INVESTMENT CASE STUDY

SUBMISSION

Group Name:

1. Ankur Mathur
2. Apurv Jain
3. Udit Sethi

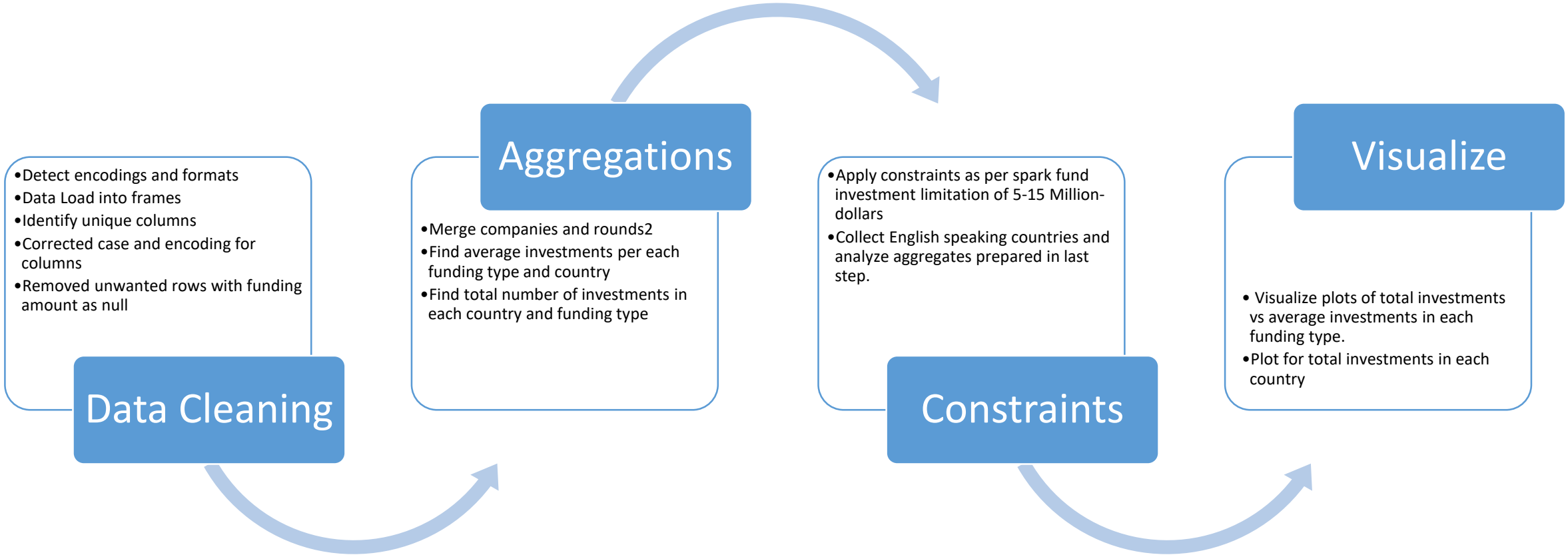
Spark Funds Case Study

Spark Funds wants to make investments in a few companies. In this presentation we are showcasing the various data insights that we were able to extract out of previous year data in order to analyse and understand the global trends in investments, so that Spark Funds can take the investment decisions effectively.

Spark Funds is able to raise between 5,000,000 $ to 15,000,000 $ for investments, and is interested in funding companies in English speaking countries only. So, as per the constraints we have identified and tried to gain insights on:

- The best funding type for funding the companies.
- The 3 best companies to fund
- The best countries to fund
- Main sectors which should be targeted for funding

Problem analysis



Checkpoint 1

1. Clean dataset companies and rounds2
2. See that the permalink column is unique in companies and can be used to merge these two datasets together
3. Values in permalink are not case-sensitive and must be in the same case so correct encoding, case and strip the values.
4. There are some missing values in *raised_amount_usd* column, so these rows are removed.
5. Use pd.merge to merge and store it in *master_frame* variable
6. Count the rows in the master_frame

Checkpoint 2

1. We need to find average funding amount of **angel, seed, venture and private equity**.
    1. Group the master_frame using column *funding_round_type.*
    2. Use the mean() function to calculate mean on this group
    3. Select only rows with indexes.
2. Since spark funds can only invest 5-15 million dollars so apply filter to the dataframe.
3. Only “venture” funding type lies in this range.

