

CIS 6930-002: Data Visualization (Spring 2018)

Project 1: Visualizing Data with Tableau

This project is expected to familiarize us with the Information Visualization system, Tableau. To learn the capabilities provided by Tableau, dataset from [Kaggle](#) is chosen, which is about [H-1B Visa Petitions](#) for the period 2011-2016. Following is excerpt of how the data is stored in csv file.

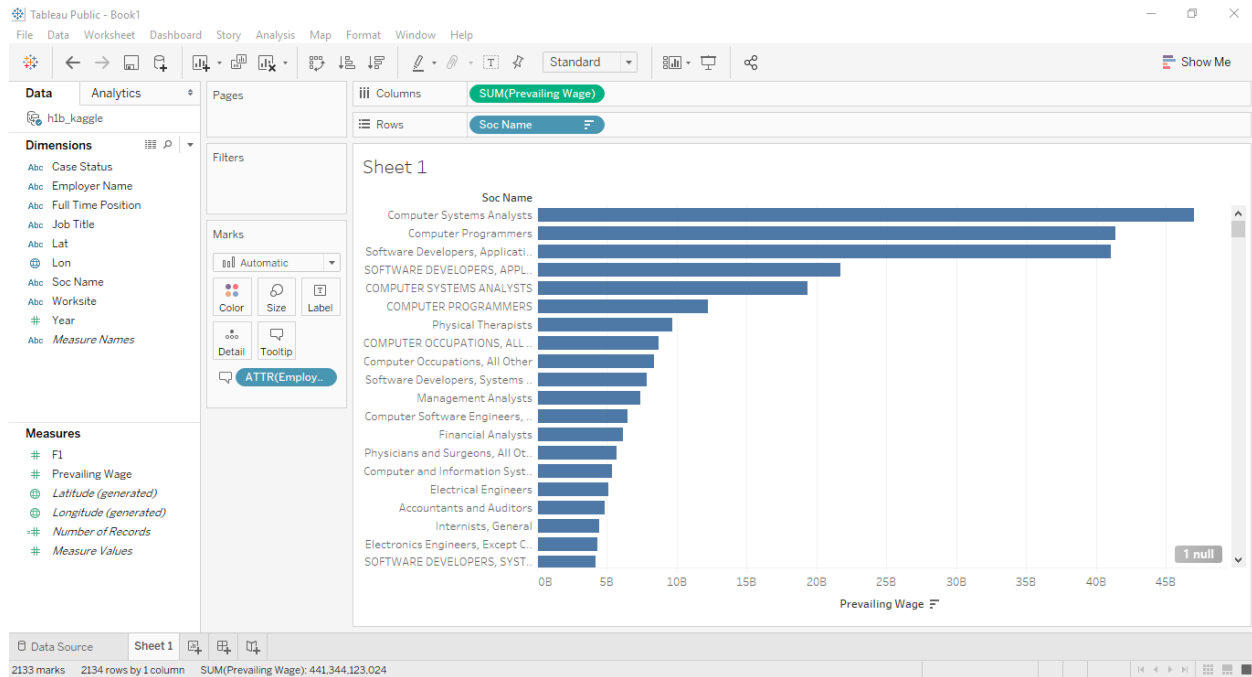
	CASE_STATUS	EMPLOYER_NAME	SOC_NAME	JOB_TITLE	FULL_TIME_POSIT	PREVAILING_WAC YEAR	WORKSITE	lon	lat
1	CERTIFIED-WITHDRAWN	UNIVERSITY OF MICHIGAN	BIOCHEMISTS AND BIOPHY	POSTDOCTORAL RESEARCH FELLOW	N	36067	2016 ANN ARBOR, MICHIGAN	-83.7430378	42.28083
2	CERTIFIED-WITHDRAWN	GOODMAN NETWORKS, INC.	CHIEF EXECUTIVES	CHIEF OPERATING OFFICER	Y	242674	2016 PLANO, TEXAS	-96.6988856	33.01984
3	CERTIFIED-WITHDRAWN	PORTS AMERICA GROUP, INC.	CHIEF EXECUTIVES	CHIEF PROCESS OFFICER	Y	193066	2016 JERSEY CITY, NEW JERSEY	-74.0776417	40.72816
4	CERTIFIED-WITHDRAWN	GATES CORPORATION, A WHOLLY-OWN	CHIEF EXECUTIVES	REGIONAL PRESIDEN, AMERICAS	Y	220314	2016 DENVER, COLORADO	-104.990251	39.73924
5	WITHDRAWN	PEABODY INVESTMENTS CORP.	CHIEF EXECUTIVES	PRESIDENT MONGOLIA AND INDIA	Y	157518.4	2016 ST. LOUIS, MISSOURI	-90.1994042	38.627
6	CERTIFIED-WITHDRAWN	BURGER KING CORPORATION	CHIEF EXECUTIVES	EXECUTIVE V P, GLOBAL DEVELOPMENT AND PRES	Y	225000	2016 MIAMI, FLORIDA	-80.1917902	25.76168
7	CERTIFIED-WITHDRAWN	BT AND MK ENERGY AND COMMODITIE	CHIEF EXECUTIVES	CHIEF OPERATING OFFICER	Y	91021	2016 HOUSTON, TEXAS	-95.3698028	29.76043
8	CERTIFIED-WITHDRAWN	GLOBO MOBILE TECHNOLOGIES, INC.	CHIEF EXECUTIVES	CHIEF OPERATIONS OFFICER	Y	150000	2016 SAN JOSE, CALIFORNIA	-121.886329	37.33821
9	CERTIFIED-WITHDRAWN	ESI COMPANIES INC.	CHIEF EXECUTIVES	PRESIDENT	Y	127546	2016 MEMPHIS, TEXAS	NA	NA
10	WITHDRAWN	LESSARD INTERNATIONAL LLC	CHIEF EXECUTIVES	PRESIDENT	Y	154648	2016 VIENNA, VIRGINIA	-77.2652604	38.90122
11	CERTIFIED-WITHDRAWN	H.J. HEINZ COMPANY	CHIEF EXECUTIVES	CHIEF INFORMATION OFFICER, HEINZ NORTH AM Y	Y	182978	2016 PITTSBURGH, PENNSYLVANI	-79.9958864	40.44062
