590DV Final Project Part 4 - Narrative Report

Based on Part 2

Final project's part 2 generates three visualization figures. Figure 1 visualizes the relationship between the awards amount and the awarded institutions distribution in 1996-2013. Figure 2 visualizes the relationship between the awards type and the awarded institution distribution in 1996-2013. Figure 3 visualizes the amount of different type of awards received in each state in 1996-2013.

The tools and libraries that are applied during the visualization include pandas, matplotlib, and plotly. The dataset is a csv document that has 43 columns.

I. Figure 1: Awards Amount and Awarded Institutions Distribution (1996-2003).

This figure picks up four columns from the document, "Institution", "AwardTotal", "Longitude", and "Latitude". The visualization displays the total of awards that each institution has received in 1996-2013, and each institution is marked in the United States map based on its specific longitude and latitude. The total amount of the awards that each institution has received is divided into one of the five levels. Each level has a different range. Each level is marked in a specific color and each level also has a corresponding size of circle. (see Table 1)

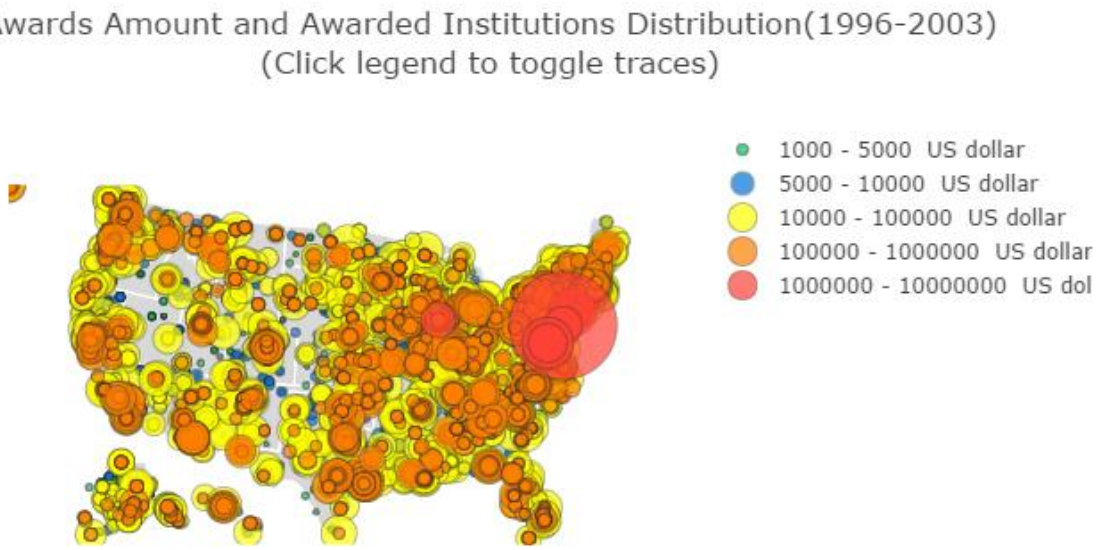
Table 1: Award Total by Level and Color

Level	Level 1	Level 2	Level 3	Level 4	Level 5
Range ($)	1,000-5,000	5,000-10,000	10,000-100,000	100,000-1,000,000	1,000,000-10,000,000
Color	rgb(0,201,87)	rgb(0,116,217)	rgb(255,255,0)	rgb(255,128,0)	rgb(255,65,54)

Then for iterator evaluates each tuple. First, the award total is selected if it satisfies the level 1 requirement, then its longitude and latitude are collected, then its size is calculated, and the level 1 color is assigned. In a same way, all the tuples are assigned into level 2, level 3, level 4, and level 5, until all the tuples have been divided into different level.

Next, a layout function is written. This function provides the title of the visualization figure, shows legends, define the scope as the United States, paints the land color as rgb(217, 217, 217). It also defines the width of country and state border line, and sets their color as rgb(255, 255, 255).

Then, data that represent all the institutions, their total amounts of awards, and all the other related information, and layout are assigned into the “fig” function. And py.iplot finally puts all the above data into the visualization.



