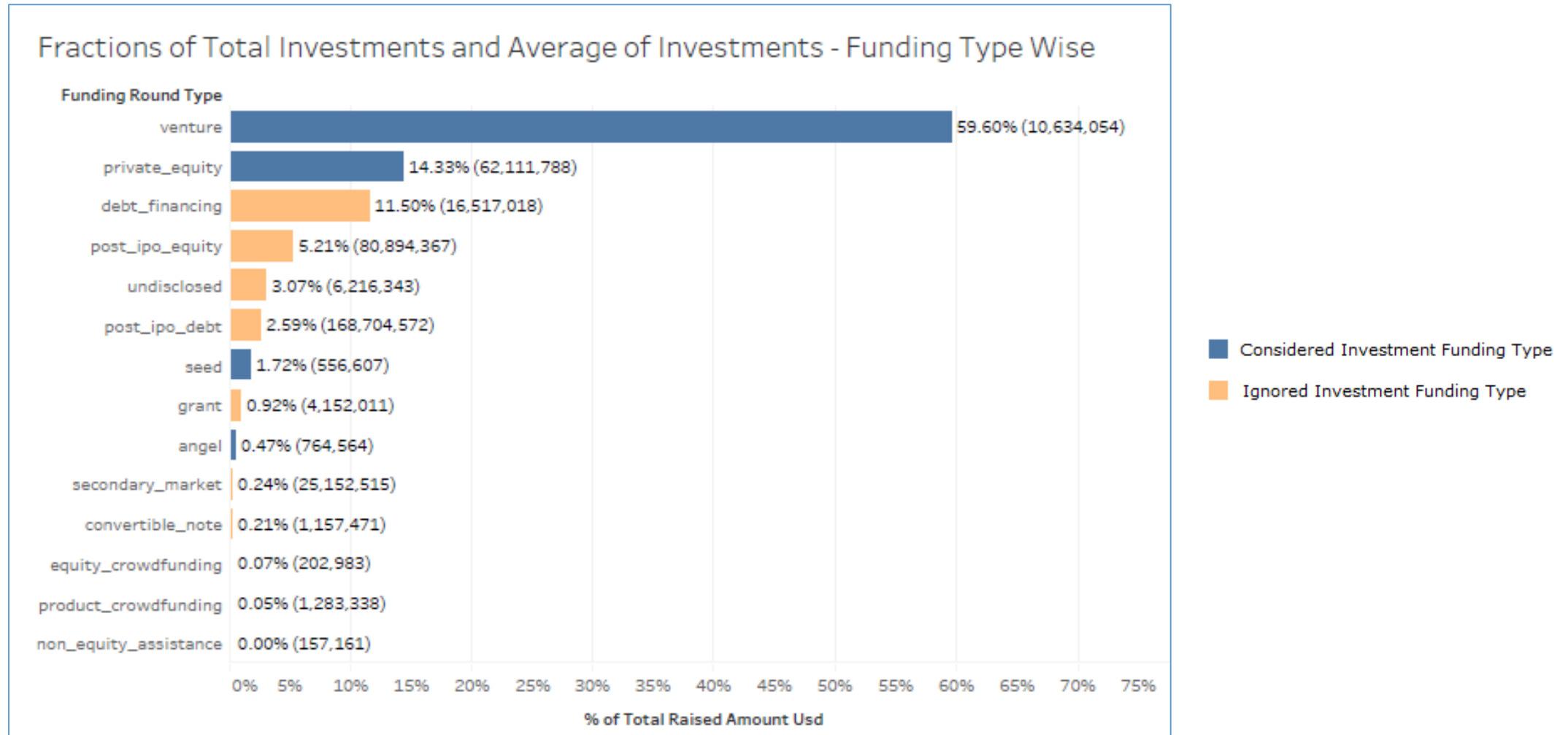
Analysis - Sector
