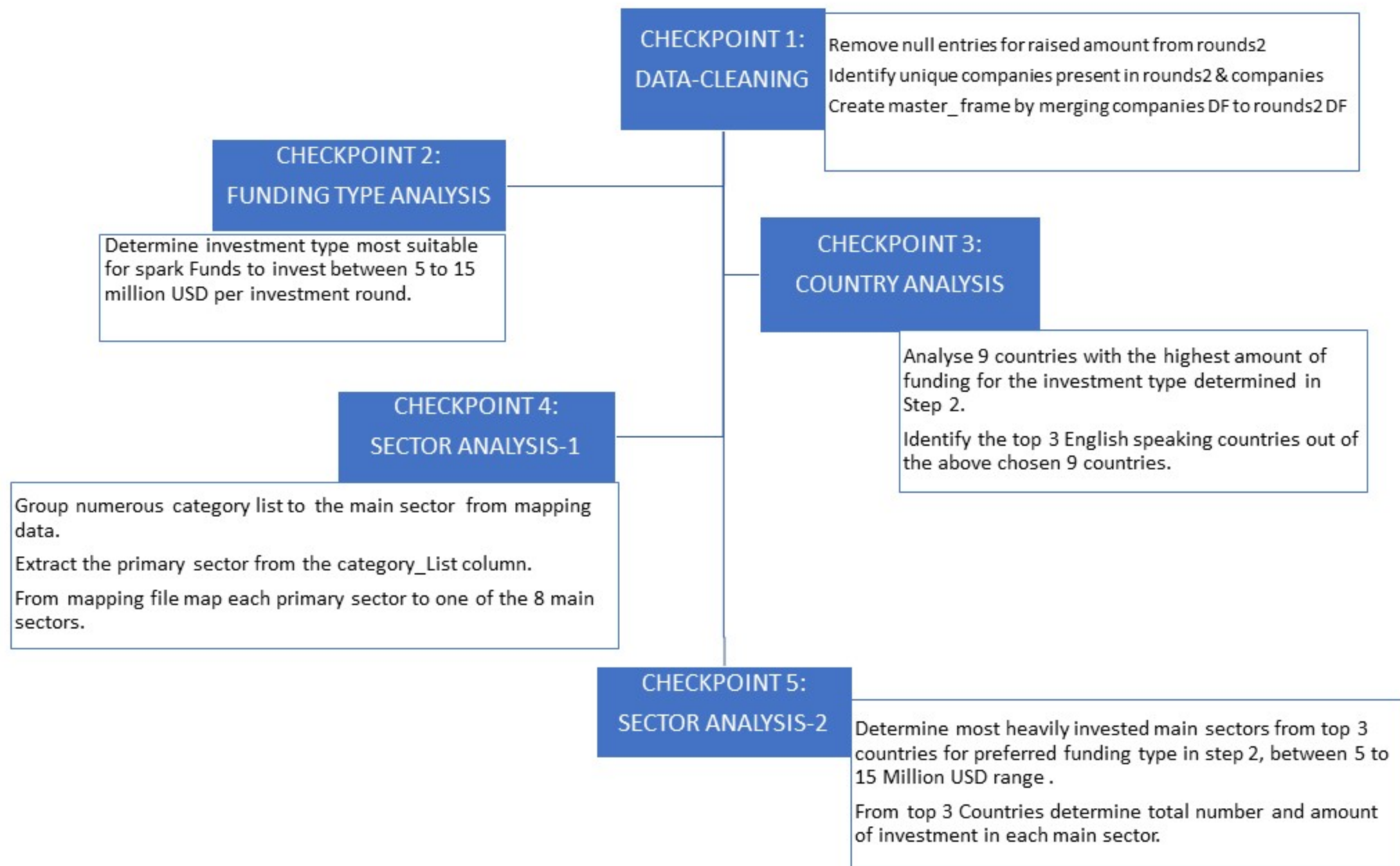# INVESTMENT ASSIGNMENT

SUBMISSION

Name: Prateek Pande

Investment Analysis for Spark Funds

Introduction - Spark Funds is an asset management company who wants to make invest in few companies based on below criteria.

1. Business Objectives – The objective is to identify the best sectors, countries, and a suitable investment type for making investments based on strategy where 'where most investors are investing'.

