Visualizing and analyzing investment opportunities for venture capitalists in the technology industry

Moreen Owino

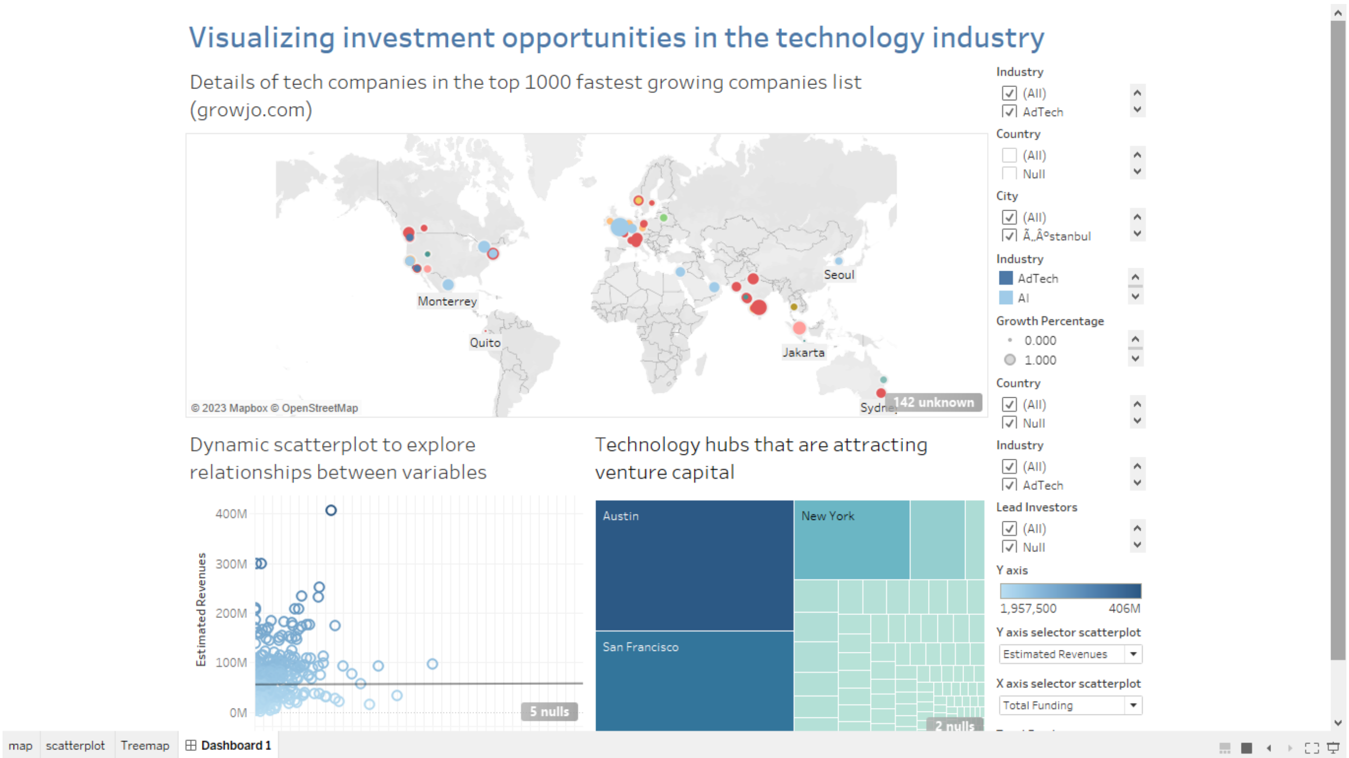

Fig. 1. Dashboard: Visualizing investment opportunities in the technology industry

Abstract— The venture capital landscape is dynamic, with constant shifts in technology hubs, valuations, and growth rates. This project aims to provide venture capitalists with a powerful tool for decision-making by analysing and visualizing data from the top 1000 fastest-growing technology companies. Leveraging maps, scatter plots, and tree maps, the paper addresses key questions related to industry trends, geographical concentrations, and other factors influencing investment success.

