

# Data Management & SQL

**Mr. Jaya Naga Venkata Satya Bobby Dangeti**

**4<sup>th</sup> November 2022**

### **Comprehensive Summary:**

The entire concept of this assignment is for us to work as a trio focusing on data gathering, collection, filtering, and presenting per the interesting such as which company possesses the highest acceptance rate in terms of H1B visas, which job in specific offers the highest rate of salary and lowest rate of salary. We used the concepts of normalization, connecting excel to other sources, uploaded the data through the command prompt, and finally ran the query.

**Host name:** sql-h1b.mysql.database.azure.com

**Username:** prof\_luis

**Password:** Hult@sql

**Video Database uploading link:** - <https://youtu.be/1vZDDMjpbM4>

### **Express Abstract and concepts:**

Our complete idea was to brainstorm the best possible conclusions from the data given. The concepts of server creation along with username and password were used. Data normalization was used to normalize the data. The concept of power bi is to filter out the columns and select what helps us in the findings which helps in the creation of schema on SQL. The concept of reverse engineering helped us deconstruct the software to distillate design information. Lastly command prompt one of the most used software was used to upload the database and further actions were performed.

### **Concise recommendations:**

The recommendations such as 45172 companies whose acceptance is ratio is 100 percent and top of it is Microsoft crop, Tesla and mastic digital technologies.

### **Insights and recommendations well connected:**

Our insights are to determine the highest acceptance rate in terms of H1B visas to find the final count of the accepted visas.

### **Facts and numbers from data analysis support the insights:**

By the data analysis, The total 56402 companies are filed for H1B visas and out of which 45172 has 100 percent acceptance rate, The top company Microsoft crop filed 9421 H1B visas and approved with 100 percent.

Second top is Tesla which filed 1099 H1B visas with 100 percent acceptance rate.

The least acceptance rate companies are 2116.

Host name: sql-h1b.mysql.database.azure.com

Username: prof\_luis

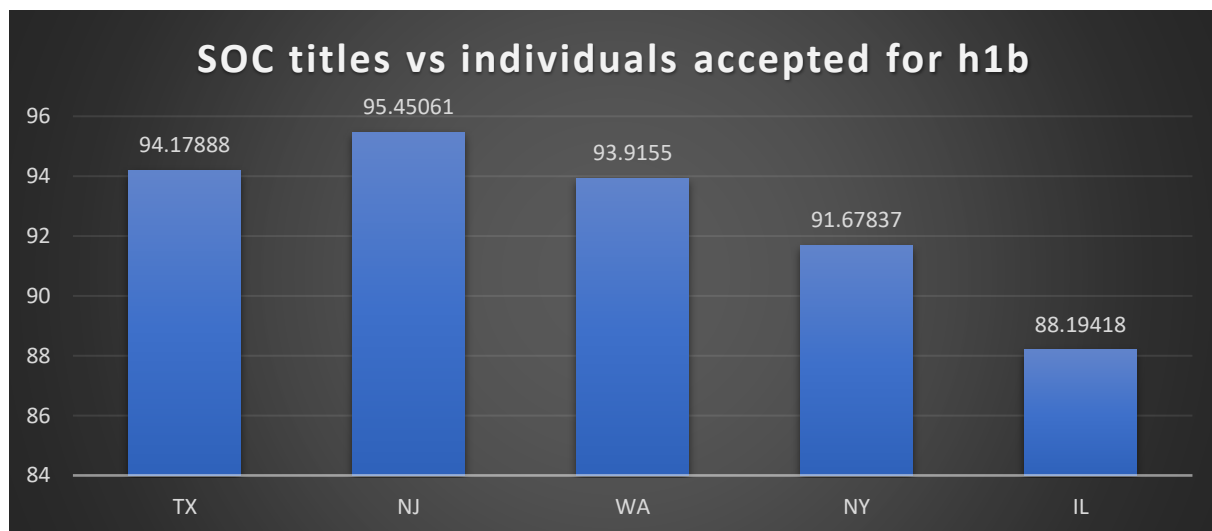
Password: Hult@sql

Video Database uploading link: - <https://youtu.be/1vZDDMjpbM4>

Based on an analysis, it has been determined that **Microsoft Corporation** has the most favorable approval rate and files the most H1B petitions out of all companies. Microsoft Corporation offers 54% of employment positions as **Software Engineers**, and the corresponding standard occupation classification is **Software Developers, Applications**. Aside from companies, the occupation with the highest acceptance rate for H1B visas is **Software Developers, Applications**. Software Engineers and Software Developers are the job titles that have the highest acceptance rate for H1B visas. Therefore, software and computer technology jobs and companies are ideal for individuals interested in obtaining an H1B visa.

