Project Report
For Census API Assessment:
Businesses and Technologies Data Analysis

By Group 4

Members:

Elwood Olson, Hugh McMurray, Michael Deehan, and Jason Singh

Submitted on:

11th October 2021

Introduction

Data from the Annual Business Survey was explored via API calls from the United States Census Bureau for the year of 2019. The survey collected information based on economic and demographic characteristics from business owners in the United States. The datasets observed included:

1. ***Company Summary*** – which provided data for employers by sector, sex, ethnicity, race, veteran status, years in business, receipts size of firm, and employment size of firm.
2. ***Technology Characteristics of Businesses*** – which provided data on technology use and production for artificial intelligence, cloud-based computing, specialized software, robotics, and specialized equipment technologies data

API calls were focused on employers having businesses under the "Professional, scientific, and technical services" of the North American Industry Classification System (NAICS) code. Trends in the overall distribution of average pay with respect to employees were observed across all states, where average pay was defined as the ratio of annual payroll to number of employees. Sales and technology adoption, specifically focused on AI and cloud-based technologies, were also observed across this industry to gauge how it was evolving with employers and affecting employees. The main questions pursued were:

1.) How well does the industry pay on a per state basis?

2.) What aspect of the industry is being sold the most on a per-state basis?

3.) What aspect of the industry is being adopted by employers/employees on a per-state basis?

4.) Where might the industry be headed, based on these trends?

Data Sources:

US Census Bureau, Annual Business Survey (ABS) APIs (2019). Retrieved from: https://www.census.gov/data/developers/data-sets/abs.2019.html

State-by-State Distribution of Pay

First, we examine the rate of pay for workers in this industry. The following histogram illustrates the state-to-state variation in average annual compensation for professional, scientific, and technical employees.

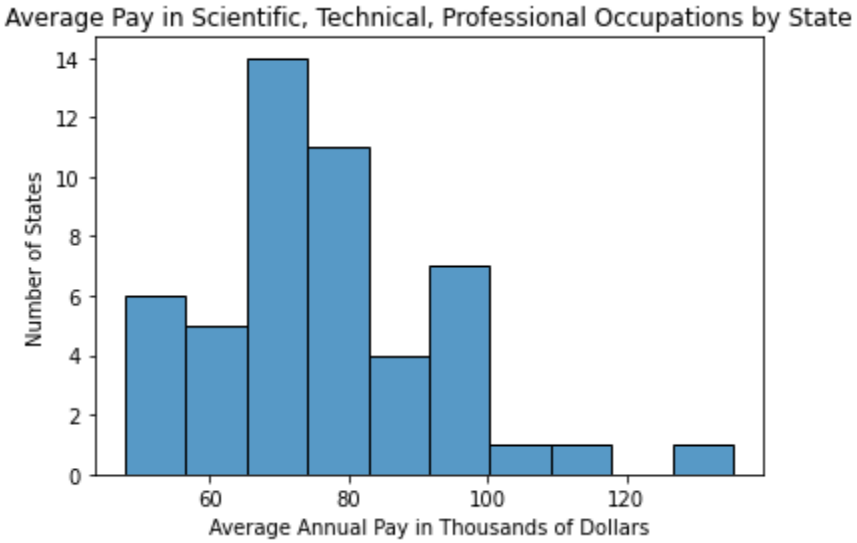

Figure 1: Average Salary for Employees in the Professional, Scientific, Technical Services Sector by State

As we can see from the peak in the above graph, many states pay workers in this sector between $60,000 and $80,000 annually on average. However, the data is widely distributed, and so many states pay workers in this sector much more and much less than $70,000 annually on average. In Kentucky, employees in these fields earn an average of about $48,000 every year, while in the District of Columbia, workers in the same field earn about $135,000 annually on average. So, what accounts for these fat tails? In the following graph, we can see the five highest paying states for the sector.