Index Terms—Venture capital, Technology companies, Valuation, Technology hubs, Top 1000 companies and startups, Industry

1 Introduction

Venture capitalists (VCs) are professional investors who manage pooled funds from various sources, such as high-net-worth individuals, institutional investors, and corporations. These funds are then strategically invested in promising startup and early-stage companies with high growth potential. Venture capitalists provide not only capital but also expertise, mentorship, and industry connections to assist these startups in achieving their growth objectives. In return, venture capitalists expect a significant return on their investment, typically through an eventual exit event such as an initial public offering (IPO) or acquisition.

As stewards of capital, venture capitalists are confronted with the formidable task of discerning promising investment opportunities amidst a landscape of rapid technological advancements. Venture capitalists employ a systematic decision-making process when evaluating potential investment opportunities, particularly in the dynamic landscape of the technology industry, Gompers et al. [2].

• *Moreen Owino is a Student at GVSU. E-mail: owinom@mail.gvsu.edu*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org.
Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx/.

The decision-making process typically involves several key stages like deal Sourcing. This is done by relying on their extensive networks and relationships within the industry to source potential investment opportunities. Another stage is due diligence where VCs conduct thorough market analysis to assess the potential size, growth, and dynamics of the target market. They evaluate the competitive landscape and identify the unique value proposition of the startup.

They also evaluate the experience, expertise, and track record of the founders and key executives and scrutinize the financial health and viability of the startup, looking at revenue projections, burn rates, and scalability. Kaplan, Sensoy, and Stromberg (2002) [3] VCs also use various screening platforms, databases, and industry reports to identify startups that align with their investment criteria and this paper uses such industry reports to try and glean insights that can aid them in making their decisions.

Using data procured from Growjo.com, a website renowned for curating information on the fastest-growing companies, this paper endeavours to explore the intersection of venture capital investment decision-making and data-driven visualizations, with a specific focus on aiding venture capitalists in navigating the complex landscape of the technology industry. These tableau visualizations, including maps, scatter plots, and tree maps aim to address critical questions, providing venture capitalists with an unprecedented toolset for deciphering trends, correlations, and growth patterns within the technological investment landscape. By immersing venture capitalists in a visually driven exploration of technology companies in the top 1000 fastest growing companies, the proposed methodology seeks to empower decision-makers to make more informed choices. This paper not only contributes to the growing dialogue on the intersection of venture capital decisions and data visualization but also seeks to pave the way for a paradigm shift in how investors leverage technology and visualization tools for strategic decision-making.

2 Literature Review

The theoretical underpinnings of this research draw inspiration from various domains, including venture capital decision-making, data visualization, and business intelligence. Venture capitalists play a pivotal role in fostering innovation and economic growth by strategically allocating capital to high-potential startups. The decision-making process of venture capitalists involves a thorough analysis of various factors, including location, valuation, growth rate, revenue, and funding among other factors. A comprehensive literature review sheds light on the significance of these factors in shaping investment choices. Geographic proximity has been recognized as a critical factor influencing venture capital investment decisions. Sorenson and Stuart (2001) [1] conducted a seminal study on syndication networks and the spatial distribution of venture capital investments. Their findings suggest that the physical location of a startup plays a crucial role in attracting venture capital funding. Proximity to established venture capital hubs and innovation ecosystems enhances a company's visibility and access to funding opportunities.

The valuation and funding levels of a startup are fundamental considerations for venture capitalists. Gompers and Lerner (1998) [2] delved into the drivers of venture capital fundraising, emphasizing the importance of valuation and funding in attracting investors. Startups with favorable valuation metrics and robust funding profiles are more likely to secure the attention and support of venture capitalists, highlighting the financial health of a company as a key determinant.

In the pursuit of identifying promising investment opportunities, venture capitalists closely examine growth rates and revenue projections. Kaplan, Sensoy, and Stromberg (2002) [3] investigated the evolution of firms from early business plans to public companies. Their study revealed that growth rate and revenue projections provided in business plans significantly influence the success of venture-backed companies. Projections demonstrating strong growth potential and sustainable revenue streams are crucial factors shaping investment decisions. For a comprehensive understanding of venture capital networks and investment performance, Chemmanur and Chen (2014) [4] offer valuable insights. Their research explores various factors, including location, valuation, and funding, in assessing the success of venture capital investments. The study provides a holistic view of the intricate network dynamics that influence investment outcomes.

