

INVESTMENT CASE STUDY

SUBMISSION

Group Name: ***Data Cubed***

1. *Ashutosh Kumar (DDA1730347)*
2. *Medhavi Shruti (DDA1730035)*
3. *Santanu Dey (DDA1730043)*

Abstract of Spark Funds Investments Case Study

Introduction

- Spark Funds, an asset management company, wants to make investments, as an early stage startup investor, 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.

Goals

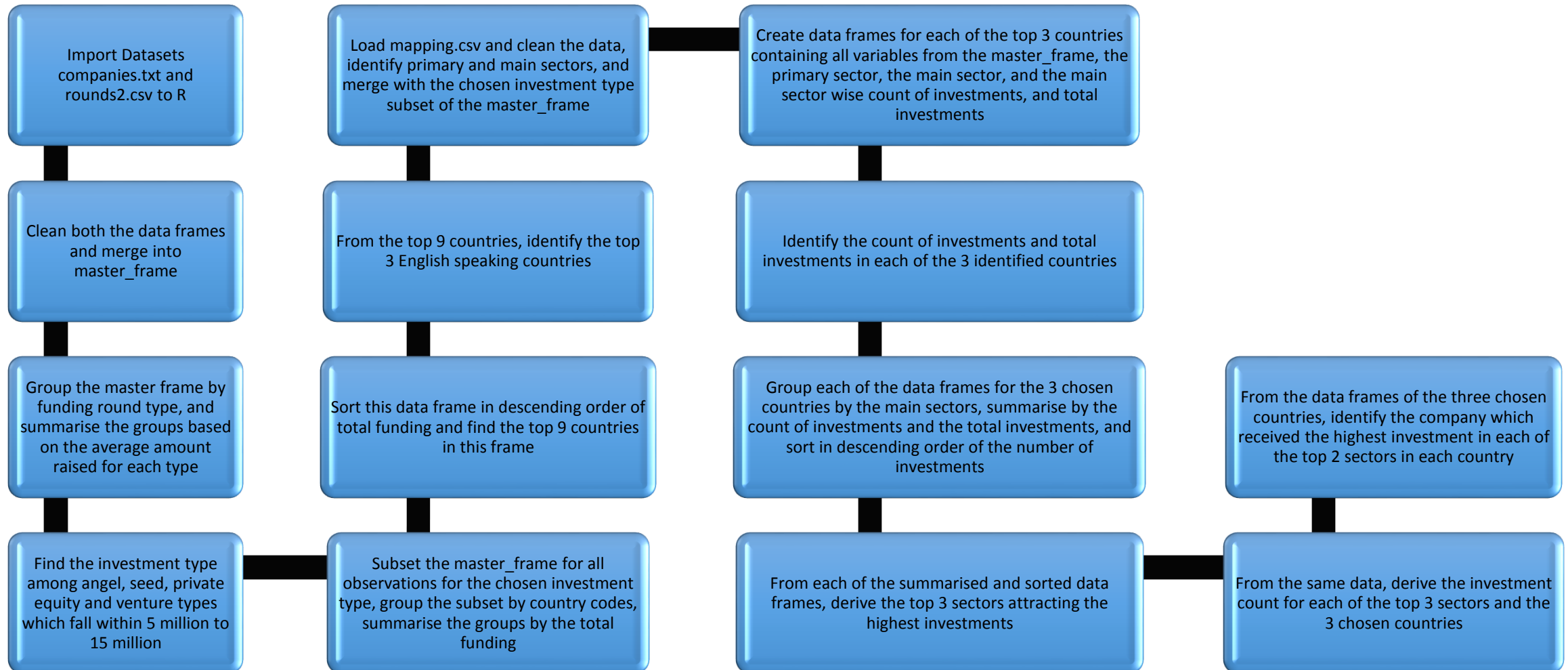
- Find the **investment type** best suited for the Spark Funds' strategy.
- Find the top **countries** which attract the most investments.
- Find the top **sectors** which attract the most investments.

Constraints

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.

Problem Solving Methodology



Analysis – Data Cleaning

- Load the datasets *companies.txt* and *rounds2.csv* into *companies* and *rounds2* data frames. While loading, convert all blank entries into NA.
- Convert the *company_permalink* column in *rounds2*, and the *permalink* column in *companies* to upper case.
- Find if there are any companies in *rounds2* which are not present in *companies* data frame, by comparing the count of unique values in each of the permalink columns.
- Convert the *company_permalink* column in *rounds2* to *permalink*.
- Merge *rounds2* and *companies* data frames by the *permalink* column including all NA values.
- Convert all NA values in the *raised_amount_usd* in the *master_frame* to zeroes.

