# H1B Applications Exploration

# Context

H-1B visas are a category of employment-based, non-immigrant visas for temporary foreign workers in the United States. For a foreign national to apply for H1-B visa, a US employer must offer them a job and submit a petition for a H-1B visa to the US immigration department. This is also the most common visa status applied for and held by international students once they complete college or higher education and begin working in a full-time position.

The following articles contain more information about the H-1B visa process:

* [What is H1B LCA ? Why file it ? Salary, Processing times – DOL](https://redbus2us.com/what-is-h1b-lca-why-file-it-salary-processing-times-dol/)
* [H1B Application Process: Step by Step Guide](http://www.immi-usa.com/h1b-application-process-step-by-step-guide/)

# Content

This dataset contains five year's worth of H-1B petition data, with approximately 3 million records overall. The columns in the dataset include case status, employer name, worksite coordinates, job title, prevailing wage, occupation code, and year filed.

For more information on individual columns, refer to the column metadata. A detailed description of the underlying raw dataset is available in [an official data dictionary](https://www.foreignlaborcert.doleta.gov/docs/Performance_Data/Disclosure/FY15-FY16/H-1B_FY16_Record_Layout.pdf).

# Acknowledgements

The Office of Foreign Labor Certification (OFLC) generates program data, including data about H1-B visas. The disclosure data updated annually and is available [online](https://www.foreignlaborcert.doleta.gov/performancedata.cfm).

The raw data available is messy and not immediately suitable analysis. A set of data transformations were performed making the data more accessible for quick exploration. To learn more, refer to [this blog post](https://sharan-naribole.github.io/2017/02/24/h1b-eda-part-I.html) and to the complimentary [R Notebook](https://github.com/sharan-naribole/H1B_visa_eda/blob/master/data_processing.Rmd).

# Inspiration

* Is the number of petitions with Data Engineer job title increasing over time?
* Which part of the US has the most Hardware Engineer jobs?
* Which industry has the most number of Data Scientist positions?
* Which employers file the most petitions each year?