12	CERTIFIED-WITHDRAWN	DOW CORNING CORPORATION	CHIEF EXECUTIVES	VICE PRESIDENT AND CHIEF HUMAN RESOURCES	Y	163717	2016 MIDLAND, MICHIGAN	-84.2472116	43.61558
13	CERTIFIED-WITHDRAWN	ACUSHNET COMPANY	CHIEF EXECUTIVES	TREASURER AND COO	Y	203860.8	2016 FAIRHAVEN, MASSACHUSET NA	NA	NA
14	CERTIFIED-WITHDRAWN	BIOCAIR, INC.	CHIEF EXECUTIVES	CHIEF COMMERCIAL OFFICER	Y	252637	2016 MIAMI, FLORIDA	-80.1917902	25.76168
15	CERTIFIED-WITHDRAWN	NEWMONT MINING CORPORATION	CHIEF EXECUTIVES	BOARD MEMBER	Y	105914	2016 GREENWOOD VILLAGE, COLU	-104.950814	39.61721
16	CERTIFIED-WITHDRAWN	VRICON, INC.	CHIEF EXECUTIVES	CHIEF FINANCIAL OFFICER	Y	153046	2016 STERLING, VIRGINIA	-77.4291298	39.0067
17	CERTIFIED-WITHDRAWN	CARDIAC SCIENCE CORPORATION	FINANCIAL MANAGERS	VICE PRESIDENT OF FINANCE	Y	90834	2016 WAUKESHA, WISCONSIN	-88.2314813	43.01168
18	CERTIFIED-WITHDRAWN	WESTFIELD CORPORATION	CHIEF EXECUTIVES	GENERAL MANAGER, OPERATIONS	Y	164050	2016 LOS ANGELES, CALIFORNIA	-118.243685	34.05223
19	CERTIFIED	QUICKLOGIX LLC	CHIEF EXECUTIVES	CEO	Y	187200	2016 SANTA CLARA, CALIFORNIA	-121.955236	37.35411
20	CERTIFIED	MCCHRYSTAL GROUP, LLC	CHIEF EXECUTIVES	PRESIDENT, NORTHEAST REGION	Y	241842	2016 ALEXANDRIA, VIRGINIA	-77.0469214	38.80484
21	CERTIFIED-WITHDRAWN	CUDDLE BARN, INC.	CHIEF EXECUTIVES	CHIEF OPERATING OFFICER (COO)	Y	117998	2016 COMMERCE, CALIFORNIA	-118.159793	34.00057
22	CERTIFIED-WITHDRAWN	WESTFIELD CORPORATION	CHIEF EXECUTIVES	GENERAL MANAGER, OPERATIONS	Y	164050	2016 LOS ANGELES, CALIFORNIA	-118.243685	34.05223
23	CERTIFIED	LOMICS, LLC	CHIEF EXECUTIVES	CEO	Y	99986	2016 SAN DIEGO, CALIFORNIA	-117.161084	32.71574
24	CERTIFIED	UC UNIVERSITY HIGH SCHOOL EDUCATI	CHIEF EXECUTIVES	CHIEF FINANCIAL OFFICER	Y	99986	2016 CHULA VISTA, CALIFORNIA	-117.084196	32.64005
25	CERTIFIED-WITHDRAWN	VMS COMMUNICATIONS LLC	CHIEF EXECUTIVES	CHIEF OPERATING OFFICER	Y	159370	2016 MIAMI, FLORIDA	-80.1917902	25.76168
26	CERTIFIED	QUICKLOGIX, INC.	CHIEF EXECUTIVES	CEO	Y	187200	2016 SANTA CLARA, CALIFORNIA	-121.955236	37.35411
27	CERTIFIED-WITHDRAWN	BOONREPRESENTATIVE CORPORATION	CHIEF EXECUTIVES	CHIEF EXECUTIVE OFFICER	Y	130823	2016 CHICAGO, ILLINOIS	-87.6297003	41.87911

