

INVESTMENT ASSIGNMENT

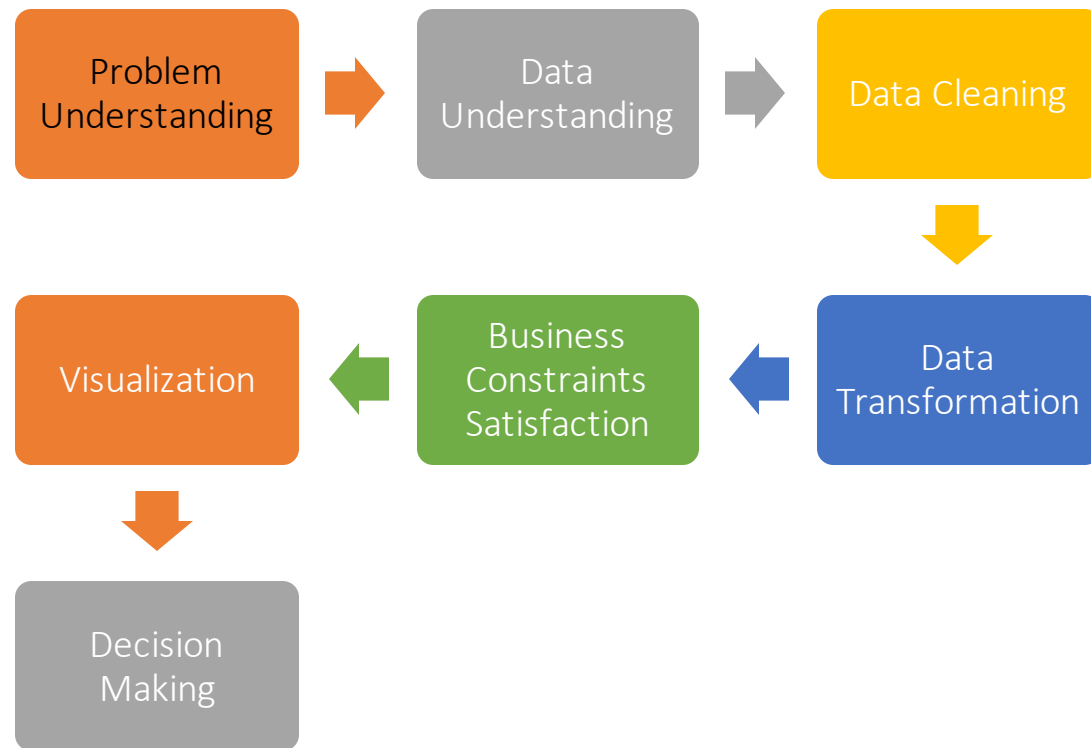
SUBMISSION

Name: Pranesh Kumar Saha

Abstract

The objective of the following analysis is to determine the best type of investment, country of investment and sector of investment for Spark Funds. It is an early stage startup investor willing to investment amount within 5-15 million USD. The company wishes to invest in sectors where the maximum investment is happening, thus reducing it's risk. Also investment will be done only in countires where english is an official language.

Problem Solving Methodology



1. Understand the problem statement and set objectives
2. Understand the given data and decide what steps are required to meet the objectives(1)
3. Clean and transform the data inline with understanding from (2)
4. Apply the constraints supplied by the business
5. Further transformations as required and construct proper visualizations to aid inference

CHECK-POINT 1 and 2 Analysis

Observation

1. Permalink column in rounds2 contains data in both upper and lower case. Company-permalink in companies has the same data in title case. Also, each permalink has '/organization/' tag preceeding the unique key
2. In merged master_frame column permalink from companies is redundant. Also, several columns like homepage_url, name, state_code etc. are not used.
3. Columns raised_amount_usd, country_code and category_list contain null values

Inference/Action

1. Split the permalink by '/' and extract the unique key, convert it to lower case and store it for both files
2. Drop all the unrequired columns from master_frame
3. Rows with null raised_amount_usd is removed, as it is the primary column required for analysis. Rows with null country_code/category_list are also removed as in further analysis they will get automatically neglected. This reduces the number of rows to process

CHECK-POINT 4 Analysis

Observations

1. Spelling mismatch between primary_sectors extracted in master_frame and category_list column in mapping. Eg- Analytics is spelled as A0lytics
2. Some categories are present in master_frame but not in mapping

Inference/Actions

1. Replace 0 with Na or na
2. By default mapped to Others. But for some assumption has been made based on similar categories. Like: Greentech – Cleantech, Racing – Automotive & Sports, Speciality Retail - Social, Finance, Analytics, Advertising and Enterprise Hardware - Manufacturing

CHECK-POINT 5

Observations

1. It is observed that the percent of closed/acquired companies from total number of companies in the top three countries USA,GBR and IND are 30.1,22.71 and 9.15 respectively.
2. Others is the top sector in each country
3. The number and amount of investments in USA far exceeds the other two countries GBR and IND
4. In USA the highest sector comes out to be Others with 2951 investments and the second highest is Social, Finance, Analytics, Advertising with 2714 investments.