**LexisNexis Risk Solutions FL**, Inc. offers the top salary of **\$7,721,498** with the highest acceptance rate for H1B visas. The company with the lowest salary offered and the highest acceptance rate is **Harmony Inc.** The salary is **\$23,504** per year.

#### Analysis with SOC:

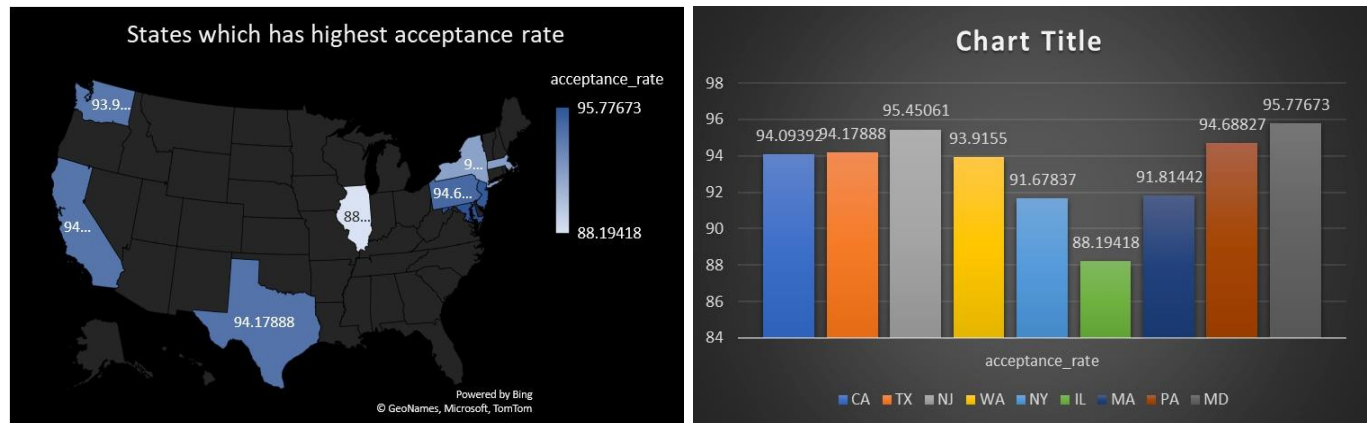


The chart depicts that the most of approval rates are for computer and software technology.

There are 116389 employers in the Software Developers, Applications category that have been approved for H1B visas, which is the highest of any SOC title.

In comparison to other SOC, the Applications SOC is more likely to approve H1Bs than any other SOC. Although there are a variety of job roles under this classification. For example, software engineer 1, software developer, etc.