2. Goals of Data Analysis –
    1. Investment Type Analysis – Spark Funds want to choose fund type that is best suited for investment.
    2. Country Analysis – Want to invest in countries which are most heavily invested in past.
    3. Sector Analysis – Want to study trend of investment made in eight main sectors to decide distribution of funds based on that.

3. Minor Constraints –
    1. Spark Funds wants to invest 5 – 15 million USD per round of investment.
    2. It wants to invest only in English

INVESTMENT ANALYSIS FLOWCHART

CHECKPOINT 1:
DATA-CLEANING

Remove null entries for raised amount from rounds2

Identify unique companies present in rounds2 & companies

Create master_ frame by merging companies DF to rounds2 DF

CHECKPOINT 2:
FUNDING TYPE ANALYSIS

Determine investment type most suitable for spark Funds to invest between 5 to 15 million USD per investment round.

CHECKPOINT 3:
COUNTRY ANALYSIS

Analyse 9 countries with the highest amount of funding for the investment type determined in Step 2.

Identify the top 3 English speaking countries out of the above chosen 9 countries.

CHECKPOINT 4:
SECTOR ANALYSIS-1

Group numerous category list to the main sector from mapping data.

Extract the primary sector from the category_List column.

From mapping file map each primary sector to one of the 8 main sectors.

CHECKPOINT 5:
SECTOR ANALYSIS-2

Determine most heavily invested main sectors from top 3 countries for preferred funding type in step 2, between 5 to 15 Million USD range .

From top 3 Countries determine total number and amount of investment in each main sector.

Investment Analysis Approach Checkpoints

Checkpoint 1 : Data Cleaning

1. Create data-frame companies and rounds2 from companies.txt and rounds2.csv respectively.

2. Remove rows with null entries for raised_amount_usd from round2 data-frame.

3. Determine unique companies from rounds2 and companies data-frame.

4. Merge the two data frames so that all variables (columns) in the companies frame are added to the rounds2 data frame in master_frame.

Checkpoint 2 : Funding Type Analysis

1. Create new data-frame after filtering master_frame for preferred funding types – venture, angel, seed and private-equity.

2. Calculate the most representative value of the investment amount for each of the four funding type.

3. For preferred funding types Find investment between 5 to 15 million USD for each investment round.

4. From above venture funding is most preferred funding type.

Checkpoint 3 : Country Analysis

1. Spark Funds wants to see the top nine countries which have received the highest total funding. Determine top 9 countries with highest total funding in the preferred funding type.

2. For the chosen investment type, make a data frame named top 9 with the top nine countries.

USA   108531347515.00

GBR   5436843539.00

CHN   5016561053.00

CAN   3647009578.00

IND   2976543602.00

ISR   2840305951.00

FRA   2112425453.00

DEU   1765026223.00

3. From above top 9 countries choose top 3 English speaking countries using -
https://en.wikipedia.org/wiki/List_of_territorial_entities_where_English_is_an_official_language

4. Thus USA, GBR and CAN are the top 3 countries for preferred investment type.

Checkpoint 4: Sector Analysis 1

1. Few of the companies have multiple category_list separated by "|". We need to split category list and take first entry and primary_sector.

2. Use the mapping file 'mapping.csv' to map each primary sector to one of the eight main sectors –
   "Automotive & Sports", "Cleantech / Semiconductors", "Entertainment", "Health", "Manufacturing", "News, Search and Messaging", "Others", "Social, Finance, Analytics, Advertising"

3. Determine merged data frame with each primary sector mapped to its main sector

Checkpoint 5: Sector Analysis 2

1. Create three separate data frames D1, D2 and D3 for each of the three countries containing the observations of funding type FT falling within the 5-15 million USD range. The three data frames should contain:
   1. All the columns of the master_frame along with the primary sector and the main sector
   2. The total number (or count) of investments for each main sector in a separate column
   3. The total amount invested in each main sector in a separate column