Inferences/Actions

1. The companies that have closed/acquired are not available for investment. But they do provide insight into trend of investment. So, these are not rejected
2. If we look into the others segment, we can observe that it includes very diverse categories ranging from IT service providers to event management companies. Thus a large chunk of companies come under it
3. USA is the perfect country for investment, as it is seeing a large number of investments from other firms
4. Although Others comes out as the top segment, we must remember it has diverse sub-sectors within it. For an **early stage startup investor** it is better to focus on a sector like Social, Finance, Analytics, Advertising which provides scope for adequate diversification of portfolio, but in very specific sectors.

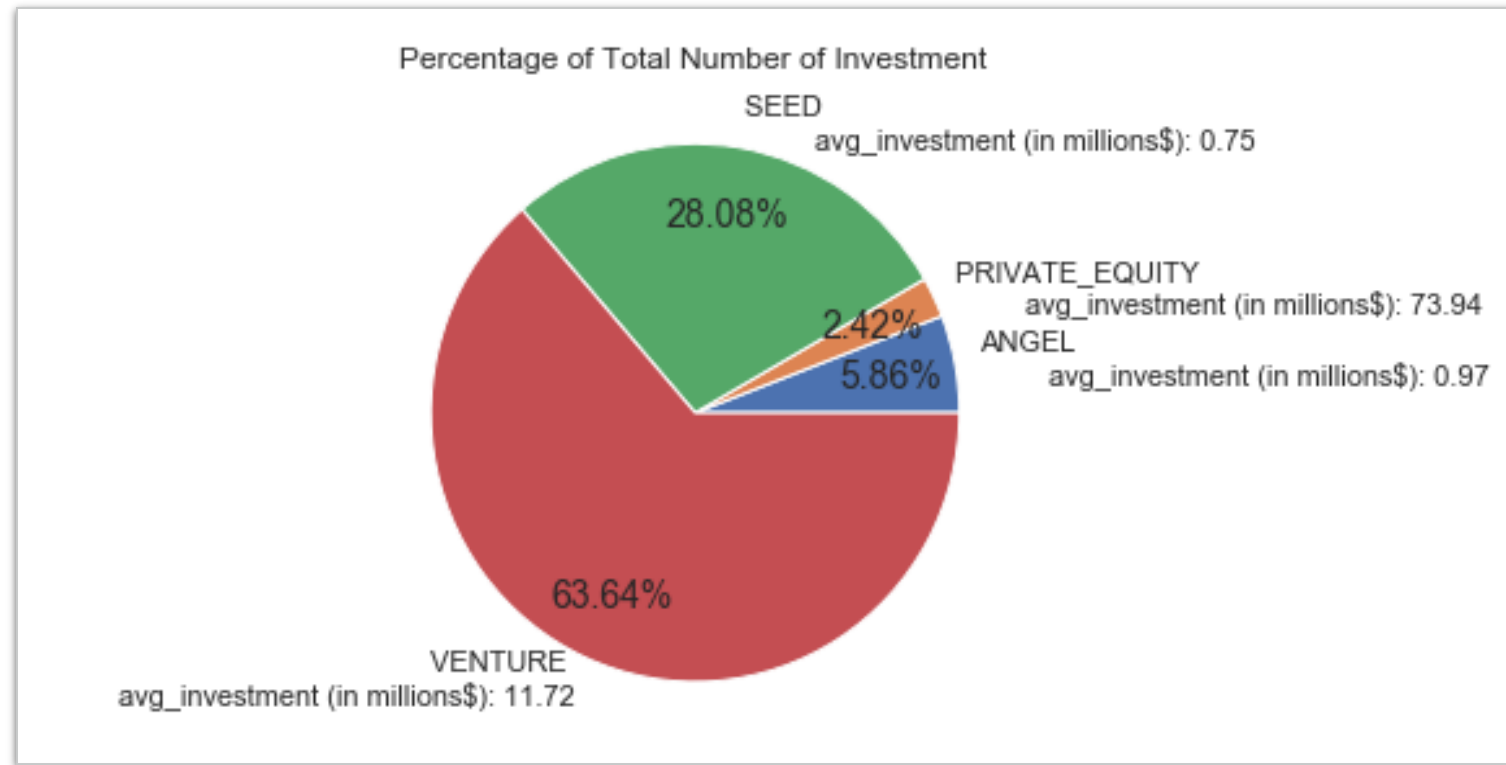
CHECK-POINT 6 Analysis

Observations

1. For plot 2, the difference in the mean amount invested in USA is much larger than the other countries. It makes the bars for other countries much smaller and hence difference in heights is not clear
2. Plot 1 is to show fraction of investment belonging to four types
3. Plot 3 is to show the number of investments in top 3 sectors in top 3 countries

