Overview

Venture capital firms are a type of investment company that helps finance innovative, early-stage startups that require funding to continue to grow. Historically, venture capital firms tend to seek out and provide more funding for startups founded and led by white men, but in recent years firms have announced efforts to focus more on diversity, equity, and inclusion (DEI) funding and initiatives to support minority, female, and LGBTQ+ founders and leaders. Our goal is to assess the true level of support and funding provided by venture capital firms to diverse startups compared to their statements. We will be exploring visualizations on firms that focus on diversity compared to those that don’t, firms with a large diversity hiring compared to those that do not, and more. This research project aims to continue throughout the 2023-24 academic year.

Methods and Data

Conducted initial data search through Duke Libraries, where we decided to use the following databases:
- Crunchbase: Provides diversity funding and investment information on public and private companies
- Preqin: Provides financial and categorical information on the alternative assets market and venture capital funded startups

We downloaded data for multiple venture capital firms, then loaded the .csv files from Excel onto Python and R for visualizations.

- ggplotly

To gain preliminary understanding of the venture capital industry, we reviewed academic literature and met with Professor David T. Robinson of Duke University and Professor Emmanuel A. Yimfor of Columbia University.

- We will be working with Professor Yimfor throughout the academic year to assist with his work on tagging the race of over 200,000 entrepreneurs to receive a final dataset.

Results

Using a small venture capital firm called Base Ventures, we developed a code template in R with their dataset in Crunchbase for future reference in analyzing larger venture capital firms. Cleaning data involved converting money into a numerical value (e.g. $125K to 125000), separating the year from the date, and cleaning the diversity type and funding round. Additionally, we plan on joining the dataset with the overall investments from Preqin.

Research Limitations

Our research team faced a few limitations this summer. For example, the diversity statistics required for Figure 4 was behind a paywall, and it took a few weeks to receive a license to access the data, delaying our time to work on visualizations. Once we had access to the website, we also had to manually convert the information into a .csv file. Additionally, diversity investments and statements for firms is a sensitive and often private subject that required us to develop our own metrics on the DEI efforts of the industry.

Future Plans

- Continue tagging entrepreneurs by race for Professor Yimfor
- Receive dataset once finished, allows us to expand our visualizations with more accurate information on entrepreneur diversity
- Develop more visualizations exploring the correlation between diverse-led firms and diverse-led startups, change during the peak of the BLM movement, etc.
- Possible cooperation with DUMAC in understanding Duke University’s investments
- Final goal is to develop a research paper or interactive website on our findings