

FINAL PROJECT - TRENDS OF H1B DATA ANALYSIS

VISUALIZATION OF COMPLEX DATA (DATS 6401)
PRESENT BY YON HO CHEONG

DATA SOURCE AND TOOL

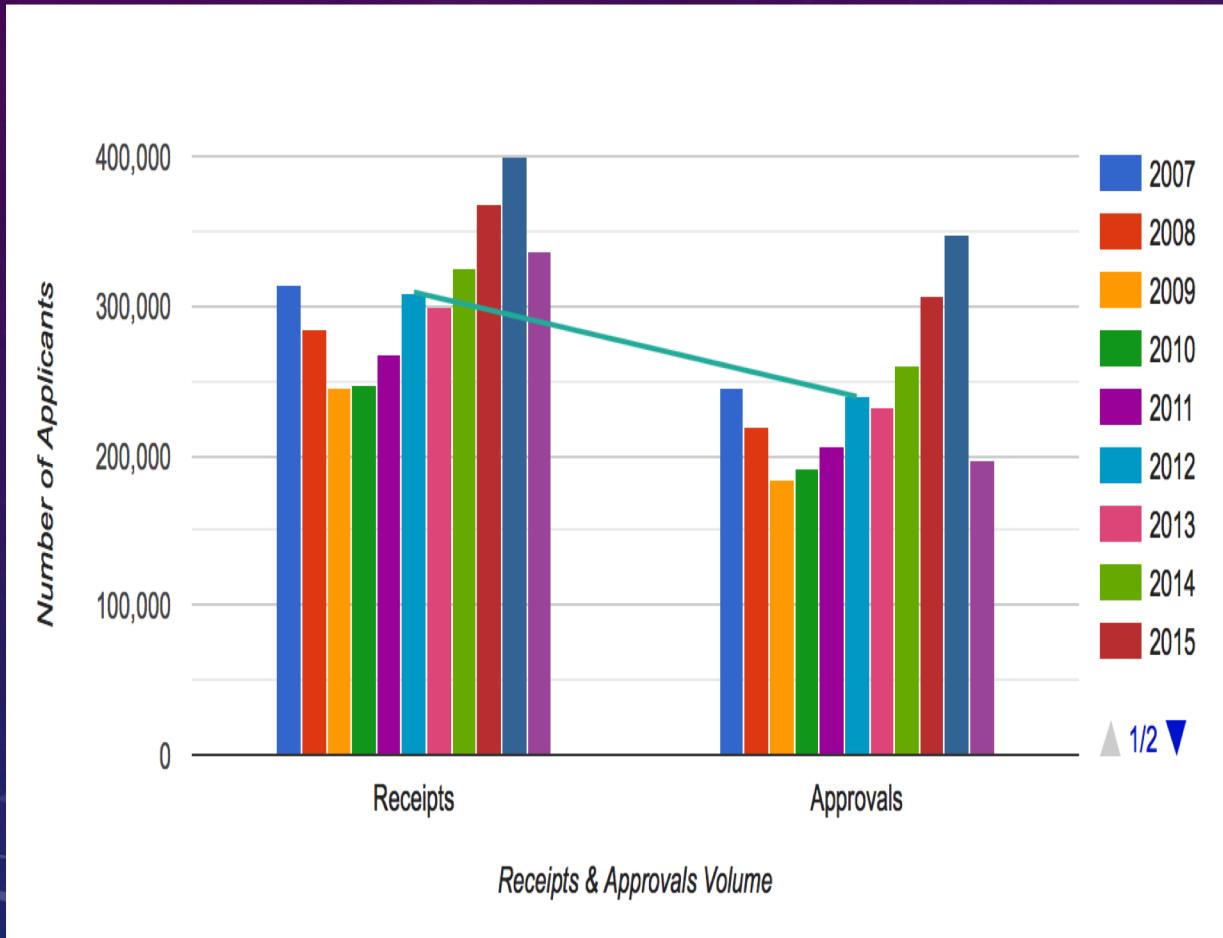
Data Source

- FY2011 to FY2016 from Kaggle
- FY2017 dataset from The Office of Foreign Labor Certification (OFLC)

