**The Evolution of India's Startup Ecosystem Funding: 2018 to 2021**

1.0 Introduction

India's startup ecosystem has witnessed unprecedented growth and innovation over the past few years, fueled by a surge in funding from various sources. From 2018 to 2021, India experienced significant changes in its investment landscape, resulting in a booming startup culture that attracted global attention.

This article provides insight into the factors that significantly impact the amount received by startups (factors such as Sector, Geographical location, What the company does and the Funding stage). This is to inform our team on what sectors to look at and tailor what they want to do to services that get the most funding. We also seek to help them find strategic location(s) to look at when starting up, and also what stage they should go seek funding. We leveraged the CRISP-DM (Cross Industry Standard Process for Data Mining) framework to provide a systematic approach for entrepreneurs interested in venturing into the Indian Startup Ecosystem.

CRISP-DM (Cross Industry Standard Process for Data Mining framework) consists of six major phases that are:

1. Business Understanding
2. Data Understanding
3. Data Preparation
4. Modeling
5. Evaluation
6. Deployment

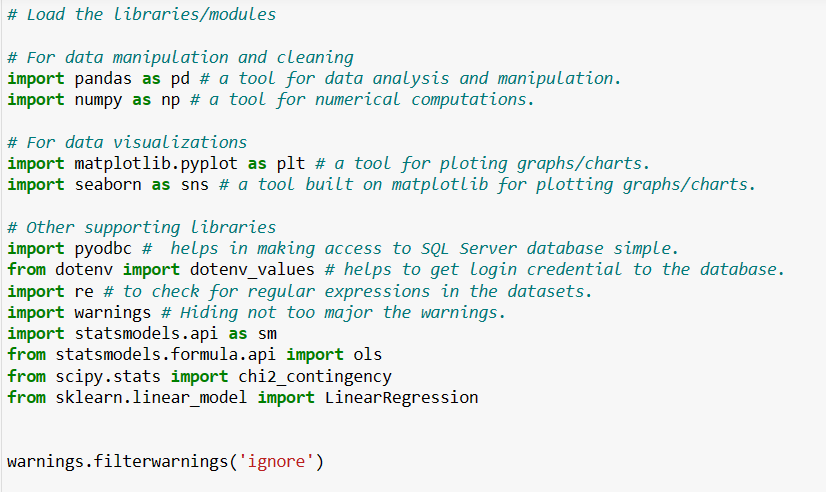
This project does not cover modeling.

## The Data

The data for the years 2018 and 2019 of funding were in separate CSV files. Datasets for 2020 and 2021 were accessed from a remote database. The four data sets were later merged after cleaning individual years separately. The four datasets had the following columns :

* Company/Brand
* Founded
* Sector
* What it does
* Founders
* Investor
* Amount ($)
* Stage
* Headquarters

# Import necessary libraries



DATA ASSESSMENT AFTER PREVIEWING DATASETS

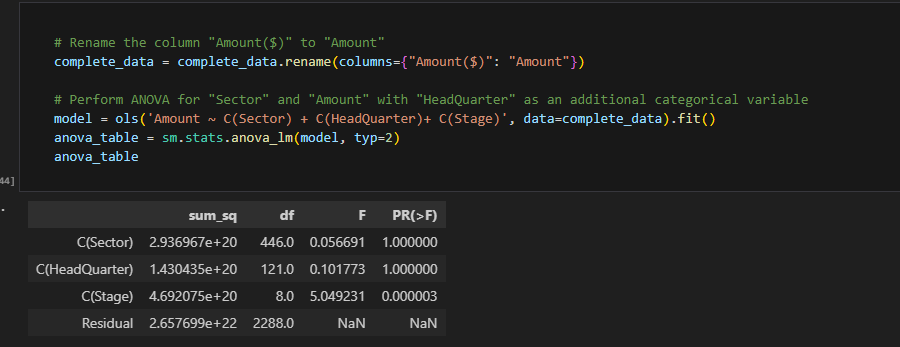
After gathering data from the sources mentioned above, we assessed each dataset for its quality issues. During our exploration of the datasets (EDA), we identified several data quality issues, which include:

* Inconsistent and missing columns: Some datasets have inconsistent column structures, with missing columns in certain cases.
* Missing columns and duplicates in the datasets: Specifically, the 2018 dataset is missing additional columns that are present in other datasets.
* Inconsistent values and currencies in the Amount column: The Amount column contains inconsistent values and different currencies.
* Inconsistent values in the Stage column: The Stage column exhibits inconsistent values across all datasets, which hampers uniform analysis.
* Missing values: Some datasets contain missing values, which need to be addressed for a complete and reliable analysis.
* Duplicated rows: Most of the datasets had information in certain rows duplicated

## 2.0 Hypothesis

Null Hypothesis(H0): There is no significant association between the funding amount and the other factors such as sector, location, and the funding stage.

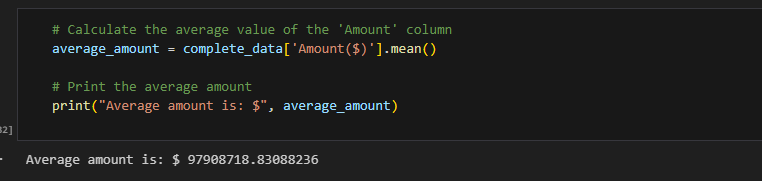
Alternate Hypothesis(H1): There is a significant association between funding amount and other factors such as sector, location, and funding stage

.

Based on the ANOVA results, we reject the null hypothesis (H0) for the "Stage" variable and conclude that there is a significant association between the funding amount and the "Stage" variable. This means that a Startup B in the more advanced "Series D" stage will receive a significantly higher amount of funding compared to Startup A in the early "Seed" stage. However, for the "Sector" and "HeadQuarter" variables, we fail to reject the null hypothesis, indicating that there is no significant association between the funding amount and these variables. Even though some sectors and companies in specific locations attracted more funding per the analysis, generally, companies in different sectors and locations get similar level of funds.

