Summary information

When examining the country of citizenship column with the ‘case status’ column, you can see there is a wide variation in visa application status. For case status, the dataset had four options including certified, certified-expired, denied and withdrawn. When sorting by case status == certified, 5 countries had 100 or more certified visas they included: Canada, China, India, Mexico and South Korea. India has the highest number by far, 4158 and China the second highest at 670. In comparison, denied cases were rare, only two countries, India and Philippines has more than 100 denied cases. India has 279 denied cases and the Philippines had 126 denied cases. Other countries with high certified cases had low denials were Canada with 19 denials, China with 34 and South Korea with 14. This suggests that certified cases doesn’t fully depend on number of applications as denials are higher for a country like the Philippines (with 100 certified cases) than China or Canada.

**import** pandas **as** pd

**import** numpy **as** np

salary\_data **=** pd**.**read\_csv("salary\_data\_states.csv")

salary\_data**.**head()

salary\_data**.**info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 167278 entries, 0 to 167277

Data columns (total 27 columns):

salary\_data['COUNTRY\_OF\_CITIZENSHIP']**.**value\_counts()

certified\_row **=** citizen**.**iloc[[0]]

certified\_row

highest\_cert **=** (certified\_row['CASE\_NUMBER'] **>** 10)

highest\_cert

cert **=** pd**.**DataFrame(certified\_row['CASE\_NUMBER'] **>** 100)

pd**.**set\_option("display.max\_rows", **None**, "display.max\_columns", **None**)

cert

denied\_row **=** citizen**.**iloc[[2]]

denied\_row

From all the applicants (6423 people) that applied for their visas who worked for Google, about 1 percent (70 people) of them were denied visas. From all the google applicants (6423 people), about 90% percent (5760 people) did not go to college. For the google employees that did study for a degree, about 660 (9.73 %) of them studied some form of computer science for their major. From the 6423 employees, 5413 employees (84%) worked in California, 542 (8%) in New York, 309 (5%) in Washington, 66 (1%) in Massachusetts,43 (.007%) in Pennsylvania, 27 (.0042%) in Colorado, 11 (.0017%) in Texas, 11(.0017%) in Illinois, 2(.0003%) in Wisconsin, 1(.00016%) in Oregon and 1(.00016%) in North Carolina.