### Analysis with state:



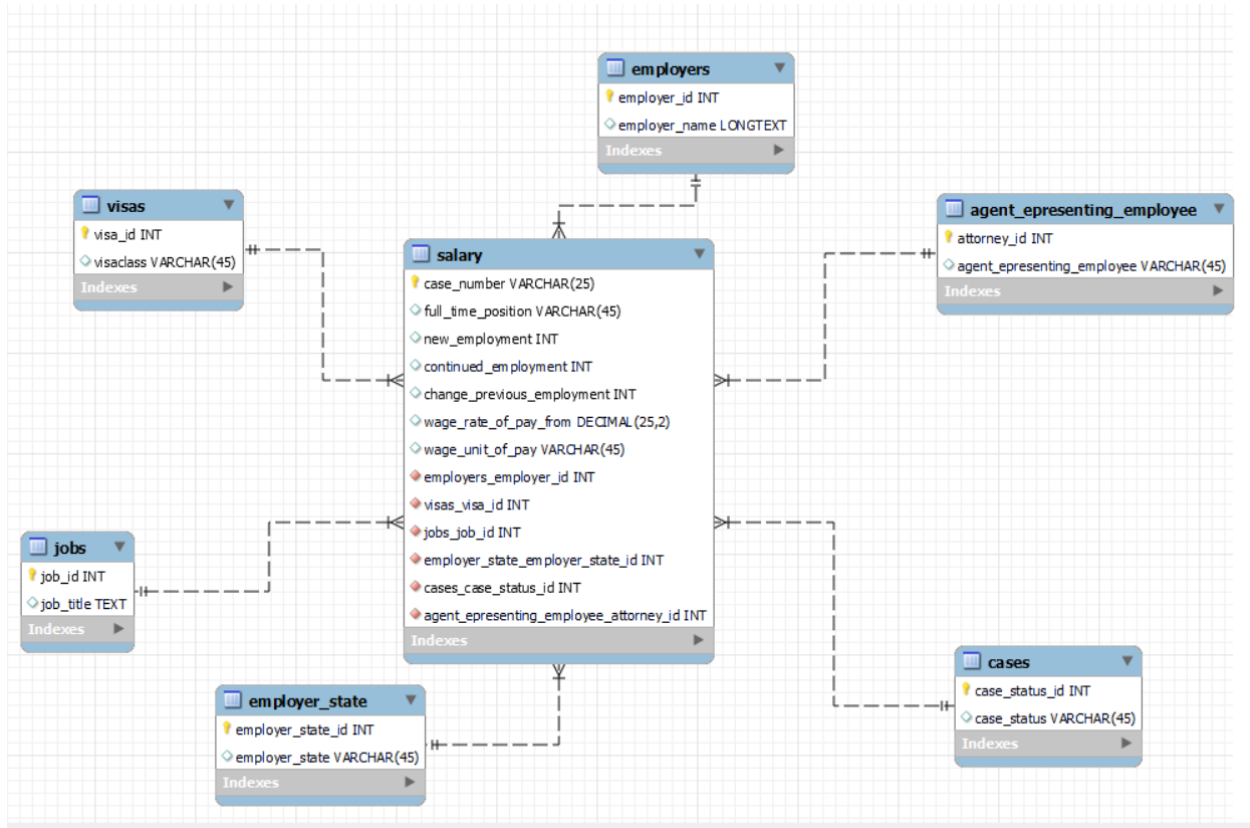
Based on the analysis, the acceptance rate varies by state. Among the states in the USA where employers file the most H1B visa petitions and have the highest H1B visa acceptance rates are MD (Maryland), NJ (New Jersey), CA (California), TX (Texas), MA (Massachusetts), etc. A petition filed with an agent by an employee is likely to be accepted 94% of the time.

### Recommendation:

As a non-immigrant, one of the best ways to obtain a H1B visa is to apply for a job in the field of software and computer technology which is extremely easy to approve. A number of technical companies favor H1b visas, including Microsoft Corporation, Tesla Inc., and Mastech Digital Technologies Inc. (These are the companies which have a high acceptance rate and file the majority of H1B visa petitions) with a job title focusing on software engineers and software developers. For part-time positions, the acceptance rate is very low. The acceptance rate also differs by state. The Maryland state has a high acceptance rate and files the most H1B visa petitions in comparison with any other state.