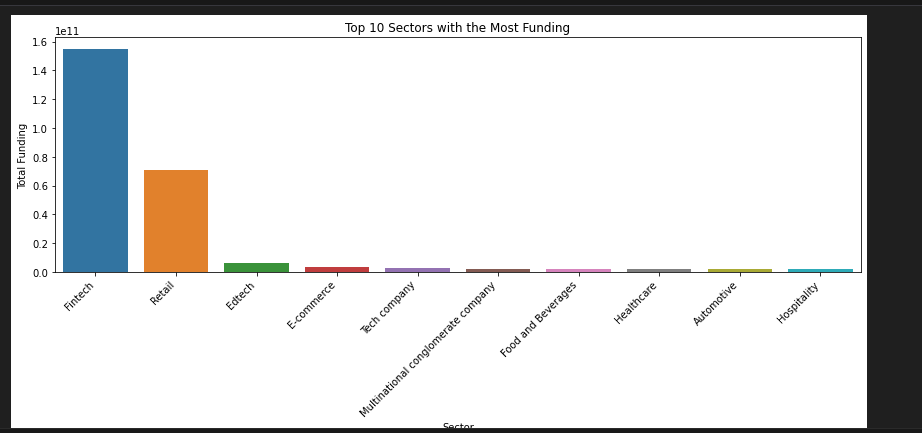
## 2.1 Questions

1. What is the average amount invested across all the years?



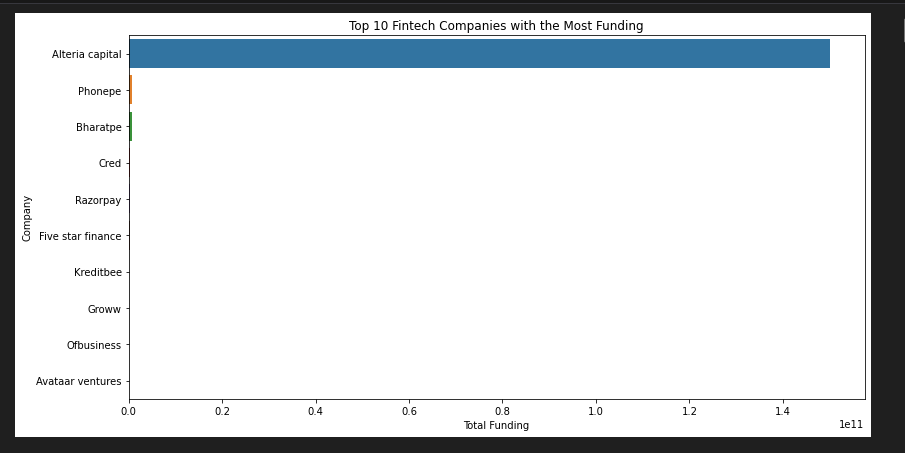
The calculated average funding amount of approximately $97,908,718.83 provides valuable information for stakeholders, aiding them in making informed decisions and gaining insights into the financial landscape of the companies represented in the dataset. For entrepreneurs and business owners, they can leverage this data to understand the typical funding expectations in their sector and plan their financial strategies accordingly.

1. What are the top10 sectors that got the most funding over the period?



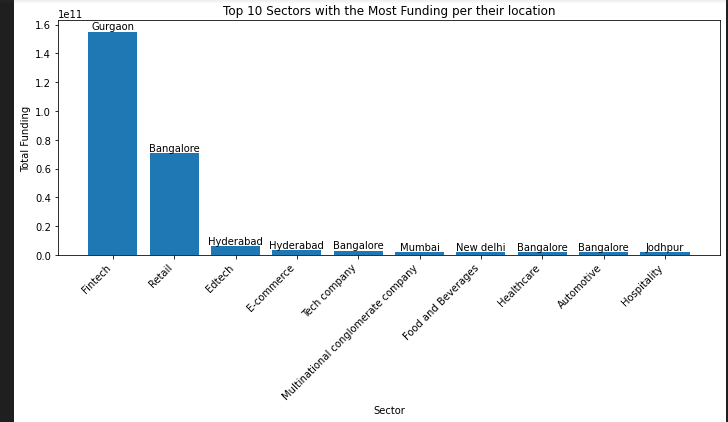
This graph reveals the diverse industries attracting substantial investments, with Fintech leading the pack, reflecting the ongoing digital transformation across sectors. By this, entrepreneurs and business owners seeking funding for their ventures can use this information to identify sectors that are currently attracting significant investments.

1. What top 10 companies under highest ranking sector got the most funding?



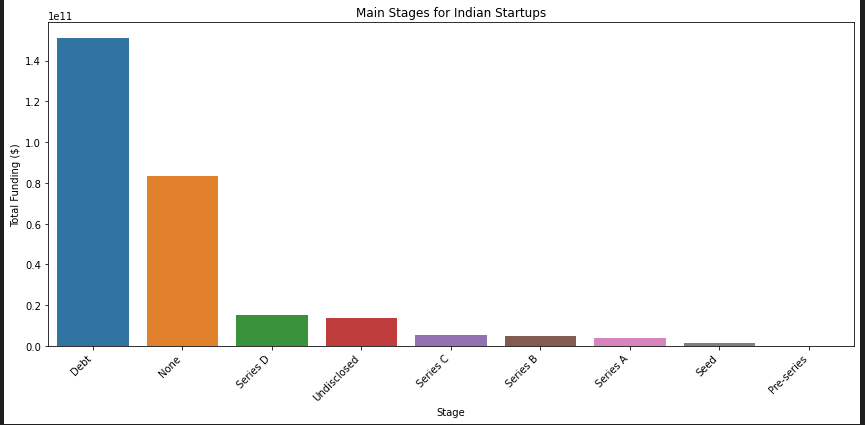
Here we see that Alteria capital, a Fintech company which is a venture debt firm received the most funding in the period of $150000000000 𝑤𝑖𝑡ℎ 𝐴𝑣𝑎𝑎𝑡𝑎𝑟 𝑣𝑒𝑛𝑡𝑢𝑟𝑒𝑠 𝑟𝑒𝑐𝑖𝑒𝑣𝑖𝑛𝑔 $100000000 which is the least amount received by the top 10 companies in the Fintech sector. Overall, this data empowers entrepreneurs in the Fintech sector to make informed decisions, fine-tune their fundraising efforts, and align their ventures with market demands. By leveraging the insights gained from analyzing the top-funded companies, entrepreneurs can position themselves for success, attract investment, and drive innovation in the dynamic world of Fintech.

