DSC 106 Final Project - Huaning Liu

In this project, I extracted the data from Michelin 2019 restaurant, includsing their postions, ratings, price levels and cuisines, etc. As a foodie, it is really enjoyable for me and potentially the users of the dashboard to visualize and pick the restaurants based on their interests. Therefore, this dashboard is basically a kick-off for poeple getting started on Michelin restaurants, and my data visualizations will be focusing on the the relation of prices and stars, the distribution of regions and cities, and those based on cuisines. Questions are proposed below for further exploration.

Question 1:

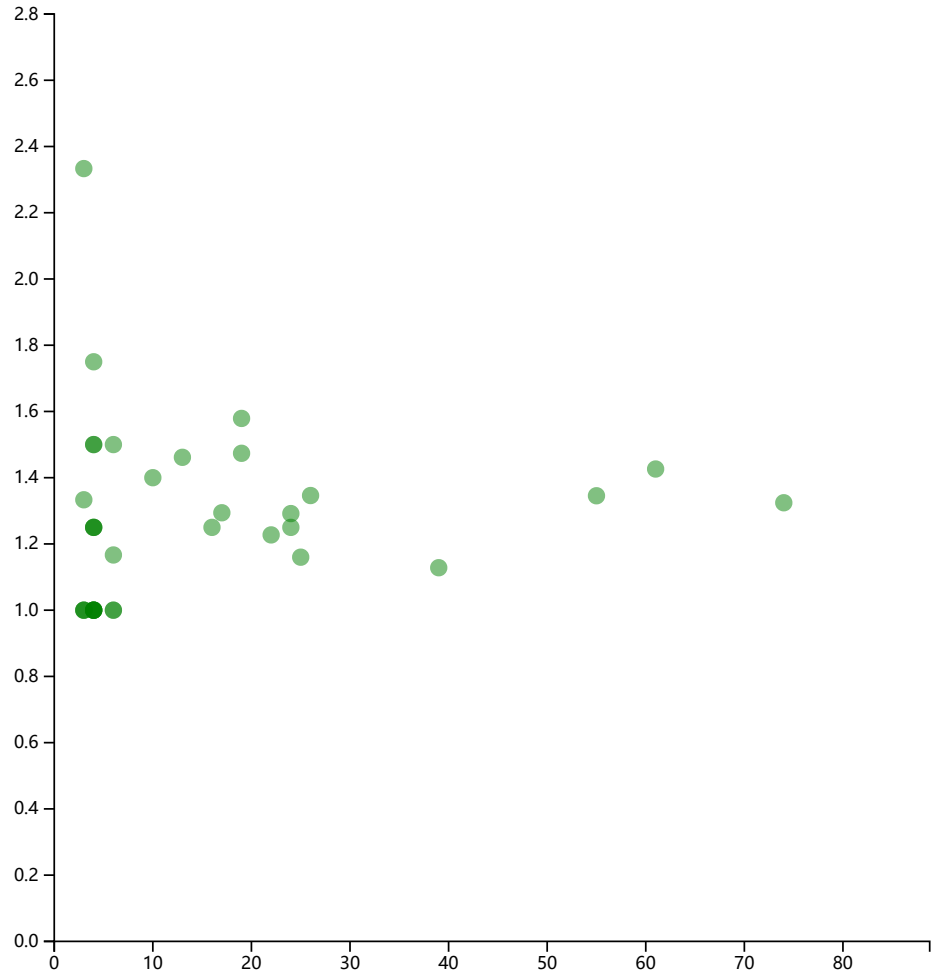
For each of the cities that has more than 2 Michelin restaurants, does more inclusion (counts) in the Michelin list indicate better average rating by stars?