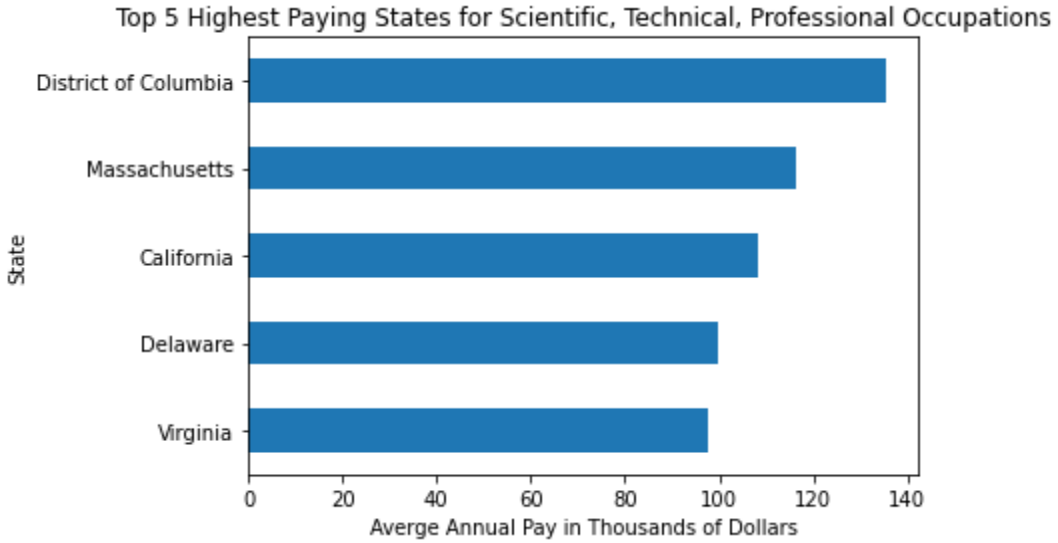

Figure 2: Top 5 States with the Highest Average Salary for Employees in the Professional, Scientific, Technical Sector

It’s difficult to discern a pattern that links these geographically disparate and economically diverse states. However, what this graph tells us is that the professional, technical, and scientific sector has a tendency toward explosive lucrativeness. In states like DC and California, where market demands for information work is high, workers in this field are likely to be well paid, no matter whether they are working on politics or consumer technology. In the rest of this report, we examine more granular variations in state-by-state usage and sales of technology which might help to explain some of the broader variations in the demand for work in this sector.

Regarding Trends in AI and Cloud Usage Across Employer Firms

The data was first analyzed to see if there were any correlations between firms that used and didn’t use the technologies. Based on a high-level trendline analysis, it seems that there is very little to no correlation between the usage or non-usage of the two technologies amongst firms, indicating that adoption of usage is mostly independent of the technology type for the firms sampled. The plots below summarize this observation:

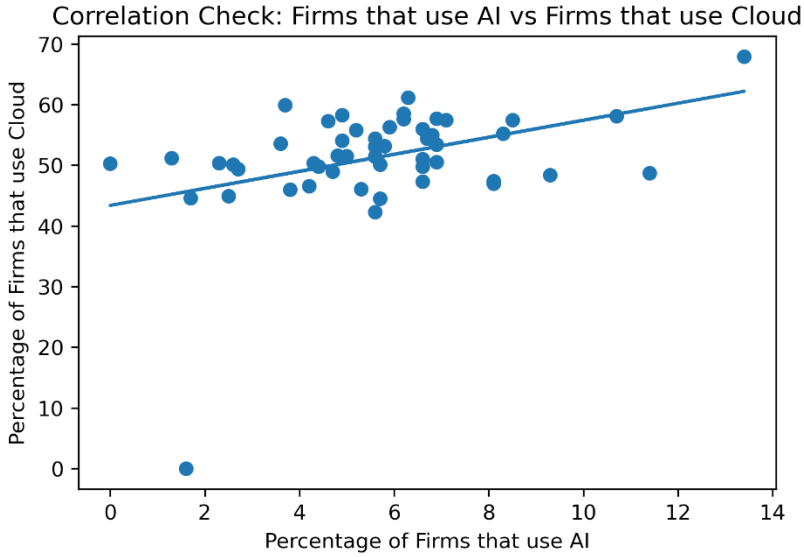

Figure 3: Checking for correlations: High-level trendline analysis: Firms that use AI vs Firms that use Cloud