Most of the papers that exist on venture capitalists and how to make investment decisions do not involve the use of visualizations but are rather based on theory and findings from data analysis. Example of papers that touched on visualization and investment decisions include Sorenson and Brath (2013), IEEE 2013 that presented a financial visualization case study involving the correlation of financial timeseries and discrete events to support investment decisions and Aigner (2013), IEEE 2013 who discussed current work practices and users' perspectives on visualization and interactivity in business intelligence. This paper therefore tries to fill that gap by providing an example of how visualizations can be used to aid venture capitalists in making their investment decisions.

3 Data

The Top_1000_Companies data set used in this project was obtained from Kaggle.com but originated from Growjo.com (www.growjo.com). Growjo.com focuses on systematically identifying and ranking the top 1000 fastest-growing companies and startups and is consistently updated and publicly accessible. The information contained within the dataset covers a wide array of critical variables for venture capitalists making investment decisions. The dataset encompasses key company details such as names, locations (including city and country), and industry types, ranging from biotech and marketing tech to AI and devops. Financial metrics, including valuation, funding levels, growth rates, and revenue, offer a comprehensive understanding of each company's financial standing. Furthermore, investment-related details, such as lead investors, funding rounds, and the respective years of funding, provide insights into the companies' funding histories. The transparency and accessibility of the dataset contribute to the credibility of the findings presented in this paper, offering venture capitalists valuable insights into potential investment opportunities within the rapidly evolving technology sector.

The Kaggle data set can be accessed using this link(https://www.kaggle.com/datasets/amritpal24/top-1000-companies-details).

This project aims to leverage this rich dataset to answer pertinent questions for venture capitalists, offering insights into the technology industry's landscape and aiding in the identification of promising investment opportunities.

4 Methodology

From the research carried out on previous papers regarding what factors were important to venture capitalists when it came to making investment decisions in the technology industry, a proposal for the project was created. It contained a list of nine questions to be answered using the data described above. This list contained questions like; Which technology hubs are attracting venture capital? Which technology companies have the highest valuations, highest funding, highest growth rate etc.? Are there any correlations between valuation, size, and revenue in the technology sector? What is the current distribution of tech companies by type (e.g., biotech, marketing tech, AI, devops)? Is there a correlation between the location (city/country) of a tech company and its growth rate or valuation? Do companies with lead investors from well-known venture capital firms tend to have higher valuations or growth rates? Can we identify any leading indicators of future success or decline in tech companies based on the data? Which companies have consistently ranked high? What are the growth trends of technology companies over different years of funding?

To answer these questions three visualizations were created in tableau using the data set. The three visualizations created were a map, scatter plot and a tree map. The visualizations, how they were created and the questions they helped answer are discussed below.

5 Discussion

A project design document was created that detailed each visualization, its encodings, and the specific questions it would answer. This made it easier to implement the visualizations in tableau since the overall design and structure had already been identified. The key was to ensure that the encodings used would enable analysis and generation of insight for the questions previously listed.

5.1 Data Preparation

The data set was filtered using the industry column and only companies in tech were selected i.e., Tech services, Support/crm tech, Software, Sales tech, Robotics, Real estate tech, Program dev, Martech, Legaltech, Internet, Insurance tech, Information technology and services, HRTech, Fintech, Edtech, Ecommerce tech, Devops, Delivery tech, Cybersecurity, Business intelligence, Bio tech, Baby tech, AI and AdTech. Some data cleaning was also done to ensure correct data formats.

5.2 Map

5.2.1 Questions to be answered by the Map

(1) Which technology hubs are attracting venture capital?
(2) Is there a correlation between the location (city/country) of a tech company and its growth rate or valuation?
(3) What is the current distribution of tech companies by type (e.g., biotech, marketing tech, AI, devops)?
(4) Which technology companies have the highest valuations, highest funding, highest growth rate, etc.?

5.2.2 Rationale for using the Map

A map can be used to give an overview of the data and show the concentration of tech companies in different cities or regions and the distribution of the tech companies by industry type. This will help identify tech hubs that attract venture capital. Filters and other encodings can be used to gain more detailed information from the data to answer the questions.