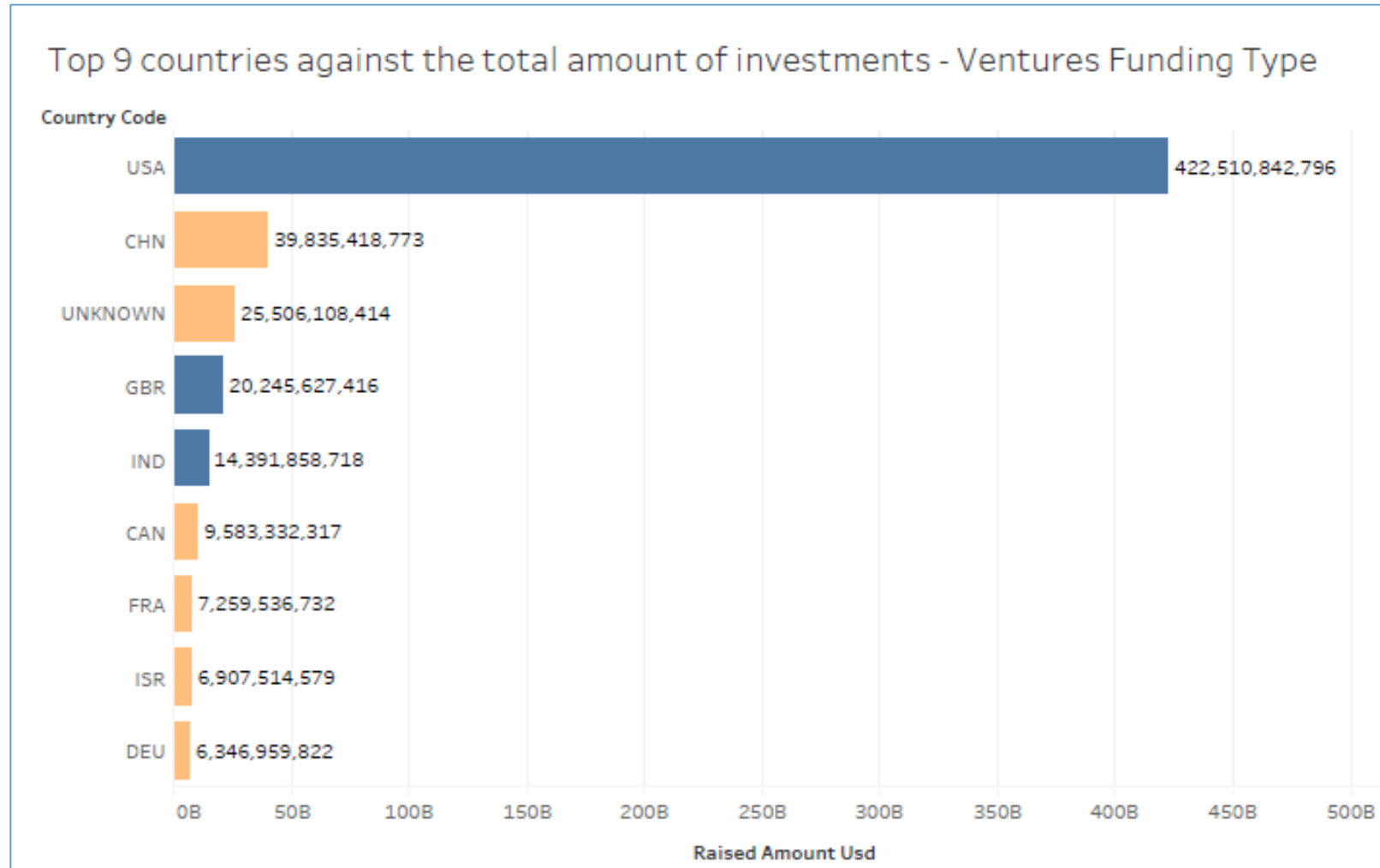
The final factor for the investment decision is "Sector"