II. Figure 2: 1996-2013 Awards Type and Awarded Institutions Distribution

Figure 2 visualizes the awards that each institution has received are assigned into 17 different program types. Each type is defined in a specific color. And each institution is marked into the United States map by using its specific longitude and latitude. Four columns are collected from the csv dataset, which are “Institution”, “ProgramType”, “Longitude”, and “Latitude”. Each program type is attributed with a particular color (see Table 2), and then each institution’s awards will be displayed in the corresponding colors on the map.

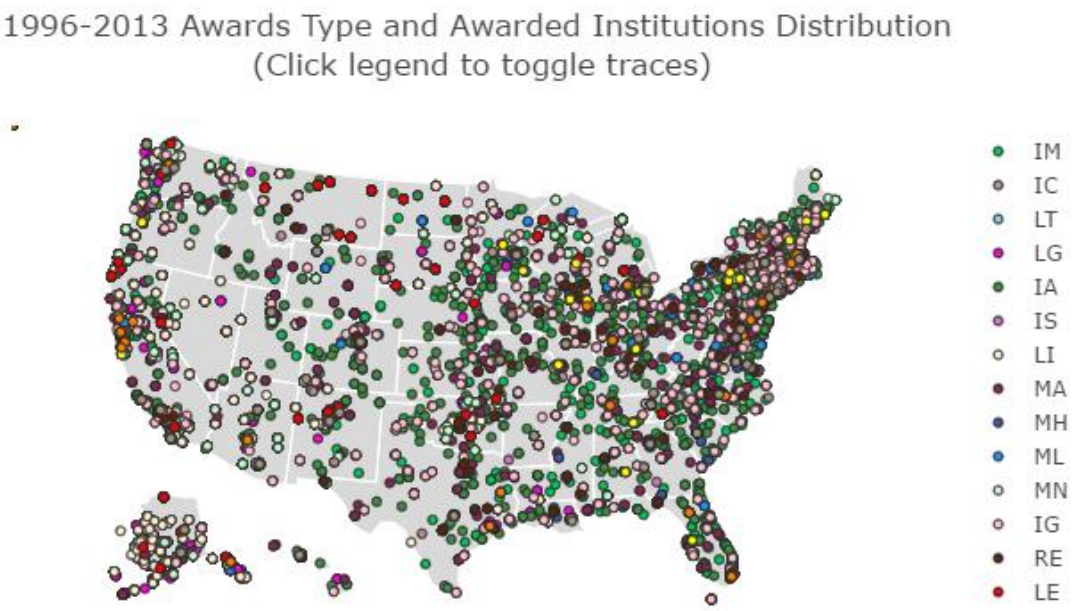
Table 2: 17 Program Types by Color

No.	Program Type	Color
1	IM	rgb(0,201,87)
2	IC	rgb(188,143,143)
3	LT	rgb(135 206 235)
4	LG	rgb(255,0,225)
5	IA	rgb(61,145,64)
6	IS	rgb(218,112,205)
7	LI	rgb(255,235,205)
8	MA	rgb(135,38,87)
9	MH	rgb(61,89,171)
10	ML	rgb(30,144,255)
11	MN	rgb(189,252,201)
12	IG	rgb(255,192,203)
13	RE	rgb(94,38,15)
14	LE	rgb(255,0,1)
15	IL	rgb(255,255,0)
16	MP	rgb(255,128,0)
17	ST	rgb(163,148,128)

A “for” iterator is used evaluate each tuple, and assigns the corresponding data into type 0, and then collected its longitude and latitude, and assigned with color 0. Then the “for” iterator will assign the corresponding data into type 1, 2, 3, …., and 16. So far, all the tuples have been evaluated and assigned into each corresponding type and color.

Then “layout” function defines the title of the visualization figure, shows legend, sets up a scope as the United States, shows the land color as rgb(217, 217, 217), defines the width and color of both state and country borders as 1 and rgb(255, 255, 255).

At last, the data and the layout are assigned into “fig” function, and py.iplot visualizes all the related data on the United States map, where each award type is displayed in a specific color and localized into each institution on the map.