Visualization Tool

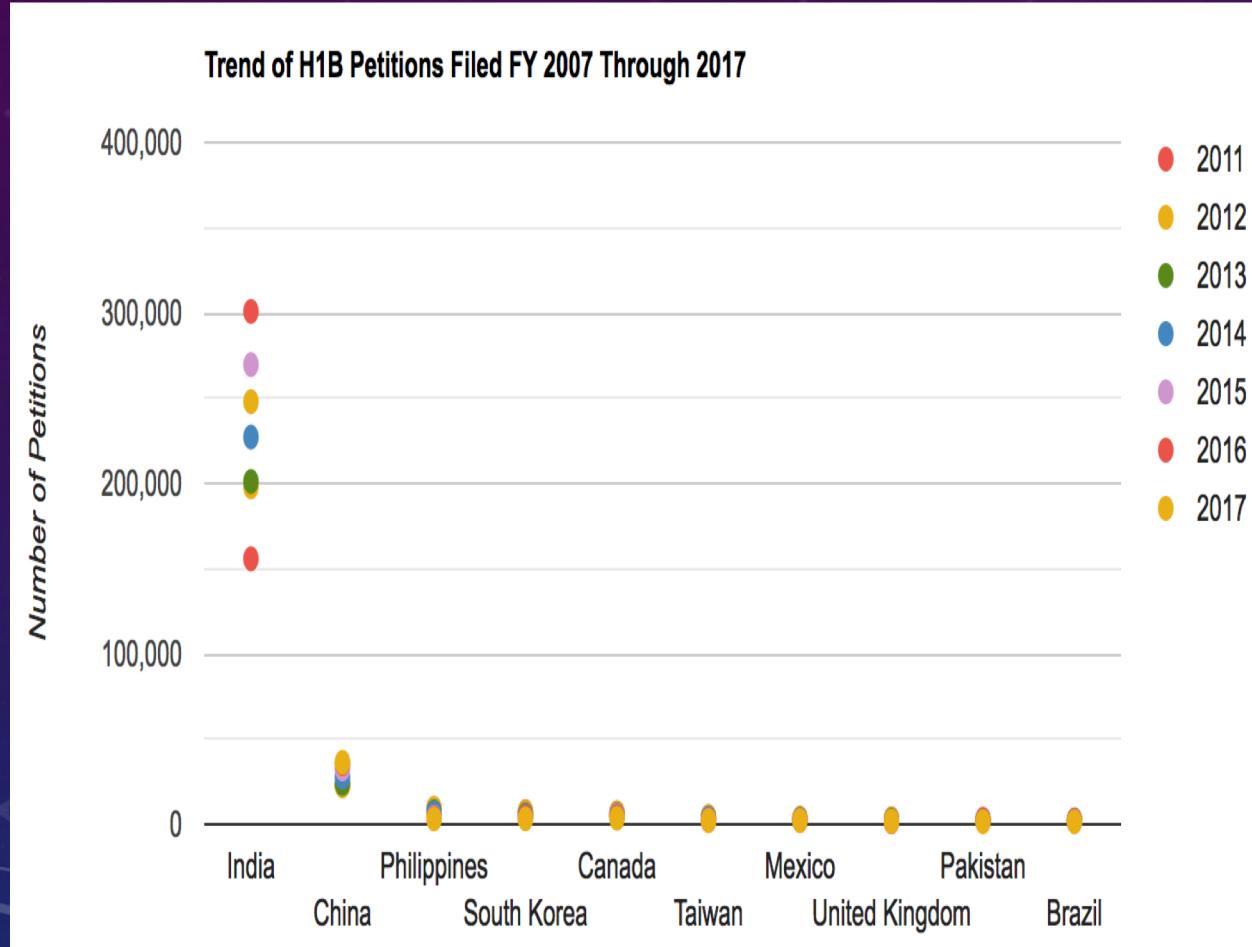
- Google API
- D3.js
- JavaScript
- HTML

