



# INVESTMENT ASSIGNMENT SUBMISSION

Name: Utkarsh Tripathi





# Abstract

- The Main objective to identify best the best sectors, countries, and a suitable investment type for an asset management company for making investments.
- The company wants to invest 5 to 15 million USD per round
- They want to invest in only English-speaking countries, reason being the ease of communication.
- The strategy is to invest where other people are investing which implies that best sectors and countries are those where most of the investors are investing.
- Data source: **crunchbase.com**





## Problem solving methodology

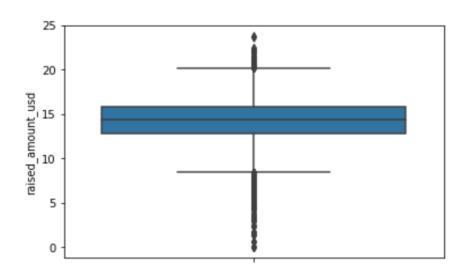
Problem Data Cleaning and Data Analysis and Data Understanding Understanding Manipulation Conclusion Missing Values Exploring the Data, Sector Analysis, Treatment, Outlier Identifying Unique **Country Analysis** Check Keys in Tables Identifying the Business Objective Identifying the Best Dropping Rows, Sectors and Country Imputation, Joining tables for Investment Changing Data Format





### Assumptions

| permalink               | 0.00  |
|-------------------------|-------|
| name                    | 0.00  |
| homepage_url            | 5.33  |
| category_list           | 2.95  |
| status                  | 0.00  |
| country_code            | 7.52  |
| state_code              | 9.49  |
| region                  | 8.81  |
| city                    | 8.81  |
| founded_at              | 17.84 |
| company_permalink       | 0.00  |
| funding_round_permalink | 0.00  |
| funding_round_type      | 0.00  |
| funding_round_code      | 72.90 |
| funded_at               | 0.00  |
| raised_amount_usd       | 17.38 |
|                         |       |

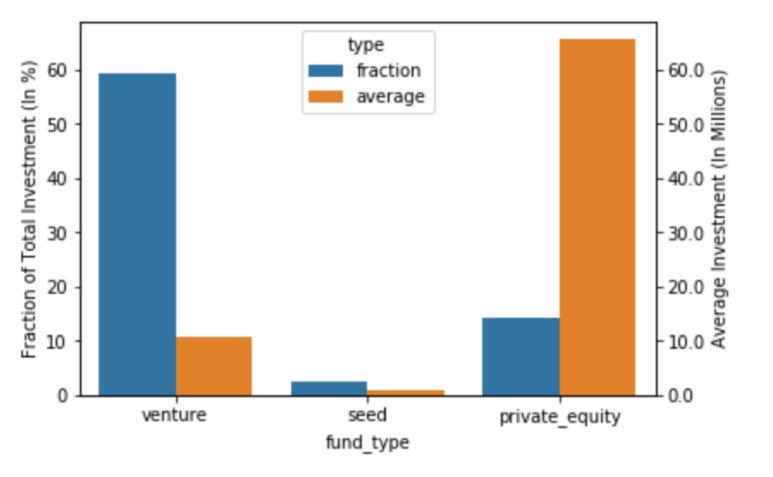


- All the missing values in `country code` column of `master` data frame are dropped since imputing them with most frequent observations result in biased data.
- All the remaining missing values in column `category list` are also dropped.
- Since 'raised amount' has lot of outliers as seen from above box plot, missing values are filled with Median of that column



### Results: Plot 1 Finding the Best Funding Type



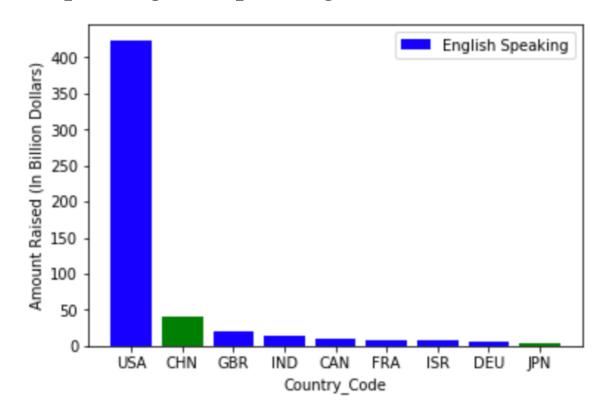


- 1.) From above graph, its clear that most of the investment is under **Venture** type (around 60% of total investment)
- 2.) Also as per requirement, Spark Funds needs to invest in range 5-15 million, average investment in that range is only available in "**Venture**" type.





Results: Plot 2 Finding the Top 3 English Speaking Countries

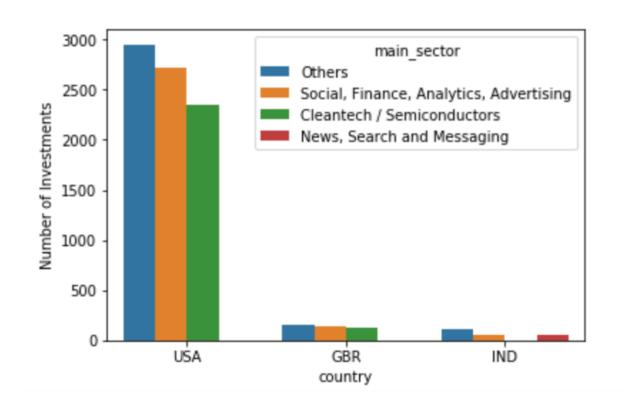


- 1.) Its clear from above graph that **USA** is the most favored nation for Investment
- 2.) The top 3 English Speaking favored nations are: USA, GBR and IND





Results: Plot 3 Finding the top 3 sectors in the top 3 English speaking country



- 1.) Most favored nation is **USA** followed by **GBR and IND**
- 2.) "Others" Sector is most favored in all three countries followed by "Social, Finance, Analytics, Advertising"





#### Conclusions

- 1.) Spark Fund should invest in Venture type of Funding since most investment happened in this type
- 2.) Most favoured English Speaking Country for Investment is USA followed by GBR and IND
- 3.) "Others" Sector is most favoured sector for Investment followed by "Social, Finance, Analytics, Advertising" in all the 3 countries.
- 4.) In USA and GBR, 3<sup>rd</sup> most favored sector is "Cleantech/Semiconductors" and in IND its "News, Search and Messaging"
- 5.) Finally, Spark Funds should invest in USA in Venture funding type and in any of the 3 most favored sectors.