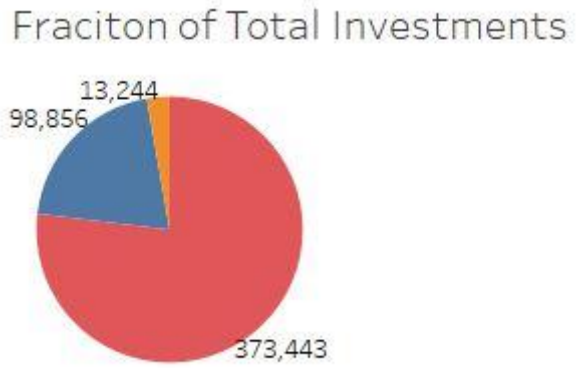
Checkpoint 3

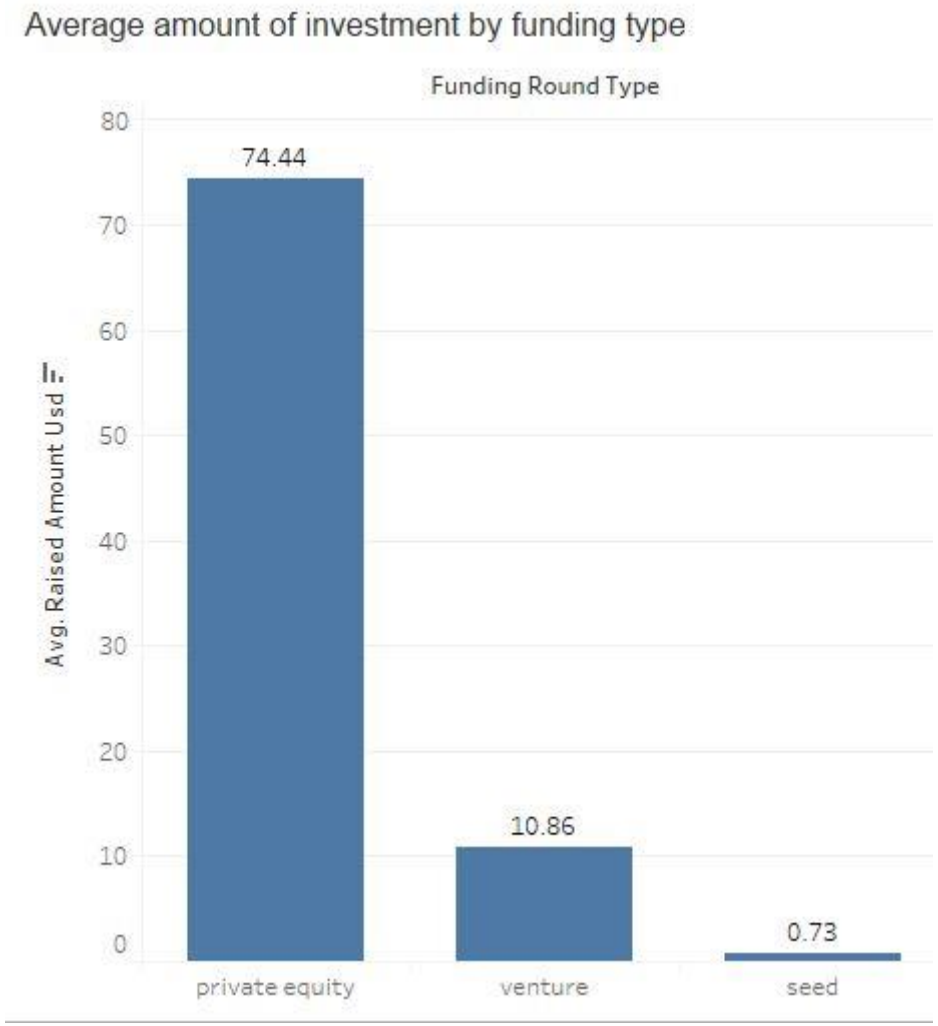
1. For country analysis we need to find English speaking countries, so we will refer to given pdf.
2. We need to filter our working set so that it only has “venture” funding type data rows.
3. We need to find top9 countries with most funding. So, We group on *country_code* and do sum() on the group.
4. We then sort the rows on ‘*raised_amount_usd*’. Take top 9 rows from this sorted dataframe.