OVERALL TREND OF H1B PETITIONS 2007 THROUGH 2017: RECEIPT VOLUME OVERVIEW



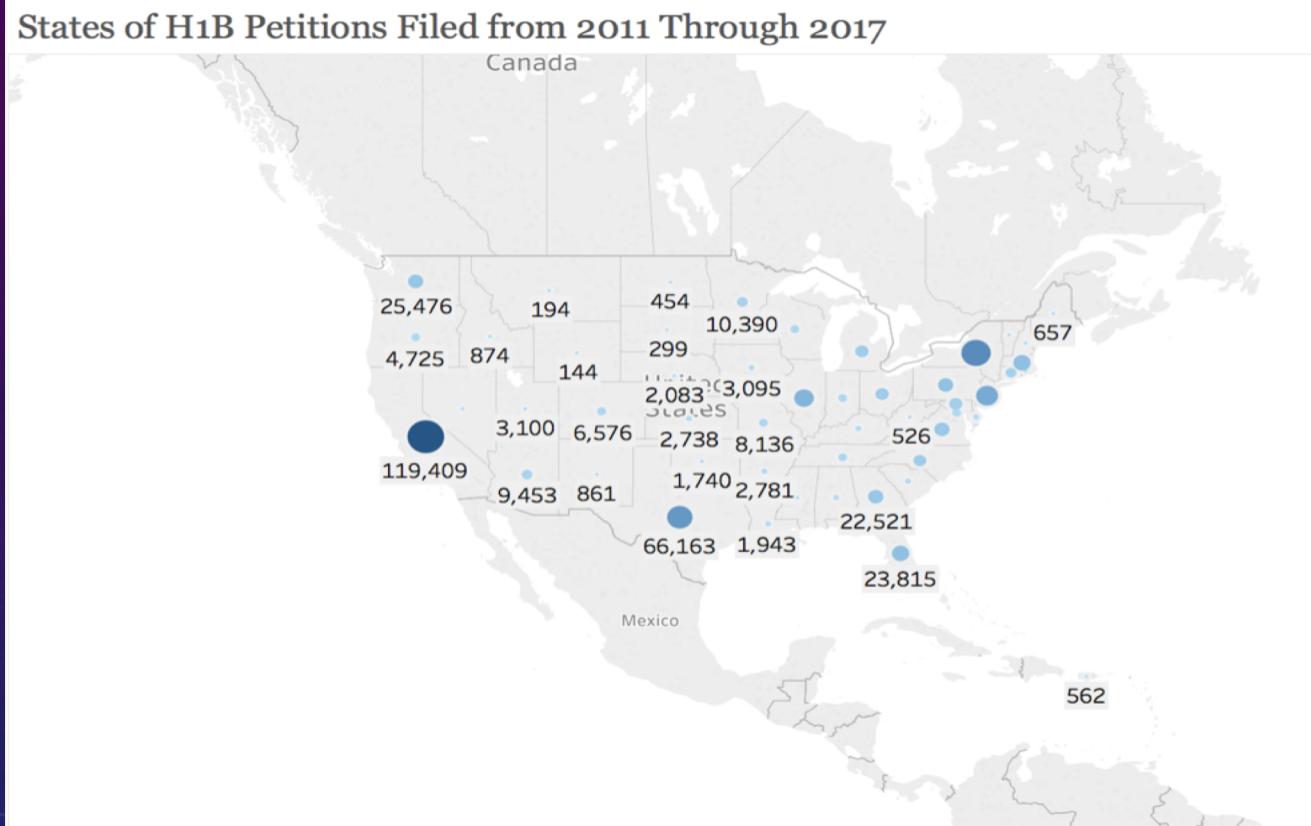
- The bar chart lists the petition status of Labor Condition Application (LCA) submitted by US employer for H1B Visa during last eleven years.
- From the additional research, the government is cutting down the number of approvals for H1B on 2017.
- As you can see the purple bar has a lower approvals rate compare to previous years.

