IBM HR Analytics Employee Attrition

Xiuping Tan

Introduction

The goal of my visualization interface is to allow users to interact and explore for analysis at different level of Human Resources Employee Attrition in IBM. In this application, it uses fictional data set created by IBM data scientists for understanding factors that lead to employee attrition such as age, distance from home, income, etc.

Overview

Before clicking anything, the interface displays an overview of the dataset with three visualizations: 1) a bar chart shows the number of employees in each department with both having attrition and no attrition, 2) a parallel coordinates chart that compares many possible factors for attrition and see the relationship between them, and 3) a sunburst chart that gives some distributions of the job roles and education fields within different departments.

Interactions

First, the attrition status selections work only on the entire dataset, so we should start with selecting an option. Upon attrition status, all the visualizations will be filtered subset of the data based on Yes, No, or both attrition. Then, selecting the departments in the bar chart can show detail views in relationship between different factors as well as distributions of job roles and education. Multiple departments can be selected as steel blue and deselected as light blue. If users selected all three departments, the chart will return to default allowing users to reselect.

Both the bar chart and parallel coordinates allow multi-selection interactive filtering. The reason I choose parallel coordinates is because there are different factors that may cause attrition and we want to see relationships between them. Users can select multiple factors with different range of values to filter the subset of data that they interested. If user click on any values not selected, it will undo the selection for that factor. For instance, if the user selects age as 50 to 60 and distance from home as 2 to 6, the subset of those data will be updated to the department bar chart and sunburst distribution of job roles.

In addition, bar chart and parallel coordinates have two-way interactions. Users can continue filter down by selecting a department after selecting multiple factors in parallel coordinates. Let's say then the user clicks on the Sales department, we can see a nice and clean sunburst chart updated with large portions on Sales Executive and Managers. From this example, one interesting finding would be within Sales Department, employees aged 50 to 60 (elder) tend to be living closer distance from home with high salary and job position. Another interesting finding is within Research & Development department, employees who have high income tend

to live closer to home and almost all of them aged above 40. They are Manager and Research Director positions.

Although the sunburst chart depends on the other two visualizations, it can clearly show the distribution of different job roles and education fields under the conditions filtered. There are more interesting analysis findings users can explore to find out.

Student Evaluation:

> "Appearance wise, the really white background can be less than ideal for large screens and there isn't much unity with the colors used. The partitions of the sunburst chart are nicely outlined to show separation and have a well-designed tooltip to display additional information if the user hovers over the right slice. The three visualizations are responsive to the size of the window such that they will adjust their size according to the size of the window. The sunburst partition chart does have an issue with this change in size as each slice must print its label and may overlap with others, making them unreadable.
>
> Filtering data is easy and simple. The bar chart only requires clicking the desired bars to filter down the results, which is easy, although it would be ideal to indicate that the bars can be clicked by changing the cursor if hovering over something that is clickable. The parallel coordinates are also easy to use but there was not much indication that the user can filter the results on each axis until after a few tries. Perhaps the visualization can be more easily understood if there were brief instructions detailing how to use the chart. Undoing the changes made on the parallel coordinates is not so easy however, so a reset button would be ideal to have. The sunburst chart is well coordinated with adjustments made on the bar chart and parallel coordinates and has a satisfying immediate effect if changes are made.
>
> Not many noticeable bugs or clear issues have been found while testing this interface, but there are desired improvements on the appearance of the visualizations."

Future Improvements

I agreed with most of the suggestions from the evaluation. I have thought of using a darker background color, but I don't think dark blue fits since I have dark blue in my charts as well. It is true that changing the cursor to indicate clickable would be easier for the user to know because the sunburst is not clickable while the other two visualizations can. I have modified my system on that. I think instructions would be helpful, but the parallel coordinates seem to be clearly for selection. Another possible improvement that I can think of is to display the text labels in the sunburst chart only for those that can fit into the slices. I haven't figure out how to do that, but once users filter down to smaller subscale, the labels should be cleaner.