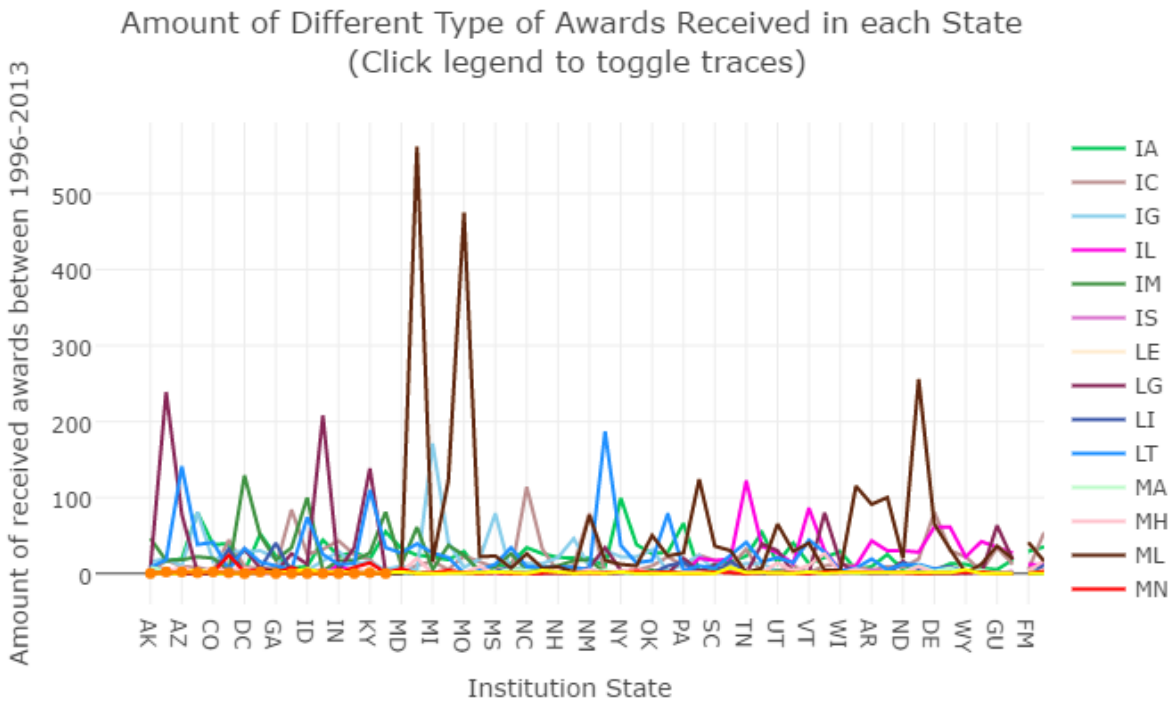


III. Figure 3: Amount of Different Type of Awards Received in Each State

This visualization figure displays the amount of each type of awards that the institutions in each state have received. The x axis lists the states, while the y axis displays the amount of received awards that the institutions in each state received in 1996-2013.

Two types of visualizations are displayed in figure 3. The first visualization displays how many awards that institutions received in the designated period. When the curser points at the vertical line of each state, a list of small boxes and numbers will appear, in which each box is assigned the color of that specific type of awards. In each box, there is a number that indicates the number of awards for the specific type, and at its right is the name of that award type. Another type of visualization is displayed by the trace lines. Each trace line represents one type of award. It varies in height by state. It can be easily seen the amount of each award type that are received in each state.

There are 17 trace lines, as there are 17 award types. Each trace line corresponds to the color of each award type. Three columns of “Longitude”, “InState”, and “ProgramType” are collected from the csv file.


Amount of Different Type of Awards Received in each State
(Click legend to toggle traces)
Amount of received awards between 1996-2013
0
100
200
300
400
500
AK
AZ
CO
DC
GA
ID
IN
KY
MD
MI
MO
MS
NC
NH
NM
NY
OK
PA
SC
TN
UT
VT
WI
AR
ND
DE
WY
GU
FM
Institution State
IA
IC
IG
IL
IM
IS
LE
LG
LI
LT
MA
MH
ML
MN