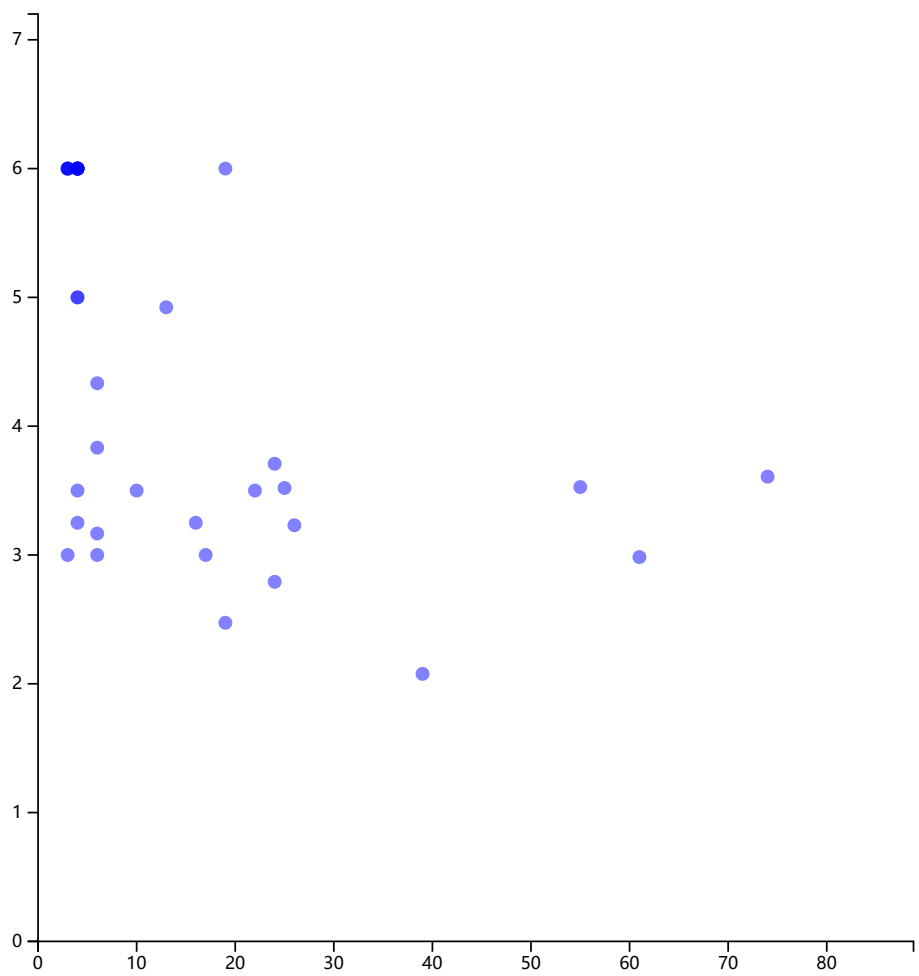
Answer 1:

Choice of color scheme: I used green for the dots as the scatterplot that display the relationship between counts and average ratings and blue for the dots as the scatterplot that display the relationship counts and average price level.
Marks: Points
Channels: Vertical Position, Horizonal Position.
Based on the question, we can detect that there actually exist a very very weak correlation between the count of Michelin starred restaurant in a city and their average price level; also, there exits an outlier which takes low count and very high rating(star). Therefore, given the "fitting line" to be roughly parallel to the x-axis, I don't think more inclusion in Michelin List hints better average rating. Also, given the brushing I am doing here, we conclude that the distribtuion of average price level for a city is a way more sparse along y-axis for more of the brushing window we experimented. This pair-brushing plot animation could also give us an intuition about both the price level and ratings for each city as a dot in the plot.

