Udacity - Data visualization

There are three plots in this project :

1. index_1.html
2. improve_coloring.html
3. index_3.html

* **index_1.html** is the main plot - where I try to bring across a narrative (Martini-glass). This has animations as well as user interactivity.

* **improve_coloring.html** is a work under progress. I've written the code you would find in the GeoMap dependency and tried to create my own Choropleth; everything is fine but for colouring. I'm experimenting with the quantizer at the moment.

* **index_3 html** I've used D3 and GeoMap dependency. This has improved readability by doing a lot of pre-processing and has additional features like - zoom on click, Tooltips based on data etc. Here I've plotted the distribution of average delays across states. There is no 'finding' that might come as a revelation, also the narrative is entirely author-driven.

---

index_1.html :

Initial design :

1. Initially, I used stacked bars to show delay by airline with the types of delays in the stack. Then I decided to plot other aspects of the data as there was no proper message to be conveyed here.

2. Then, I used bar-graph to plot average delays by airline (taking mean of all the possible delays). After this, I proceeded to use Dimple.JS and Storyboard controls to animate the data of delay by airline through the years. - https://discussions.udacity.com/t/mini-project-2-take-two-dand/25145/169?u=thiduck. Although this helped to show how airline performed through the years, from the feedbacks I received I found out a better aspect of visualization I could use and decided against using this iteration.

3. Finally, I decided to plot delays by the months of the years using Storyboard as I felt it would be nice to observe how the delays occur given any time of the year and that this might be a useful narrative.

index_1.html -> which is the main plot. This is a martini-glass type narrative - where I animate through the data but users can pause and get their own insights and understandings from the data.

Reasons for choosing the visualization elements in index_1.html:

For ‘Delays through the months’ I feel using **Scatter + Line** plot was the best choice, since a line graph is always associated with time-series data and it is easier to show *trends* using a line graph; scatter plot adds points on the graph to show distinct and discrete data points (without which the line may look like a continuous trend , without corresponding points for months).

I have also added highlights to draw focus to part of visualisation I want the user to note and changed colors of the circle using css.

Does the visualization have a clear finding?

I'm trying to convey that the delays are more likely to happen during time of

a. Summer Vacation (June- July)

b. Christmas and new-year (December - January)

since the air-traffic would be much higher than expected. This is especially clear from the increasing trends in Nov-Dec and May-June , in all the observed years.

Does the visualization focus on its finding?

I've tried to use colours as preattentive processing and I’ve used line-graph to show the variation with time and trends. I've also used different colours and size to highlight the two trends I want the users to focus on.

Feedbacks :

1. Suggestion to show trends with time instead of airline. (Used line-graph and changed the plot x-axis to months of a year).

2. Add a legend as the message wasn't really straightforward from the 'months' x-axis (I have added a legend).

3. Change font-size of X and Y-axis labels (Font size increased).

4. Change position of animating storyboard control box (Box now on the bottom left corner - without overlaying the line-graph).

5. Suggestion to use bootstrap to display the title (I have used bootstrap).

6. Suggestion to use the data to also plot a choropleth and show distribution of delays across states(I have created two choropleths).