TRENDS OF H1B PETITIONS FILED FROM 2011 TO 2017:BENEFICIARY COUNTRY OF BIRTH



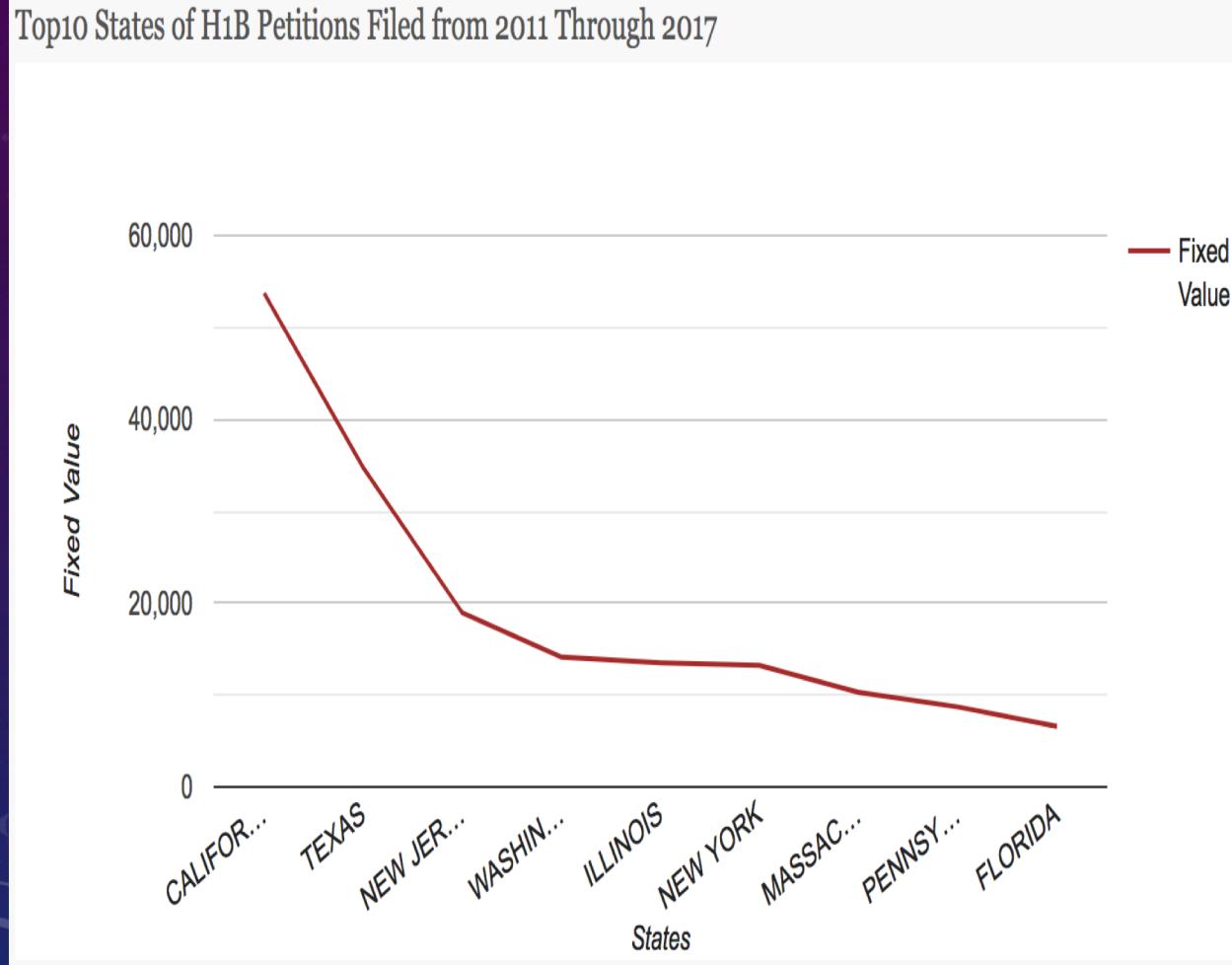
- In the past decade, due to the nature of demand for high-skilled workers, visa holders have clustered in STEM fields and come mostly from countries in Asia such as China and India.
- India has a highest volume of Petitions filed (247,927) on FY2017 and China also higher rate (36,362) compare to other countries.
- Volume has decreased from 2011 to 2017.

ALL STATES OF H1B PETITIONS FILED FROM 2011 THROUGH 2017



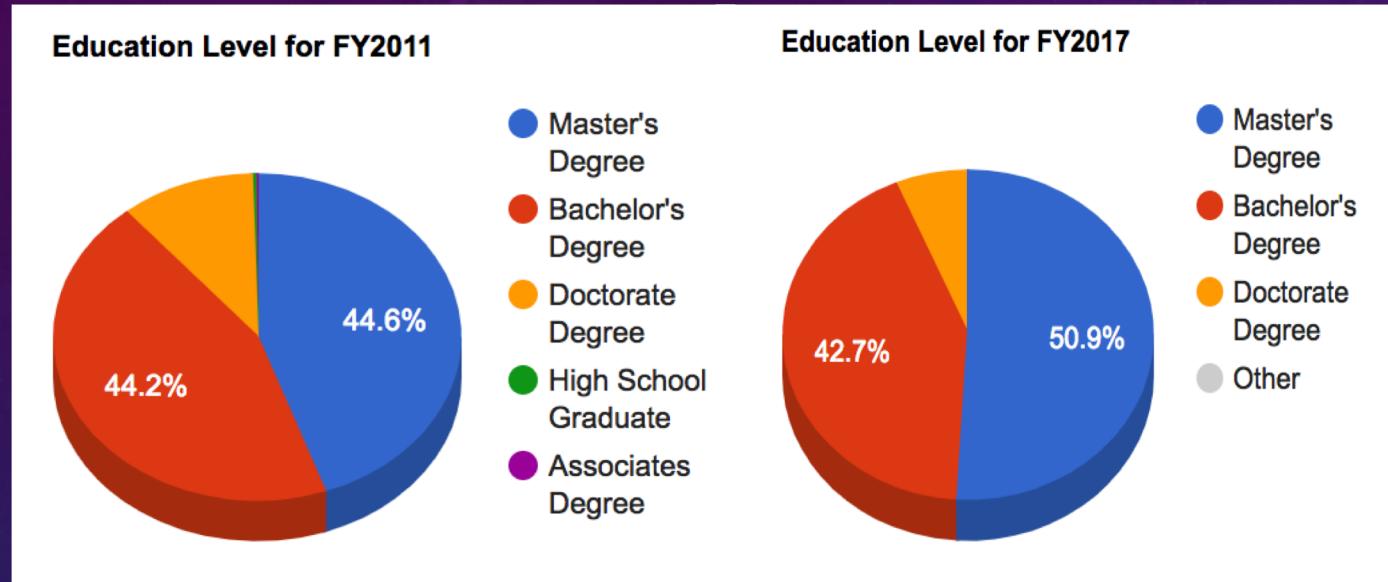
- Overall distribution across united states show that, Central region is comparatively less when compared to East, West or South regions,
- Increase chances of finding an employer that sponsor work-visa are more likely to be found in the darker regions of the map.

