

## Methodology & FAQs



### Methodology

PledgeLA's venture data report analyzes **proprietary industry data from Pitchbook** to learn more about VC portfolio investments. Altogether, PledgeLA annually codes and analyzes more than 6,301 transactions per year representing 44,802 people, 4,441 companies, and 69 firms. We are the only effort of its kind that annually monitors the advancement of regional venture data, holding accountable an industry vital to the future of all Angelenos. This cross-section of data allows us a holistic look at the Los Angeles tech and venture industry - including the investments in inclusion from institutional funders and the demographic makeup of start-up founders.



### FAQs

#### **Q: How did PledgeLA calculate the portfolio diversity of its VC firm members?**

- **Pitchbook** is an online platform that provides information on private and public companies, including investments, funding information, founding members, and industry news. Using data-scraping tools, we collected publicly available information on the investments of all 80 of our VC members to establish a complete list of companies they'd invested in from January 1, 2016 - December 31, 2021. Some firms had no investments during this time horizon and were excluded from the analysis, leading to a final data set of 57 firms. We then retrieved information on each portfolio company, including their headquarters location and the names of their founding team.



#### **Who We Track:**

L.A. Based Firms



#### **Sources:**

Pitchbook (2021) &  
Crunchbase (2019-2020)



#### **Method:**

Algorithmic Imputation  
via NAMSOR + Human  
Research Team



#### **Research Team:**

Prof. Jasmine D. Hill,  
UCLA Luskin School



➡ To assess the diversity of each portfolio companies' founding team we used two sources. First, **Pitchbook** offers demographic data (gender and race) on many founders in a growing effort to assess representation in tech (read about their methodology here). Where available, we scraped Pitchbook's data on founders' race and gender. Because Pitchbook does not have diversity data on every founder, we used a second technique to assess diversity: algorithmic imputation. We used the **NAMSOR API**, which is a peer-reviewed instrument that compares first and last names to decades of sociolinguistic data to recognize the likely cultural origin and gender category. Finally, a trained researcher combed through the data to reconcile discrepancies between Pitchbook and NAMSOR, and correct any other irregularities in the data. While not as perfect as self-report, as a methodology vetted and commonly used also by top academics to impute demographics, we are confident in the reliability of our estimates.

### **Q: How accurate is this data? What if transactions aren't listed on Pitchbook?**

➡ Pitchbook is the industry-standard that investors use themselves to track investment data and conduct their own research. Pitchbook's team of 1,500 data operations staff spends more than 800,000 person hours per year collecting, validating and organizing investment data through more than 100 proprietary processes. For more on Pitchbook's research process, [click here](#). PledgeLA's two-phase, secondary scrub (used to further validate founders' gender, race and location) improves upon this base and fills in any gaps. While a small number of transactions can be delayed or excluded from the platform, the vast majority are included, meaning that they are unlikely to change the overall trends we're seeing in portfolio diversity.

### **Q: How well does this data reflect the Los Angeles venture ecosystem?**

➡ Nearly all venture firms in the Los Angeles area (including all major, and many emerging firms) are PledgeLA members, and thus included in this analysis. This means that the analysis of overall portfolio diversity is reflective of LA venture as a whole. In future years, we hope to expand this analysis, comparing Los Angeles-based VCs with others operating in Southern California or other regions.