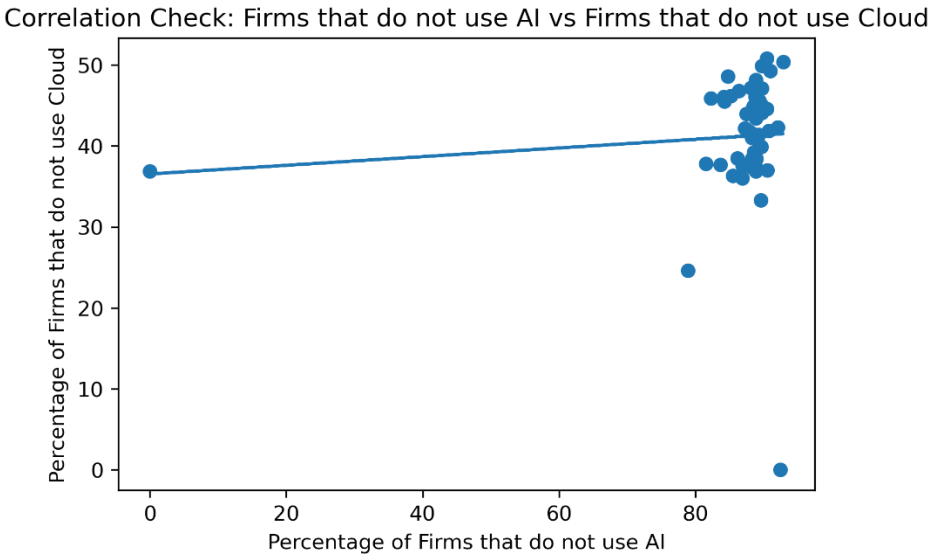

Figure 4: Checking for correlations: High-level trendline analysis: Firms that do not use AI vs Firms that do not use Cloud

The top 15 states for each technology-adoption were found and further trends were observed at a high-level. Four cases were observed, viz.:

1.) High AI Usage vs. No AI usage in the top 15 states for firms that have high AI usage (based on the sample observed).

Employees per firm was deduced based on the ratio of the total number of employees to total number of firms for the dataset and this ratio was used to 'normalize' the distribution with respect to the states observed. It was found that high AI-usage was more dominant in larger companies (firms having more employees on average) compared to firms that did not use the technology, as summarized in the figure below:

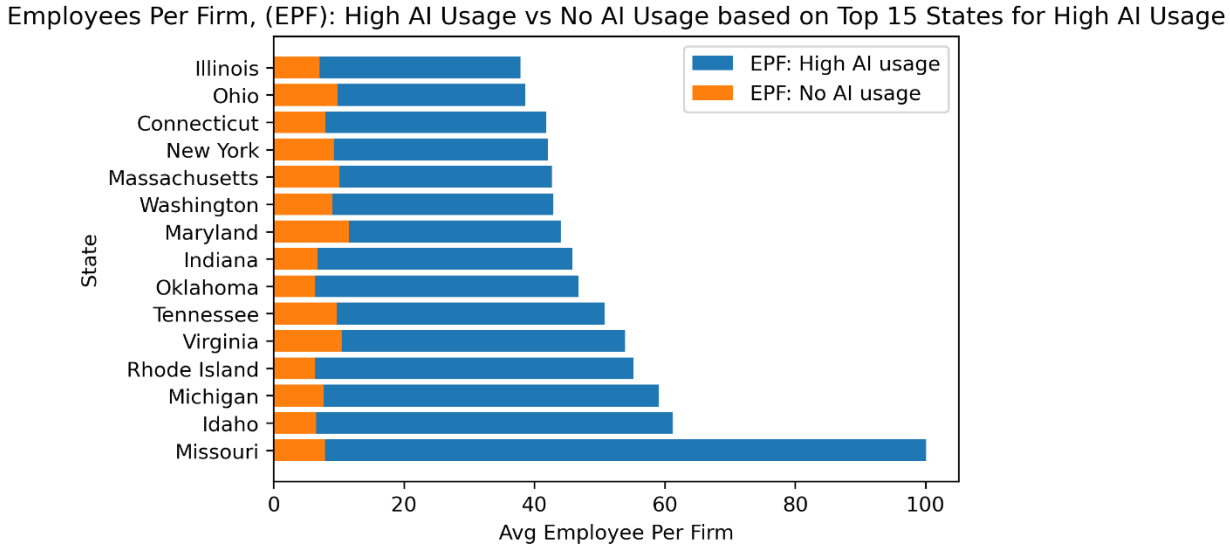

Figure 5: AI usage and non-usage comparison amongst top 15 states for High AI usage

2.) Companies that had High Cloud usage were also observed in these top 15 states (for high AI usage), and it was found that, in general firms that used AI extensively tended to be much larger than those that used cloud-based technology extensively, as summarized below.

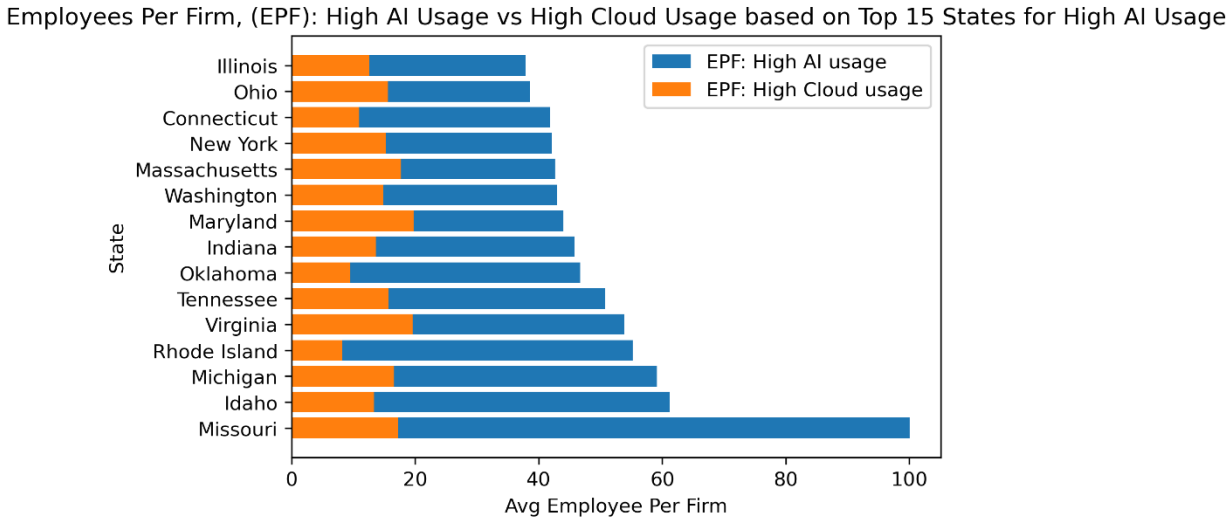

Figure 6: AI usage and Cloud usage comparison amongst top 15 states for High AI usage

3.) High Cloud usage vs no Cloud usage in the top 15 states for firms that have high Cloud usage (based on the samples observed).
Employees per firm for firms that use Cloud-based technologies extensively vs firms that don't use Cloud-based technologies were observed, as summarized below:

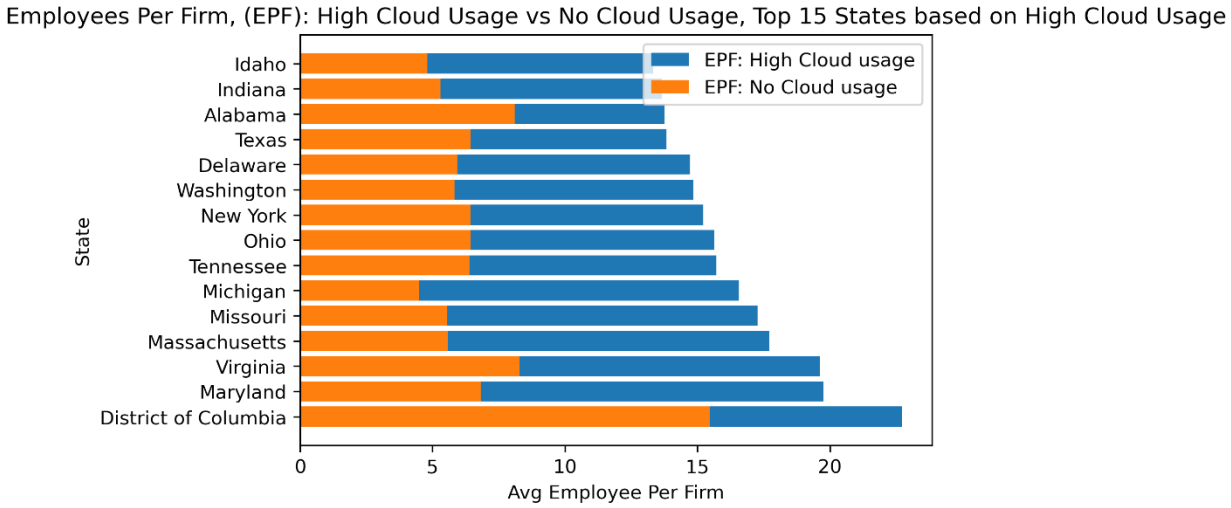

Figure 7: Cloud usage vs No Cloud usage amongst the top 15 states for high Cloud usage

It seems that there was a less polar split between companies that don't use Cloud tech and companies that do (as compared to AI usage/non-usage), indicating that there might not be as distinct a need for cloud-based tech as compared to AI-based tech amongst the firms observed in the survey (i.e. cloud-based companies are more general purpose in their application of their technology) Most of the top 15 AI states re-appeared in the findings as well, indicating that these tech-states aren't limited to a specific technology in general (but there seems to be little correlation amongst the technologies adopted amongst firms, as implied by the scatter plot).

4.) Firms that had extensive AI usage were also observed for the top 15 states that had extensive cloud usage, and it was found that in general, AI usage was more adopted than cloud-based tech in larger companies, re-enforcing the notion that AI-usage seems to be more prevalent in larger companies whereas cloud-usage is more common in medium to smaller companies (and may not be as necessary or as applicable for large scale as compared to AI technologies, at least for the firms in the observed dataset). The plot below summarizes the findings:

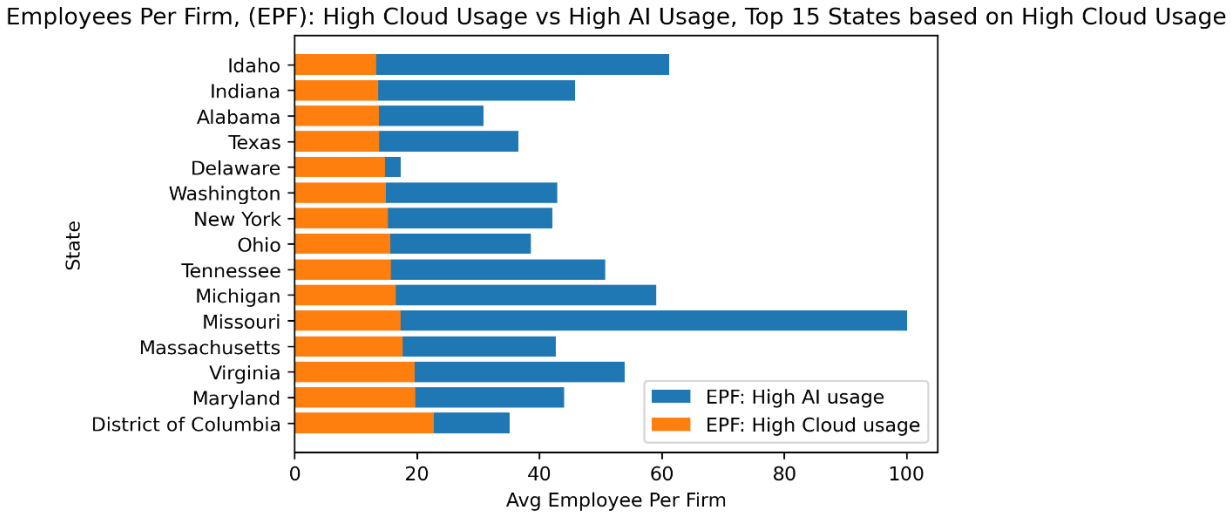

Figure 8: Cloud usage vs AI usage amongst firms in the top 15 states for high Cloud usage

Analyzing Sales by State:

We wanted to look at sales of different technologies by states, focused on cloud-based and AI technologies. The following bar chart shows the top 10 states based on what percentage of their sales consisted of AI technologies.

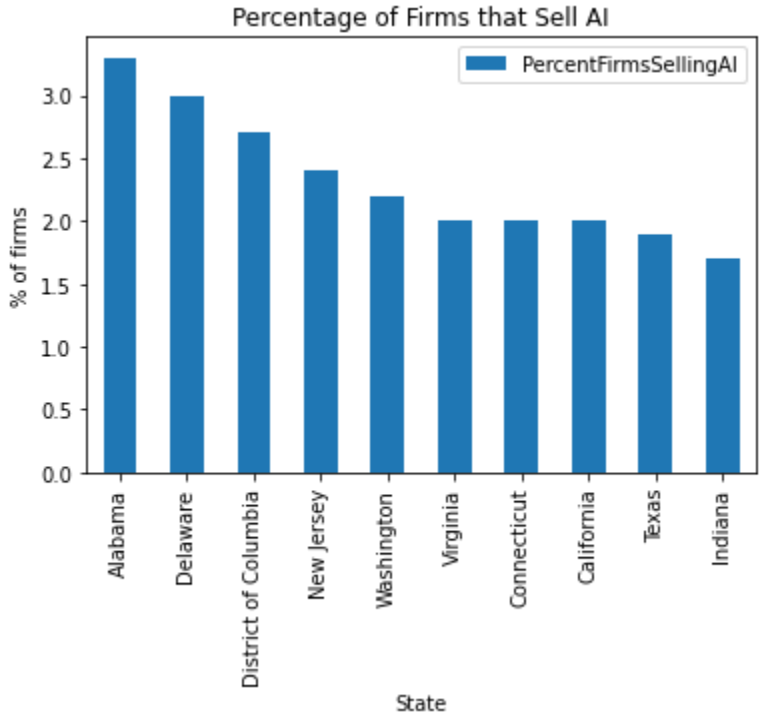

Figure 9: Percentage of Firms that Sell Artificial Intelligence by State (Top 10)

As we can see, the top states ended up being Alabama, Delaware, District of Columbia, and New Jersey. While DE, DC, and NJ being included in this were not surprising, Alabama is quite surprising, as the tech sector in Alabama is not as big as it would be in other states. This result is probably because there just is less data in Alabama verses a state like California, where there is a significantly larger tech sector.

Next, we have another bar chart for the Percentage of tech sales that are cloud-based technology, again grouped by state.