# Appendix:

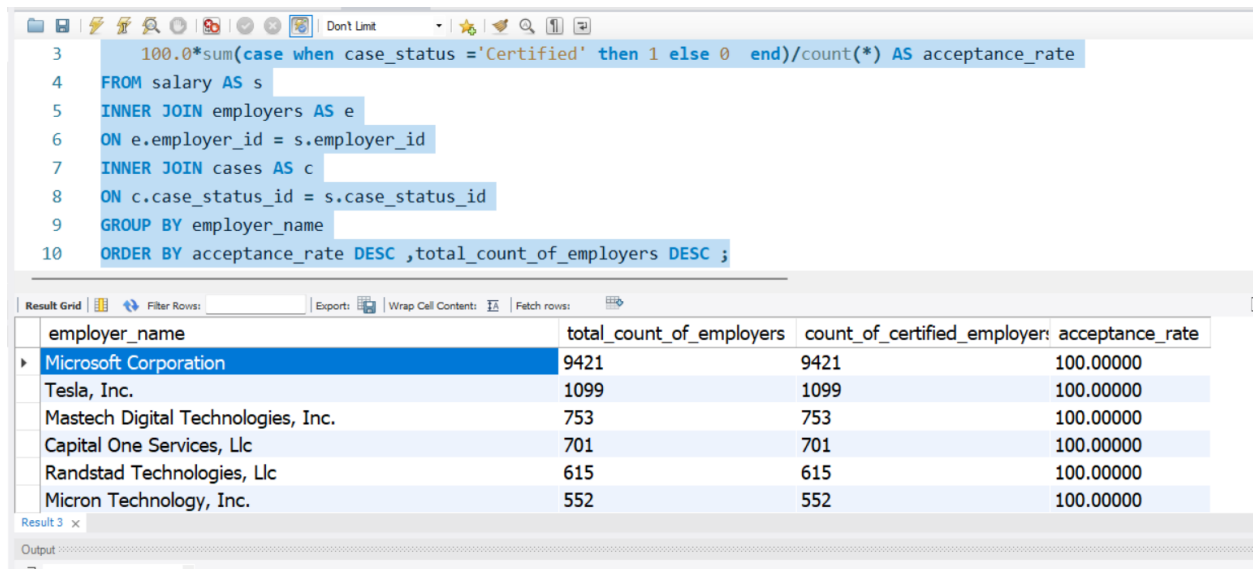
## Schema:



## Code for finding highest accepted h1b visas employer name :

```
select employer_name, COUNT(employer_name) AS total_count_of_employers,  
       sum(case when case_status = 'Certified' then 1 else 0 end) AS count_of_certified_employers,  
       100.0*sum(case when case_status = 'Certified' then 1 else 0 end)/count(*) AS acceptance_rate  
FROM salary AS s  
INNER JOIN employers AS e  
ON e.employer_id = s.employer_id  
INNER JOIN cases AS c  
ON c.case_status_id = s.case_status_id  
GROUP BY employer_name
```

ORDER BY acceptance\_rate DESC ,total\_count\_of\_employers DESC ;



The screenshot shows a SQL query in a text editor and its corresponding results in a grid. The query calculates the acceptance rate for employers based on certified cases. The results grid shows the top employers with an acceptance rate of 100.00000.

```
3      100.0*sum(case when case_status = 'Certified' then 1 else 0 end)/count(*) AS acceptance_rate
4 FROM salary AS s
5 INNER JOIN employers AS e
6 ON e.employer_id = s.employer_id
7 INNER JOIN cases AS c
8 ON c.case_status_id = s.case_status_id
9 GROUP BY employer_name
10 ORDER BY acceptance_rate DESC ,total_count_of_employers DESC ;
```

employer_name	total_count_of_employers	count_of_certified_employers	acceptance_rate
Microsoft Corporation	9421	9421	100.00000
Tesla, Inc.	1099	1099	100.00000
Mastech Digital Technologies, Inc.	753	753	100.00000
Capital One Services, Llc	701	701	100.00000
Randstad Technologies, Llc	615	615	100.00000
Micron Technology, Inc.	552	552	100.00000

**Code for finding state which has highest acceptance rate:**

Select employer\_state , sum(case when case\_status ='Certified' then 1 else 0 end) AS  
count\_of\_certified\_employers,

100.0\*sum(case when case\_status ='Certified' then 1 else 0 end)/count(\*) AS acceptance\_rate