TOP 10 STATES OF H1B PETITIONS FILED



- California is the state with most H1B petitions because there are many technical companies.
- States like California, New York, Washington, New Jersey, Massachusetts, Illinois, and Texas are the prime location for foreign workers and provide many job opportunities.
- The main areas of specialization that employers intend to fill using foreign workers comprise of technical jobs with a great emphasis on STEM programs, such as engineering and computer science.

COMPARISON OF EDUCATION LEVEL FY2011 VS FY2017



Education level has been changed for some reasons.

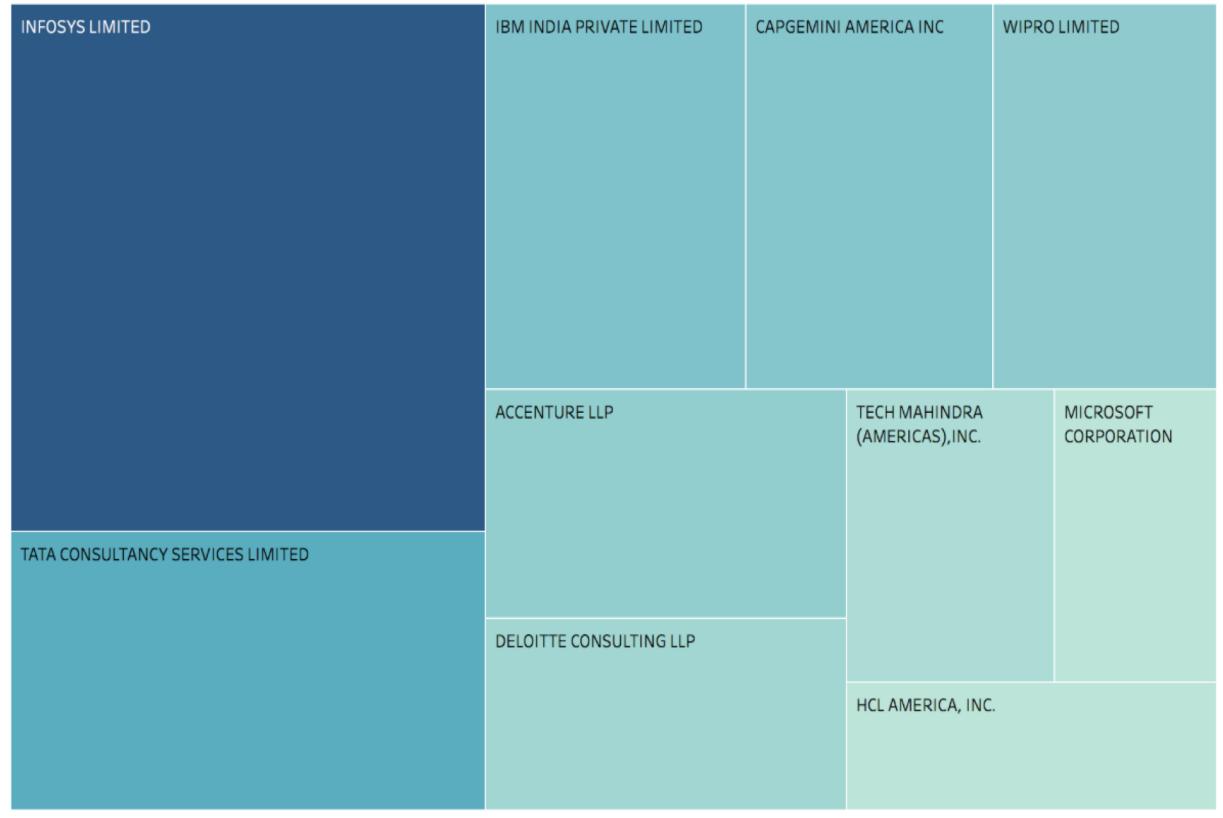
- Some job positions required higher level of Education.
- Tuition & Time

Master degree holders increased from 44.6% to 50.9%.

Bachelor's degree holders decreased.