Analysis – Investment Type Analysis

- Group the *master_frame* by the *funding_round_type* column.
- Summarise the groups based on the average amount raised per investment round.
- Find the average funding amount for *angel*, *seed*, *private_equity*, and *venture* funding types.
- Find the investment type among *angel*, *seed*, *private_equity*, and *venture* funding types which fall within the range 5-15 million USD.
- We find that *Venture* funding type is the best investment type for Spark Funds' investment goals.

Analysis – Country Analysis

- Subset the *master_frame* for all observations for the *Venture* investment type into the *df_venture* data frame.
- Group the subset by *country_code*.
- Summarise the groups by the total amount of investments.
- Discard the NA row in the summarized data frame.
- Sort this data frame in descending order of total amount of investments.
- Find the top 9 countries in this data frame.
- From the top 9 countries, identify the top 3 English speaking countries.
- We find that *USA*, *Great Britain* and *India* are the top 3 English speaking countries attracting the highest amount of investments.

Analysis – Sector Analysis

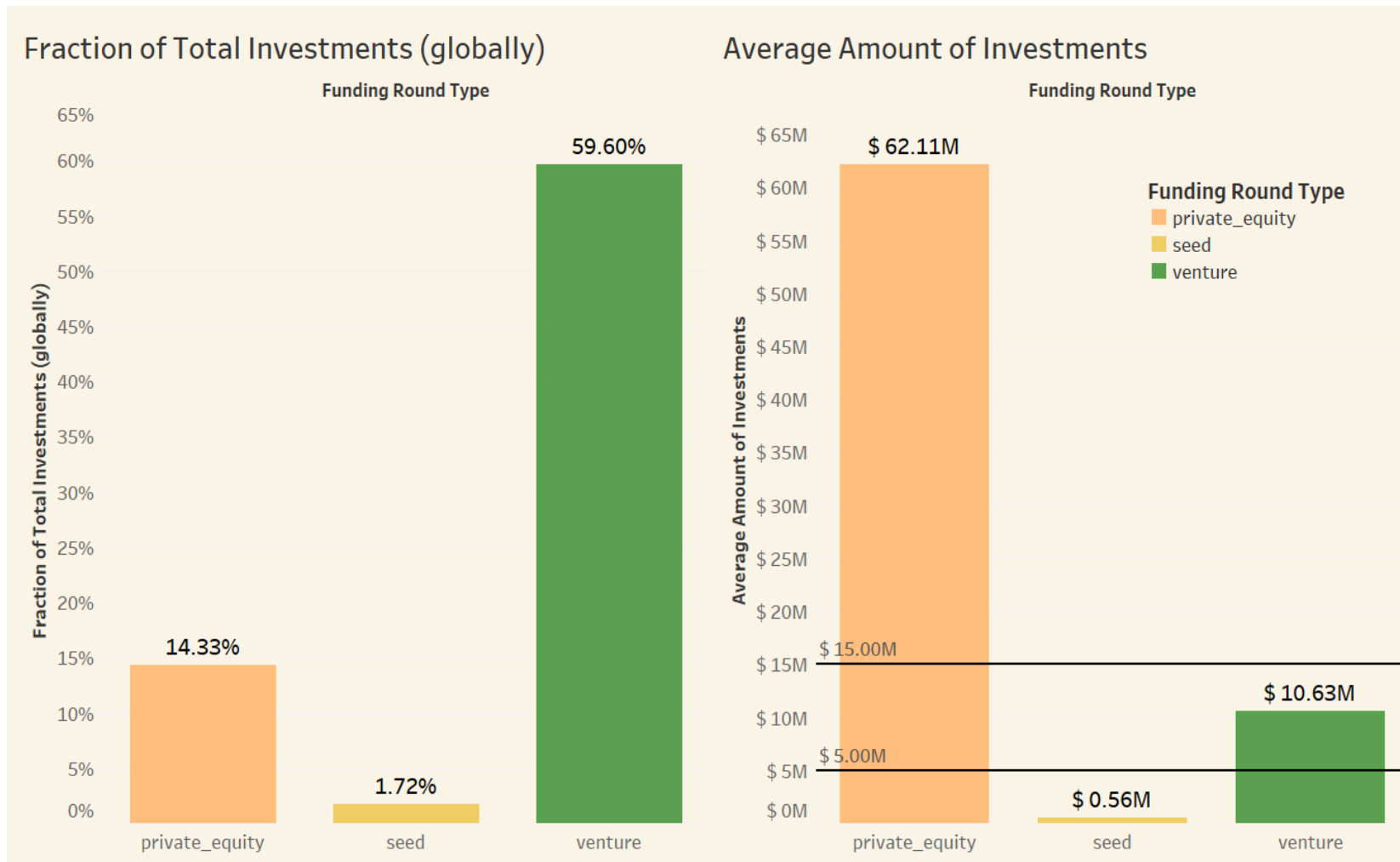
Part 1:

- In *df_venture*, for *category_list* values separated by a pipe (vertical bar |), the first string before the pipe is considered to be the *primary_sector*. A new column named *primary_sector* is added to *df_venture* containing the converted *category_list* values.
- Load *mapping.csv* and clean the data: During analysis of the mapping data, it was found that it contains dirty data, which was cleaned to have proper names for the *category_list*. Remove the *Blanks* column, and convert wide data into long data, by mapping each primary sector to one of the eight main sectors. Rename the *category_list* column to *primary_sector*. We now merge the cleaned *mapping* data with *df_venture* by *primary_sector*.
- We now have *df_venture_sectors*, a merged data frame having all the columns of *master_frame* mapped to its corresponding primary and main sectors.

Part 2:

- Get the *subset* of *df_venture_sectors*, for *raised_amount_usd* values in the range 5-15 million USD. From this subset, create 3 further subsets for each of the top 3 countries *USA*, *Great Britain*, and *India*, called D1, D2, and D3, respectively. Add the main sector wise number of investments and total investments to each data frame.
- We now have data frames for each of the top 3 countries containing all variables from the *master_frame*, the *primary_sector*, the *main_sector*, and the main sector wise number of investments and the total investments.
- Identify the count of investments and total investments in each of the 3 identified countries
- Group each of the data frames for the 3 chosen countries by the *main_sector*, summarise by the number of investments and the total investments, and sort in descending order of the number of investments.
- From each of the summarised and sorted data frames, derive the top 3 sectors attracting the highest investments
- From the same data, derive the investment count for each of the top 3 sectors and the 3 chosen countries
- From the data frames of the three chosen countries, identify the company which received the highest investment in each of the top 2 sectors in each country.