FROM salary AS s

INNER JOIN employer\_state AS emp

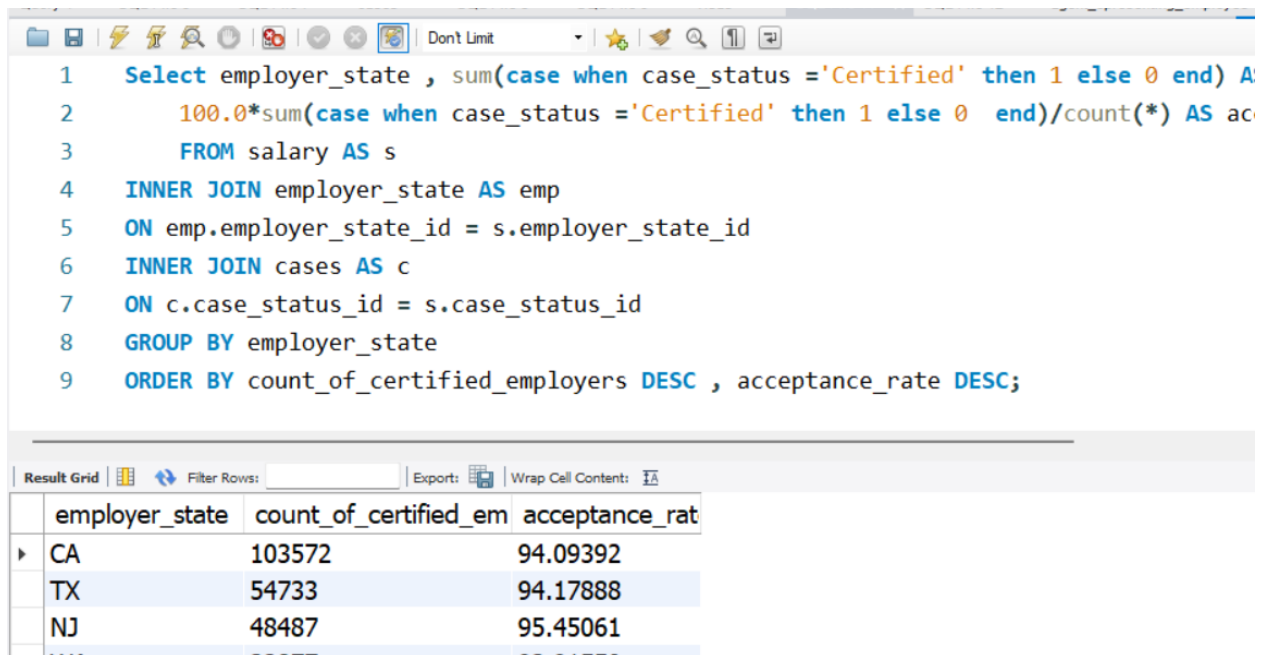
ON emp.employer\_state\_id = s.employer\_state\_id

INNER JOIN cases AS c

ON c.case\_status\_id = s.case\_status\_id

GROUP BY employer\_state

ORDER BY count\_of\_certified\_employers DESC , acceptance\_rate DESC;



The screenshot shows a SQL IDE window with a query editor and a results grid. The query is a SQL statement that calculates the acceptance rate for certified employers by state. The results grid shows the output of the query, with columns for employer\_state, count\_of\_certified\_em, and acceptance\_rat.

```

1  Select employer_state , sum(case when case_status = 'Certified' then 1 else 0 end) AS
2      100.0*sum(case when case_status = 'Certified' then 1 else 0 end)/count(*) AS ac
3      FROM salary AS s
4      INNER JOIN employer_state AS emp
5      ON emp.employer_state_id = s.employer_state_id
6      INNER JOIN cases AS c
7      ON c.case_status_id = s.case_status_id
8      GROUP BY employer_state
9      ORDER BY count_of_certified_employers DESC , acceptance_rate DESC;

```

employer_state	count_of_certified_em	acceptance_rat
CA	103572	94.09392
TX	54733	94.17888
NJ	48487	95.45061
WA	33377	93.81550

#### **Code for finding percentage of acceptance rate with agent representing attorney:**

Select agent\_epresenting\_employee , sum(case when case\_status ='Certified' then 1 else 0 end) AS  
count\_of\_certified\_employers,

100.0\*sum(case when case\_status ='Certified' then 1 else 0 end)/count(\*) AS acceptance\_rate