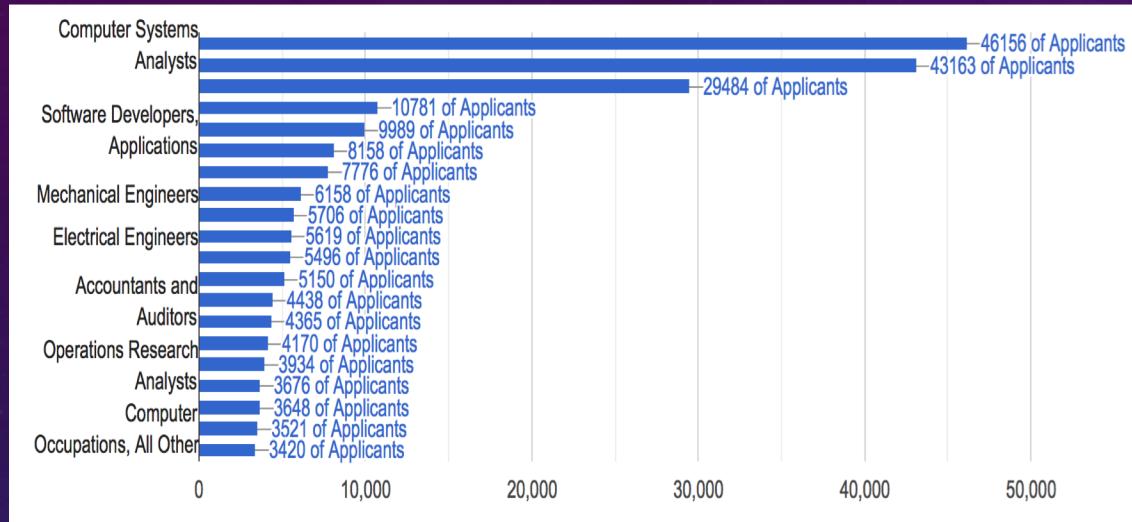
OVERALL TOP10 EMPLOYERS WHO SUBMIT THE GREATEST NUMBER OF APPLICATIONS DURING FY2011 TO FY2017

Overall Top10 Employers who submit the greatest number of applications during FY2011 to FY2017



- Through the treemap, INFOSYS LIMITED stands at the top followed by TCS, Wipro and others.
- INFOSYS LIMITED(n=69,267) has submitted highest H1B Visa Applications in last six years.
- TATA CONSULTANCY SERVICES LIMITED (n=36,654) also has high volume of applications submitted.
- Top three Companies that submit most H1B Visa Applications are companies based in India associated with software and IT services.

TOP 10 OCCUPATIONS, TECH VS BUSINESS



- Technical Jobs fill up the majority of Top 10 Jobs among foreign workers.
- For Examples, Accountant($n=4,438$) vs Computer Systems Analyst($n=46,156$)
- Computer Systems Analyst and Software Developers are the most popular occupation from the dataset.
- We will talk about more detail of Data Science field.

COMPARISON OF DATA ANALYST VS DATA SCIENTIST



Mean Wage of a DA
vs DS with Case
Status Certified

Mean Wage of a DA vs ..

CALIFORNIA	69,275
COLORADO	58,568
CONNECTICUT	63,981
DELAWARE	62,641
DISTRICT OF COLUMBIA	62,619
GEORGIA	69,332
KANSAS	62,913
MICHIGAN	60,487
MINNESOTA	59,236
NEW HAMPSHIRE	62,598
NEW JERSEY	61,018
NEW MEXICO	60,409
NEW YORK	59,740
NORTH CAROLINA	61,888
OHIO	59,650
OREGON	59,872
SOUTH DAKOTA	61,630
VERMONT	58,587
VIRGINIA	59,321
WASHINGTON	67,689

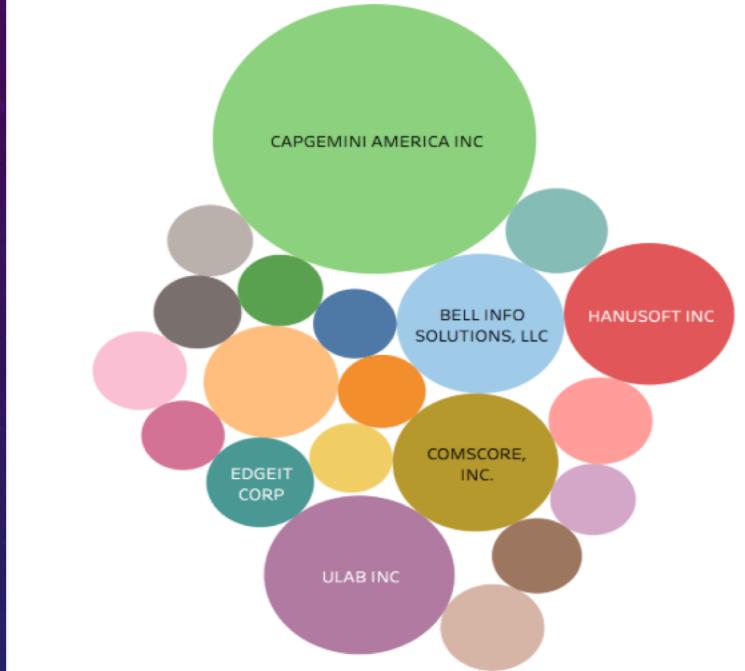
Mean Wage of a DS
vs DA with Case
Status Certified

Mean Wage of a DS vs ..