One can also use color and size encoding like in this case to represent the industry and growth rate and add tool tip information as well as filters to enable the venture capitalists identify companies that meet the criteria that they are interested in.

Fig. 2. Map: Dynamic map with filters showing the geographical location of the various top tech companies.

5.2.3 Visual encodings and the rationale behind them

I. Data points: Company

II. Color Encoding for data points: Industry
Rationale: Color helps viewers quickly distinguish between different values and identify trends or patterns.

III. Size Encoding for data points: Growth percentage-dimension
Used the size of data points (e.g., circles) to represent the growth percentage. Larger circles represent bigger/higher growth.
Rationale: Size encoding helps in understanding the scale of companies' growth at a glance.

IV. Label: City
This will enable the user to identify the technology hubs that have lots of tech companies

V. **Details:** Country, City, Company name
This will provide information on the country, city and company name, valuation

VI. Tooltip Information:
Tooltips to display detailed information about each data point. Include company names, valuations, growth rates, country, city, total funding.
Rationale: Tooltips provide additional context and specific data points, allowing users to explore the details of each company.

VII. **Filters:** Industry, Country, City, Growth percentage, Total funding, Valuation.
Rationale: Filters will allow the venture capitalists to filter based on the specific criteria that they are interested in to ease their decision-making process.

5.3 Scatter Plot

5.3.1 Questions to be answered by the scatter plot

The following were the questions to be answered using the scatter plot.

(1) Are there any correlations between valuation, size, revenue etc. in the technology sector?
(2) Do companies with lead investors from well-known venture capital firms tend to have higher valuations or growth rates?
(3) Can we identify any leading indicators of future success or decline in tech companies based on the data?
(4) Which technology companies have the highest valuations, highest funding, highest growth rate, etc.?
(5) What are the growth trends of technology companies over different years of funding?

5.3.2 Rationale for using the scatter plot

Scatter plots are ideal for exploring correlations between two quantitative variables and help identify patterns and trends in your data. One can create scatter plots to compare valuation and revenue or valuation and size, helping the user determine if there's a positive or negative relationship between these factors. With filters for lead investors one can analyze if there's a correlation between having lead investors from well-known firms and valuations or growth rates etc. It can also help the users identify the companies with the highest valuations and funding by plotting these values on the axes. The spread and clustering of data points in the scatter plot helps one quickly identify the companies that excel in these specific metrics.

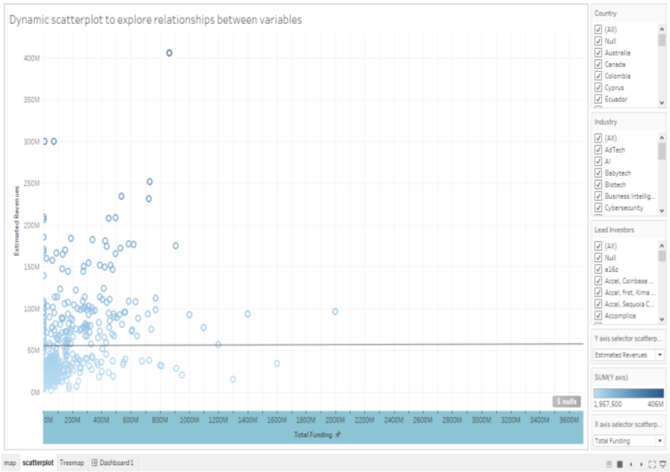

Fig. 3. Scatter plot: Dynamic scatter plot with filters showing the relationship between total funding and estimated revenues. In this case there seems to be no correlation between the two.

It can also be used to search for indicators or characteristics that are associated with future success or decline. The scatter plot's visual nature makes it easier to spot potential leading indicators.

5.3.3 Visual encodings and the rationale behind them

I. **X and Y Axes:** For the X and Y axes parameters were created and calculated fields that allow the user to choose which variables they want on the x and y axis so that they can explore the relationships between different pairs of variables on the scatter plot. There is a visible filter for them to choose the x and y variables. So, the venture capitalist can choose x axis variables such as total funding, valuation, and estimated revenues. For y axis variables they can choose variables such as growth percentage, employees, founded year, total funding, valuation, and estimated revenues**.** The X and Y axes represent the variables one is comparing, making it easy to visualize relationships.

