

# CT5100-Assignment 4: Part 1

## Summary

The plot shows the comparative ranking of students across the Organization for Economic Co-operation and Development (OECD) member states.

Visualisation decisions of the plot are as follows:

1. Data : Variables to be plotted on X and Y axis are OECD member countries and the comparative rankings respectively. The data is ordered in the increasing order of rankings from low to high and filtered according to year 2018.
2. Aesthetics: Two groups are shown here Boys and Girls, Boys are represented using filled circles with a shade of blue and Girls are represented using an unfilled diamond shape with same shade of blue as border colour. Data points for Ireland and its X axis label is coloured red so that its easily distinguishable.
3. Geometrics: For geometrics line, points and segments are used. A segment from the X-axis touches the bottom data point (Boys in this case) and a line between Boy and Girl data point is present for the same X value.
4. Themes:
  - Background: A tint of blue colour is used as background colour. The plot doesn't appear to have a margin and title.
  - Axes: The X axis line was removed, and the axis text were tilted to 45-degree angle so that the labels or country names could be shown clearly. The Y axis was also removed and a continuous scale break of 20 were given with lower limit 340 and upper limit 560. The X and Y axis title were also removed.
  - Gridlines: The panels X major and minor gridlines, ticks along with Y minor and ticks were removed from the axes. The horizontal grid lines were reduced in size and had given a white color running from left to right.
  - Legend: The legend orientation was changed to horizontal and its position were changed to bottom left corner. The legend title was also removed.

How to reproduce this plot with ggplot in R:

1. For plotting, we need X values i.e. country (location) names, Y axis values and subject – Boy or Girl. **Group** the data by year, location, subject and **summarise** their performance value. Keep only the information for 2018, order it according to performance and given a **flag** value for observations for Ireland as TRUE and rest false. An observation of OECD average was also found and added to the dataset using **mean** on the performance values and **rbind** on the dataset to add a row.
2. Give location as **X aesthetic** and Values as **Y aesthetic**, and **shape** as subject.
3. Use **geom\_point()** with group as subject and **geom\_line()** to plot the line between the subjects. The custom colors and shapes are given by **scale\_color\_manual()** and **scale\_shape\_manual()**, **geom\_segment()** is used to draw line from X axis to the data point.
4. **scale\_y\_continuous()** with **labels** and **limits** were added to the plot.
5. Themes were added to the plot using **theme** with the background colour and plot margins were removed. The **axis.line.x**, **axis.ticks.x**, **axis.line.y**, **axis.ticks.y**, **axis.title.x** and **axis.title.y** were given an **element\_blank()** to remove them. To tilt the text **axis.text.x** was updated using **element\_text()** and **angle**.
6. The **panel.grid.major** and minor for x and y were also given **element\_blank()** to remove them.
7. The legend was manipulated by giving **legend.title** blank, changed position and direction by **legend.position** and **legend.direction = "horizontal"**.
8. To this plot add a separate **geom\_point**, **line**, **segment** and **text** to show the Ireland data point in red colour the same dataset with flag TRUE should be used for its aesthetics.