

# INVESTMENT CASE STUDY

## SUBMISSION

Group Name:

1. Ankit Singhai

## <Abstract>

As an asset management company we would like to understand the global trends of investment across different sectors & countries.  
So an informed decision can be made on where to invest.

Constraints on investments are:-

1. Investment range 5M – 15M USD.
2. Only English speaking countries.

Start

## Business Understanding

1. Business Goal – Invest in sector(s) & countries where most investment are made wr.t. funding type.
2. Data Analysis Goals are:-
  1. Funding type best suited for Spark Investment.
  2. Country analysis – Top 3 countries for the selected funding type.
  3. Sector analysis – Top 3 sectors where most investment are done in each top 3 countries.

## Understanding Data

1. Companies – Data about company inception, geo-location & primary sector.
2. Rounds – Data about funding(s) & it's type received by company.
3. Mapping – Information on sector(s) mapped to primary sector.

## Data Preparation

1. Companies – Removing rows that doesn't have :-
  1. primary category (companies.category\_list) – cannot be mapped to sector.
  2. Geo location – country, state, region & city – can't be mapped to country.
2. Rounds – removed rows that have NULL value in raised\_amount\_usd – because analysis is to find suitable investment.
3. company.permalink & rounds2.company\_permalink data are in different cases changed both them to lower case.

## < Problem solving methodology >

### Analysis

1. On merged data calculated mean / median of raised\_amount\_usd across different funding types.
2. Based on 5 – 15 Million USD investment range figured out Venture – suited type.
3. For Venture funding type summed all investment grouped by country to determine top 3 English speaking nations.
4. Mapped the above data to get main sector - determined top 3 sector for each top 3 countries.
5. For top 2 sectors for each country determined top company for investment.

