

# INFO 536: Project Proposal

Quan Gan, Yuxuan chen

October 28, 2021

## 1 Title

H-1B LCA petitions analysis from 2016 to 2020

## 2 Group Members

We will finish the final project in a group. Quan Gan and Yuxuan chen are the group members. Generally, we will use the Wechat application to communicate with each other, a Chinese multi-purpose instant messaging, social media and mobile payment app developed by Tencent. The google docs is another tool that we used to finish the written portion. In the following project, we might create a GitHub repository to manage the code portion. We decide to have a weekly meeting for each project phase to discuss how to arrange the project.

## 3 Introduction

Non-US nationals who wish to work in the United States must first get valid visas. The H-1B visa is a non-immigrant visa that permits a foreign citizen to work in the United States in a speciality vocation. This is practically the only approach that international students could settle in the United States. The US employers will fill petitions on behalf of the high skilled foreign employees to apply for the H-1B visas. LCA refers to Labor Condition Application, which is the first process of H-1B applications [1]. LCA is essential for ensuring that you are paid properly as a foreign worker and that you are not exploited by US companies. H-1B applications are approved by draw but LCA are approved primarily by Wage that must be higher than the same level US employees' minimum wage. In this project, I'm investigating at what factors impact the approval of an H-1B LCA application. we wish US companies could reference this research to recruit high skills workers and help them gain H-1B and also the foreign job seekers or international students could follow this research to find qualified US companies. The larger questions the project addresses are how vocations' needs evolve in the US market. The wage change and number of applications reflect the market's needs in the United States. The study will reveal what factors contribute to the denial of LCA applications and which jobs are most likely to be approved. The significance of this research is that it may represent the challenges of acquiring a work permit in the United States as well as trends in the US market. The project's societal problems are that it will affect international students' or foreign job seekers' decisions about whether or not to study or settle in the United States.

Furthermore, the project will affect the US companies about whether or not to hire foreign employees. We'll emphasize on data science-related professions like data scientist and data analyst after analyzing the total LCA applications.

## 4 Research Question

From 2016 to 2020,

1. What factors impact the approval of LCA applications?
2. What companies have high expectations of international workers?
3. What are the changes of data science jobs' demands in the US market?

## 5 Methods

The data set is LCA Programs (H-1B, H-1B1, E-3). It contains LCA applications xls files from 2016 fiscal year to 2020 fiscal year. For the 2020 fiscal year, there are four xls files so in the later data wrangling, We need to merge them into one table. Every year, the number of variables grows. There are 41 variables in 2016, but there are 95 variables in 2020. Furthermore, the names of the variables differ. As a result, We'll use variables that are present in all fiscal years and have the same meaning. The variables listed below will be used.

Variable name	Definition
CASE_NUMBER	Unique identifier
EMPLOYER_NAME	Employer's name
STATUS	Status associated with the last significant event or decision
CASE_SUBMIT	Date and time the application was submitted
VISA_CLASS	type of temporary application
SOC_NAME	Title of the SOC occupational group
JOB_TITLE	Job title
WAGE_RATE_FROM	Employer's proposed wage rate
WAGE_RATE_TO	Maximum proposed wage rate
WAGE_RATE_UNIT	Unit of pay for proposed wage rate
FULL_TIME_POS	Y = Full time; N = Part time position
WORKLOC1_CITY	job location
WORKLOC1_STATE	job location
PW_1	Prevailing wage rate
PW_UNIT_1	Unit of pay

This data set is extracted from non-immigrant and immigrant application tables within the office of foreign labor certification's case management systems. The U.S. DEPARTMENT OF LABOR publishes the disclosure data on their website [2].

## 6 Results

We intend to visualize the results once the data has been analyzed. The graphs and charts might help to better illustrate my views. To test the variables, We may utilize

statistical methods such as the t-test. The confidence interval and p-value will be used to support the data, which will corroborate my findings. In the last, We will highlight institutions. Applicants often submit LCA applications through mail, however errors occur when the US Department of Labor transforms data to a digital format. Furthermore, persons with bachelor's or master's degrees from the United States dominate the LCA applications. However, some candidates acquire their degrees outside in the United States, and some applicants file for LCA many times each year. The data biases are caused by this phenomena.

## **7 Other Considerations**

There is a problem with which we have no idea how to address, and it is not unique to this project. When a variable involves the different units especially the salary units, how should we convert them to the same units? The salary units are "hours", "weeks", "bi-weekly", "Month", and "year". We could not just multiply 12 to convert month units to year units. We could not promise that everyone would work the entire year without taking any vacations.

## **References**

- [1] H-1b specialty occupations, dod cooperative research and development project workers, and fashion models. <https://www.uscis.gov/working-in-the-united-states/h-1b-specialty-occupations>. Accessed: 2021-10-21.
- [2] U.S. DEPARTMENT OF LABOR. Performance data. <https://www.dol.gov/agencies/eta/foreign-labor/performance>. Accessed: 2021-10-21.