Information Visualization

Homework 8

Sourabh Shetty

When it comes to high dimensional data, it becomes very important to find tools that help visualize it without compromising on too many of the dimensions.

Andromeda is a great tool in that regard, since it helps condense the data into two dimensions that can be visualized and interacted with effectively, by implementing weighted multidimentional scaling, and allowing the user to play with the input and the output.

Most importantly, it does this automatically, doesn't require the user to write code, and hides the calculations so that the user can focus on the data.

Once the class data was loaded into Andromeda, right off the bat we notice that a few people stand out significantly. Chris in particular, gets pushed way to the right. Khadijah also is pushed to the top, although not as significantly, which indicates that even though she is an outlier, her answers were still much closer to the norm than Chris' was.

Similar observations can be made about Hu and Frank, wherein despite them being outliers, their disparity to the norm is not as severe as Chris or Khadijah's.

Since Chris was such an extreme outlier, it was fun to play around and manipulate the weights of each dimension to try and force him to not be as much of an outlier. I adjusted the sliders on some of the dimensions on which his data was relatively high or low, to give more or less weightage depending on how it would affect his distance to the other people.

This let me bring his marker closer to others without having to click and drag it myself.
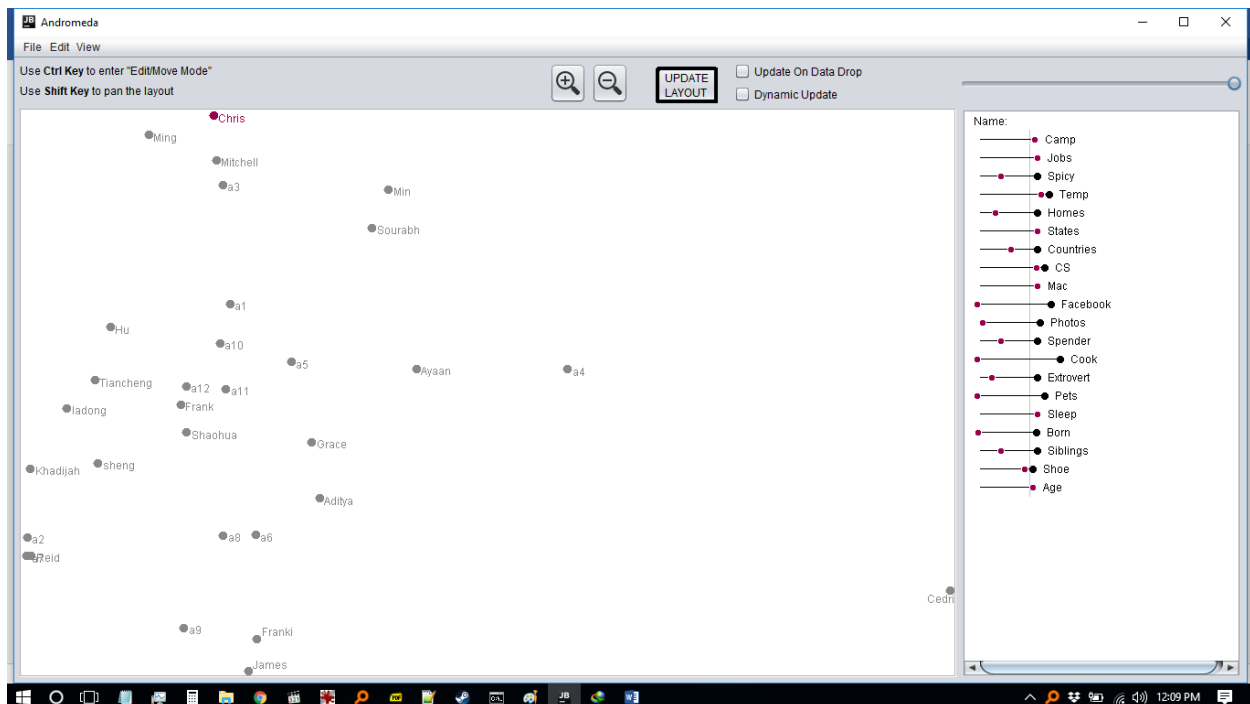
Next I held the Ctrl key, and clicked and dragged on Chris to bring him closer to the others. On the dimensions sliders, Andromeda showed the new values his data would have had to have to make that new position possible

When I clicked the Update Layout button, it readjusted the figure based on the new data:
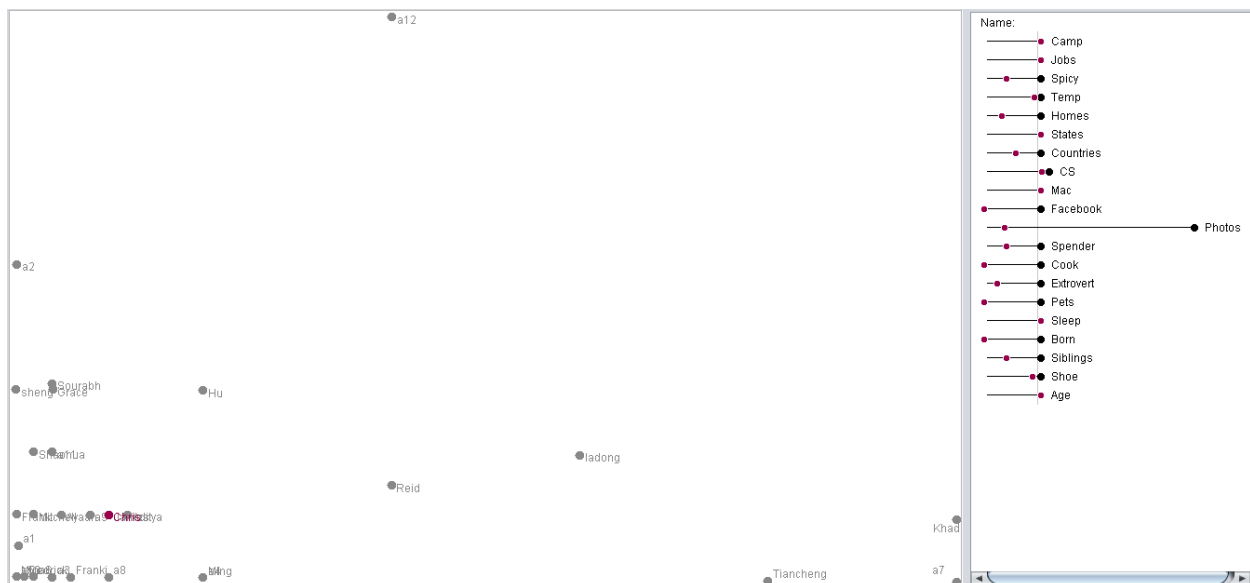
This also updated the sliders and their weights. This provided the insight that since Chris' photos taken was the data value closest to the group, Andromeda increased the weight and thus the importance of the Photos dimension, to force Chris to be more central to the group.

This also had the side effect of pushing Khadijah further away from the group, since her data value for the Photos dimension was one of her more outlying data values, and giving that dimension more weightage separated her from the group even more.

I conducted my analysis by initially just blindly playing around with all the buttons on Andromeda. Once I had a feel of the application, I first started moving the sliders and trying to understand how they affected the figure. Once I got that, I started to manually move the marker itself around, to see how the weights for each dimension would be modified. In my first method, I had to use my own intuition to guess which sliders to move to make it possible, but using the second method, I was able to precisely see which ones were needed

Thus using Andromeda to analyze this data set helped me visualize and understand this data set more, and helped me understand weighted multidimensional scaling in much more detail than I did previously.