

## **In this issue >>>**

This dataset offers comprehensive data on H-1B visa applications submitted in the US between 2011 and 2018.  
Type: Categorical Dataset



***Investigating Impact on Job Market, Salary, & Approval***

# H-1B Non-Immigrant Labour Visa



## ***Current topic >>>***

**The dataset contains 25 columns and covers a total of 3.36 million records spanning from 2011-2018.**

The Office of Foreign Labor Certification processes case certification decisions and collects data from employers' Labor Condition Applications for use in this project (OFLC). When an H1B company wants to hire non-immigrant employees for a particular job activity, the Labor Condition Application (LCA) is a document that is submitted to the U.S. Department of Labor Employment and Training Administration (DOLETA). The data is retrieved from the Department of Labor website.

## ***Analysis on Non-Immigration Working Visa in US***

**For more than two decades, the H-1B visa has been the main route for foreign students and job applicants to get employment in the United**

***H-1B: Specialty Occupation Person: to seek a particular profession. Requires a college degree or its equivalent. includes government-to-government research and development, or co-production programmes managed by the Department of Defense.***

### **About This Dataset**

This dataset includes details on the H-1B visas that US companies have secured to hire foreign workers in specialised positions like software developers, engineers, and scientists. This dataset includes a variety of features of the visa petition procedure, including thorough information about the petitioners, their occupations and wages, and whether or not their applications were accepted. We have a rare chance to examine the effect that these visas have on job options for talented individuals in America, from current salary levels to visa application success rates. We may learn more about how labour dynamics, which are frequently governed by laws and regulations relating to immigration, by analysing this dataset. Looking at H-1B visas that are refused or approved with different income levels is an attractive potential for study on wage inequities. It will provide light on salary disparities between domestic workers and overseas workers, enabling policymakers to better comprehend why some businesses may struggle to find domestic talent or attract foreign talent based on wage levels.



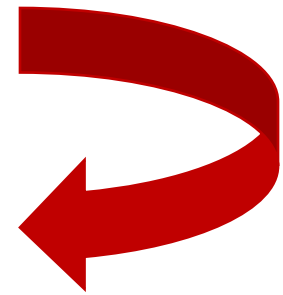
License CC0: Public Domain

<https://www.kaggle.com/datasets/thedevastator/h-1b-non-immigrant-labour-visa>

Set the stage >>>

The Relevant Columns in the Dataset include:

Column name	Description
Case Status	The status of the case, either approved or denied. (String)
Case Year	The year in which the case was submitted. (Integer)
Case Submitted	The date on which the case was submitted. (Date)
Decision Date	The date on which the decision was made. (Date)
Employer Name	The name of the employer. (String)
Employer City	The city in which the employer is located. (String)
Employer State	The state in which the employer is located. (String)
Employer Zip	The zip code of the employer. (Integer)
Employer Country	The country in which the employer is located. (String)
Job Title	The title of the job for which the visa is being applied. (String)
SOC Name	The name of the Standard Occupational Classification for the job. (String)
SOC Code	The Standard Occupational Classification code for the job. (Integer)
Full-time Position	Whether the position is full-time or not. (Boolean)
Prevailing Wage	The prevailing wage for the job. (Integer)
Prevailing Wage Unit	The unit of the prevailing wage. (String)
Prevailing Wage Level	The level of the prevailing wage. (String)
Wage From	The minimum wage for the job. (Integer)
Wage To	The maximum wage for the job. (Integer)
Wage Unit	The unit of the wage. (String)
Wage City	The city in which the job is located. (String)
Wage State	The state in which the job is located. (String)
Employer H1B Dependent	Whether the employer is H-1B dependent or not. (Boolean)
Employer Willful Violator	Whether the employer is a willful violator or not. (Boolean)
Longitude Information	The longitude of the job location. (Float)



