

# INVESTMENT CASE STUDY

## SUBMISSION

Group Name:

1. Lipsa Satapathi
2. Prabhudatta Praharaj
3. Abinash Panda
4. Ankit Gupta

# Abstract

Spark Funds, an asset management company looking forward to invest in certain country , company and segment where maximize profit .

## **Data Sets Available for Analysis**

- Company : Data Set containing company Id , name , Url , county code , state etc
- Rounds : Data Set containing funding id , Funding amount , Funding Type , date , company etc
- Mapping : Data Set containing sector classification

## **Spark Funds wants to perform the following Analysis**

- Investment type analysis: Understanding investments in venture, seed/angel, private equity categories, etc. To decide strategy.
- Country analysis: Understanding which countries have had the most investments in the past and could be favorable for Spark Funds
- Sector analysis: Understanding the distribution of investments across the eight main sectors per mapping file

## **Spark Funds wants to identify on 3 parameter to process with investment**

- Best English Speaking county to Invest
- Best Sector to Invest
- Best Company to Invest

# Problem Solving Methodology

## Business Understanding

- Sparks wants to find best country Segment and Company to invest

## Data Understanding

- No of unique companies = 66368
- Funding Type, Funding Amount
- Country, sectors, category etc

## Data Preparation

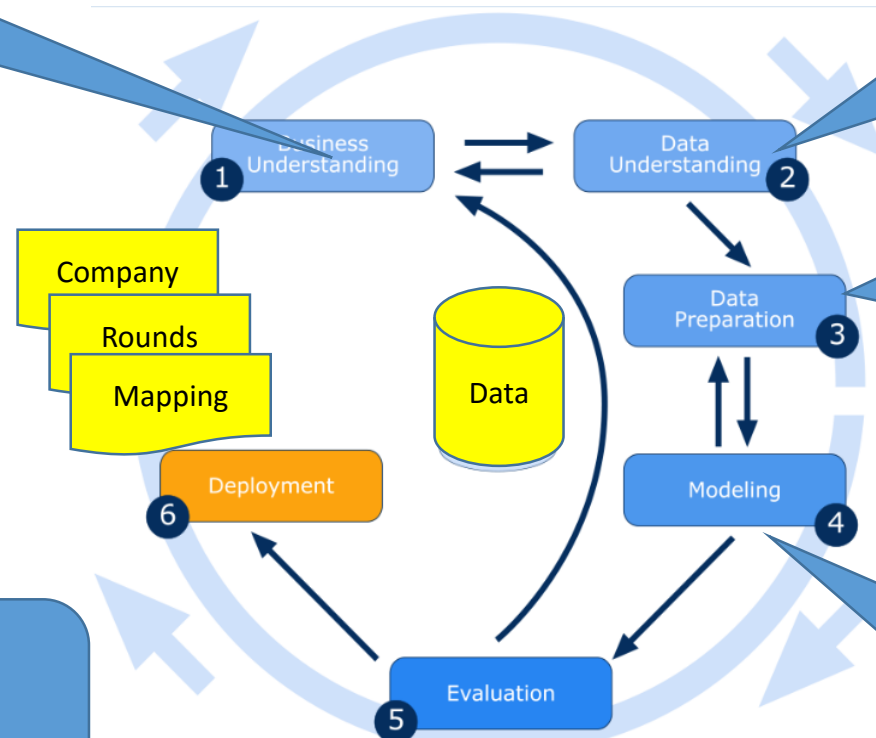
- Convert 0 to Na in mapping file for category list without affecting enterprise 2.0
- Remove Sector names as blank
- Converted permalink to Lower case in 2 data set
- Wide to long format conversion in mapping file
- Get English speaking county name from PDF

## Data Modeling / Manipulation

- Merge Company & Rounds 2
- Filter investment between 5 – 15 million
- Filter Funding type = Venture
- Filter English speaking country
- Create 3 data set D1, d2, D3 for USA, GBR, India

## Evaluation

- Using D1, D2, D3 find out best country, best sectors & best company



CRISP-DM Framework is referred to Analyze and solve this problem

# Assumptions

- We have kept the blank & NA rows and have used `rm.na = True` during calculation.
- We have removed Blanks from the main sectors because that is not one of the main sectors.
- We have considered the English speaking countries as given in the pdf provided on the case study page and nothing apart from it.

