

INVESTMENT ASSIGNMENT SUBMISSION

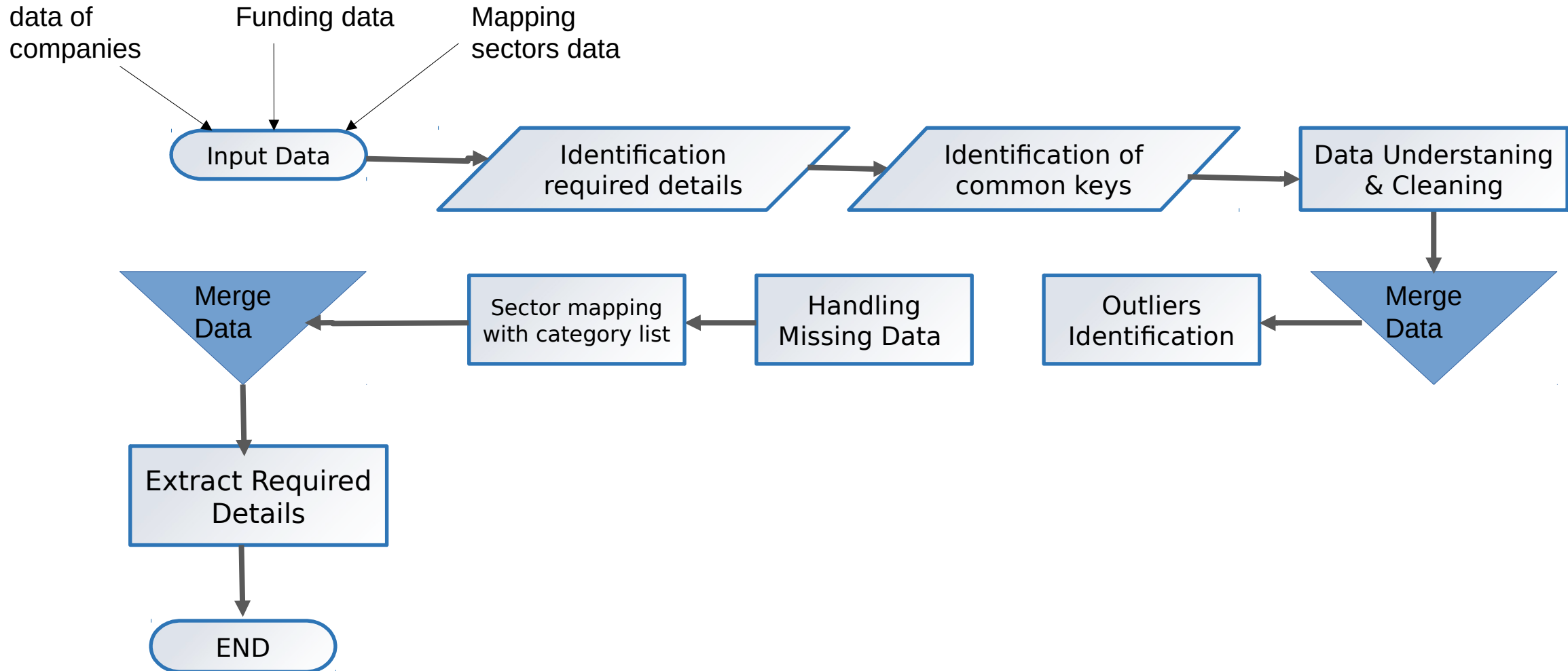
Name: Jitendra Kumar Kushwaha

Identifying the best sectors, countries for making investments

- **Goals**

- Find out the details of the amount invested in the venture, seed, angel and private equity and the best sector to invest amount.
- Find out the most invested countries and best choice of the investors.
- More accurate results of the investment amounts in the venture, seed, angel and private equity.
- Data Cleaning
- Graphical Representations

Problem solving methodology



Business and Data Understanding

- Invest only in English-speaking countries and invest between 5 to 15 million USD per round of investment.
- Load the companies and rounds data in the dataframe using “ISO-8859-1” encoded format.
- Explored and understand data using pandas methods like .info(), .describe(), .head(), .tail(), .shape etc.