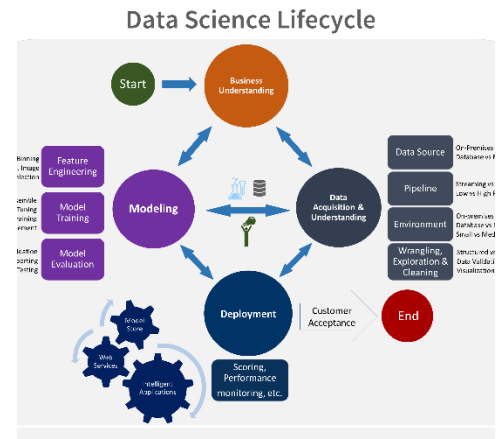
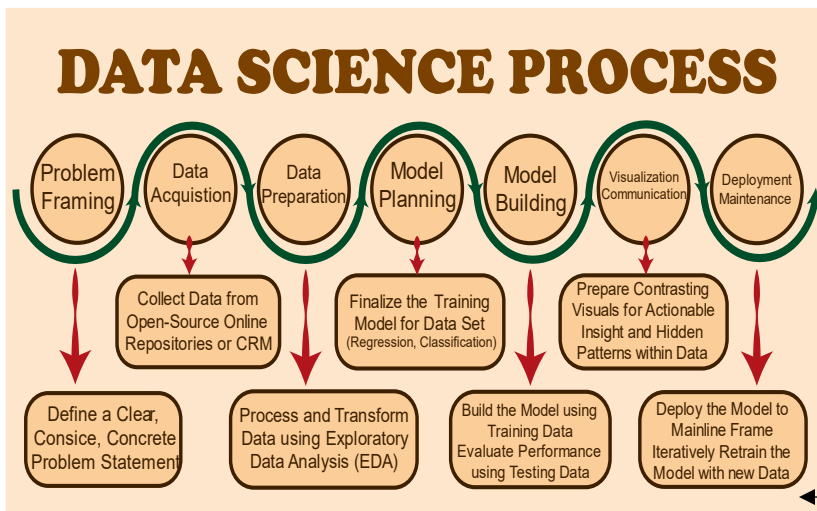
**The Description  
of the Columns  
are as Follows**

# Essential Questions for Data Analysis and Visualisation

To analyze the benefits of the H1B visa in the united state using a large amount of data which gives insights into the visa application and job trends in the United State. The new platform like big data application was used.

## Binary and unitary for data analysis Questions:

- Q1). Analysis of the number of applications year-wise. And the applicant's growth rate per year.
- Q2). Summary of Applicants according to Case status.
- Q3). What are the top 10 desirable Job Titles and their average salary?
- Q4). Top 20 Companies who sent the highest H-1B visa applications.
- Q5). Comparison between the average salary of the top 10 companies who are in demand AND the top 10 companies who are paying the highest salary.
- Q6). Out of all states which 10 states are more desirable state by employees? Which top 5 states have denied most? And we have to find out if is there any correlation between these two results
- Q7). The ratio of Full time and Half time positions.
- Q8). How many applicants applied for the Data Science job profile? Data Science job profile includes Job Titles like Data Scientist, Data Analyst, Data Engineer, Machine Learning Engineer, and Business Analyst.
- Q9). which trending job market in the United State?
- Q10). Finding the top job position each year?
- Q11). The average prevailing wage for the job each year?
- Q12). Total number of petitions in each year?



## Steps of Data Analysis

The data analysis process includes **defining a business problem, understanding and collecting data, extracting data from various sources, applying data quality for data cleaning, selecting features through exploratory data analysis, and identifying outliers.** It follows specific stages such as **and removal, data transformation and data creation. Visualize with charts and graphs, apply statistical analysis, and machine learning models.**

**Last step: Deploying the Model:** Based on test results, the model with the best results is finalized and deployed into production when the desired results are achieved through appropriate testing according to business needs. The data science modelling process is now complete.