❑ For the top 3 countries, we have identified from "Country Analysis" we will consider the top three sectors based on **COUNT** of investments

❑ Below is the observation from this analysis

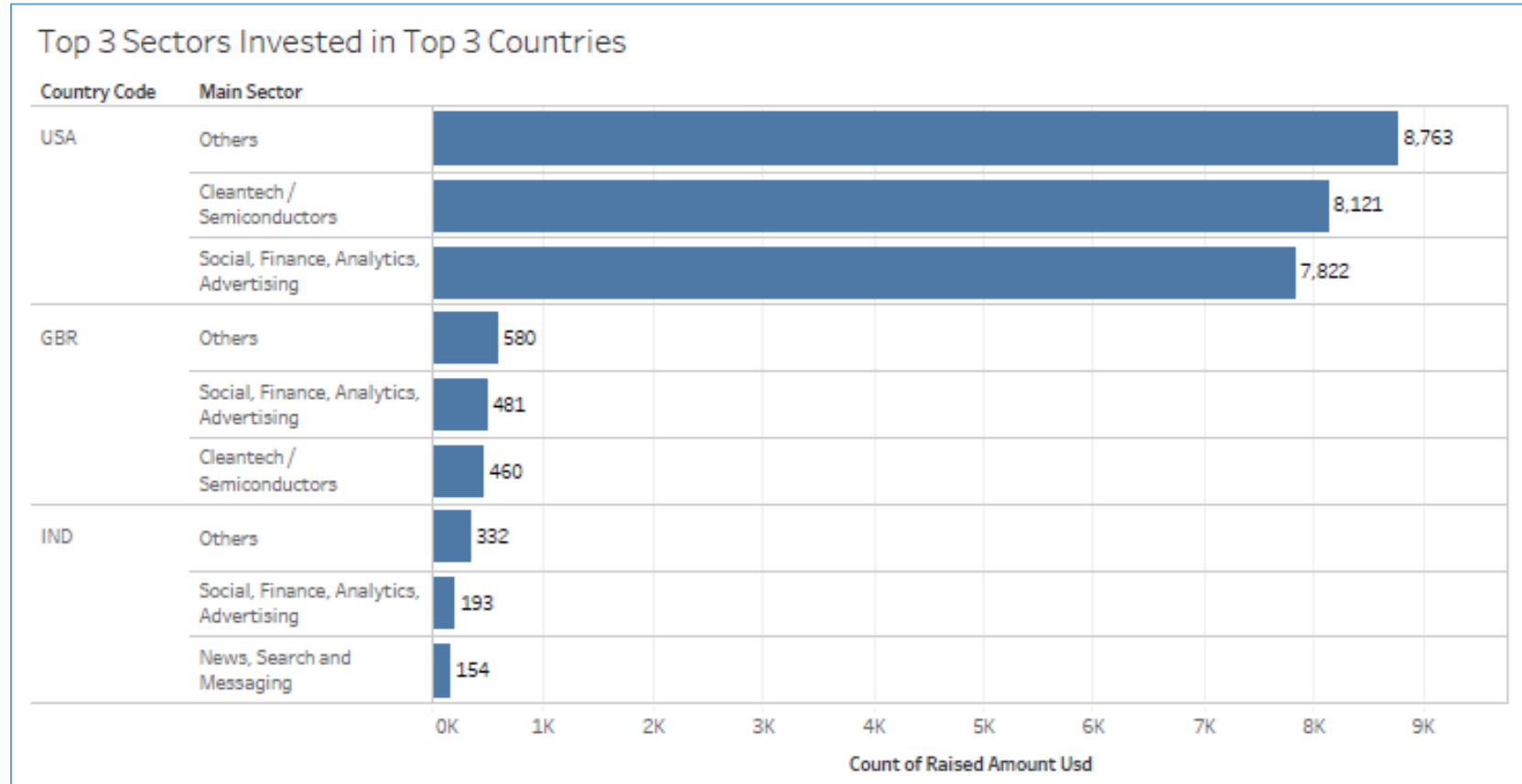| Country Name	Top First Sector	Top Second Sector	Top Third Sector
United State of America (USA)	Others (8763)	Cleantech, Semiconductors (8121)	Social, Finance, Analytics, Advertisement (7822)
United Kingdom (UK)	Others (580)	Social, Finance, Analytics, Advertisement (481)	Cleantech, Semiconductors (460)
India (IND)	Others (332)	Social, Finance, Analytics, Advertisement (193)	News, Search and Messaging (154)

Results – Plot 1

Fractions of Total Investments and Average of Investments - Funding Type Wise

Funding Round Type

Funding Round Type	% of Total Raised Amount Usd
venture	59.60% (10,634,054)
private_equity	14.33% (62,111,788)
debt_financing	11.50% (16,517,018)
post_ipo_equity	5.21% (80,894,367)
undisclosed	3.07% (6,216,343)
post_ipo_debt	2.59% (168,704,572)
seed	1.72% (556,607)
grant	0.92% (4,152,011)
angel	0.47% (764,564)
secondary_market	0.24% (25,152,515)
convertible_note	0.21% (1,157,471)
equity_crowdfunding	0.07% (202,983)
product_crowdfunding	0.05% (1,283,338)
non_equity_assistance	0.00% (157,161)

Legend:
- Considered Investment Funding Type
- Ignored Investment Funding Type

% of Total Raised Amount Usd

Results – Plot 2

Top 9 countries against the total amount of investments - Ventures Funding Type

Country Code

Country	Raised Amount Usd
USA	422,510,842,796
CHN	39,835,418,773
UNKNOWN	25,506,108,414
GBR	20,245,627,416
IND	14,391,858,718
CAN	9,583,332,317
FRA	7,259,536,732
ISR	6,907,514,579
DEU	6,346,959,822

Legend: Considered Countries / Ignored Countries

Results – Plot 3

Top 3 Sectors Invested in Top 3 Countries

Conclusions

Below are the conclusion drawn out of this Case study

❑ "VENTURE" is the identified investment funding type

❑ United States of America (USA), United Kingdom (GBR) and India (IND) are the countries identified as top three English speaking countries

❑ Below sectors are the ones chosen to go with the investments country wise

USA   "Others", "Cleantech / Semiconductors" and "Social, Finance, Analytics, Advertisement"

GBR   "Others", "Social, Finance, Analytics, Advertisement" and "Cleantech / Semiconductors"

IND   "Others", "Social, Finance, Analytics, Advertisement" and "News, Search and Messaging"

❑ Below companies are identified as top companies (with highest count of investments) in top two best sectors country wise

USA   "Others"   Soft   "Cleantech / Semiconductors"   Free Scale Semi Conductors

GBR   "Others"   OneWeb   "Social, Finance, Analytics, Advertisement" Powa Technologies

IND   "Others"   Flipkart   "Social, Finance, Analytics, Advertisement" Shopclues.com