**Parallel Coordinates**

* **Data**
  + The data has several different types of variables or attributes. Name and country is nominal and then there are several quantitative score. For the parallel coordinates, we decided to show the 5 quantitative scores, i.e., teaching, income, citation and international and to to not show student staff ratio, number of students and sex ratio because the 5 scores seem to be at same conceptual level from a point of view of comparing the universities. Another reason for choosing these were because they had same scale and hence comparable across attributes. Apart from the quantitative scores, we also decided to include name and country, which are nominal variable.
  + We also decided to cut down the data and separate it by year. So only the top 30 universities are plotted in the graph. I will go over the justification for this in modification section.
* **Visualization**
  + Parallel coordinate is an excellent visualization for showing many discrete entity values across many different attributes. We decided to keep 7 attributes from the data set, each represented in a different axis. This is in contrast to radial charts, which can only show a very limited number of entities at a time and are better suited for aggregated values rather than discrete values.
* **Modification**
  + For the modification we decided add an axis for the name of the university and create a mouse over event for these so that when one hovers over the name of a university they can see the path followed by (and hence the scores in different attributes) the line of that university. This the reason we could not add all the data point from each year into the visualization. Since having about 800 universities on one axis would have either made the visualization too long vertically and introduced a long scroll or would have meant that each university name become too small and incomprehensible. In both cases the graph would have been pretty much unusable for any meaningful comparison.
  + We also decided to use color-coding for lines representing universities in different countries. This is useful if you want compare the general performance of different countries. We decided to go with very disparate colors for all the countries so that the difference is marked when we brush for some score in the quantitative scale. The key on the side provides guide on the color usage for different countries
  + We divided the dataset by year. The parallel coordinate only plots one year at a time and the year to be plotted can be selected from the dropdown on top. When the page is loaded for the first time the data from 2016 is plotted.
* **Analysis**
  + The parallel coordinate is good for showing multiple values at a time and could have been potentially used for showing the whole dataset. However, we made a call to cut down the data because we wanted to implement the mouse over on university name the individual scores. A lot of time, in parallel coordinates, the individual scores are lost because of the sheer quantity of lines being plotted between axises. In those cases the parallel coordinate just ends up showing a general shape of the path taken by the lines. In this case it seems that it would be more interesting to compare individual university scores rather than looking at overall trend as it is pretty varied and inconclusive. Therefore we decided to keep the mouse over for particular university and limit the data.
  + We decided to plot different years in different graphs. The drawback here is that viewers could only compare universities in same year but not the performance over time. We made the decision that we would support the former comparison, as we had to choose from either of the two options.
  + We decided to use color-coding the paths by countries. The drawback here is that it makes the visualization look cluttered. The solution could have been either using a color progression or not using color codes at all. We decided to use different colors (and not progressive colors) as the comparison while brushing is more effective with different colors then in case of progression. In progression it would be more difficult to tell one color from another.
* **References**
  + <http://bl.ocks.org/jasondavies/1341281>
  + <http://stackoverflow.com/questions/24811850/color-coding-parallel-coordinates>
  + <http://bl.ocks.org/syntagmatic/4020926>