Stop

## <Analysis>

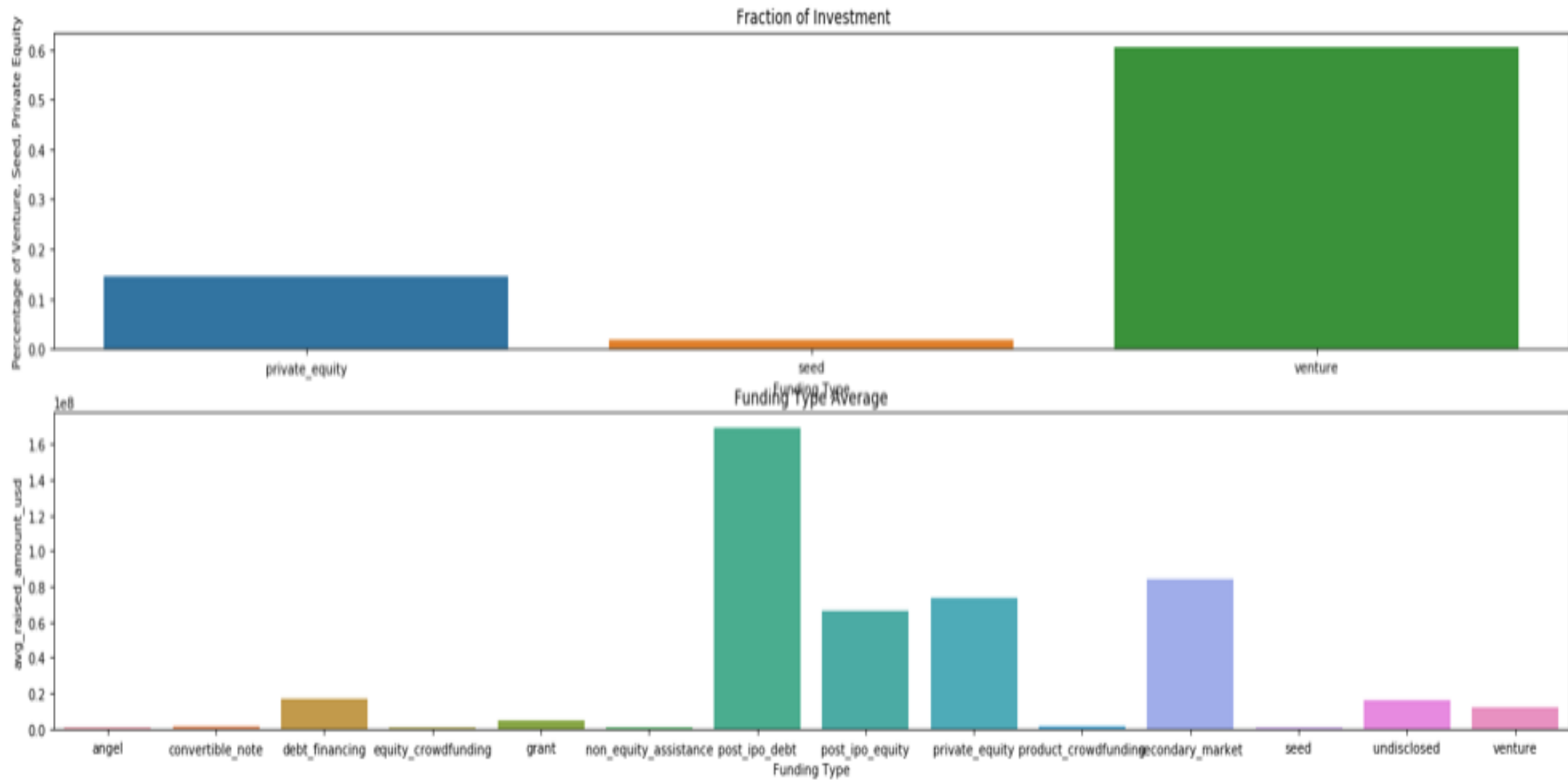
Steps followed to derive the results:-

- Load the data from companies.txt & rounds2.csv into companies & rounds2 data frame.
- permalink & company\_permalink were common & unique series in this 2 data frames but were in different cases.
- Converted permalink & company\_permalink to lower cases, both the series used to merge the data from companies & rounds2 to create master\_frame.
- On initial looking of master\_frame data, few columns were having missing values – started treating them:-
  - funding\_round\_code series has 72% missing values – dropped it from master\_frame.
  - raised\_amount\_usd series has 17% NaN values, filtered out those (because of huge Standard Deviation NaN were not replaced with mean).
  - category\_list which would be used to map to sectors (main category) having null values are also discarded.
  - rows where country\_code is missing & country can't be determined – filtered out.
- Using pivot table determined the funding round type (i.e. investment type) that is best suite for investments (using mean & median statistics to arrive at conclusion) – i.e. venture.