Results – Plot 1

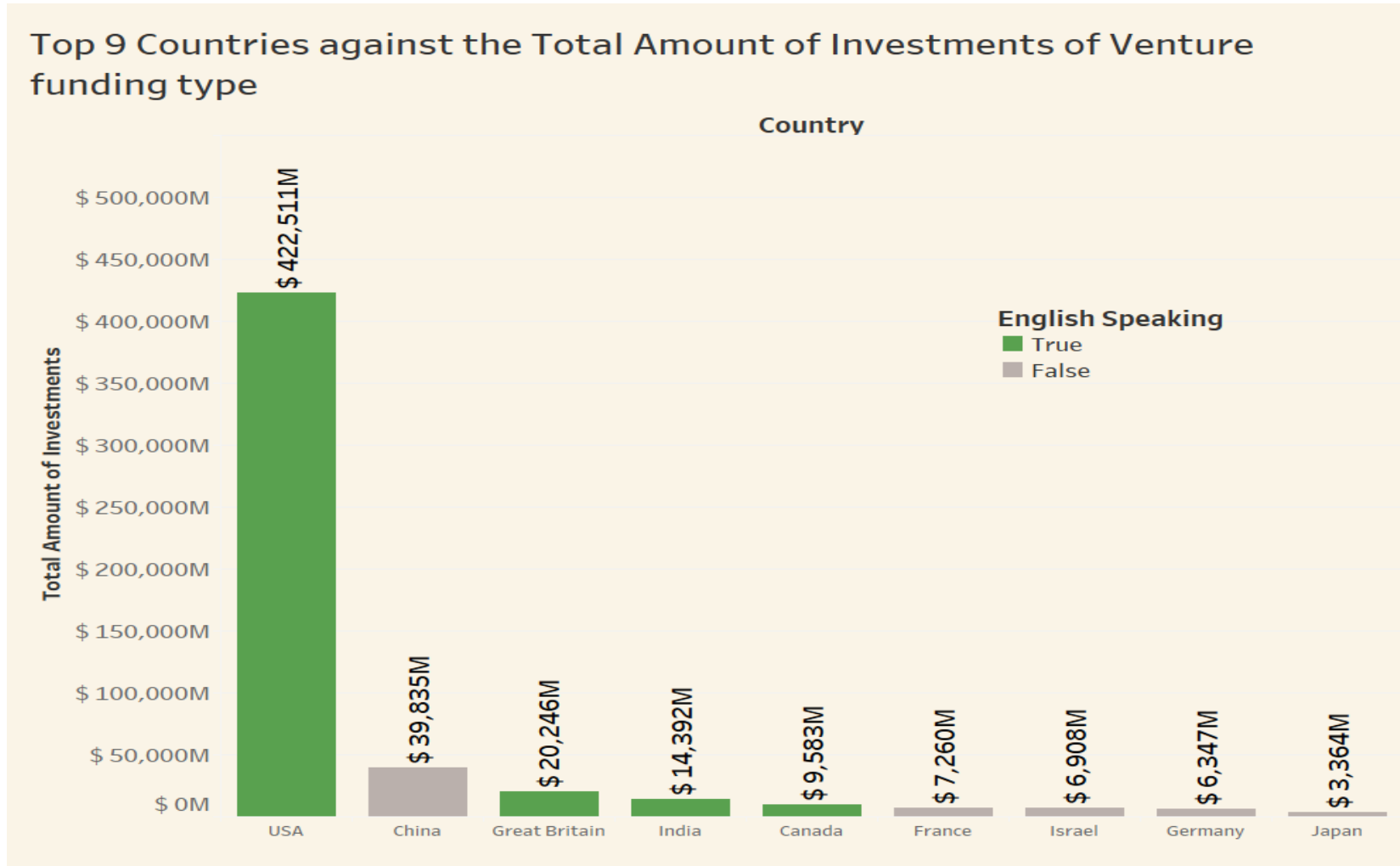


Funding Type Analysis

Among *private_equity*, *seed*, and *venture* funding types:

- In terms of *Fraction of Total Investments (globally)*, ***Venture*** has the highest fraction (percentage).
- In terms of *Average Amount of Investments* per funding round within the range 5-15 million USD, *Venture* is the only investment type.
- **So, *Venture* is the best investment type for Spark Funds' investment goals.**