In [243]: `companies.describe()`

Out[243]:

	permalink	name	homepage_url	category_list	status	country_code	state_code	region	city	founded_at
count	66368	66367	61310	63220	66368	59410	57821	58338	58340	51147
unique	66368	66102	61191	27296	4	137	311	1092	5111	3978
top	/Organization/Trion-World-Network	Spire	http://www.askforoffer.com	Software	operating	USA	CA	SF Bay Area	San Francisco	01-01-2012
freq	1	4	5	3995	53034	37601	12900	8804	3526	2730

- Identifying the common keys column.

Data Cleaning

- Companies Datasets

```
companies.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 66368 entries, 0 to 66367
Data columns (total 10 columns):
permalink      66368 non-null object
name           66367 non-null object
homepage_url   61310 non-null object
category_list  63220 non-null object
status         66368 non-null object
country_code   59410 non-null object
state_code     57821 non-null object
region        58338 non-null object
city           58340 non-null object
founded_at     51147 non-null object
```

- Identifying the keys column
- Change the key in the lowercase and decode in 'ISO-8859-1' format and decode ASCII characters.
- Merge companies datasets to the rounds datasets.

- Rounds Datasets

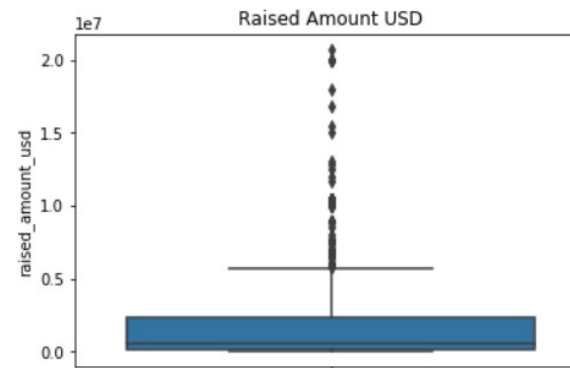
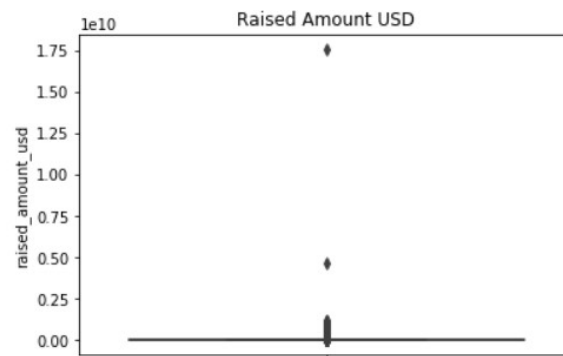
```
rounds2.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114949 entries, 0 to 114948
Data columns (total 6 columns):
company_permalink      114949 non-null object
funding_round_permalink 114949 non-null object
funding_round_type     114949 non-null object
funding_round_code     31140 non-null object
funded_at              114949 non-null object
raised_amount_usd      94959 non-null float64
```

- Identifying the keys column
- Change the key in the lowercase and decode in 'ISO-8859-1' format and decode ASCII characters.
- `rounds2['company_permalink'].isin(companies['permalink'])`

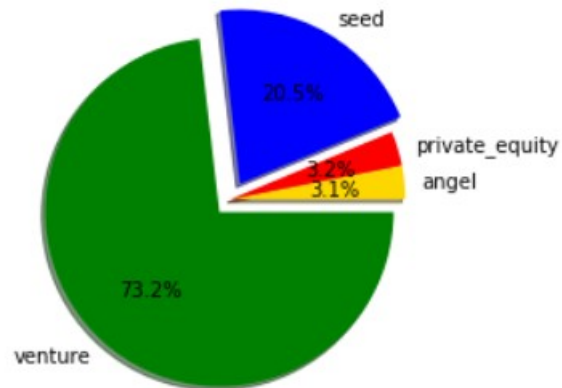
Data Cleaning Cont.