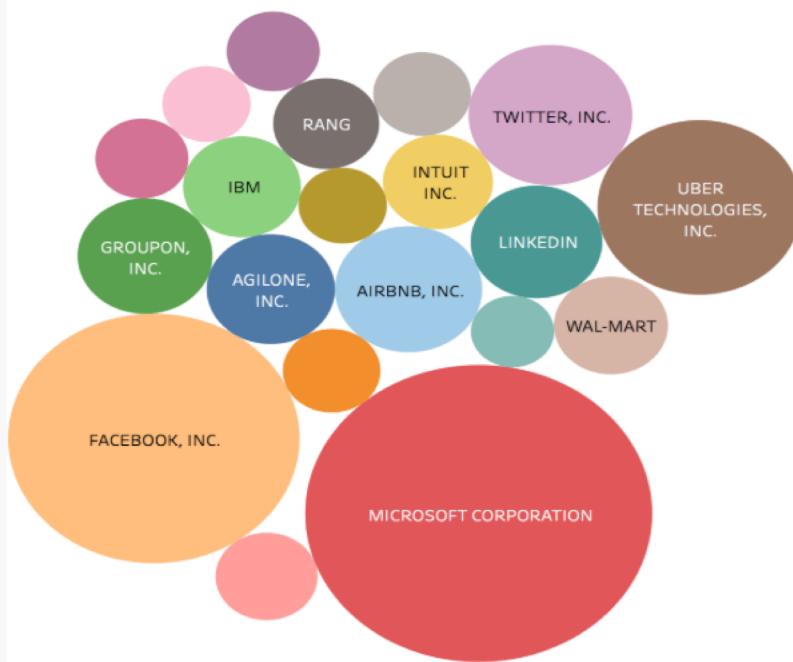
ARIZONA	78,558
CALIFORNIA	100,803
CONNECTICUT	80,928
DISTRICT OF COLUMBIA	90,996
HAWAII	74,880
KANSAS	79,893
MARYLAND	76,987
MASSACHUSETTS	78,981
MICHIGAN	74,191
MINNESOTA	78,173
MISSOURI	73,878
NEVADA	78,073
NEW JERSEY	82,281
NEW YORK	82,354
OREGON	84,150
PENNSYLVANIA	68,717
SOUTH CAROLINA	99,327
TEXAS	74,060
VIRGINIA	75,254
WASHINGTON	84,157

TOP 10 COMPANIES APPLY FOR H1B (DATA FIELD) DATA ANALYST VS DATA SCIENTIST

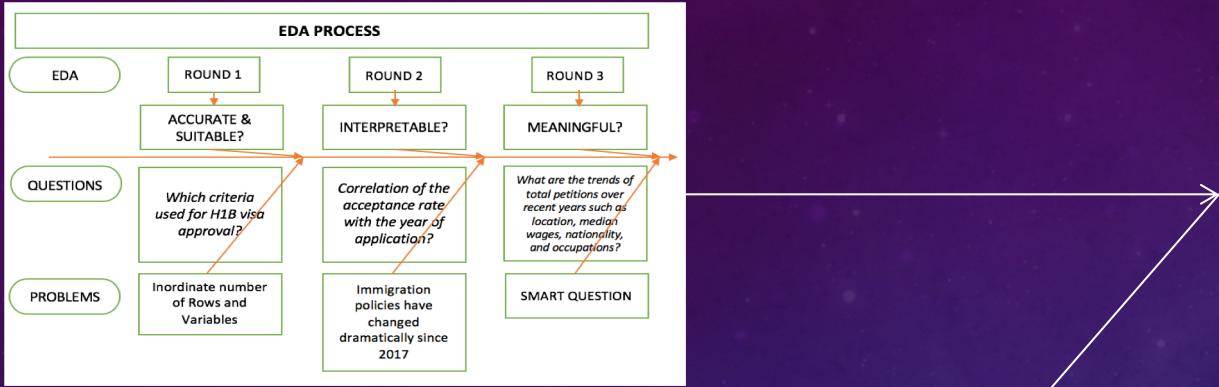
Top 10 Companies apply for H1B application for the Data Analyst



