



INVESTMENT ASSIGNMENT SUBMISSION

Name: Abdulazeez Basheer





Abstract

- Objective:
 - To find the best investment sector for Spark Funds to invest
- Constraints for investments:
 - To invest between 5 to 15 million USD per round of investment
 - To invest only in **English-speaking countries**
- Data Availability:
 - **companies**: A table with basic data of companies
 - rounds2: A table with funding round details.
 - Sector Classification: The purpose is to simplify the analysis into eight sector buckets
- Analysis
 - To choose one of these four investment types for each potential investment (venture, angel, seed, and private equity)
 - The top three English-speaking countries.
 - The total number (or count) of investments for each main sector
 - The total amount invested in each main sector



Problem solving methodology



 Spark Funds wants to invest where most other investors are investing

- Load data from files
- Check data quality

- Describe()
- Info()
- Isnull()

- Cleaning data
- Selecting relevant data
- Integrating data

Business Understanding

Load Data

Understand Data

Massaging Data

Funding Type Analysis

Country Analysis

Sector Analysis

Plots

 Choose best funding Venture, angel, seed, Private Equity w.r.t Investment Amount

- Top nine countries
 Received the highest total funding
- Top three Englishspeaking countries

- Impute primary sector of each category list from the category list column
- The total number & amount (or count) of investments for each main sector

- Total investments (globally) in venture, seed, and private equity
- The number of investments in the top 3 sectors of the top 3 countries





Steps Performed:

- Load the dataset into pandas with appropriate encoding format
- Understand data values in each columns
- Retrieving columns which are required for analysis
- Handling missing values
- Treating errors in columns values such as typo error(replacing 0 with na).
- Updating primary columns with lower case in each data frame for further analysis.
- Merging data frames to make master data frame.
- Adding new columns into master dataFrame.
- Calculating aggregate values to know the count, mean, sum etc..





Assumptions:

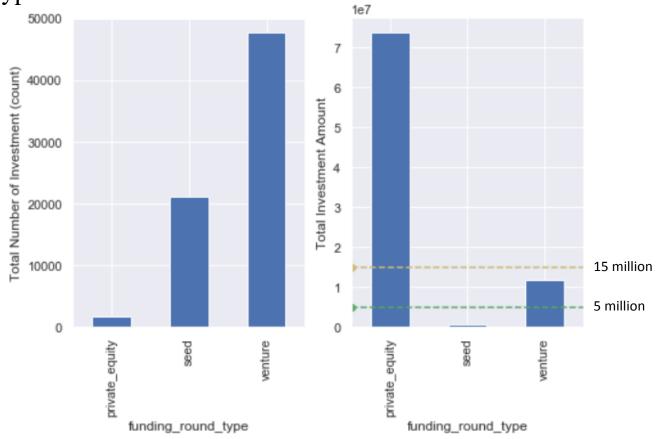
- Encoding Format
- Dropping Columns with high null value percentage.
- Dropping Rows where the min, max and mean difference is high and also the data loss is less than 15%
- Finding Country's official language manually.
- Presence of 0 in category_list of mapping file is typo error, replacing 0 with na (excluding valid values like Enterprise 2.0)
- As per my analysis, got **others** as one of the top 3 sectors for investment sector, Excluding Others sector in the final calculation of finding the top 3 investment (Since I cannot say to investors, to invest on **Others sectors** instead of specific.)
- While finding the most invested company among the top first and second sector, found multiple results with same investment amount





Total investments (globally) in venture, seed, and private equity, and the average amount of investment in each funding type

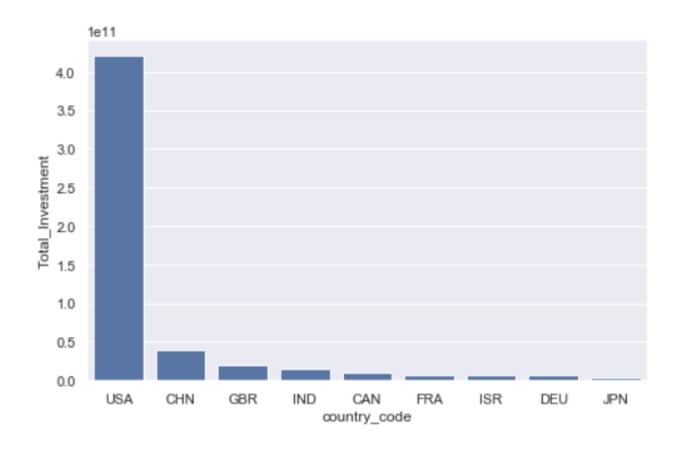
Venture Funding round type is the best







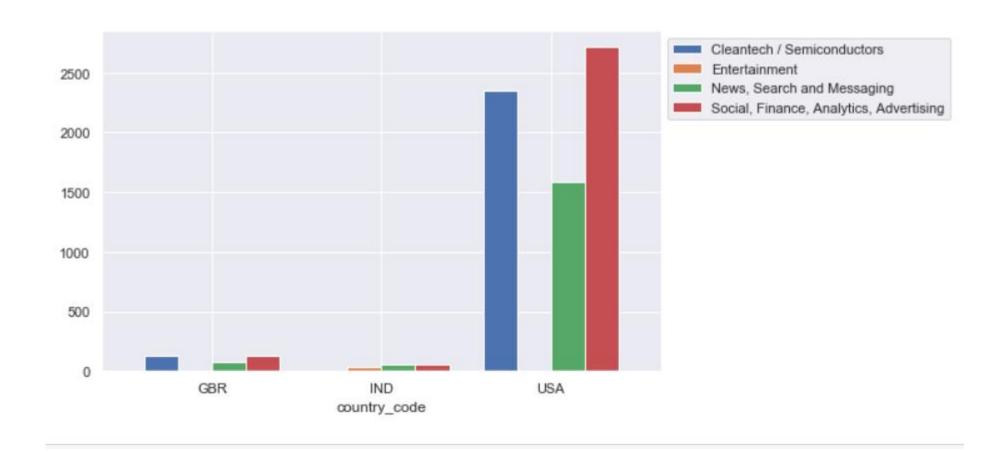
Top 9 countries against the total amount of investments of funding type Venture Sector USA is the Top country







The number of investments in the top 3 sectors of the top 3 countries for Venture Sector Main Sector in USA is Social , Finance, Analytics, Advertising







Conclusions

The Best investment for Spark Funds will be in

Country: USA

Main sectors: Social, Finance, Analytics, Advertising and

Funding Round Type is Venture for 5 to 15 million USD