- Merged two datasets
- `master_frame = pd.merge(rounds2, companies, on='permalink')`
- Funding type analysis in the master_frame
- Removed the column which are not useful for the analysis
- Handled missing data for the selected investments type ['venture', 'angel', 'seed', 'private_equity']
- Detect outliers with z-score based on the investments type and after performed data aggregation and updated NaN value with mean, median or mod.
- Below boxplot chart for the angel investmetn type.

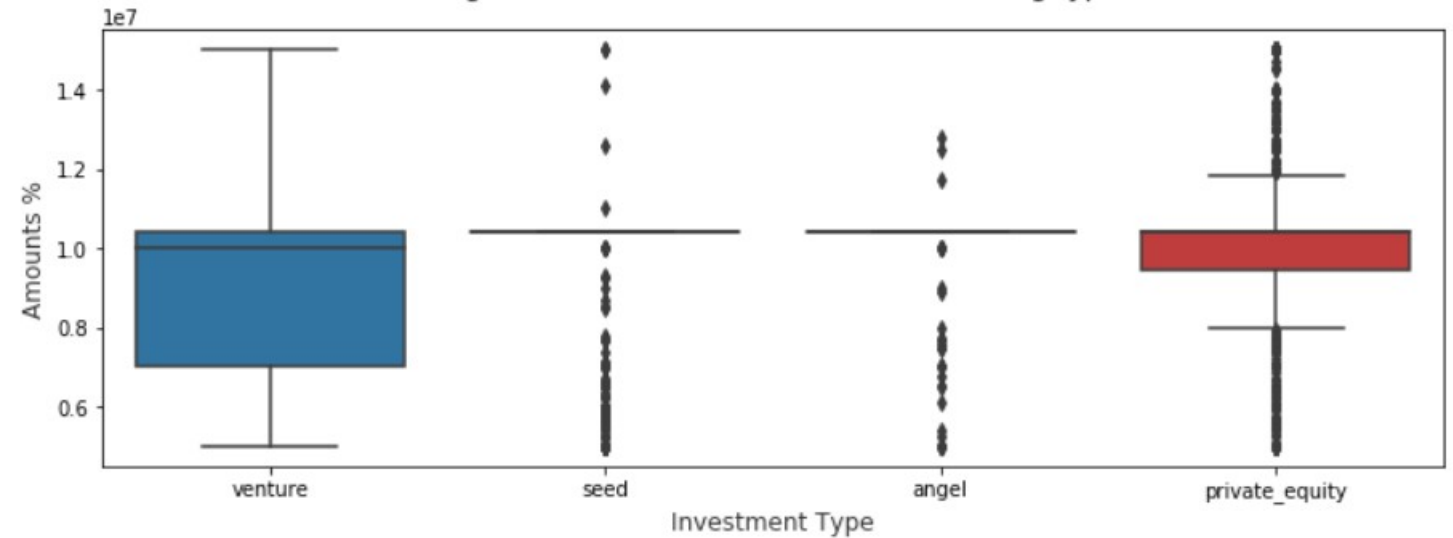


Total investments (globally) in angel, venture, seed, and private equity

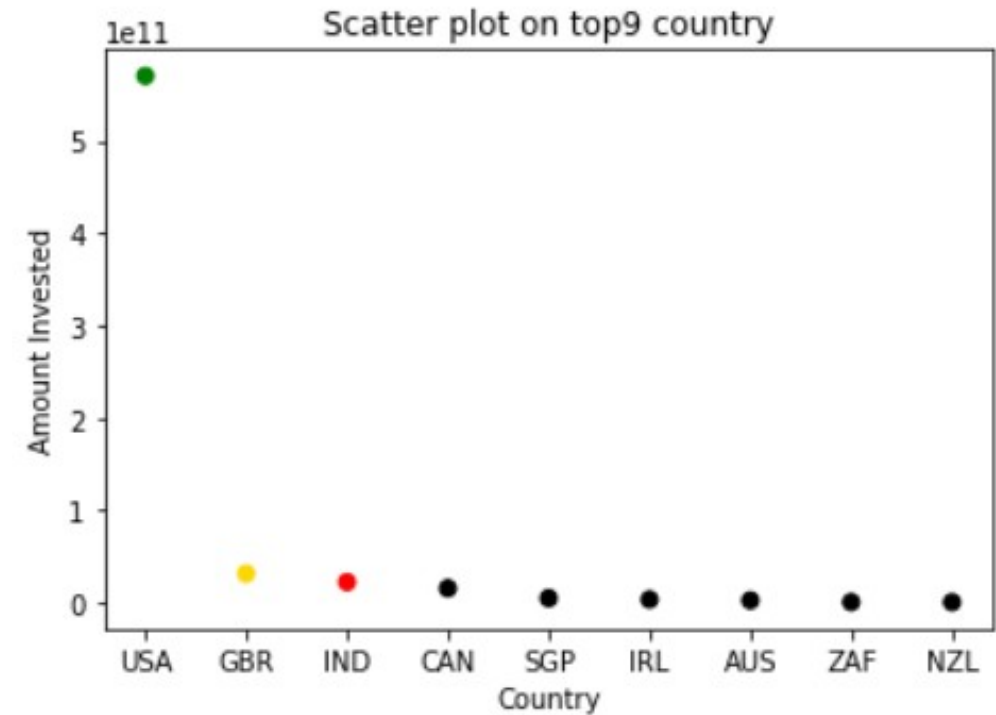
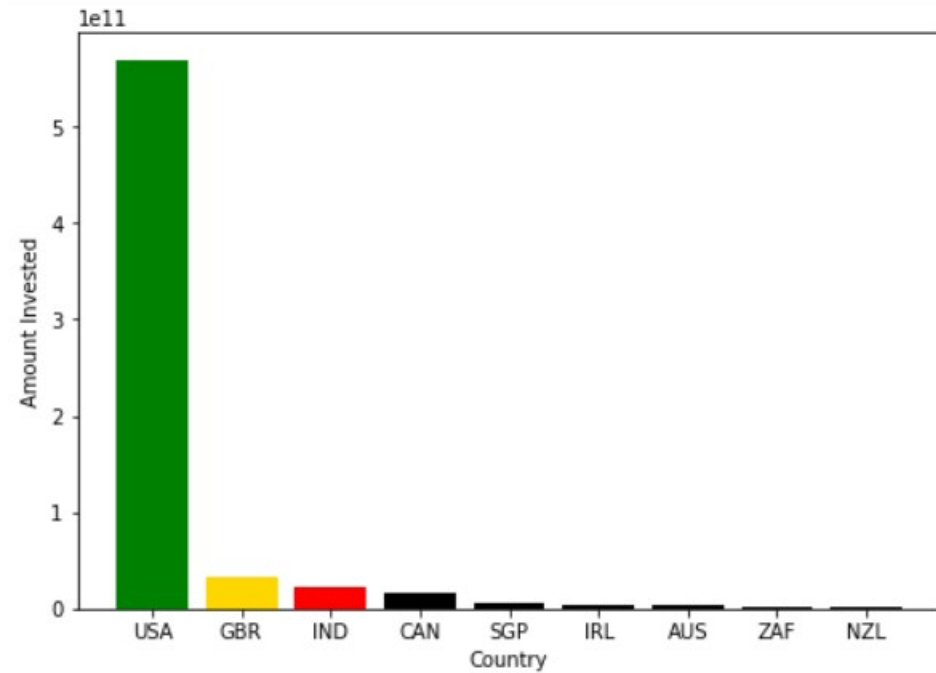
Total amount of investments of funding type FT