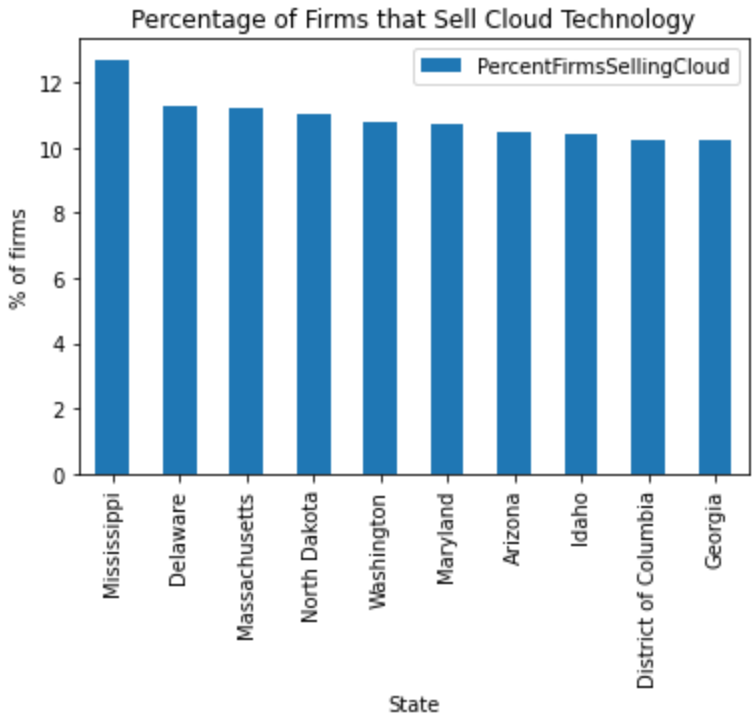

Figure 10: Percentage of Firms that Sell Cloud Technology by State (Top 10)

Once again, we have a mix of states that are both surprising (Mississippi and North Dakota) and not surprising (Delaware and Massachusetts). This is likely for similar reasons to the AI sales, as different states have larger or smaller tech sectors respectively.

Conclusions:

Businesses classified as "professional, scientific, and technical services" by the North American Industry Classification System include a multitude of different firms. For example, some of the companies included in this business sector include legal services, accounting companies, consulting firms, and computer services. Most of these business activities require a high level of expertise, training, and skill. One of the trends we noticed from our datasets was that “professional, scientific, and technical services" companies located in places like California, District of Columbia, and New York tend to pay higher employee salaries than companies in places like South Dakota, Kentucky, and West Virginia. We also noticed places close to the District of Columbia (Virginia, Delaware, and Maryland) yielded higher-than-average employee salaries as well. We deduced that salaries were higher in regions where services like legal representation, business consultants/analysts, and computer science professionals were higher in demand. For example, companies like Apple and Google are based in California. These companies demand high quality software engineers, legal teams, consultants, accountants, and other business professionals, driving salaries higher. The same could be inferred about the regions around Washington D.C., where government officials, legal professionals, and business experts are also prevalent.

As technology continues to advance, many firms have started to utilize technologies like artificial intelligence, cloud computing, robotics, virtual reality, and cyber security. We decided to investigate the trends regarding artificial intelligence and cloud computing technology usages among firms in the “professional, scientific, and technical services" sector. One of the trends we noticed was that firms with more employees have started to commit to using artificial intelligence and cloud computing technologies. We believe that bigger companies have the resources to effectively grow and adapt to the changing technological world, while smaller companies may lack the resources and will adapt to newer to technology at a slower rate.

When looking at companies that are currently selling technologies to consumers, we noticed the trend that a higher percentage of firms (by state) sell cloud technology than AI technology. While the state with the highest percent of firms selling AI barely eclipses the three percent mark, multiple states have double-digit percent of firms selling Cloud technology. Cloud technology currently seems to be more popular in sales as companies like Amazon, Microsoft, and Google, along with many others, have begun to sell Cloud computing products. Considering this data, we believe that entering the business of cloud technology could be difficult because it has become more and more saturated. We believe that the artificial intelligence industry is much less saturated and could be an interesting market for businesses to infiltrate.

Technology continues to advance every day. Our world will continue to adapt, and data will continue to accumulate. We believe that as more and more businesses seek data analysis, automation, and new technology, the landscape of this sector will certainly change. Certainly, we believe that more firms will begin leveraging or selling technologies like cloud technology, artificial intelligence, robotics, and many more technologies in the future. We also believe that this sector is loaded with many essential business services like accounting, tax preparation, legal services, and engineering, so the demand for many of the services of this sector will remain high and salaries should remain quite consistent.