

Assignment 1

Zhuoyang CHEN 50005220

1. With the original canvas size of 600×500 px, while 10,000 points look relatively distinguishable, when scaling to 1,000,000 the figure is too dense and all points look like a block of green. Besides since the data are generated using d3.js after the website is opened each time instead of using existing dataset, it takes quite a long time for the figure to show up.
2. (1) To reduce the density of the scatterplot, aggregation method could be utilized to summarize points within an area. For example, a heatmap could be created to represent the number of points in a 0.1×0.1 square area (10 by 10 grid), with a range of 0 to 1 for both x and y axis. Or drawing circles which the area is proportional to the number of points in the corresponding area.
(2) Discretization could be another method to better visualize the data. A barplot-based density plot could be derived with proper binning to investigate the distribution of the both the x and y values of all points.
3. **(1) For aggregation:**
Benefits:
The aggregation method efficiently reduces the number of points to be drawn and thus reduce the density of the figure. It is quite flexible in terms of aggregation methods and scales to be chosen.
Drawbacks:
Since all points in a certain area are summed up and presented as one, the resolution of the whole figure decreases and losing the information that how those points are distributed in the grid. Whether they are evenly distributed like a block or gaussian distributed like a ball would exhibit exactly the same when they have the same number of points.
(2) For discretization:
Benefits:
It is especially suitable to investigate the distribution of data, in this case, the x and y values. Similar to the aggregation method, for discretization method different scales or binning strategies could be chosen case by case.
Drawbacks:
Since discretization summarizes points as the count of each bin, it may not straightforward to see how the data are actually distributed in a 2D presentation compared to the aggregation method.
4. An aggregation method is implemented using d3.js based on the template provided.