## <Analysis>

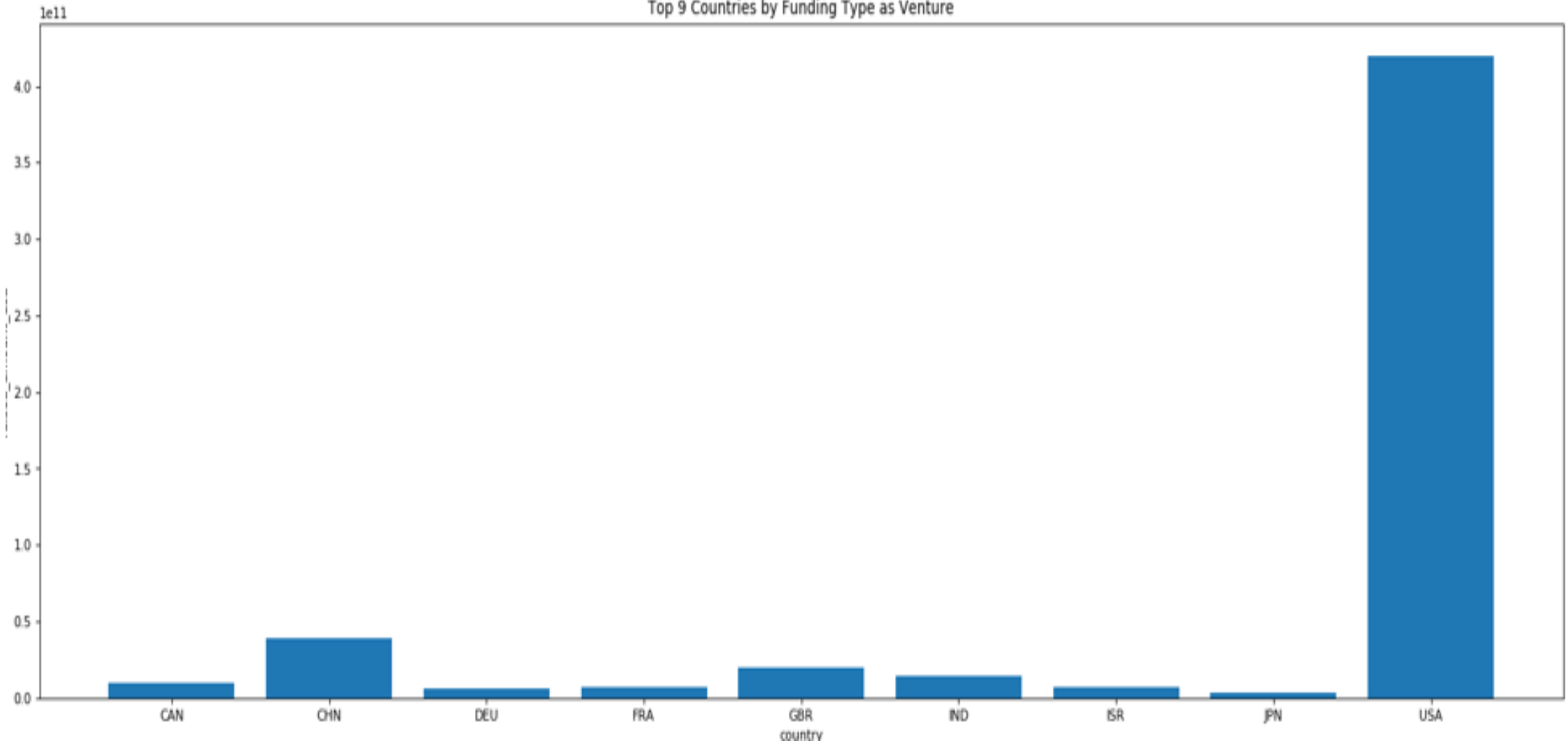
- Using pivot command & slicing determined top 9 countries for venture investment type.
- Using the web link figured out English Speaking countries.
- Filtered the data from master\_frame into 3 data frames namely d1, d2, d3 by country\_code (USA, GBR & IND), funding\_round\_type (venture) & investment range (5 – 15 Million USD).
- Using pivot & group by commands determined:-
  - Top 3 sectors by count wise.
  - Total no. of investment for venture funding type.
  - Top company for top 2 sectors with highest investment.
  - Total investment for each country.
- Plotted graph using matplotlib & seaborn libraries.