Plot:

Relation of number of restaurants and average rating/price level in cities





Question 2:

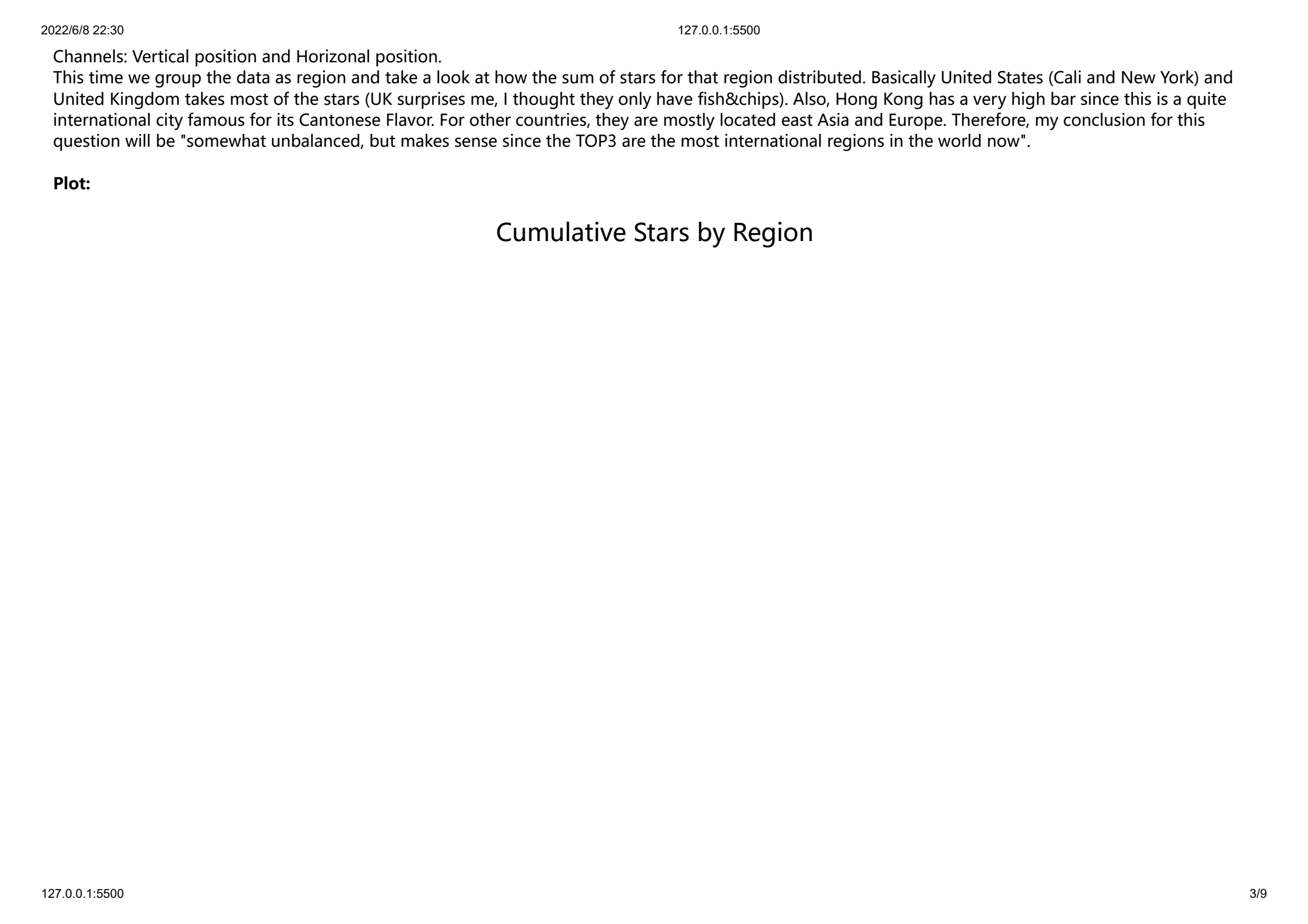
How are the stars of Michelin restaurants 2019 distributed by region?

