

Exploration

I chose to investigate Question (1) from the brainstorming section, i.e. 'which job categories are the most prevalent in the VISA sponsoring data'.

As a first step it is interesting to see what job categories are most prevalent nationwide. The results for the top 10 (this number can be changed in the provided python code) is shown in Figure 1:

Job Title	Number of LCA cases
Computer Systems Analysts	92984
Software Developers, Applications	77878
Computer Programmers	72353
Computer Occupations, All Other	39573
Software Developers, Systems Software	15726
Management Analysts	11654
Accountants and Auditors	9488
Financial Analysts	9316
Network and Computer Systems Administrators	8182
Mechanical Engineers	7402

Figure 1.

It comes to no surprise that software related jobs many of the top places due both because of the abundance of high tech, software providing companies in the US economy as well as the great demand in many and diverse fields.

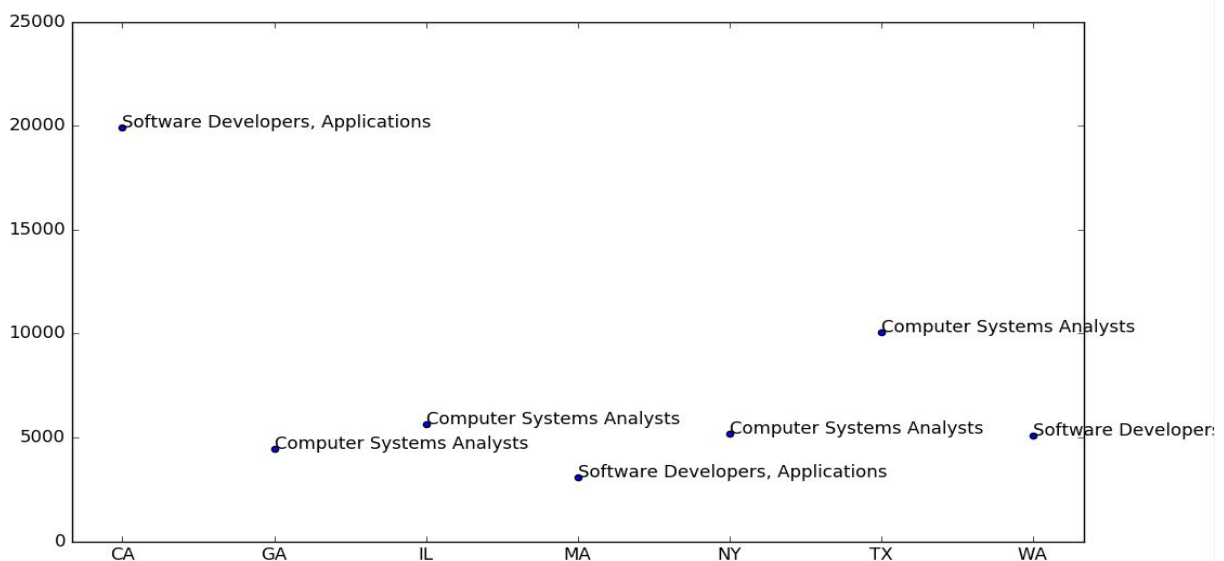


Figure 2.

To provide similar data at the state level ones has to decide what state to assign each LCA. For the the purposes of my exploration I think the state of each employer represents the place where each position has originated. In real world this is a discussion that needs to take place between the data scientist and the consumer of the results of the investigation.

A summary of the results in presented in Figure 2. To make the figure readable, only a handful of the US states are resented. If needed one can change that set as needed/desired in the provided python code.

The python code prints a table the entire list of states and their most prevalent job category.