



## Digital Receipt

This receipt acknowledges that **Turnitin** received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author: Nakul Joseph  
Assignment title: Data Visualisation Assignment 4: Pa...  
Submission title: Assignment4\_Part2&3.pdf  
File name: Assignment4\_Part2\_3.pdf  
File size: 339.74K  
Page count: 7  
Word count: 1,621  
Character count: 8,673  
Submission date: 23-Mar-2020 12:53PM (UTC+0000)  
Submission ID: 1280355974

### CTS100-Assignment 4: Part 2

#### 1. Explanation

This section contains the explanation of how to recreate the OECD – PISA Reading performance 2018 graph with R's ggplot library.

##### 1.1 Data

To begin with, The OECD-PISA.csv file was read in and its rows were filtered over the year 2018 and unwanted columns were removed.

- o A new data frame was created with Location/Country ordered in ascending by subject 'BOY's value.
- o A new **Location** column was added with full name of each location so that it can be easily manipulated while plotting.
- o A **Segment** column was also added with Subject 'BOY's value which was used to draw the segment lines between subject 'Boy' and 'Girl'.
- o For the countries to appear in the **custom order**, the Location factor was transformed to the order we required.
- o Finally, custom color patterns (red = "#E8272C", blue="#359EC8") were created to distinguish Ireland's data points with other locations.

##### 1.2 Aesthetics

- o Location on X-Axis and Value on Y-Axis.
- o SUBJECT was given as shape aesthetic.

##### 1.3 Geometric

- geom\_point() with color and fill as custom colours were given with a size of 3.
- Scale\_shape\_manual() is used to give shapes 16 for Boys and 23 for Girls. Shape 16 is filled round shape and 23 is diamond shape with option to fill. The labels are also specified using this method.
- geom\_segment() with colour white running from the X axis to Boys data point for each location.
- geom\_segment() with grey colour running from Boys data point to Girls data point for each location.
- scale\_y\_continuous() used to limit Y axis between 340 and 560, breaks were also added between the limits every 20 values.
- scale\_x\_discrete() used to add space between the Y axis and the first data point.

##### 1.3 Themes

- theme\_classic() was set as the plot theme.
- Axes:
  - o Axis line, axis ticks and titles were removed for both the axes.
  - o X axis text were tilted to 45°, given vertical and horizontal adjustments, font family set to Sans, font size to 7 and finally the text custom text colour was given.
  - o Y axis text were given a margin to move the text inside the plot using axis.text.y().
- Background:
  - o Using panel.background() the background colour ("F2F2F2") was added.
  - o panel.grid.major.y was given an element\_line() with size of 0.4, type "solid" and colour white.
  - o The plot margin was set using plot.margin() such that 0.9 margin was left in the bottom area of plot.
- Legend:
  - o Legend title was removed by giving legend.title as element\_blank().
  - o The legend position was changed to horizontal by setting option in legend.direction.
  - o Finally, Legend was positioned using legend.position option.