# Analysis

- We have converted perma link in both the dataset in company and rounds2 to lowercase.
- Merging both the dataset to master\_frame.
- As per the requirement we need to find out suitable funding type from 4 types between 5 to 15 million range
  - venture
  - Seed,
  - Angel
  - private equity
- We created a subset with it and find the average investment amount(raised\_amount\_usd) for each of 4 funding type
- We have taken the maximun to find out the suitable funding type
- We have concluded **venture** is the suitable funding type

## Analysis ..

- Top 9 countries as per investment we have found out for suitable funding type venture and investment amount 5 to 15 million USD.
- We have manually created the vector for English speaking country referring to the PDF .So filtering those from top 9 ,We got ur top 3 countries where Maximum investment happened.
- Top3 countries are found top be **USA , GBR , IND.**
- After that loaded the mapping file. Remove all 0 to na in category list and making sure it will not affect other values. As we need to consider only 8 sector so removed 'blank' as a sector.
- Converted wide to long format.
- By using '|' separator from the master frame we have got the primary\_category and then we have merged with the mapping file on main\_category.So each primary\_category is mapped to main\_category.

## Analysis ..

We have created D1,D2,D3 for each each of the three countries containing the observations of funding type FT falling within the 5-15 million USD

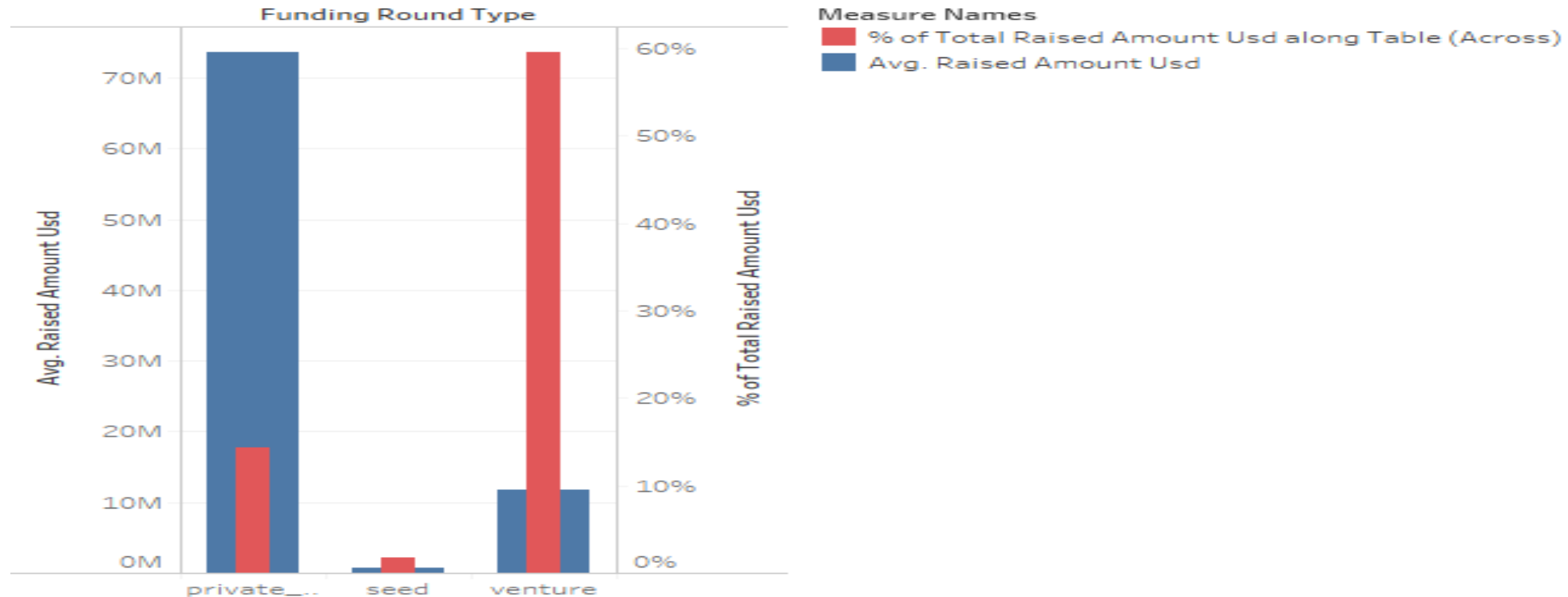
It also contains

- 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

From this we have calculated top 3 sectors for each of the countries and no of investment for each one and company received highest investment

# Results - Part 1

Checkpoints 6: Plot 1

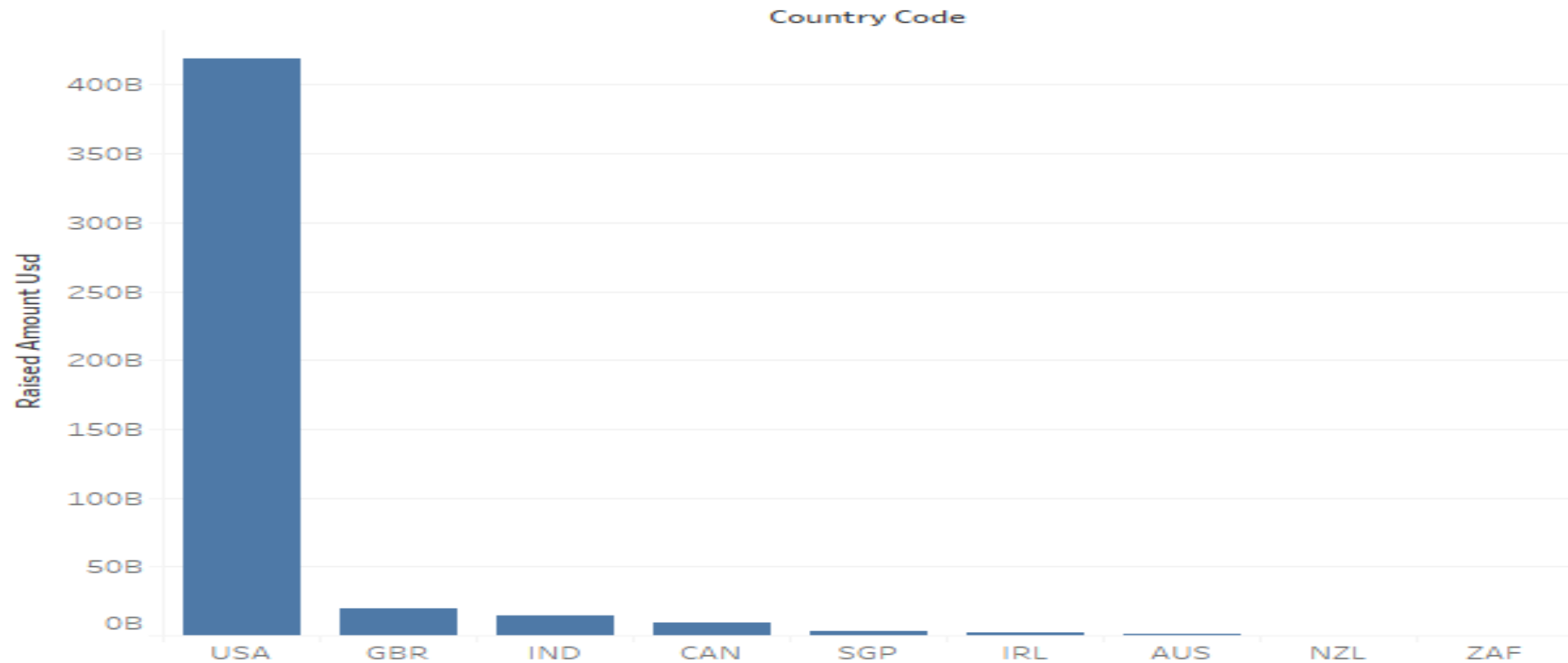


Avg. Raised Amount Usd and % of Total Raised Amount Usd along Table (Across) for each Funding Round Type. Color shows details about Avg. Raised Amount Usd and % of Total Raised Amount Usd along Table (Across). The data is filtered on LOOKUP(ATTR([Funding Round Type]),0), which keeps private\_equity, seed and venture.