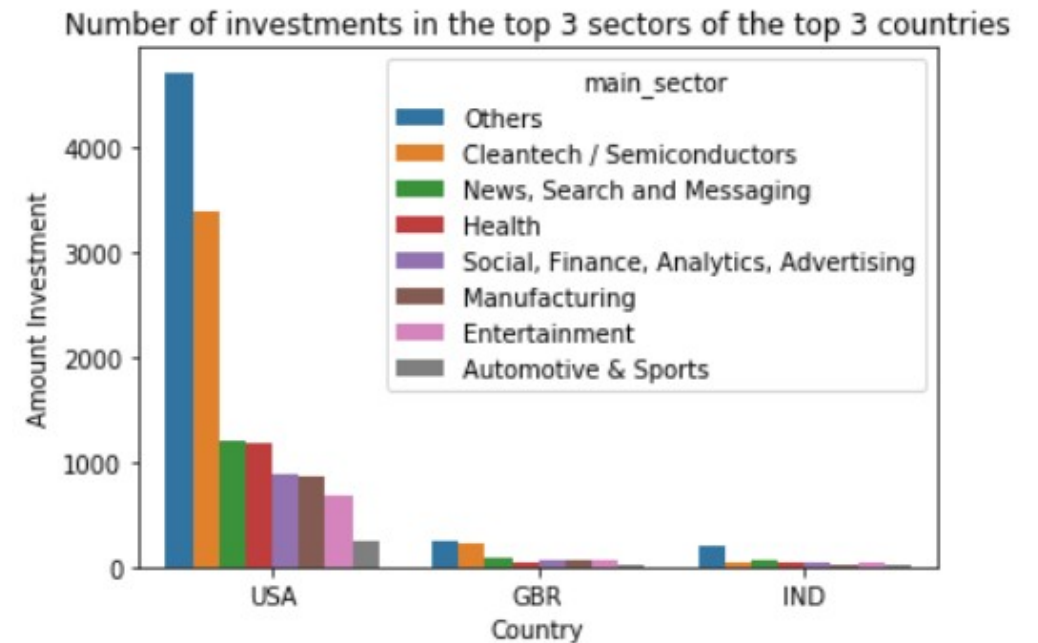
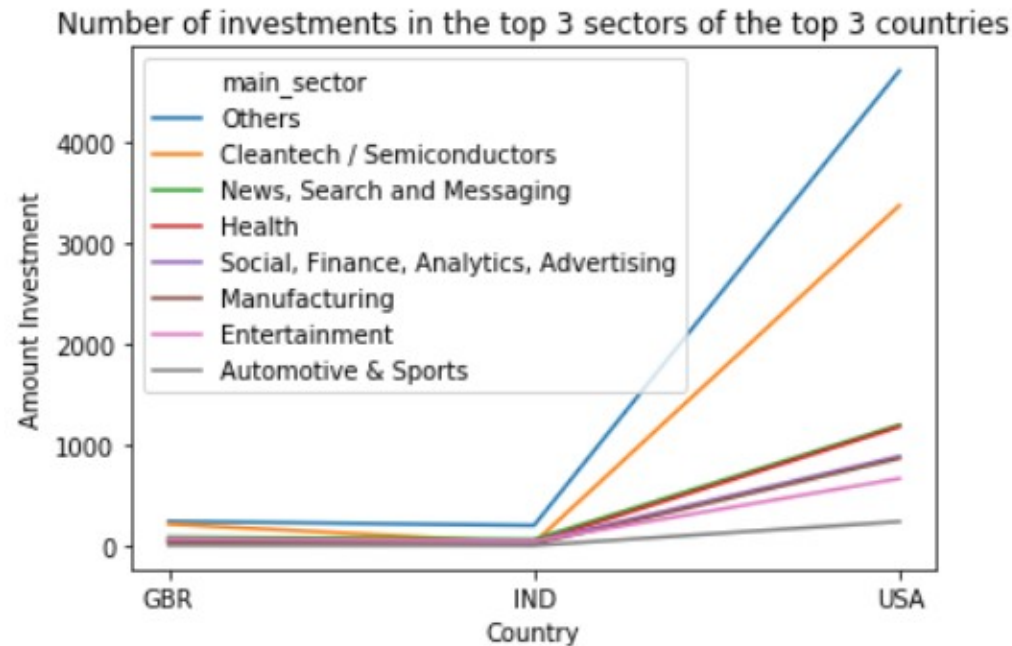
Average amount of investment in each funding type



Top 9 countries against the total amount of investments of funding type FT



Number of investments in the top 3 sectors of the top 3 countries



Conclusions

- As per the analysis.
 - USA, GBR & IND are the top three countries where investors are investing.
 - USA is a top country where investors are investing and which has highest investment comparison to any other countries.
 - Venture type investment is in the top from the choosen investment type.
 - Others & Cleantech / Semicondutors are the top 2 sectors where investors are investing.

.