2. Using the three data frames, you can calculate the total number and amount of investments in each main sector.

3. Create Data-Series which contains required amount and count for all three countries.

4. Determine total number of investments in all three countries.

5. Total amount of investments in all three countries.

6. Top Sector name (no. of investment-wise)

7. Number of investments in top sector.

8. For top sector count-wise which company received the highest investment

9. For second best sector count-wise which company received the highest investment.
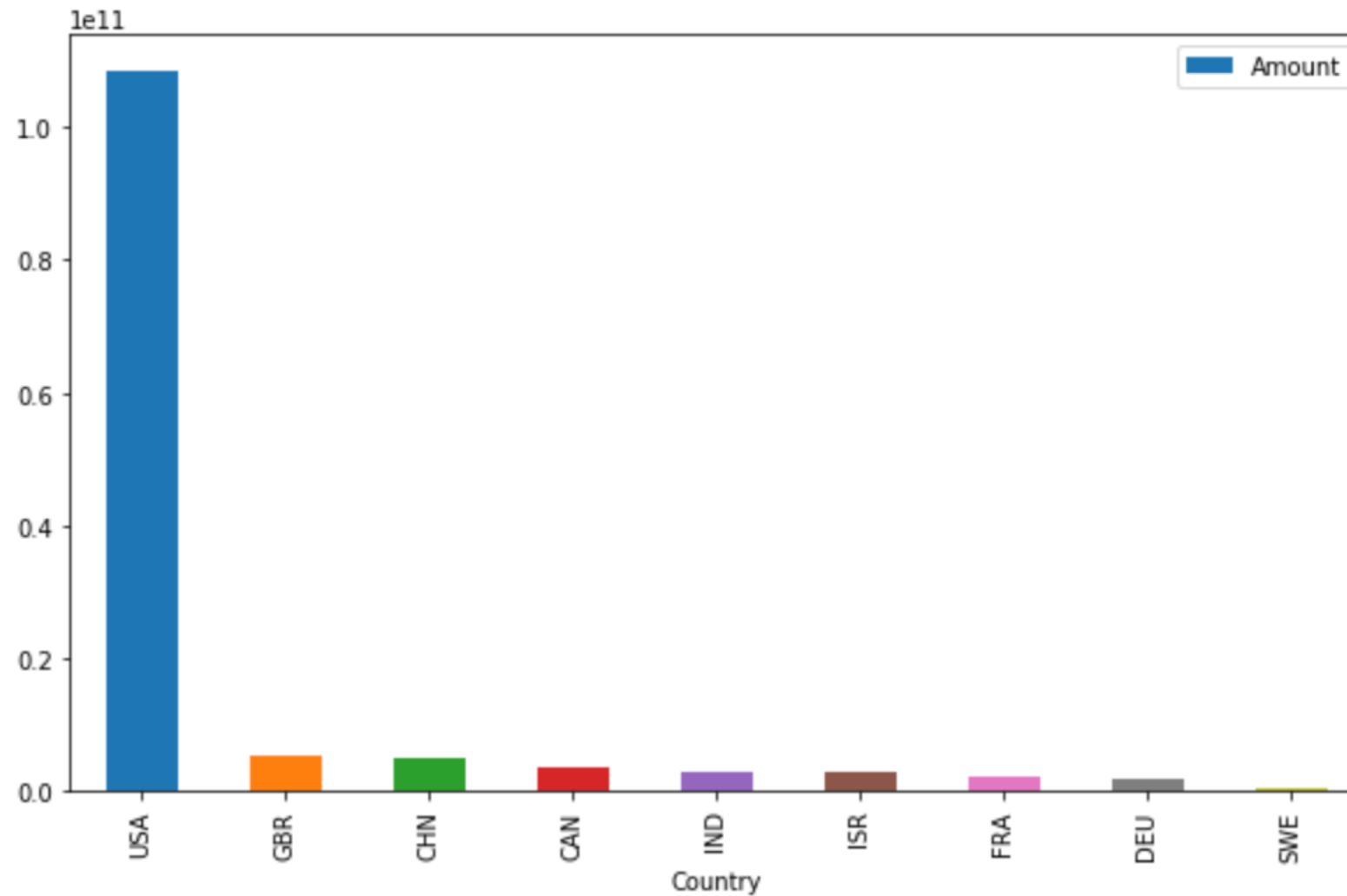
A plot showing the fraction of total investments (globally) in venture, seed, angel, and private equity, and the average amount of investment in each funding type.
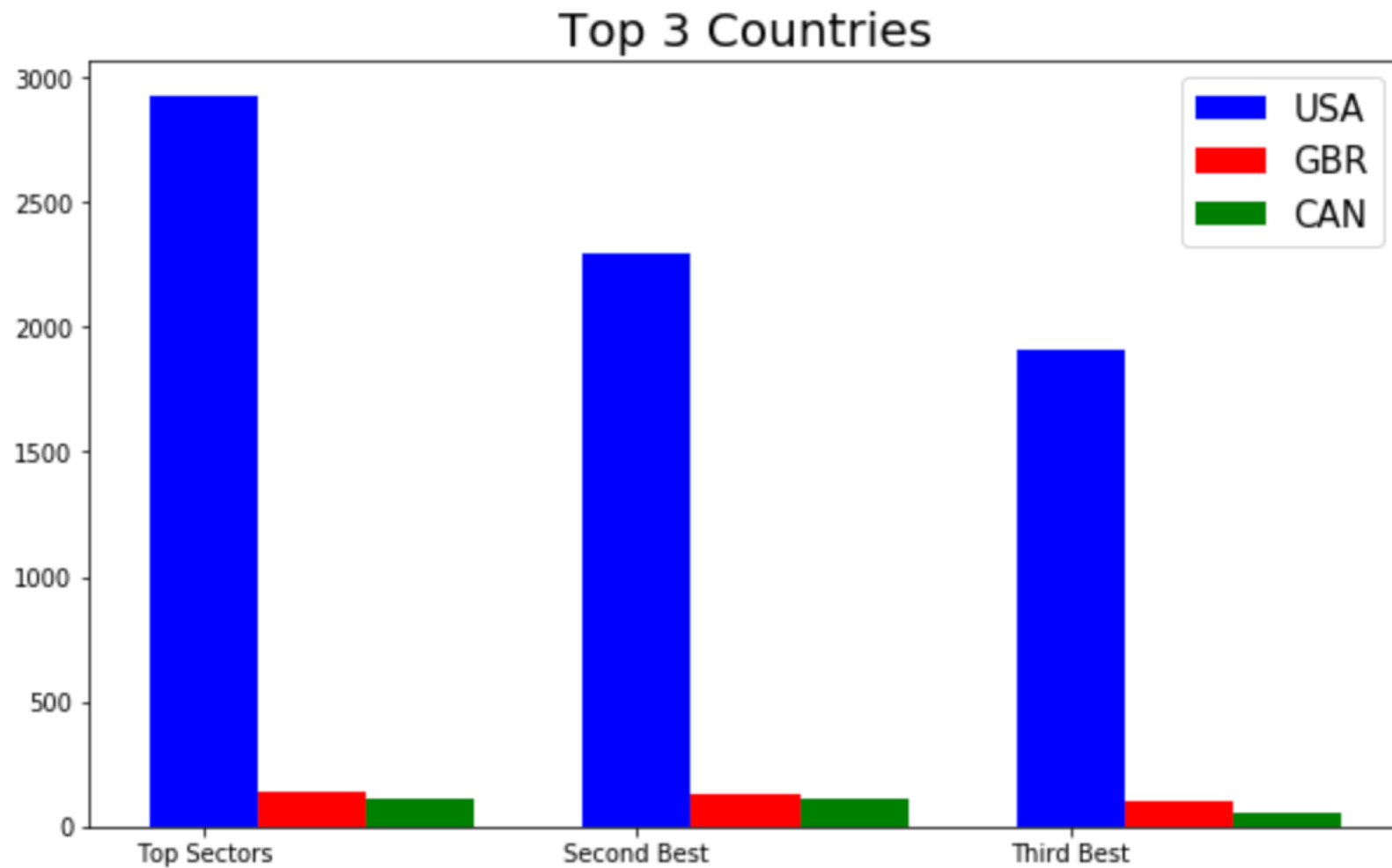
Plot 1
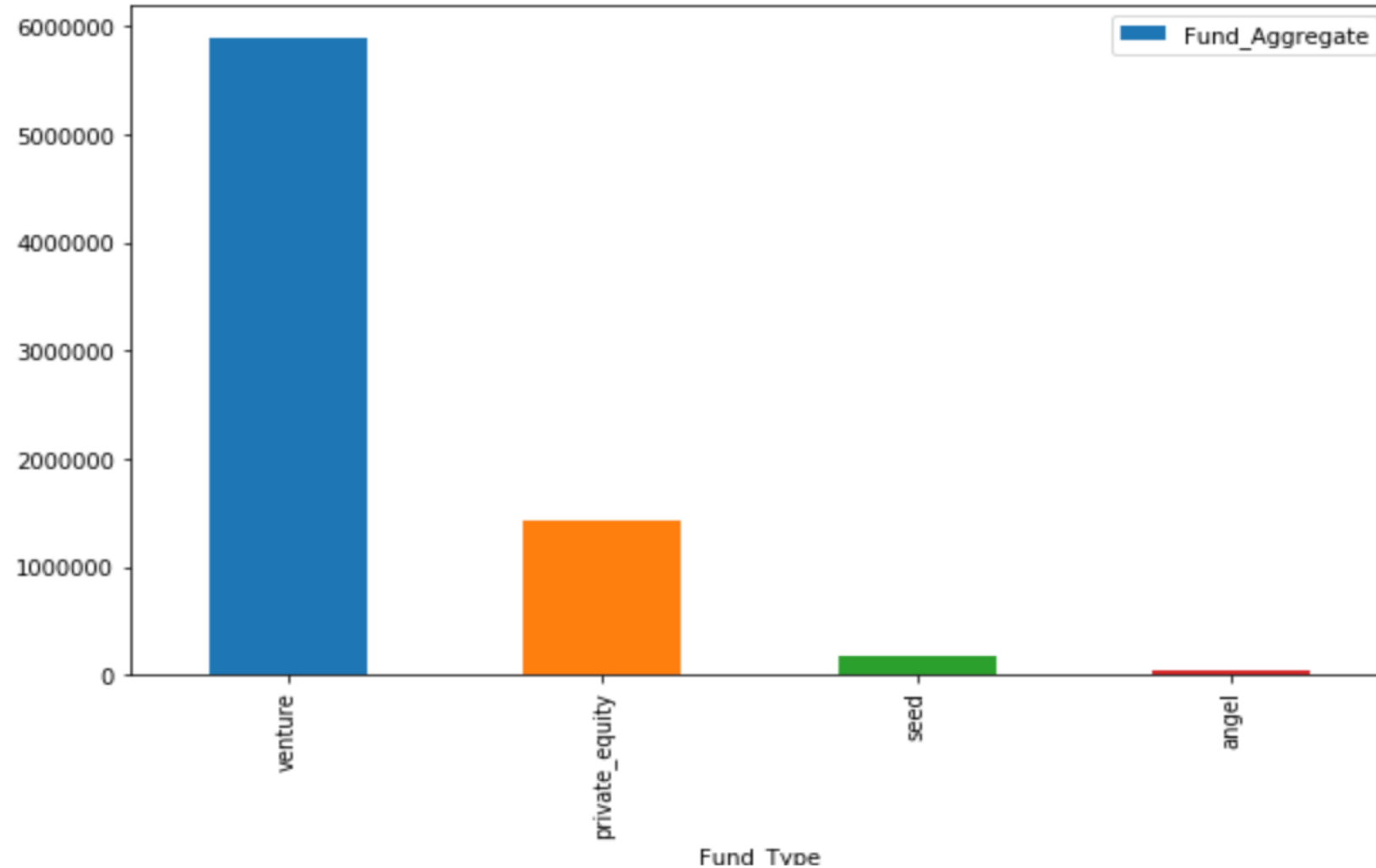
A plot showing the top 9 countries against the total amount of investments of funding type FT.

Plot 2

A plot showing the number of investments in the top 3 sectors of the top 3 English speaking countries on one chart

Plot 3

Conclusion

Following conclusion can be drawn from analysis for Sparks Funds -

1. Best Fund Type for investment in range of 5 to 15 million USD – Venture

2. Top 3 English speaking countries to make investment in venture fund – USA, GBR and CAN.

3. Best Company for USA count wise –
    1. Virtustream
    2. Biodesix

4. Best Company for GBR count wise –
    1. Electric_cloud
    2. EUSA Pharma

5. Best Company for CAN count wise –
    1. Fresco Microchip
    2. QuickPlay Media