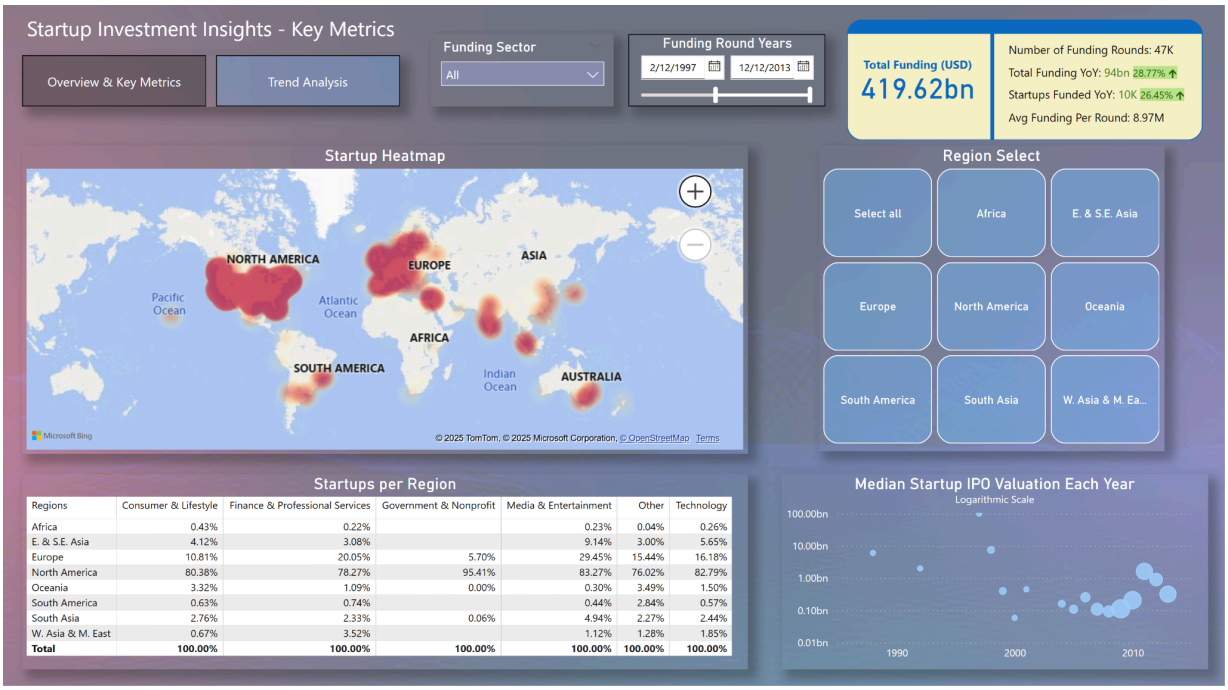
Venture Open Data Dashboard

Overview

This dashboard examines and provides insights on both Venture Capital Investors and Startups with data taken from the [Startup Investment Dataset] via Kaggle.



Startups per Region

Regions	Consumer & Lifestyle	Finance & Professional Services	Government & Nonprofit	Media & Entertainment	Other	Technology
Africa	0.43%	0.22%		0.23%	0.04%	0.26%
E. & S.E. Asia	4.12%	3.08%		9.14%	3.00%	5.65%
Europe	10.81%	20.05%	5.70%	29.45%	15.44%	16.18%
North America	80.38%	78.27%	95.41%	83.27%	76.02%	82.79%
Oceania	3.32%	1.09%	0.00%	0.30%	3.49%	1.50%
South America	0.63%	0.74%		0.44%	2.84%	0.57%
South Asia	2.76%	2.33%	0.06%	4.94%	2.27%	2.44%
W. Asia & M. East	0.67%	3.52%		1.12%	1.28%	1.85%
Total	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**

Business Questions

The dashboard can be used to answer the following questions:

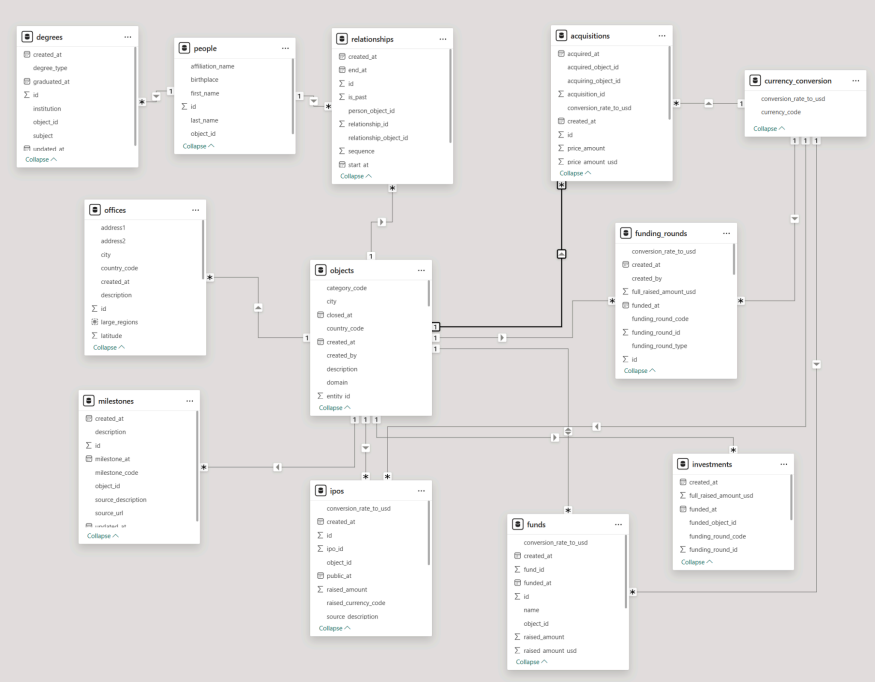
- How much has been invested in Startups? How has this changed over time?
- Where are Startups primarily located and for which sector(s)?
- What is the expected valuation of a Startup that successfully completes an IPO? How has this changed over time?
- What are the differences in investments and success between specific sectors?
- Which investors have been the most active within specific sectors?
- How has Startup success changed, in terms of acquisitions, given increasing funding?

Data Source and Preparation

The dataset, as mentioned previously, is the [Startup Investment Dataset] via Kaggle. The dataset consists of 11 different tables providing information about names of companies from 1960 to 2013, Venture Capital Funds and Investors, people associated with specific Startups, products related to Startups, along with information regarding funding rounds and how much

money a startup raised over time, whether they reached IPO and the valuation of the company when they did, and if they were acquired, among other statistics.

A star schema design was created to establish appropriate links between the tables to draw information regarding Startups to answer the above questions. An additional table named `currency_conversion` was created in order to assist with converting all values to the same base currency (more information regarding this later). Below is the overview of the schema:



The `objects` table was used as the main fact table considering it had individual records for each company, person, financial organization, and product linked to a single column from which all other tables could link to in one form or another. All tables save for `people`, and `degrees` were directly connected to `objects`.

Data Cleaning and Calculated Fields

All tables were loaded into Power BI where reviewing, analyzing, and transformations of all the data took place. From reviewing each table and how they related to each other, null values were removed from all used tables as no real imputations could be derived from any of them after inspection. A new table, `currency_conversion` was created based on currency codes within dataset containing conversion rates to USD based on historical 2010 figures (oldest reliable figures found) to handle converting raised amount columns within `acquisitions`, `funding_rounds`, `funds`, and `ipos` tables to allow for 1 to 1 comparisons.

Apart from that, the following Measures and columns were created to aid with creating visualizations:

Measures

- `Average Funding Per Round` - simple average of total funding
- `Total Funding` - simple sum of overall funding
- `Prior Funding` - sum of all previous funding to the current selected year
- `YoY` - total funding subtracted by prior funding
- `YoY Percent` - formula for using YoY as a target and showing the percent change while also displaying an arrow dependent on direction of change (positive/negative)
- `Startups Prior Funded`, `Startups Funded YoY`, `Startups YoY Percent` are the same but for the count of Startups within `funding_rounds`
- `Fund Name`, `IPO Company Name`, `Company Name` are within `funds`, `ipos`, and `offices` respectively to tie `objects[name]` correctly to the respective tables
- `Company Count` within `objects` acts as a quick filter for selecting “Company” from `objects[entity_type]`
- `Funding Per Category` allows for linking funding totals from `funding_rounds` to the category types in `objects`
- `Sectors` creates higher level groupings for values in `objects[category_code]`

Added Columns

- `offices[valid_location]` - flag created for displaying only values with valid latitude/longitude values
- `offices[large_regions]` - column of manual buckets for country codes based on global regions

Dashboard Layout

The dashboard consists of two pages, each consisting of a header area (Title, page navigation, slicers, and a KPI card) and the main chart areas.

