Salary Data Set Description

When a US company wants to hire someone from outside of the United States for a technical position, they have to file an application to the United States government to get a green card or visa for the foreign applicant. These applications allow the US government to track who is entering and leaving the country for work-related reasons, and ensure that immigrants are neither being taken advantage of nor causing adverse effects for U.S. workers. To ensure equity for US and non-US workers, companies have to state how much they are planning on paying the employee every time they submit a visa or green card application. They also have to state the average amount an employee with similar skills and background typically gets paid for the same position, a figure called "the prevailing wage." This publically available data provides a unique view into what types of salaries you might encounter for different data–related jobs in the US.

Source
The original data was compiled by the US Department of Labor's Office of Foreign Labor Certification (http://www.foreignlaborcert.doleta.gov/performancedata.cfm).

Other reported uses of the data:
http://data.jobsintech.io/
http://blog.bigml.com/2013/10/01/using-text-analysis-to-predict-h1-b-wages/

Overall Description
We assembled 167,278 records from the Labor Condition Applications (which include H1-B and other specialty visas) and the permanent resident applications from 2008 to 2015. The number of H1B and permanent visa applications considered in the dataset can be seen in the table below.

Year	Number of visa applications	Number of permanent residency applications
2008	0	5
2009	103	4
2010	962	19
2011	8,877	80
2012	9,448	146
2013	47,426	1,600
2014	58,289	9,050
2015	31,080	189

Variable Descriptions
For each visa application in the dataset, the following information was prepared. Data was discarded or cleaned when necessary.

FIELD NAME IN DATASET	DESCRIPTION
Case Number	Unique identifier assigned to each application submitted for processing to the ETA National Processing Center. A = Atlanta NPC; C = Chicago NPC.

FIELD NAME IN DATASET	DESCRIPTION
Case Status	Status associated with the last significant event or decision in the case application. Valid values include "Certified," "Certified-Expired," "Denied," and "Withdrawn."
Case Received Date	Date the application was received by the ETA National Processing Center.
Decision Date	Date on which the last significant event or decision was recorded by the ETA National Processing Center.
Employer Name	Name of employer requesting certification.
Prevailing Wage Submitted	Prevailing wage for the job being offered.
Prevailing Wage Submitted Unit	Unit of Pay. Valid values include "Hourly (hr)," "Weekly (wk)," "Bi-Weekly (bi)," "Monthly (mth)," and "Yearly (yr)."
Paid Wage Submitted	Lower range of the wage offered to applicant
Paid Wage Submitted Unit	Unit of Pay. Valid values include "Hourly (hr)," "Weekly (wk)," "Bi-Weekly (bi)," "Monthly (mth)," and "Yearly (yr)."
Job Title	Payroll title of the job being offered, as entered on the application.
Education Level Required	Minimum level of education required to adequately perform the duties of the job being offered.
College Major Required	Major field of study required based on the education requirement.
Experience Required Y N	Identifies whether experience in the job offered is a requirement (Y/N).
Experience Required Num Months	The number of months experience that are required for the job (if applicable).
Country Of Citizenship	Country of citizenship of the foreign worker being sponsored by the employer for employment in the United States.
Prevailing Wage Soc Code	Occupational code associated with the job being requested, as classified by the Standard Occupational Classification (SOC) System.
Prevailing Wage Soc Title	Name associated with the Prevailing Wage SOC Code.
Work City	City information of the foreign worker's intended area of employment.
Work State	State information of the foreign worker's intended area of employment.
Work Postal Code	Zip Code information of the foreign worker's intended area of employment.
Full Time Position Y N	A binary variable indicating whether the position is full-time (Y) or part-time (N).
Visa Class	Indicates the type of application submitted for processing. Options include "H-1B," "E-3 Australian," "H-1B1 Chile," "H-1B1 Singapore," "greencard."
Prevailing Wage Per Year	Prevailing wage per year, calculated for this Coursera course from the "Prevailing Wage Submitted" and "Prevailing Wage Submitted Unit" fields.
Paid Wage Per Year	Paid wage per year, calculated for this Coursera course from the "Paid Wage Submitted" and "Paid Wage Submitted Unit" fields.
Job Title Subgroup	Subgroup label made for this Coursera course based on the "Job Title" field. Options include assistant professor, attorney, business analyst, data analyst, data scientist, management consultant, software engineer, teacher.