In the above excerpt it may be observed that, for a specific petition, its case status, employer name, standard occupational system name, job title, prevailing wages ang worksite etc. has been stored. In an attempt to visualize this data better and answer few question pertained to it, Tableau is used and following three queries have been satisfied,

- Which industry or SOC_Name has been offered most prevailing wages during the period 2011-2016?

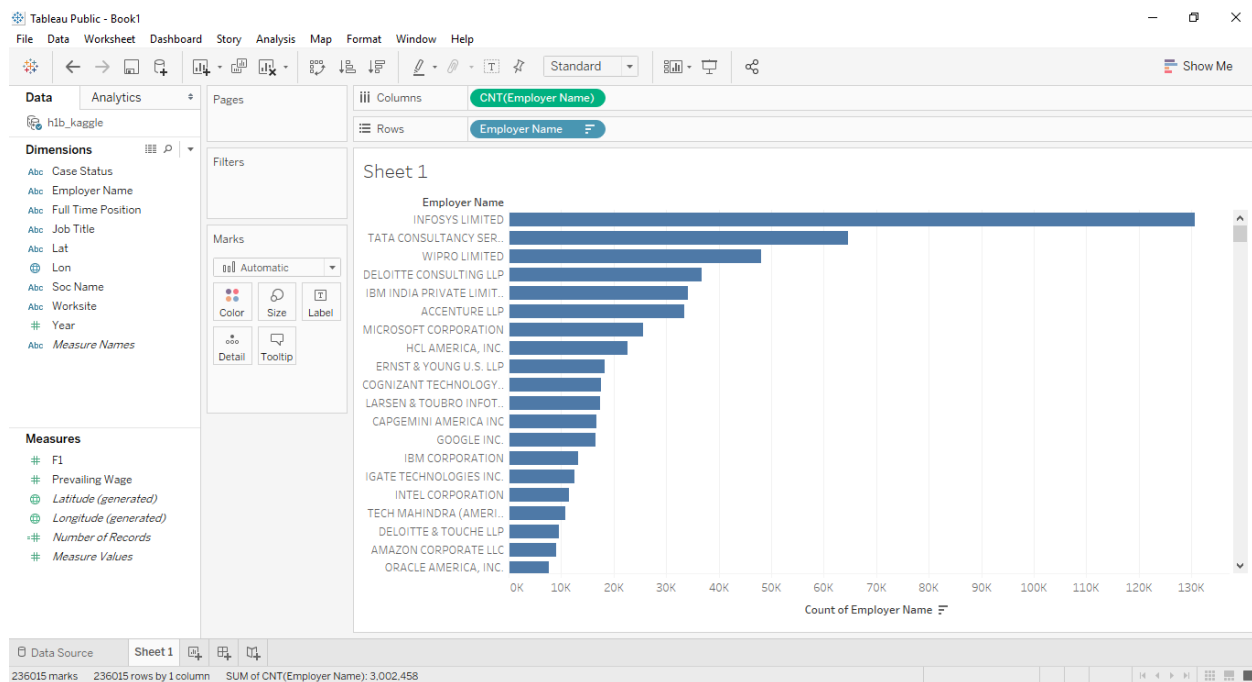
In order to find out which SOC_Name tag has been offered most wages in 6 years from the given data, focus will be on using Soc Name as dimensions alongside varying prevailing wages. To obtain horizontal bars representing sum of all the prevailing wages associated with that specific Soc Name, SUM of prevailing wages has been taken. Once that result is obtained, we may change the view in descending order. This provides an insight letting us know the Soc name for which maximus prevailing wages have been offered in 6 years i.e., Computer System