1. Does location affect the funding amount given to the top 10 sectors?



From the graph above, companies Bangalore are most likely to receive funding irrespective of the the sector they operate in. The next city that realized funding was Hyderabad which had Edtech and E-commerce as the most funded in the sectors. However, Fintech received the most funding as a sector operated in Gurgaon city of India

1. Does the stage affect the amount of funding given?

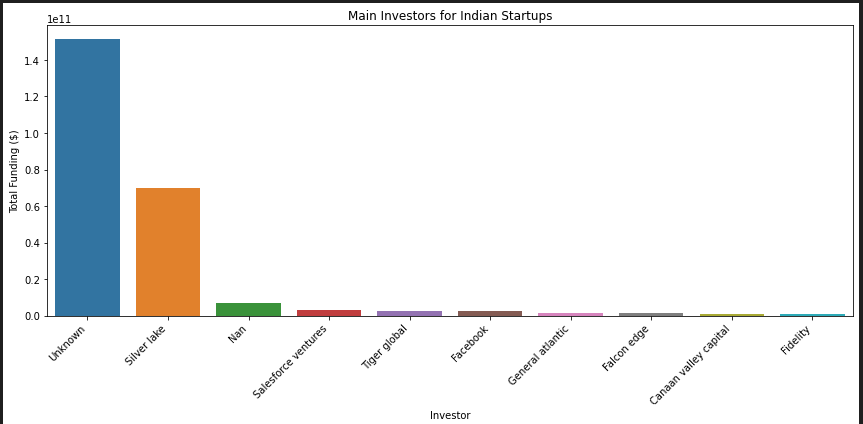


Funding Amounts by Investment Stage:

* Debt: The Debt stage stands out as the most well-funded investment category, with a substantial funding amount of $151,171,232,000. Debt funding refers to capital raised by companies through loans or other debt instruments, which they commit to repay with interest.
* None: Interestingly, there is a significant funding amount of $83,496,625,874 categorized under 'None.' This might represent cases where the stage of investment is not explicitly specified which is also seen because the data is incomplete.
* Series D: Among the venture capital stages, Series D funding ranks third, with a notable amount of $15,332,145,360. Series D funding typically occurs in mature companies looking to expand and further solidify their market position.
* Undisclosed: The 'Undisclosed' stage follows closely with $13,874,248,980 in funding. This category includes situations where companies do not disclose the specific investment stage publicly.
* Series C: Series C funding received $5,426,678,700, indicating continued investor interest in companies that have demonstrated growth and a scalable business model.
* Series B: Companies in the Series B stage secured $4,782,518,350, reflecting investors' confidence in these companies' potential for expansion and development.
* Series A: Series A funding amounted to $4,102,898,196, representing the initial significant round of funding for promising startups.
* Seed: The Seed stage, which marks the early funding phase for startups, accumulated $1,411,953,521 in investments.
* Pre-series: The Pre-series stage received the smallest funding amount of $29,000,000, as it represents the early fundraising phase before a formal series funding round.