## <Results>



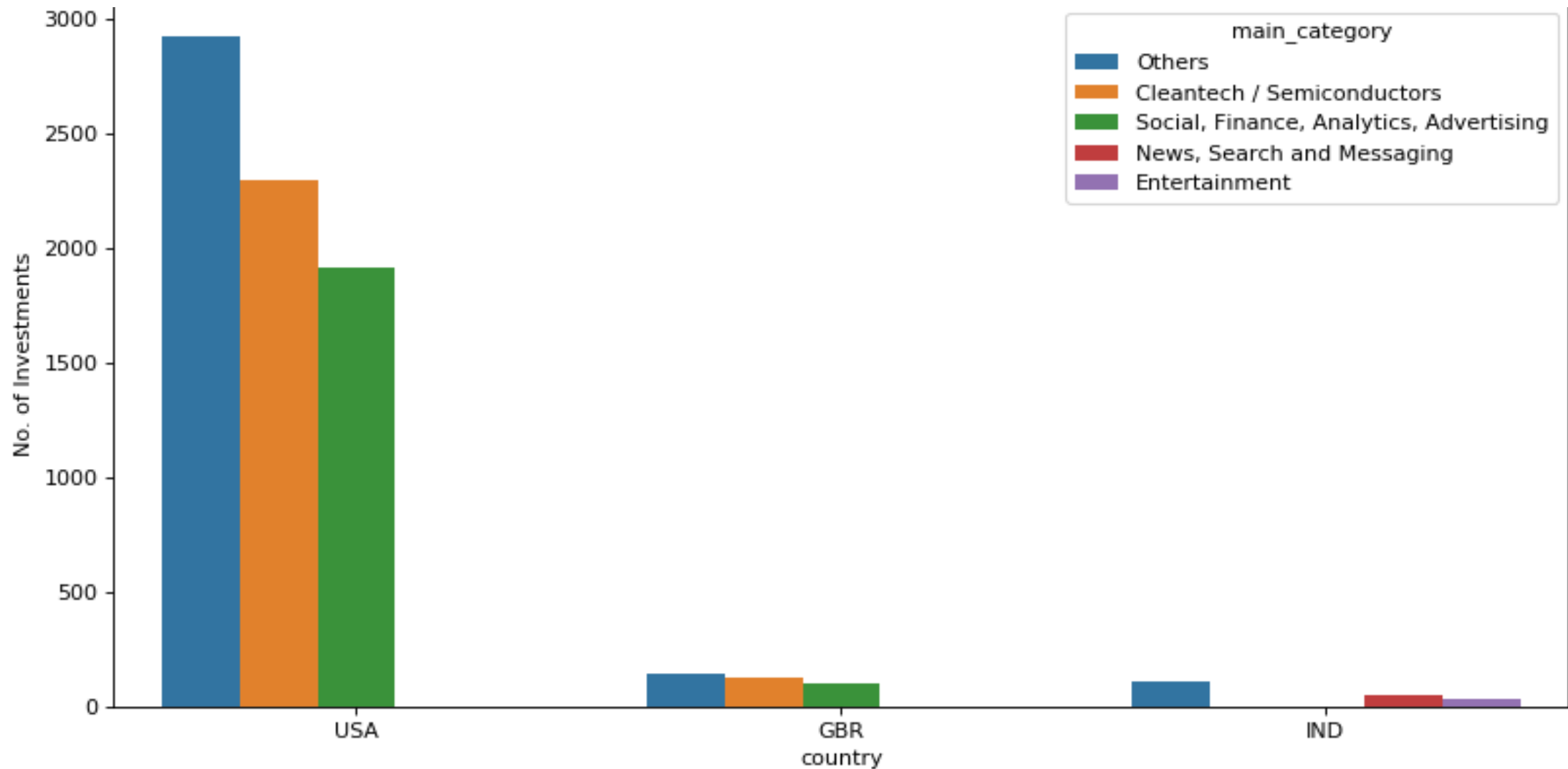
## <Results>

Top 9 Countries by Funding Type as Venture





## <Results>



## <Conclusions>

1. Funding type - Venture
2. Top 3 English speaking countries USA , GBR & IND.
3. Top 3 Sectors by count of investment:-
  - USA – Others, Clean Tech & Semiconductors, Social Finance Analytics & Advertisement
  - GBR – Others, Clean Tech & Semiconductors, Social Finance Analytics & Advertisement
  - IND – Others, News Search & Messaging, Entertainment
4. Highest investment in companies for top 2 sectors :-
  - USA – Virtual Stream, BioDesix
  - GBR – Electric Cloud, EUSA Pharma
  - IND – FirstCry.com, Gupshup
5. Total no. of investment:-
  - USA – 11149
  - GBR – 577
  - IND – 299