Inferences/Actions

1. As the main objective of the plot is to discern the top three countries based on investment, the value of y-axis is not of quantitative use. So, a logarithmic y-axis is selected which gives much better comparison between heights
2. The pie chart is selected because only four segments are there, hence no cluttering
3. Grouped bar chart is selected as it gives a comparative view of the numbers both for segments within a country and with other countries

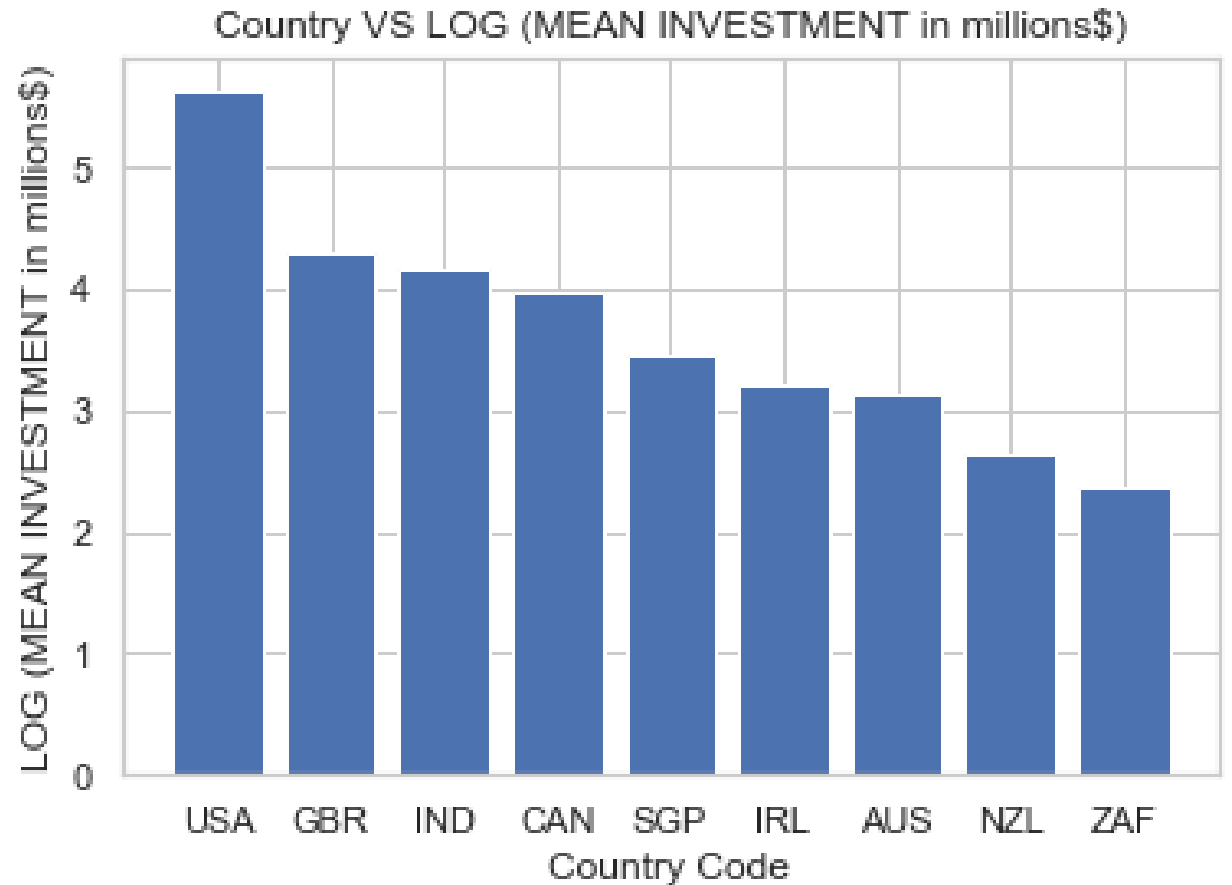


Percent Of Total Number Of Investment Belonging To Each Investment Type

Clearly the investment type best suited for given constraint is Venture with average investment of 11.75 million dollars and 62.30% share.