II. **Marker Encoding:** Used markers (points) to represent individual tech companies and color to distinguish between different industry types. This helps to convey additional information about the companies.

III. **Tooltip Information:** tooltips to display detailed information about each data point, including company names, valuations, Total funding, industry, lead investor and other relevant data. This helps to provide specific details when users hover over data points.

IV. **Reference Lines and Trend Lines:** Added trend line to the scatter plot to help users quickly identify patterns or relationships in the scatter plot.

V. **Filter or Parameter Controls:** Implemented filters or parameter controls to allow users to select specific criteria (e.g., company industry type, country, x and y axis variables, lead investors), which will dynamically update the scatter plot. This interactivity allows users to explore the data based on their specific variable interests or questions.

5.4 Tree Map

5.4.1 Questions to be answered by the Tree Map

(1) Which technology hubs are attracting venture capital?

5.4.2 Rationale for using the Tree Map

A tree map is a suitable choice for visualizing the popularity of different technology hubs based on venture capital investments because tree maps excel in representing Part-to-Whole Relationship. In this case, the "whole" is the total venture capital investment, and the "parts" are the individual tech hubs. The users can easily see how each hub contributes to the total. It is also effective for displaying hierarchical data, where the hierarchy is represented by the top-level categories (tech hubs) and their subdivisions (e.g., companies or venture capital investment within each hub). Other benefits are space efficiency, color encoding for differentiation between groups and interactivity by allowing users to click and get more information about the hubs.

5.4.3 Visual encodings and the rationale behind them

I. **Size (Area):** The size of each rectangle represents the total venture capital invested in companies within each technology hub. Larger rectangles indicate higher levels of investment. This allows users to visually compare and assess the importance /size of each hub relative to the others.

II. **Color**: Color intensity is used to represent the popularity of each hub or to distinguish different hubs from each other. Lighter or darker colors can indicate low or high levels of venture capital investment. This helps users visually identify the hubs within the tree map quickly. Color also adds visual appeal to the tree map, making it more engaging and user-friendly.

III. **Label**: The label encoding displays the name of each technology hub, providing users with a clear identifier for each rectangle. This helps users understand which hub each rectangle represents. This labeling ensures that the audience can navigate the tree map with clarity.

IV. **Tooltip**: Tooltips provide interactive information when users hover over a specific rectangle. One can include details such as the exact amount of venture capital investment for the hub, number of companies in the hub etc., making the visualization informative and user-friendly. The tooltip rationale is that it provides an interactive element by offering additional information on demand, allowing more in-depth exploration of the data by users

V. **Filters**: City, Country
Rationale: allows the venture capitalist to view specific cities countries within the tree map.

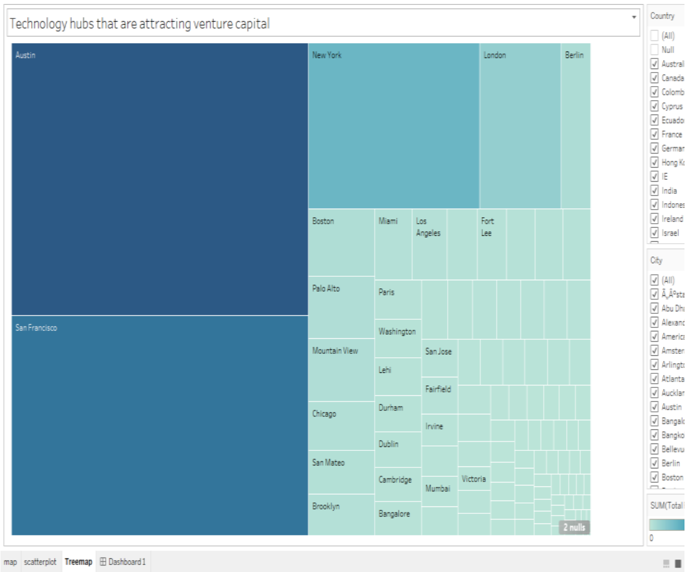

Fig. 4. Tree Map: Dynamic tree map with filters comparing the technology hubs based on the amount of funding or venture capital they are attracting.