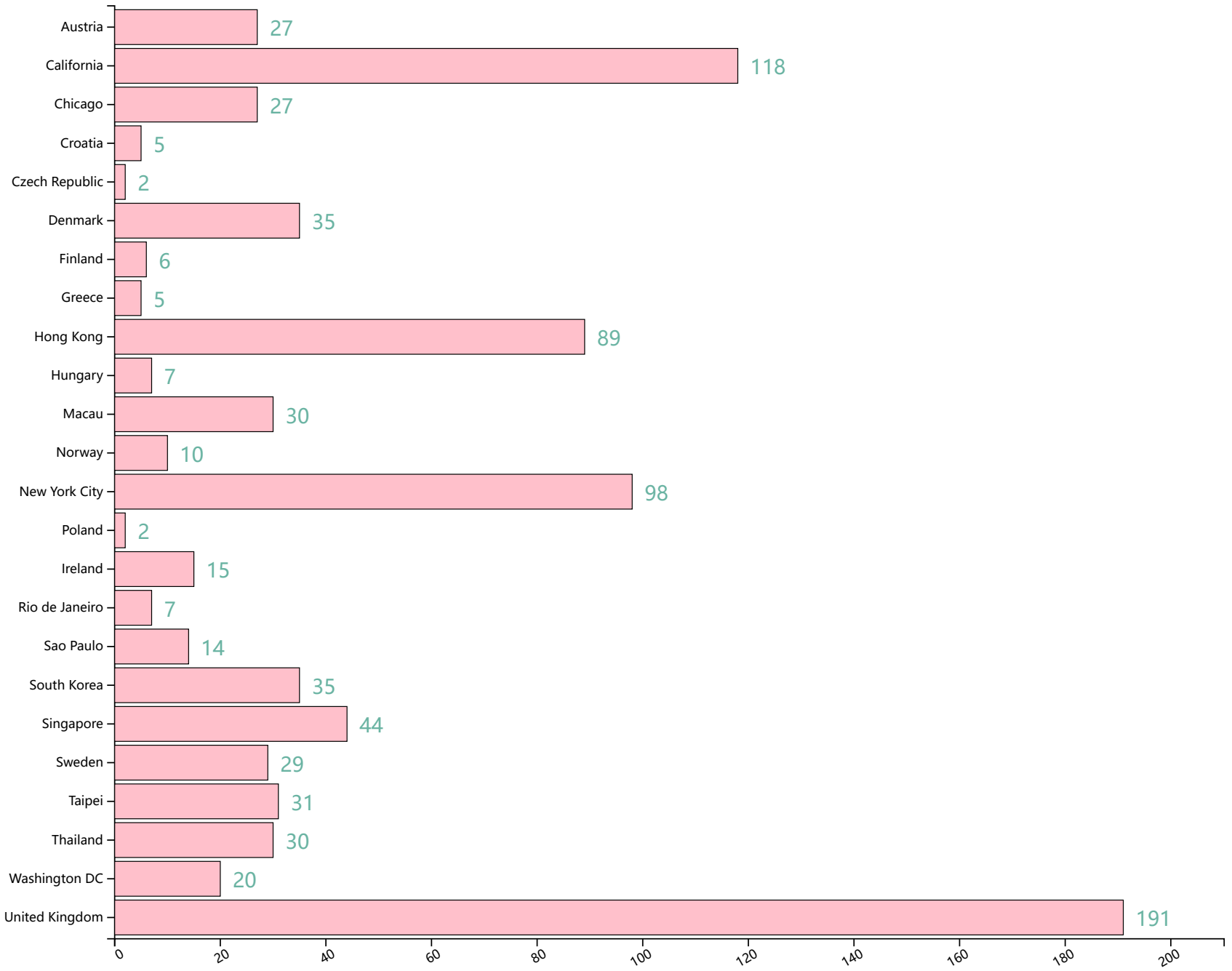
Answer 2:

Choice of color scheme: Pink is used as the overall color of the bars.
Mark: Line

Channels: Vertical position and Horizonal position.
This time we group the data as region and take a look at how the sum of stars for that region distributed. Basically United States (Cali and New York) and United Kingdom takes most of the stars (UK surprises me, I thought they only have fish&chips). Also, Hong Kong has a very high bar since this is a quite international city famous for its Cantonese Flavor. For other countries, they are mostly located east Asia and Europe. Therefore, my conclusion for this question will be "somewhat unbalanced, but makes sense since the TOP3 are the most international regions in the world now".