Country-wise Mean Investment

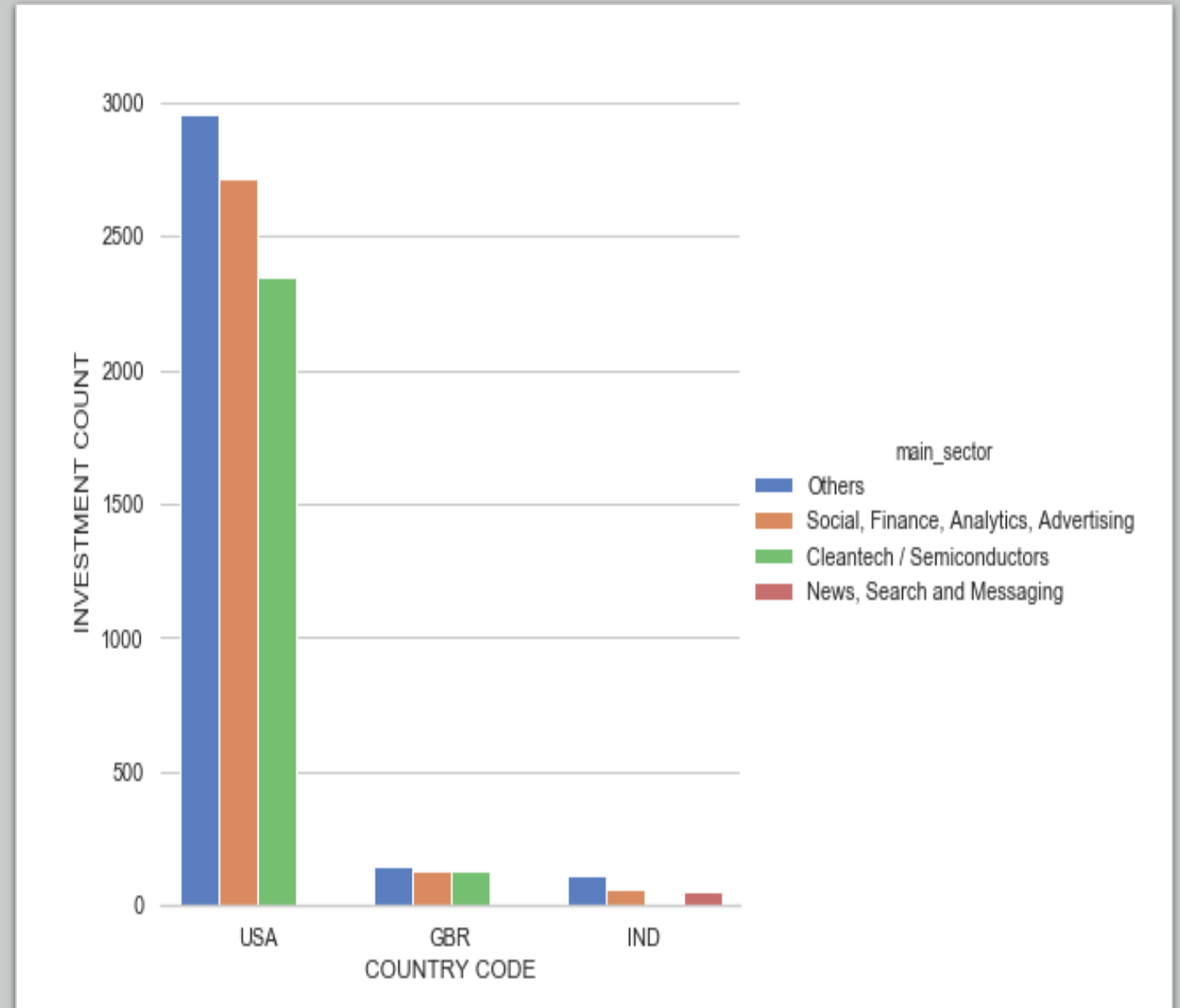
- The y-axis is logarithmic scale of base 10 over investment in millions of dollars
- Clearly the top three countries are USA, GBR and IND





Investment Count For Top Three Sectors In Top Three Countries

- Clearly the number of investments in USA exceeds that of GBR and IND combined
- Others emerges as the top investment sector in all three countries
- Social, Finance, Analytics emerges as second most invested sector in USA and IND



Conclusions

- The best investment type is venture, as it has maximum market share and the representative amount falls perfectly within the range of Spark funds target range
- The best country to invest is the USA, as it has both highest number and amount of investments
- The best sector to invest is the Social, Finance, Analytics, Advertising sector. It will provide adequate scope for diversification of portfolio without introducing the curse of too many options