Visualizations of Four Major Sports League Titles in the U.S.

Ashton Hsu (amh333) Zheng Yao (zy87) Alicia Wu (yw344)

** Note: needs to be on a web server to work. Also it takes about 8 seconds to load all visuals.

A. A description of the data.

We got the data from Wikipedia table “List of US cities by number of professional sports championships” (link). The raw data from Wikipedia contains a list of US cities that has won professional sports championships. It has the number of championships won by each city in the 4 major league sports: MLB, NBA, NFL, and NHL, as well as the totals by city. Additionally, it also has the number of championships from MLS, AFL (Arena Football League), MLL (Major League Lacrosse), and NLL (National Lacrosse League). We believe that the 4 major sports are the most crucial components of this dataset, since they are followed by the most people. So we omitted the columns MLS, AFL, MLL, and NLL, and focused only on MLB, NBA, NFL, and NHL. As a caveat, the numbers are championships won by cities, not franchises. So for example, the Syracuse Nationals’ 1 NBA championship is counted for Syracuse, even though the franchise later became the Philadelphia 76ers. Some of our visualizations need state information, and we just manually appended another column for that.

For mapping, we used this class note to retrieve the precise longitude latitude of the cities. We linked the two dataset together with their shared city name. Additionally, the original dataset does not contain the state in which each city locates in; we manually added them.

The variables in our set of visualizations include:

1. Name of the city (together with its location)
2. Name of the state (together with its location)
3. Number of MLB championships each city wins
4. Number of NBA championships each city wins
5. Number of NFL championships each city wins
6. Number of NHL championships each city wins
7. Total number of titles each city / state wins

B. A description of the mapping from data to visual elements.

In index.html, we include four major types of visualizations. The first two maps contain circles proportional to the total number of championships won by each city, positioned at the city’s location. The circles have various opacities since some circles overlap and would be hard to see. The radius of the circles had to be calculated by square rooting the total number of championships and then linearly scaling it by some constant factor. The circles are filled in green, since sports in general are related with the human beings and nature.

Originally we thought of pinpointing the pie chart of each city on the map. However, later we found that it would be too crowded even if we did the zoom-in version and set up the opacity. When the pie charts overlap with each other, it is hard to tell the distribution. We realized that it is not necessarily true that we should reflect three elements (distribution of of titles, count of titles, locations) on a single visualizations. In the end, we illustrated the count and locations on the first set of maps, and count combined with distribution in the end. Count of the titles certainly matter the most - the winner takes all. We made our third chart for the exact same reason. If a city has won a title in all four major leagues, that certainly means something. That's why we denote shem using stars instead of circles.

The single league set of maps is colored by state per sport. The darker the shade a state is, the more championships it has in that particular pro sport. We use the color() function in D3 to implement that. The color scale ranges from white to either red (MLB, indicating the red lines on a baseball, since it is hard to use white here), orange (NBA, indicating the color of a basketball), blue (NFL), or gray (NHL, indicating the color of a hockey puck). Along with that is the visualization of the total number of championships, aggregated by states. For consistency, the total number of titles are filled in green

The last set of visualization is a set of pie charts for each city that illustrate the ratio of championships in the four sports. Similarly, we used red color for MLB, orange for NBA, matching a basketball; blue for NFL, matching a football; and light gray for NHL.

For us, the major takeaway from project 1 is that we learned to interpret our data from different angles, and separate them up in various visualizations instead of crowding them in one. Following the TA's advice, we also learned that legends are extremely important for the users to understand the visualizations.

C. The story.

Our visualization displays the distribution of championships in the 4 major league sports by cities. Our first map shows the totals of all 4 sports for each city, where the larger the circle, the more championships the city has won. Our second map is a zoom-in version of the east coast, since the cities are more dense and it would help the viewer see each east coast city more clearly. This visualization would be useful for viewers who want a general overview of which cities have successful sport franchises.

Our third chart pinpoints the cities that have won all four titles in the four major leagues. This gives the viewer a quick summary of where the pro sports powerhouses are located.

Our fourth visualization is a set of maps, where each is attributed to a different major sport. It gives a visualization of how many championships each *state* has won in any of the major league sports. The darker the state, the more championships it has won. These maps allow the viewers to geographically understand where each sport thrives or has thrived. This would be useful for viewers studying the history of sports or attempting to predict sports trends. For example, in the NBA map, the obviously darker colored Massachusetts and California can be partly due to the 1980s duopoly of Magic Johnson's LA Lakers and Larry Bird's Boston Celtics. We can also see that on the NHL map, the northern, colder states, such

as Michigan and New York, are darker--this can be attributed to the fact that ice hockey is more popular in colder places and because they are more influenced by Canada, where ice hockey the national sport. The surprise also comes here: although everybody knows that the eastern and western coast are better at sports, it is still hard to believe that the center has literally nothing.

Our fifth visualization is a set of pie charts, one for each city, that shows the ratio of the different sports championships. The size of each pie is proportional to the total number of championships in all sports.

Acknowledgement:

- Part of the zoom-in code is adapted from TA Josh C's work posted on Piazza.
- We got help from examples provided by Mike Bostock and Datamap