Plot:




Austria
27
California
118
Chicago
27
Croatia
5
Czech Republic
2
Denmark
35
Finland
6
Greece
5
Hong Kong
89
Hungary
7
Macau
30
Norway
10
New York City
98
Poland
2
Ireland
15
Rio de Janeiro
7
Sao Paulo
14
South Korea
35
Singapore
44
Sweden
29
Taipei
31
Thailand
30
Washington DC
20
United Kingdom
191
0
20
40
60
80
100
120
140
160
180
200

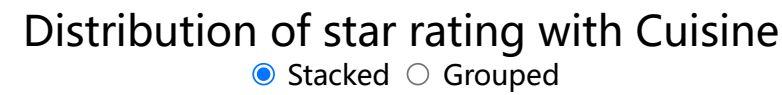
Question 3:

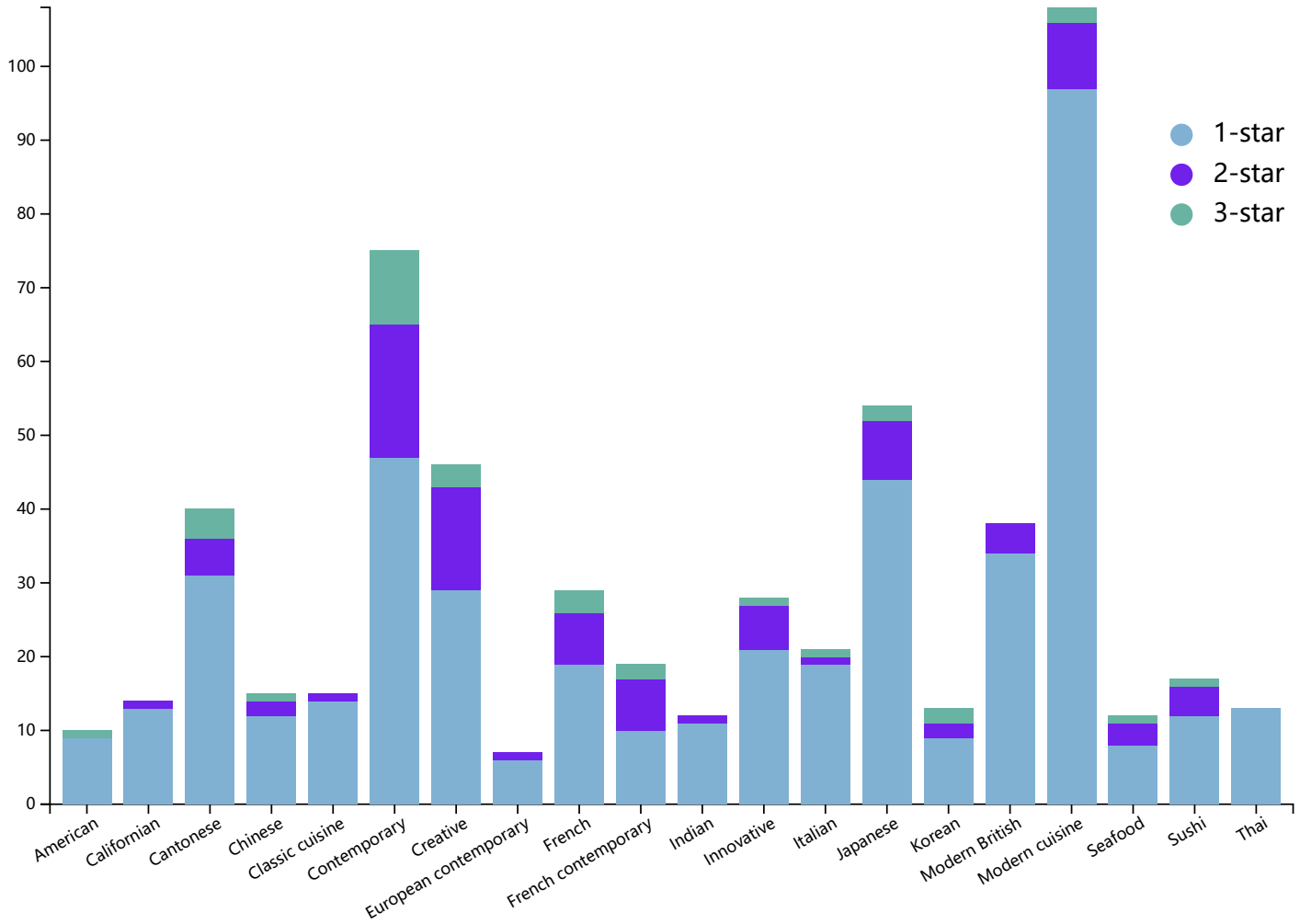
How are stars distributed among restaurant cuisines within TOP 20 popular cuisines?