Checkpoint 4

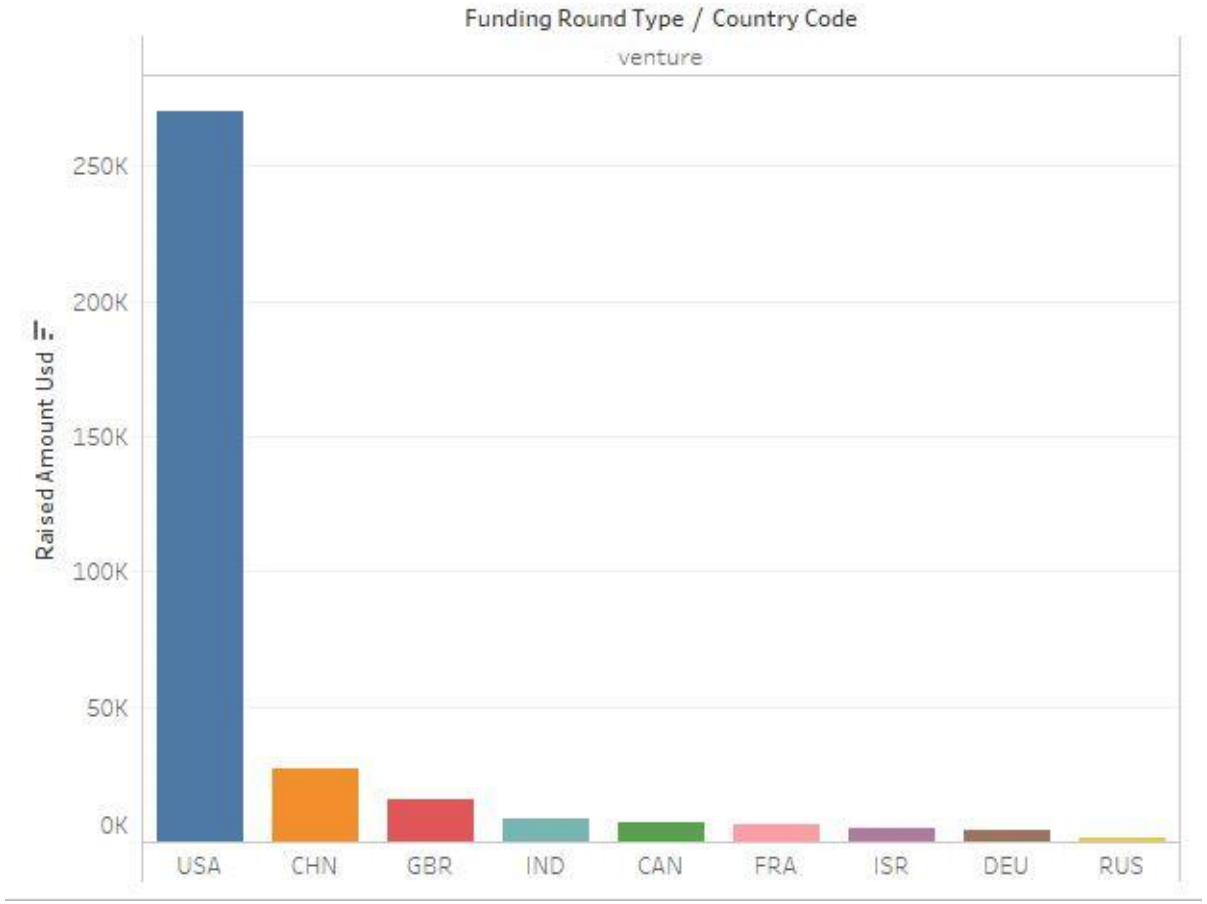
1. For sector analysis we need to load mapping.csv file.
2. We need to remove *Blanks* column and rows that have *category_list* as NULL.
3. We need to extract *primary_sector* from category_list using split on the column and getting the first item.
4. We need to Unpivot the mapping.csv file rows, so that we have only 3 columns instead of 10.
5. Using pd.melt() function we specify *id_vars*
6. We then merge the dataframe into *master_frame* dataset.