Top 10 Companies apply for H1B application for the Data Scientist



CODE REVIEW



EDA Process:
Explore to view the outline of the dataset

```
<body>
<div id="main">
<div id="header">
<div id="logo">
<div id="logo_text">
    <!-- class=logo_color allows changing of color -->
    <h1><a href="index.html">H1B Data Analysis Project <span class="logo_color"> </span></a></h1>
    <h2>visualization of Complex Data (DATS 6401)</h2>
</div><!-- logo_text -->
</div><!-- logo -->
<div id="menubar">
<ul id="menu">
    <!-- put class="selected" in the li tag -->
    <!-- for the selected page to highlight -->
    <li class="selected"><a href="index.html">Home</a></li>
    <li><a href="objectives.html">Objectives</a></li>
    <li><a href="RelatedResearch.html">Research</a></li>
    <li><a href="Visualization.html">Visualization</a></li>
    <li><a href="Conclusions.html">Conclusions</a></li>
    <li><a href="Contact.html">Contact</a></li>
    <li><a href="contact.html">Aboutme</a></li>
</ul>
</div>
```

To make appropriated Visualizations, variables should be Developed and compiled into visualization utilizing programs.

```
In [13]: #Top 10 Companies that apply for H1B application for the Position of Data Scientist
df2[df2['JOB_TITLE']=='DATA ANALYST']['EMPLOYER_NAME'].value_counts().head(20)

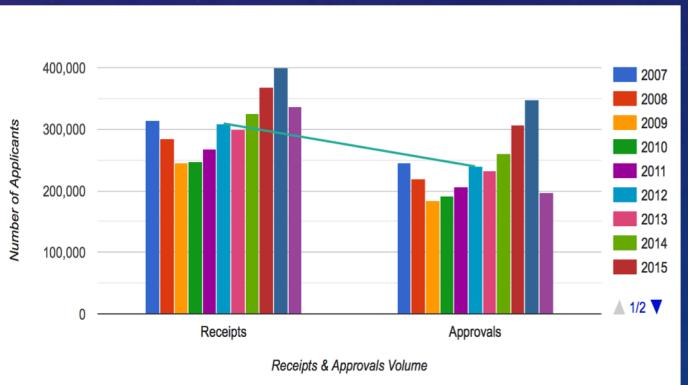
Out[13]:
```

EMPLOYER_NAME	Count
CAPGEMINI AMERICA INC	271
ULAN INC	94
HANUSOFT INC	75
BELL INFO SOLUTIONS, LLC	73
COMSCORE INC.	71
ALCOTRADING LP	47
EDGEIT CORP	30
VIGILANT TECHNOLOGIES LLC	28
IBM INDIA PRIVATE LIMITED	28
GROUPEWORLDWIDE LLC	27
TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA	23
VERISYS HEALTH, INC.	21
MICROINFO INC	20
NETSCOUT SOLUTIONS,LLC	20
VASTIKA, INC.	19
NEO PRISM SOLUTIONS LLC	19
CALIFORNIA CREATIVE SOLUTIONS, INC.	19
EDGE IT CORP	18
PROGRESSIVE CASUALTY INSURANCE COMPANY	18
SRIKETER INC	18

```
In [14]: #Creating a New Column State from Column WORKSITE
df2['STATE']=df2['WORKSITE'].apply(lambda x:x.split(',')[1])
df2.head() #Now a new column State has been added at the end

Out[14]:
```

Python for data cleaning and mining:
Doing some data preprocessing to transform the raw data into an understandable format.



Extract some aspects, and analyze the changes in employers' preference as well as forecasts for the future trends.

CONCLUSION FOR H1B DATA VISUALIZATION PROJECT

- The trends and measurements in this report provide vital information on the demand for technical jobs that are mostly filled by foreign workers.
- States like California, New York, Washington, New Jersey, Massachusetts, Illinois, and Texas are the prime location for foreign workers.
- Average Salary of Data Science related job in the current decade is higher than overall mean for Data Analyst that lands to foreign workers.
- From the additional research, the government is cutting down the number of approvals for H1B on 2017.
- For future forecasts for the H1B visas related with Data Science position, I found that the Data Scientist position has experiences an exponential growth in terms of H1B visa applications.
- Interestingly, the Data Scientist jobs are clustered in West region especially in San Francisco having the highest number.
- In Virginia, Amazon, Microsoft, and Facebook Data center are coming in next couple of years, so Virginia will be one of the States which has lot of tech job opportunities.

THANKS FOR YOUR PATIENCE.
LEAVE SOME YOUR THOUGHTS
AFTER THIS ANALYSIS.