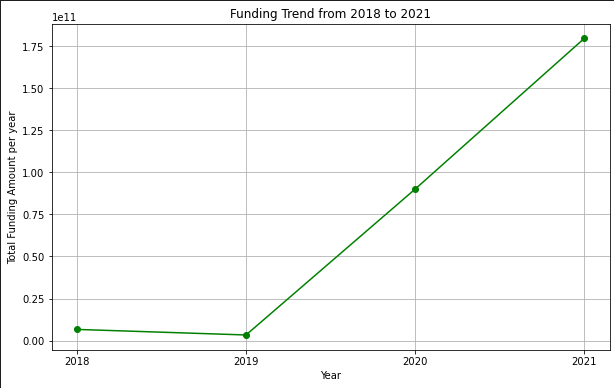
This analysis allows entrepreneurs to understand the funding patterns at different stages of growth. It helps in strategically planning their fundraising efforts and aligning funding needs with investors' preferences at specific stages.

1. Who are the top investors?



From the graph above, "Unknown" emerges as the top investor with a substantial funding contribution of around $151.34 billion, indicating strong interest sepite undisclosed identities.𝐹𝑜𝑙𝑙𝑜𝑤𝑖𝑛𝑔 closely is Silver Lake ,a prominent venture capital firm, with investment of $70 billion in Technology-driven companies. Additionally, the "Nan" category represents significant funding of approximately 6.6 billion drom undisclosed investors which could have resulted from improper cleaning of our data or incomplete data used for analysis. 𝑁𝑜𝑡𝑎𝑏𝑙𝑒𝑝𝑙𝑎𝑦𝑒𝑟𝑠𝑙𝑖𝑘𝑒"𝑆𝑎𝑙𝑒𝑠𝑓𝑜𝑟𝑐𝑒𝑉𝑒𝑛𝑡𝑢𝑟𝑒𝑠"𝑎𝑛𝑑"𝑇𝑖𝑔𝑒𝑟𝐺𝑙𝑜𝑏𝑎𝑙"𝑎𝑙𝑠𝑜𝑚𝑎𝑘𝑒𝑛𝑜𝑡𝑎𝑏𝑙𝑒𝑐𝑜𝑛𝑡𝑟𝑖𝑏𝑢𝑡𝑖𝑜𝑛𝑠,𝑖𝑛𝑣𝑒𝑠𝑡𝑖𝑛𝑔𝑎𝑝𝑝𝑟𝑜𝑥𝑖𝑚𝑎𝑡𝑒𝑙𝑦6.64 3.01 billion and $2.36 billion, respectively, in Indian startups, with potential impacts on innovation and market advancements.

1. What is the funding trend in the Indian startup ecosystem between 2018 to 2021?

The funding given to startups in the Indian ecosystem has shown a consistent upward trend over the years. There was a slight reduction in funding from 2018 to 2019, followed by a significant and steady increase from 2019 to 2021.

## Recommendations

1. Consider sectors like Fintech, Edtech and Retail when venturing into these ecosystems, however be mindful of the competition in such sectors.
2. Conduct thorough market research to identify potential investors interested in your sector and offerings as the main investors are unknown to the public.
3. Tailor fundraising efforts based on the investment stage. Debt seems to be the main source of funding for startups in the Indian business ecosystem. due diligence must be done to find out what stage suits the growth phase.
4. Companies in cities such as Gurgaon, Bangalore, Hyderabad,Mumbai, NewDelhi and Jodhbur received quite significant funding within the period. In these cities except for Gurgaon, getting funded irrespective of Sector is highly attainable.

## Conclusion

The analysis of funding patterns in the Indian startup ecosystem provides valuable insights for entrepreneurs and investors alike. The data indicates a positive trend of increasing funding over the years, with specific investors who are Unknown and Silver Lake playing significant roles in supporting startups. However, there are still challenges represented by the "None" and "Nan" categories, which may require further investigation and data improvement efforts.

## References

<https://www.ibm.com/docs/en/spss-modeler/saas?topic=dm-crisp-help-overview>

2022 EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal | [www.eprajournals.com](http://www.eprajournals.com/) | Journal DOI URL: <https://doi.org/10.36713/epra2013>

<https://www.neusourcestartup.com/startupindia/startup-problems-india-facing>

<https://www.theofficepass.com/toppings/common-problems-entrepreneurs-face-starting-startup-india.html>