Analysts have been offered a total of \$47B over the period of 6 years. Following is the horizontal bar graph representing the same.



The next question which is to be addressed is,

- Which employer has filed most petitions?



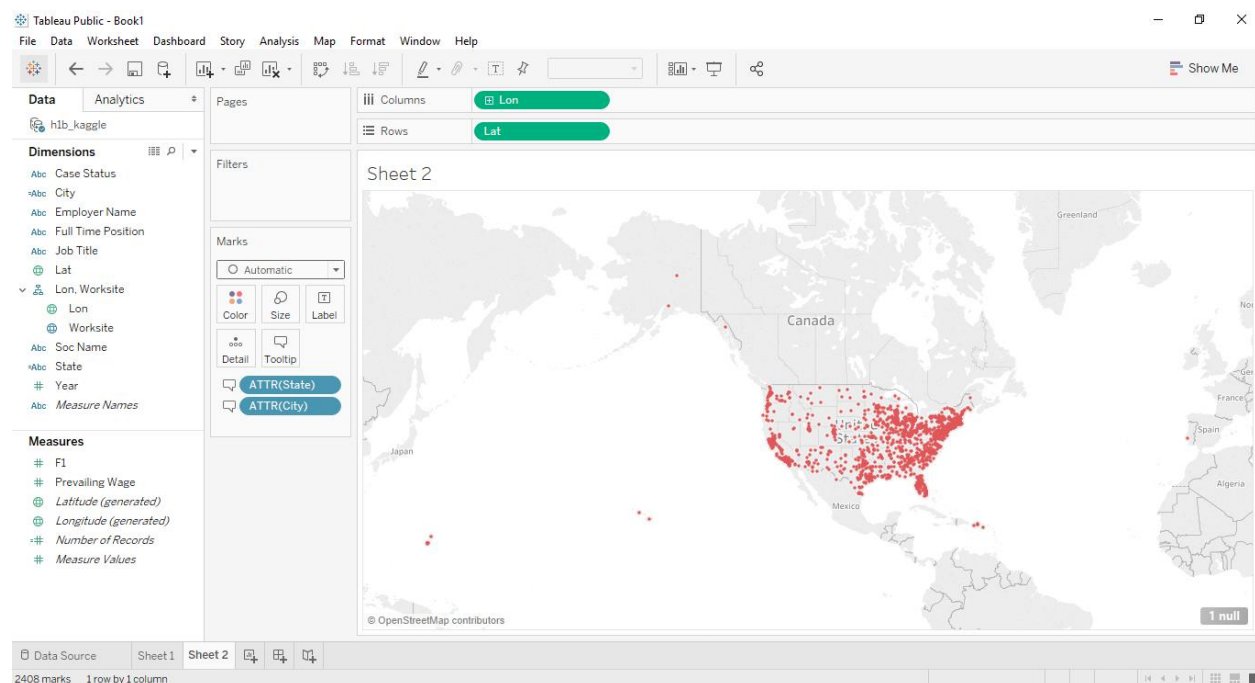
To find out which employer has filed most number of petitions, first total number of petitions are calculated using SUM function on employer name. And further, it is plotted against each employer. This gives a horizontal bar graph representing number of petitions filed by each employer.