KPI Card

The KPI card consists of five different KPIs, all of which are updated based on the timeframe:

- ***Total Funding (USD)*** - total funding for Startups
- ***Number of Funding Rounds*** - total number of funding rounds
- ***Total Funding YoY*** - most recent year of funding and the percent change from the previous year
- ***Startups Funded YoY*** - same as previous but for the count of Startups funded
- ***Avg Funding Per Round*** - displaying the average funding per round(s) selected

Slicers

There are two global slicers that affect the KPI card and visualizations on both pages and one localized slicer to the map graph on the first page:

- ***Funding Sector*** - a dropdown to filter between different sectors of Startups

- ***Funding Round Years*** - a slider to enable filtering between different timeframes for showing changes over time in the data
- ***Region Select*** - a card select filter for honing in on a particular section of the map automatically

Main Charts

The main charts are on two separate pages, Overview & Key Metrics and Trend Analysis:

Overview & Key Metrics

- ***Startup Heatmap*** - a Map visualization employing geolocation data from `offices` with the heatmap functionality enabled for displaying concentrations of startups globally. Both the *Region Select* and *Funding Round Years* slicers can update the data displayed. The map can be manually zoomed into as well
- ***Startups per Region*** - a matrix for displaying percentage of startups per region based on column totals. Shows the distribution of startups globally. Functions with the *Funding Round Years* slicer and is meant to give a deeper picture of the information displayed on the heatmap
- ***Median Startup IPO Valuation Each Year*** - a scatterplot for displaying the information described in the title. The y-axis values are the valuation totals within the data displayed on a logarithmic scale. This was done to normalize the data and present a more meaningful metric of evaluation without removing the extreme values. It shows a general upward trend of valuation for Startups that reach an IPO. The size of the points is determined based on the count of startups that reached an IPO in each year. It is not filterable by slicers

Trend Analysis

- ***Top 5 Investors by Total Funding*** - a horizontal bar chart of the top 5 Venture Capital Investors in terms of funding for a particular sector. The default is for all sectors, however the *Funding Sector* slicer can update and change the financial organizations displayed. It is not filterable by the year slicer
- ***Top 5 Investors per Startups Funded*** - a line graph displaying the same information as the previous chart but in terms of the number of Startups they have funded over time. It is also filterable by *Funding Sector* but not by *Funding Round Years*. A non-adjustable filter was placed on the date range, selecting the year 2000 as the base given the lack of data available in years prior
- ***Startup Acquisitions Over Time*** - a stacked column chart showing the number of Startups that were acquired over time along with showing the distribution of the sector of the startups acquired. The previous chart's stipulations also apply here
- ***Total Funding by Sector*** - a stacked area chart that displays the overall total funding raised by startups over time, broken down by sector. This chart is fully able to use both the available global slicers

Analysis and Conclusion

From reviewing, analyzing, and visualizing the data, it can be seen that the Startup world both started and continues to be focused mainly on the technology sector.This can be seen from viewing the sector breakdowns in the various charts within the dashboard. This would make sense, given how technology is ever evolving and how it can be applied to many different forms of life. This is also corroborated by the comparative increase in funding for Startups specializing in technology as it takes up the main share of all funding within the data. Given that this data runs to 2013, the trends seen within the data can be confirmed as correct when compared to how technologically focused and driven society has become in 2025.

Apart from this, however, the data shows that while there was a significant increase in Startups founded and funded, the rate of funding started to slow down towards the end of the timeframe within the dataset. Given the context of current events, one possible cause for this could have been the subprime mortgage crisis within the United States in 2007-2008 leading to a significant recession. Seeing as how the vast majority of Startups are within the United States, given the economic downturn it would stand to reason that this would have a negative effect on the willingness of Venture Capitalists to invest in Startups. Looking at the last year available within the data, 2012-2013, a significant increase in year over year funding can be seen in the data, which would coincide with the upturn of the economy in the United States, giving credence to this insight.

Overall, the data confirms a general increase in both the number of and funding for Startups, but not necessarily an increase in the overall success rate of companies reaching the IPO stage. The increased competitiveness of the Startup world paired with the selectiveness of Venture Capital Funds to invest shows that it is as tough as ever to reach success. What can be shown, as previously stated, is that your best bet is likely to employ technology in a creative way given the overall proclivity of investment towards that sector.