FROM salary AS s

INNER JOIN agent\_epresenting\_employee AS a

ON a.attorney\_id = s.agent\_representing\_employer

INNER JOIN cases AS c

ON c.case\_status\_id = s.case\_status\_id

GROUP BY agent\_epresenting\_employee;

```
Query 1  SQL File 5*  SQL File 7*  cases  SQL File 8*  SQL File 9*  visas  SQL File 11*
[Icons] Don't Limit
1 • Select agent_epresenting_employee , sum(case when case_sta
2     100.0*sum(case when case_status ='Certified' then 1 el
3     FROM salary AS s
4     INNER JOIN agent_epresenting_employee AS a
5     ON a.attorney_id = s.agent_representing_employer
6     INNER JOIN cases AS c
7     ON c.case_status_id = s.case_status_id
8     GROUP BY agent_epresenting_employee;
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	agent_epresenting_employee	count_of_certified_em	acceptance_rat		
▶	No	120583	90.28377		
	Yes	353930	93.78536		

Used Pivot for analysis of wage to get easier:



B6		23504	
	A	B	C
1			
2	WAGE_UNIT_OF_PAY	Year	
3			
4	Max of WAGE_RATE_OF_PAY_FROM	Column Labels	
5	Row Labels	Certified	Grand Total
6	Highest Harmony Inc	23504	23504
7	Mount Carmel School, Inc.	23562	23562
8	Shazia Zafar MD LLC	24149	24149
9	West side alliance soccer club	24500	24500
10	Missouri State Univeristy	25760	25760
11	PINCC CORP	26320	26320
12	Sweet Tree Arts	26520	26520
13	BABY BOOMERS INTERNATIONAL PRESCHOOL INC	26582	26582
14	BABY BOOMERS INTERNATIONAL PRESCHOOL	26582	26582
15	CENTER FOR NATIONAL INDEPENDENCE IN POLITICS	26811	26811
16	Fort Scott Community College	26930	26930
17	Coyote O&P Inc	27000	27000
18	Amendment 2 Consultants, LLC.	27500	27500
19	Colegio Santa Gema, Inc.	29270	29270
20	Midwest University	29670	29670
21	Globridge Group Corporation	29786	29786
22	ARTLABBE USA	30000	30000
23	Calvary Baptist Church Lansdale PA	30000	30000
24	Bossier Parish Community College	30000	30000
25	Fleet of North Miami & Auto Service Inc.	30000	30000
26	Golden Land Ginseng Inc.	30000	30000

Pivot tabs for soc:

CASE_STATUS	Certified	
Row Labels	Count of CASE_NUMBER	
Software Developers, Applications	153431	
Software Developers, Systems Software	30599	
Computer Systems Analysts	25984	
Computer Systems Engineers/Architects	17387	
Information Technology Project Managers	11455	
Software Quality Assurance Engineers and Testers	11390	
Computer and Information Systems Managers	11134	
Business Intelligence Analysts	11101	
Operations Research Analysts	9194	
Mechanical Engineers	8275	
Computer Programmers	7610	
Computer Occupations, All Other	7584	
Electronics Engineers, Except Computer	7393	
Statisticians	7330	
Financial Analysts	7001	
Electrical Engineers	5949	
Accountants and Auditors	5932	
Management Analysts	5802	
Database Administrators	4839	
Industrial Engineers	4465	
Computer and Information Research Scientists	4286	
Information Security Analysts	3961	
Market Research Analysts and Marketing Specialists	3877	

### PivotTable Fields

Choose fields to add to report:

Search

- ☒ CASE\_NUMBER
- ☒ CASE\_STATUS
- ☐ RECEIVED\_DATE
- ☐ DECISION\_DATE
- ☐ ORIGINAL\_CERT\_DATE
- ☐ VISA\_CLASS
- ☐ JOB\_TITLE
- ☐ SOC\_CODE

Drag fields between areas below:

<b>Filters</b> CASE_STATUS	<b>Columns</b> 
<b>Rows</b> SOC_TITLE	<b>Values</b> Count of CASE_NUMBER

☐ Defer Layout Update Update

**Graphs are done in excel :**

**State vs acceptance Data:( graph 2 & graph 3)**

employer\_state acceptance\_rate

CA	94.09392
TX	94.17888
NJ	95.45061
WA	93.9155
NY	91.67837
IL	88.19418
MA	91.81442
PA	94.68827
MD	95.77673

**Soc vs approval Data: ( graph 1)**

Software Developers, Applications	153431
Software Developers, Systems Software	30599
Computer Systems Analysts	25984
Computer Systems Engineers/Architects	17387
Information Technology Project Managers	11455