Total Investments vs Avg. Funding

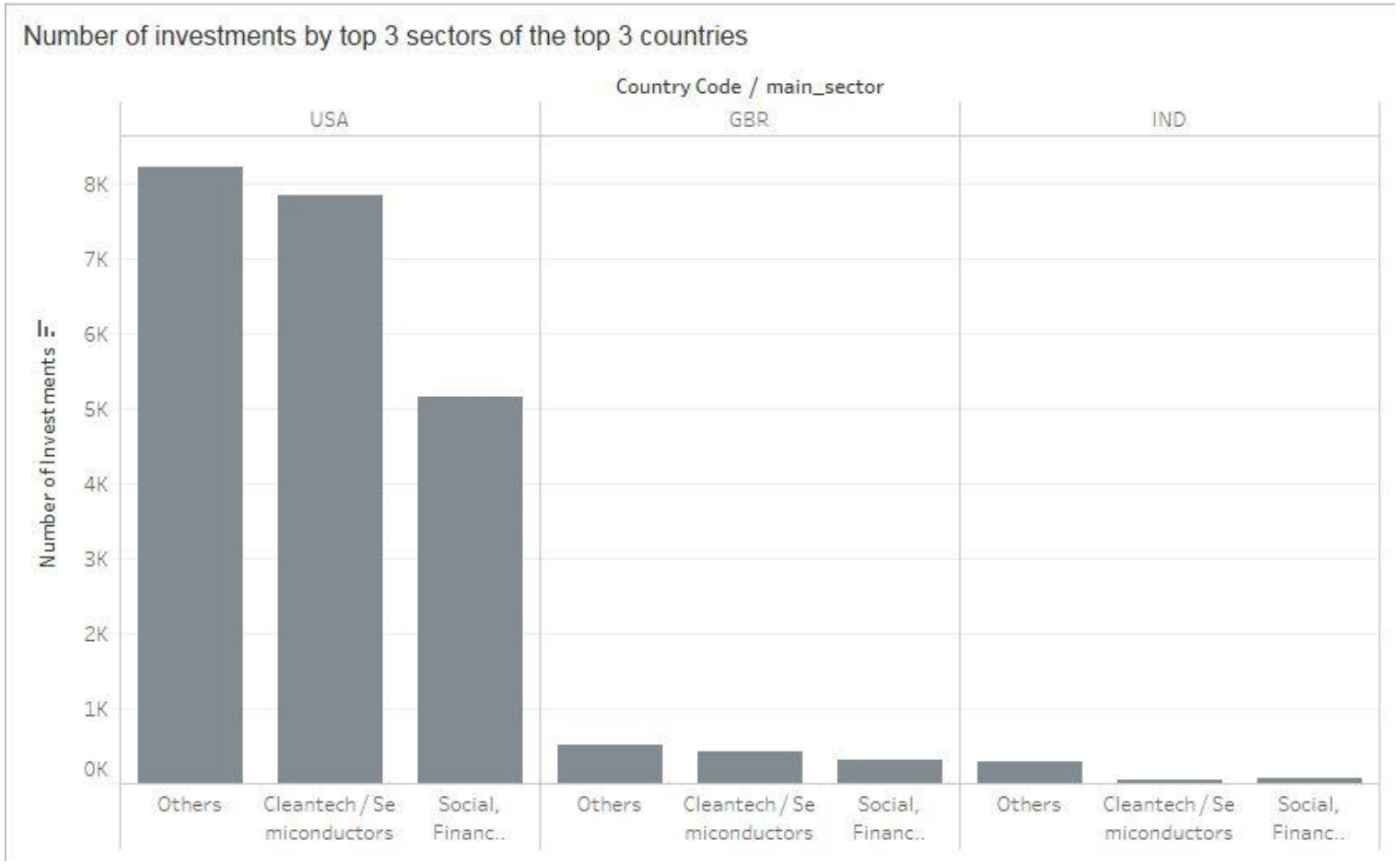




Top 9 countries vs The total amount of venture investments



No. of venture investments in the top 3 sectors of the top 3 countries



Conclusion – Main insight

So according to our analysis:

- Venture funding type is best suited for spark funds strategy
- Unites States (USA), United Kingdom (GBR) and India (IND) are top 3 countries to invest in.
- Main sectors that are best for investment are:
  - Others
  - Cleantech/Semiconductors
  - Social, Finance, Analytics, Advertising
  - News, Search and Messaging