Answer 3:

Choice of color scheme: I use green, purple and blue to represent the three levels of rating
Marks: line
Channels: Vertical position and Horizonal position and color
This time we group retaurants in the most popular 20 cuisines by their stars and accordingly have a stacked barplot. We could see that Modern absolutely takes the dominant amount here, but interestingly it stil does not have a very high proportion of 2 and 3 star ratings. Therefore, a informative point about the distribution will be that 2-star and 3-star are most evenly awarded based on cuisine. Also, Modern , contemporary and Japensese as the three dominant cuisines takes overall the most "award-likely" positions.

Plot:





Question 4:

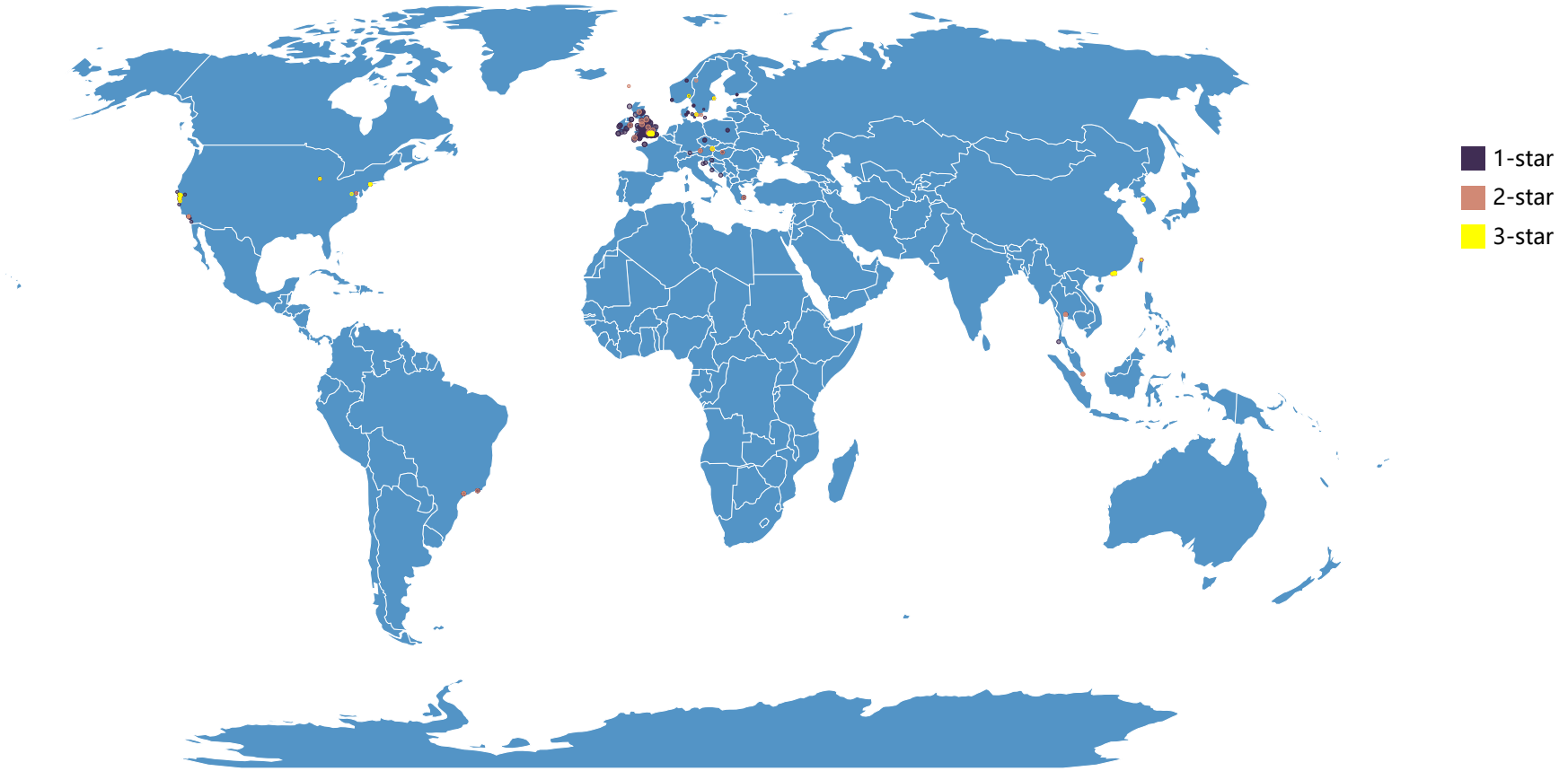
Specifically, how are Michelin 3-star restaurants distributed over the world? And how does their prices intervenes here?

Answer 4:

Choice of color scheme: blue for the continents in the geo-map for the world, yellow for 3-star restaurants, light orange for 2-star restaurants and dark purple for 1-star restaurants.
Marks: Area as continents, point

Channels: Color, size
We make an multi-functional map to answer the question, which includes zooming, button and labeling. Note that if the user cannot see the points clearly, zoom as a functionality is added so that they could easily see the details of a specific region. Also, buttons for star level is added. If they zoom in, they could also seem the difference of radius that shows the scale of price level. Back to the question, we actually notice that the 3-star restaurants is quite evenly distirbuted along United States, Asian, and Europe, with other continents missing. For the Europe, most of the 3-star restaurants locates at the western side. For United States, they located at east and west coasts. And for east Asian, they locates at specifically HK, Taiwan and Japan. Then adding 1-star and 2-star ones back, I could see that now Europe takes a dominantly leading place, that a cluster forms there. Interestingly we notice the emergence of good restaurants at South America. If you zoom in, labels related to the restaurant's name, city and cuisince will also pop up. I hope this will be a good demo as a platform for the foodies:).

Plot:



Question 5:

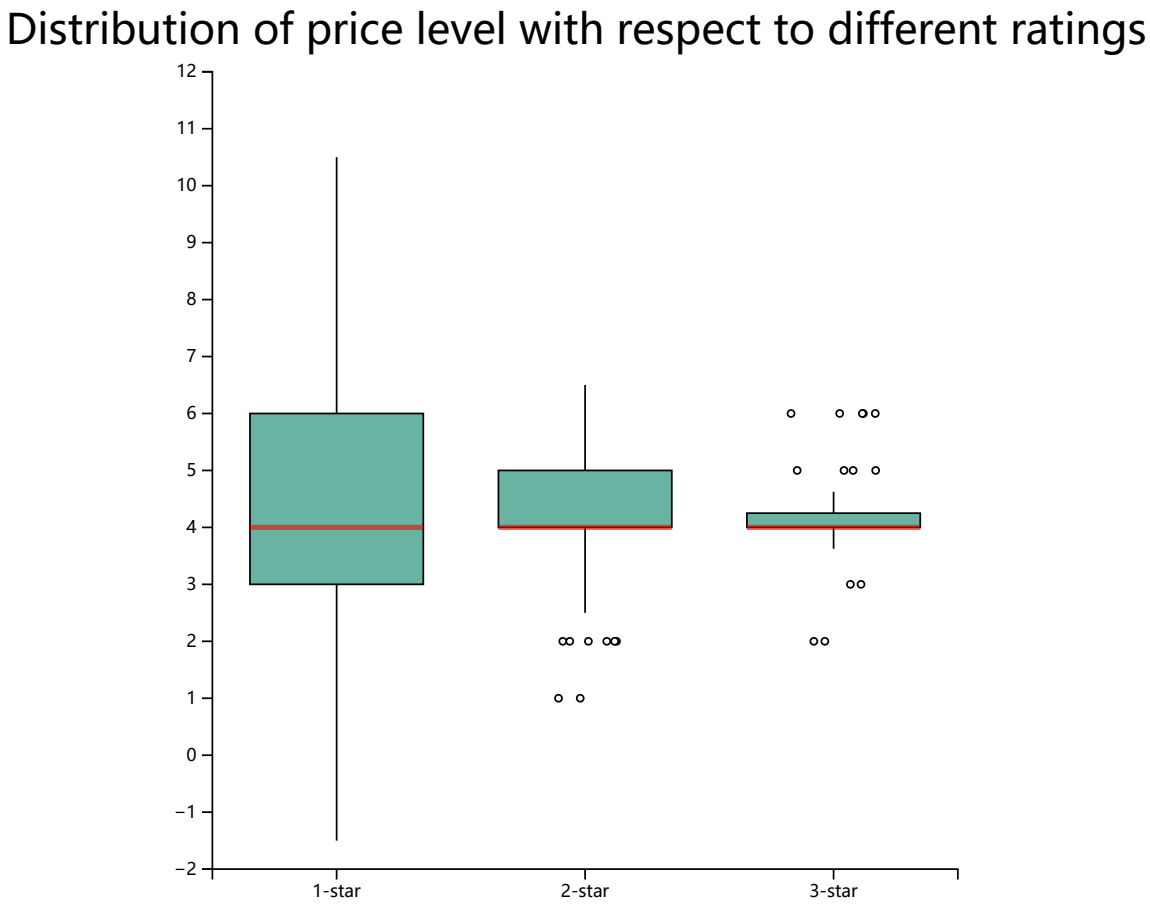
How's the price level for Michelin Restaurants 2019 distributed w.r.t star?

Answer 5:

Choice of color scheme: I use red to notify the median of each distribution, and use light green to display the box.
Marks: line, point
Channels: vertical position, horizonal position

In this plot, I display the boxplot in order that we cannot see how the price levels are distributed clearly in neither in map or the plots before. Therefore, it will be good to display this for this question. We notice that The three median of prices are roughly the same, proving that this does not follow our "convention" that three-star restaurants must be pricy - they win for their great great flavors. Also, we see that the higher rating is, the most unstable/sparse (given the outliers) the distribution of price levels is, which is reasonable since I think for three-star ones, their food usually consists of seasonal seafood whose price varies a lot along the year. Back to the question, the three series of prices w.r.t ratings for Michelin restaurants 2019 is quite approaching to each other, with the variance/sparsity to be different.

Plot:



In conclusion, I hope this project gives a good view statistically and visually about the information of Michelin Restaurants 2019. Good luck foodies!