# Results Part 2

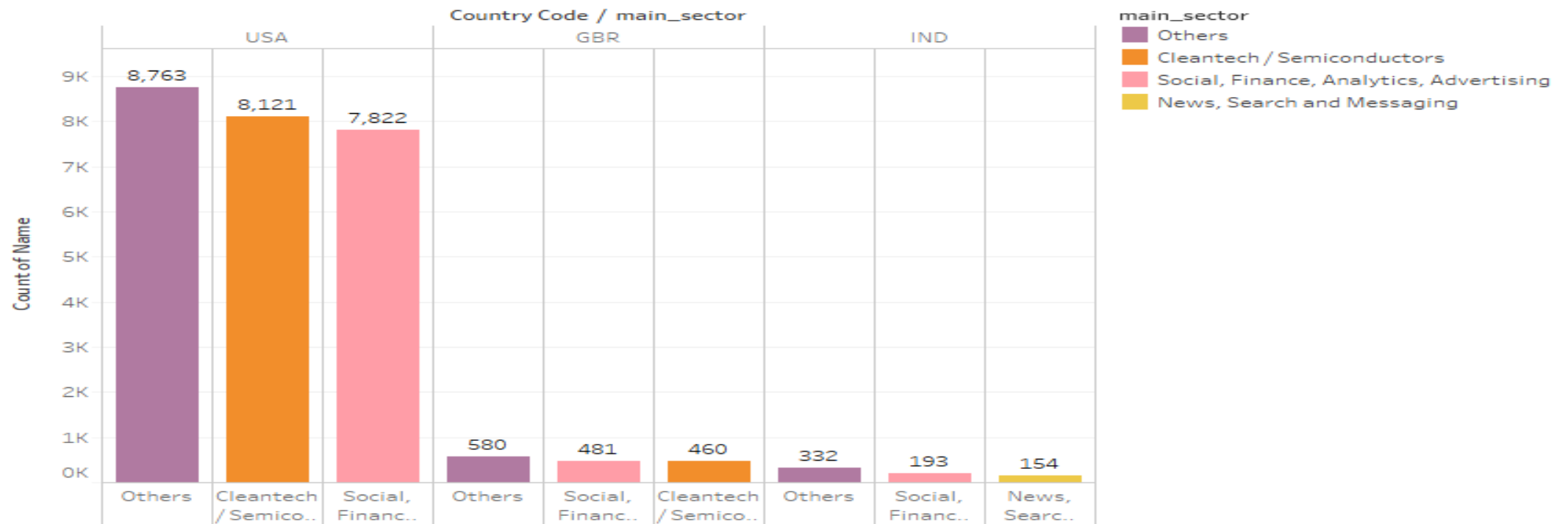
Checkpoint 6: Plot 2



Sum of Raised Amount Usd for each Country Code. The data is filtered on Funding Round Type, which keeps venture. The view is filtered on Country Code, which has multiple members selected.

# Results Part3

## Checkpoint 6: Plot 3



Count of Name for each main\_sector broken down by Country Code and Country Code & main\_sector (Combined). Color shows details about main\_sector. The marks are labeled by count of Name. The data is filtered on Funding Round Type, which keeps venture. The view is filtered on Country Code and Country Code & main\_sector (Combined). The Country Code filter has multiple members selected. The Country Code & main\_sector (Combined) filter keeps 9 members.

## Conclusions

1. From plots the countries USA, Britain and India looks quite promising for investment in venture category.
2. We can also see that the sectors **Others, Cleantech...Semiconductors and Social..Finance..Analytics..Advertising** are the ones getting more investments. So these are recommended sectors to be invested in.
3. Among the companies SoFi, FreeScale Semiconductor, OneWeb, Powa Technologies, Flipkart and ShopClues are few that Sparks Fund should invest on.