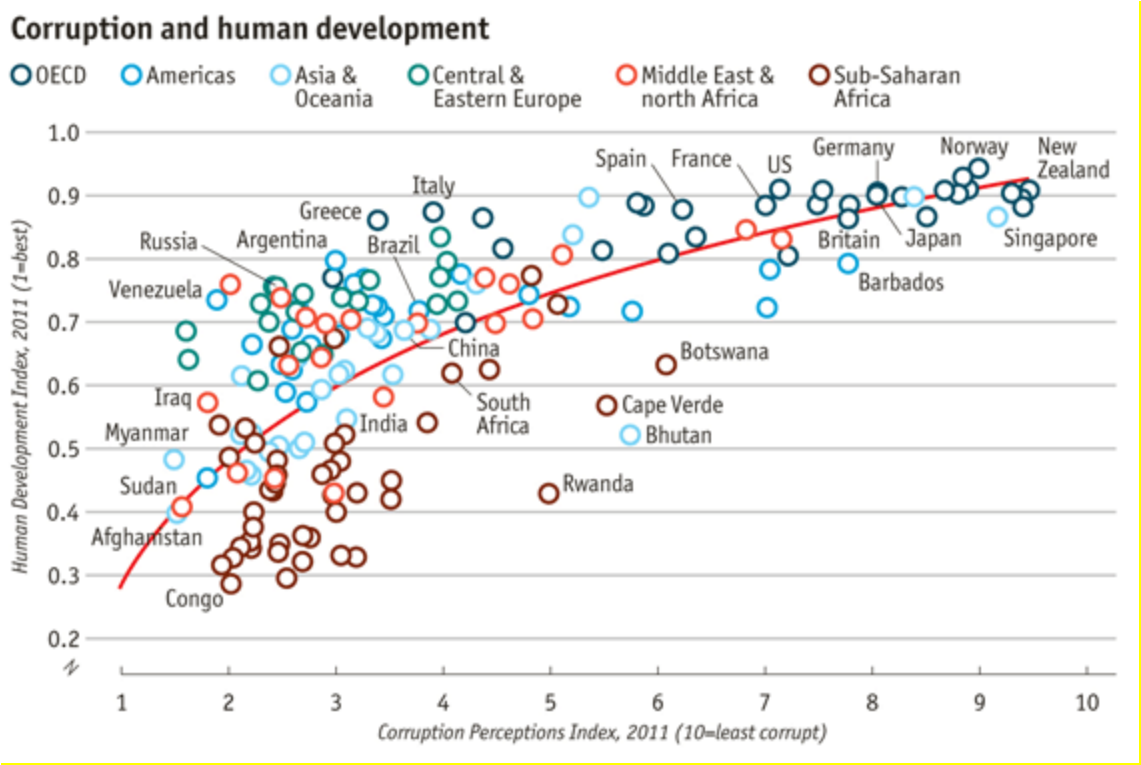
Learning Outcomes

In this session, you will primarily learn to use the package ggplot2 (short for grammar of graphics plot) in R. After the session, you will have an understanding on the following:

1. How to draw insights from a plot
2. Difference between a good plot and a bad plot
3. Plotting charts using ggplot2 by following instructions.
4. Experimenting with data and generating plots to derive interesting insights.

Exercise

Make a scatter plot to show the relationship between the 'Human Development Index' and the 'Corruption Perceptions Index' of countries. Use the Economist Data.csv and try create a graph as close as possible to the below graph.



Read "EconomistData.csv" in R. This data consists of Human Development Index and Corruption Perception Index (CPI) scores for several countries. A high CPI indicates a lower level of corruption.

1. Create a scatter plot with CPI on the x axis and HDI on the y axis.
2. Map the color of the points to Region.
3. Add a trend line.
4. Add country labels (select only a few)
5. Modify the x, y, and color so that they have more easily-understood names (e.g., spell "Human development Index" instead of "HDI").

6. Add chart title
7. Change the theme of plot obtained:
   a. Change the background theme to no color.
   b. Change the color of values on x and y axis to “blue”

Choose a data of your choice and build interesting plots

If you are unsure, choose a data from the default R datasets and pick one which interests you and build at least two plots which convey something meaningful. If you are still unsure, use the HR_attrition.csv file to come up with stats about attrition.