Finally, the third and last question to be solved is,

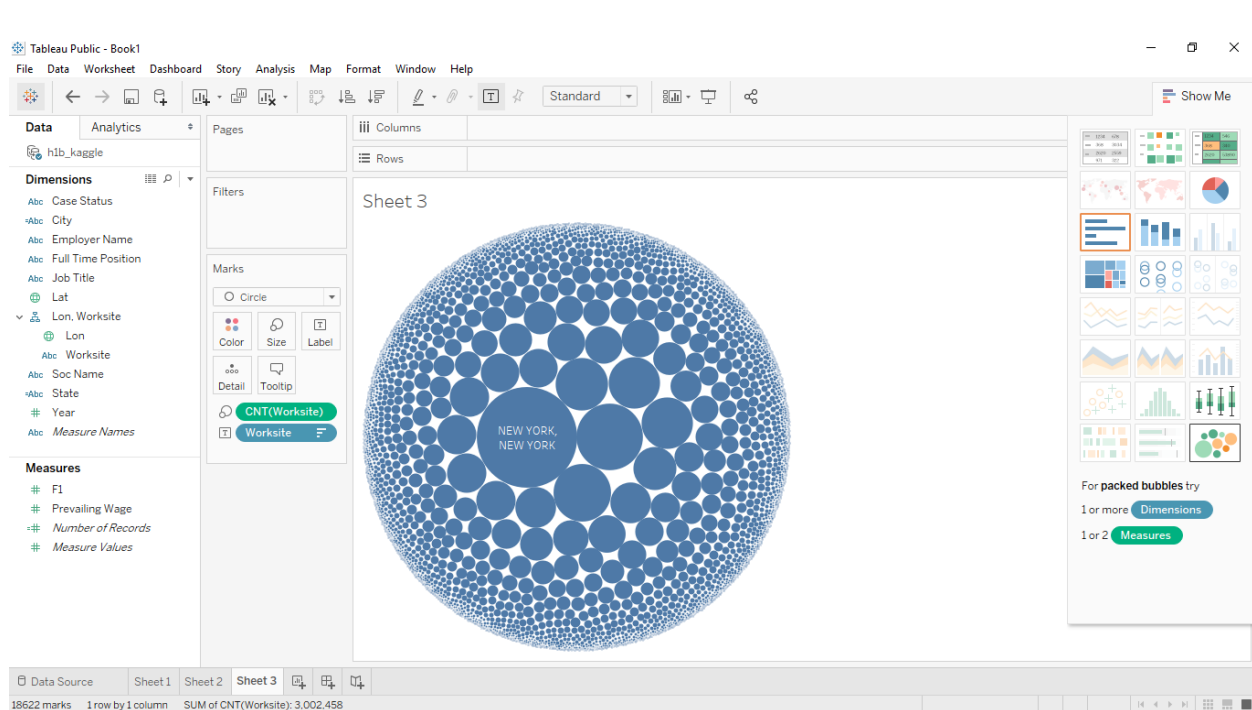
- Which demographic has attracted most H-1B petitions over the period of 6 years?

To find out what place attracted most H-1B petition, the map can be used by using latitude and longitude data given while custom splitting the worksite in city and state and using it in toolkit.

To display the number of employer from each city, count of employer name in the toolkit can be used.



Since the above map will not show the city with most filed petitions, we can use packed bubble type graph which represents worksite where largest petition filed with the help of largest bubble among various bubbles, each of their sizes are analogous to number of petition filed. Following is the packed bubble type graph implying New York City, New York is the one where largest numbers of H-1B petitions have been filed.



Thus, addressing each of the above questions with the help of Tableau and trying to answer them helped understand data visualization tool's utility while analyzing data.