6 Conclusion

In the ever-evolving landscape of the technology industry, making informed investment decisions is paramount for venture capitalists seeking to capitalize on emerging opportunities. Leveraging the power of data visualization tools, such as Tableau, can provide a comprehensive understanding of the market dynamics and assist venture capitalists in navigating the complexities of investment choices. Through the analysis of data sourced from Growjo.com on the fastest-growing tech companies, the paper aimed to address crucial questions and provide actionable insights for venture capitalists.

The exploration began with a map visualization, shedding light on the geographical distribution of tech companies and uncovering potential technology hubs that attract venture capital. By incorporating color and size encodings, it facilitated the identification of industry types and growth rates, enabling venture capitalists to discern patterns and trends at a glance. The map serves as a powerful tool for assessing the concentration of tech companies in specific regions, aiding in strategic decision-making regarding investment destinations.

The scatter plot further delved into intricate relationships between various factors such as valuation, funding, and growth rates. By allowing users to dynamically choose variables for the X and Y axes, we provided flexibility in exploring correlations that are vital for investment decisions. The scatter plot's visual nature aids in identifying outliers, understanding patterns, and assessing the impact of lead investors from well-known venture capital firms on valuations and growth rates.

Finally, the tree map provided a hierarchical view of technology hubs based on venture capital investment, offering insights into the popularity and significance of each hub. The size and color encodings facilitated a quick assessment of the relative importance of hubs, allowing venture capitalists to prioritize their focus based on investment potential.

In conclusion, the visualizations not only answered the specific research questions but also equips venture capitalists with a dynamic toolkit to navigate the complexities of the technology investment landscape. The interactivity and depth provided by these visualizations empower users (venture capitalists) to make informed decisions, identify emerging trends, and seize opportunities in the fast-paced and competitive world of technology investments. As technology continues to reshape industries, the integration of data visualization tools becomes indispensable for venture capitalists aiming to stay ahead in the ever-changing investment landscape. The paper therefore shows that visualizations can be used to aid venture capitalists in their investment decision process.

Acknowledgments

The author wishes to thank Professor Jonathan Leidig for providing instructional guidelines for the paper.

References

[1] Sorenson, O., & Stuart, T. E. (2001). Syndication Networks and the Spatial Distribution of Venture Capital Investments. American Journal of Sociology, 106(6), 1546–1588.

[2] Gompers, P., & Lerner, J. (1998). What Drives Venture Capital Fundraising? Brookings Papers on Economic Activity, 2, 149–192.

[3] Kaplan, S. N., Sensoy, B. A., & Stromberg, P. (2002). Should Investors Bet on the Jockey or the Horse? Evidence from the Evolution of Firms from Early Business Plans to Public Companies. Journal of Finance, 57(3), 1051–1078

[4] Chemmanur, T. J., & Chen, Z. (2014). Venture Capital Networks and Investment Performance. Journal of Financial and Quantitative Analysis, 49(1), 103–135.

[5] W. Aigner, "Current Work Practice and Users' Perspectives on Visualization and Interactivity in Business Intelligence," *2013 17th International Conference on Information Visualisation*, London, UK, 2013, pp. 299-306, doi: 10.1109/IV.2013.38.

[6] E. Sorenson and R. Brath, "Financial Visualization Case Study: Correlating Financial Timeseries and Discrete Events to Support Investment Decisions," *2013 17th International Conference on Information Visualisation*, London, UK, 2013, pp. 232-238, doi: 10.1109/IV.2013.31.

[7] E. Sorenson and R. Brath, "Financial Visualization Case Study: Correlating Financial Timeseries and Discrete Events to Support Investment Decisions," 2013 17th International Conference on Information Visualisation, London, UK, 2013, pp. 232-238, doi: 10.1109/IV.2013.31.

[8] W. Aigner, "Current Work Practice and Users' Perspectives on Visualization and Interactivity in Business Intelligence," 2013 17th International Conference on Information Visualisation, London, UK, 2013, pp. 299-306, doi: 10.1109/IV.2013.38.

[9] Growjo. (2023). Fastest Growing Companies and startups Data Set. Retrieved from https://www.growjo.com

[10] Kaggle data set from https://www.kaggle.com/datasets/amritpal24/top-1000-companies-details).