

Visualizing Innovation Activity in the U.S

Research Goals:

As part of my dissertation, I am developing a novel visualization application that can enable relevant stakeholders visualize and explore innovation activities in the U.S and analyze the role of university-industry collaborations in bolstering innovation. In this context, innovation activities include research and development activities (R&D), venture capital activities and entrepreneurial activities.

Who are the potential users of the application?

- Universities (technology transfer centers such as CoMotion)
- Researchers/Academics
- Investors looking to invest in a region/technology
- Startups looking to partner with universities
- Innovation labs in large companies
- Policymakers/government officials interested in looking at regional collaborative activities

Preliminary Results:

I chose the metropolitan statistical area (MSA) as my regional unit of analysis. Even though conducting a state-level analysis would be easier, as a lot of federal data is easily accessible at the state level, it would be based on the false assumption that university-industry collaborations occur uniformly across a state. In reality, these collaborations (for all three dimensions: education, research and entrepreneurship) are centered in urban, metropolitan regions. Therefore, an analysis at the MSA level will provide a more detailed and specific picture of these interactions. Furthermore, as a starting point, I decided to choose 2019 as my year for the analysis (with some exceptions, as noted below). This is because data for 2020 and 2021 will likely be influenced by the COVID-19 pandemic. For a preliminary analysis, I am only looking at cross-sectional data. To commence my preliminary tool design, I chose three data sources for each MSA in the U.S:

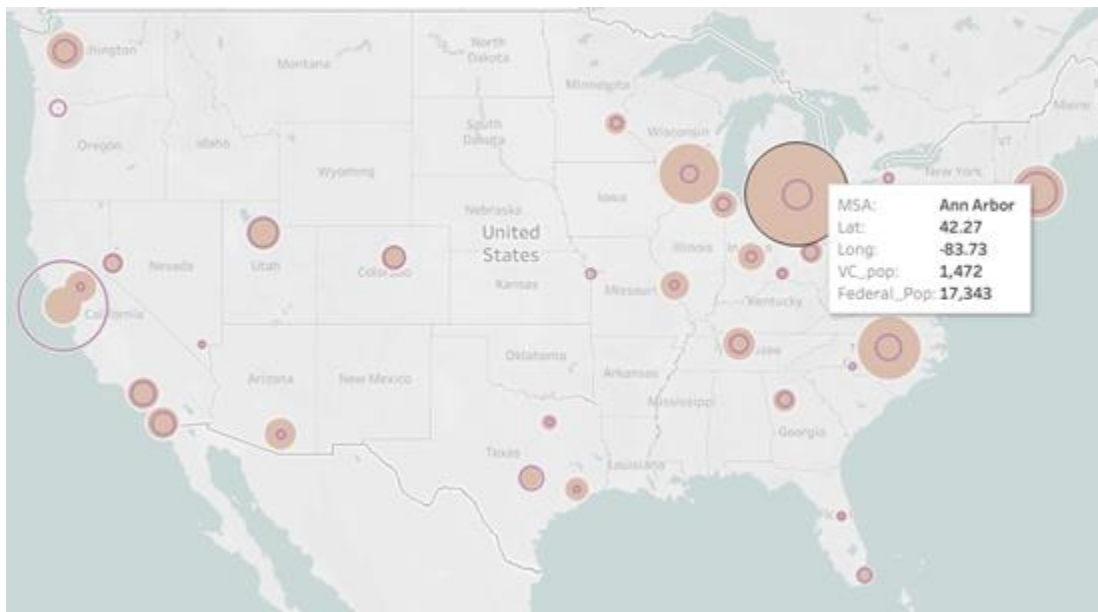
Type of Data (unit)	Year	Reason
Research grant funding (\$)	2019	An indicator for research
Small Business Innovation Research (SBIR)/Small Business Technology Transfer (STTR) (\$)	Averaged across 2017, 2018 and 2019 because of high year-to-year fluctuations	An indicator for research commercialization from universities to industry
Venture Capital (VC) disbursed (\$)	2019	An indicator for entrepreneurship

Given below are the steps I followed to filter the data and obtain regions of interest:

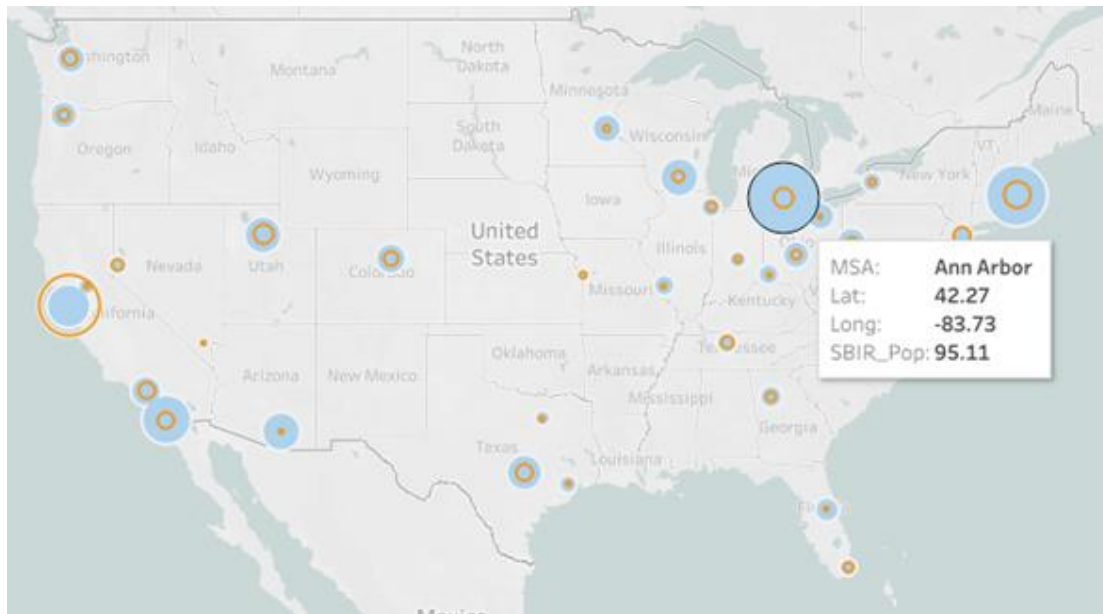
1. The total funding for each category was sorted from highest to lowest.
2. All the null/NA values were removed.
3. It was noted that in all three cases the funding values decreased exponentially and reached their asymptotic value around the 40th index. Therefore, the top 40 MSA regions were considered for further analysis for each indicator category.
4. Among the top 40 in each list, those MSA regions were selected that appeared in all three lists. This narrowed the list of regions being considered to 37

Therefore, out of 384 MSAs in the U.S, 37 of them received three-quarters of the research funding (75.0%), over half of the SBIR/STTR (67.1%) funding, and most of venture capital funding (95.0%).

Preliminary Design:



The solid brown circles depict the research funding in an MSA and the pink rings depict the VC funding.



The solid blue circles depict the SBIR/STTR funding in an MSA and the orange rings depict venture capital funding

Research Goals That can be Fulfilled through the project:

1. Extending the visualization application to include:
 - a. University information (Top priority)
 - b. Industry/sectoral information (Top priority)
 - c. Entrepreneurship activity information (See: <https://indicators.kauffman.org/>)
 - d. Any other direction you may have in mind
2. Designing an application that is ready for Phase 0 user testing
3. Potential user testing (stretch goal)

*****Proposal feedback*****

Feedback1

I like how the data on funding types is visualized using a heat map. The single design displays a lot of information without being cluttered.

I like the data bar at the bottom, but I was hoping it would open a new window when clicked with some data information related to the data bar title.

The entire dashboard is highly interactive, with multiple screens for visualizing data.

Feedback 2:

I like the heatmap's tooltip feature. Very informative.

I wish there was more explanation about new establishments broken down by industry sector.

I wish the vertical name on venture capital funding by technology was removed or renamed.

I also like the map of federal R&D expenditure by university. It was interesting to see how clicking on each city in the pie chart changed its discipline area.

Major Takeaways

For previous assignments, I had mostly used the Altair or Plotly library features for visualization, but this time our team decided to use the tableau. I was able to create a preliminary design of the dashboard with the help of the tutorial and self-learning through internet resources. People who like to code, including myself, tend to make visualizations in python or javascript, but being able to use tableau makes a lot of visualization easier and saves me time.

I also learned how important data wrangling is when creating a good visualization. For the final project, we spent a significant amount of time gathering the necessary data and then customizing the data format to meet our design requirements. Next, we first created our initial dashboard on Altair, which took nearly 3 hours to create a basic map design, and then we moved to design a dashboard with tableau, which gave us a lot of flexibility in creating different types of visualization.

Another thing I learned from this project is that making a tableau dashboard public using desktop tableau is not straightforward as it may seem. As a first-time user, I had to use Google and watch a YouTube to make the tableau dashboard public.

Design Rational:

Preliminary dashboard design rational ([link](#)):

Map design (cities with highest R&D Funding): My first choice was to use a map to see which cities received the most funding. As I believed that visualizing those cities on a map would make it easier for users to understand, I used the tooltip feature to show users the funding distributed by industry in percentage format, as well as the capital invested for 2019 year.

Grant received by Technology: Our data contained information on how funding is distributed by industry for all cities. Instead of looking at this for each city, I wanted the user to see the overall picture of funding distribution across US cities. As is clear, healthcare receives close to half of all funding.

SBIR and Federal bar chart: Both types of funding are illustrated using a bar chart with the funding amount on the y axis and the city name on the x axis. I wanted the user to be able to see which cities received the most and least funding right away, which is why I used a sorted bar chart.

Cities with most funding vs employments: I wanted the user to see if there was any correlation between the cities that received funding and the employment rate. Therefore, I thought a vertical bar chart would be the best way to show the trend. This chart reduces the amount of space at the bottom of a dashboard. I tried several times to change the width of the dashboard, but it didn't work. In the future, I will do more research to determine how to set the dashboard width to fit the design.

Poster design rational

Our team mostly followed the format provided by the instructor, with a few changes. Since our data was so large, it was imported to include preliminary data analysis and data sources information. For the milestone review, we used a Google form to collect feedback from the potential users, which we then shared on LinkedIn and with family and friends. The data from the feedback was then used in the evaluation/feedback section. For design goals, we listed key questions that we hoped to answer through this project. We used the diverge converge principle for the approach section, with each member creating their own dashboard design and then converging the design ideas into one. The visualization in the approach section ranges from the preliminary to the final stages.