Results – Plot 2

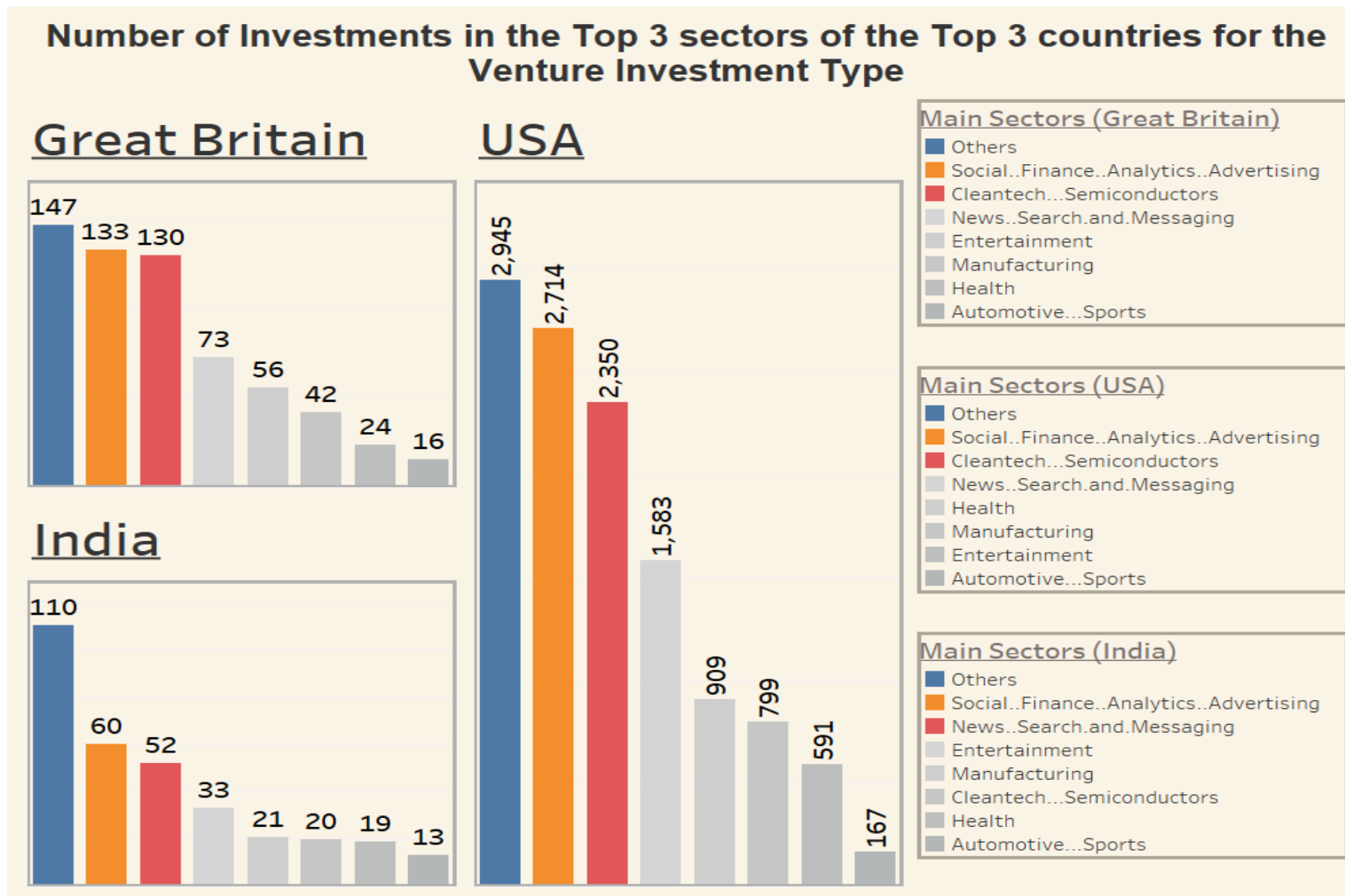


Country Analysis

Among the top 9 countries attracting the total amount of *Venture* investments:

- *USA* is the top English speaking country.
- *China* can be discarded as it is **not** an English speaking country.
- *Great Britain* (GBR) and *India* (IND) are the second and third top English speaking countries respectively.
- **So, the top 3 English speaking countries suitable for Spark Funds' investment goals are USA, Great Britain and India.**

Results – Plot 3



Sector Analysis

The top 3 sectors attracting the highest number of **Venture** investments in **USA** are:

1. **Others**
2. **Social..Finance..Analytics..Advertising**
3. **Cleantech...Semiconductors**

The top 3 sectors attracting the highest number of **Venture** investments in **Great Britain** are:

1. **Others**
2. **Social..Finance..Analytics..Advertising**
3. **Cleantech...Semiconductors**

The top 3 sectors attracting the highest number of **Venture** investments in **India** are:

1. **Others**
2. **Social..Finance..Analytics..Advertising**
3. **News..Search.and.Messaging**

Conclusions

Sl.No.	Goal	Result
1	Find the investment type best suited for the Spark Funds' strategy.	Venture
2	Find the top countries which attract the most investments.	USA, Great Britain, and India
3	Find the top sectors which attract the most investments (USA)	<ol style="list-style-type: none"> 1. Others 2. Social..Finance..Analytics..Advertising 3. Cleantech...Semiconductors
4	Find the top sectors which attract the most investments (GBR)	<ol style="list-style-type: none"> 1. Others 2. Social..Finance..Analytics..Advertising 3. Cleantech...Semiconductors
5	Find the top sectors which attract the most investments (IND)	<ol style="list-style-type: none"> 1. Others 2. Social..Finance..Analytics..Advertising 3. News..Search.and.Messaging
6	Top 2 companies in the USA (Total investment wise)	<ol style="list-style-type: none"> 1. Virtustream 2. SST Inc. (Formerly ShotSpotter)
7	Top 2 companies in the GBR (Total investment wise)	<ol style="list-style-type: none"> 1. Electric Cloud 2. Celltick Technologies
8	Top 2 companies in the IND (Total investment wise)	<ol style="list-style-type: none"> 1. FirstCry.com 2. Manthan Systems