Questions

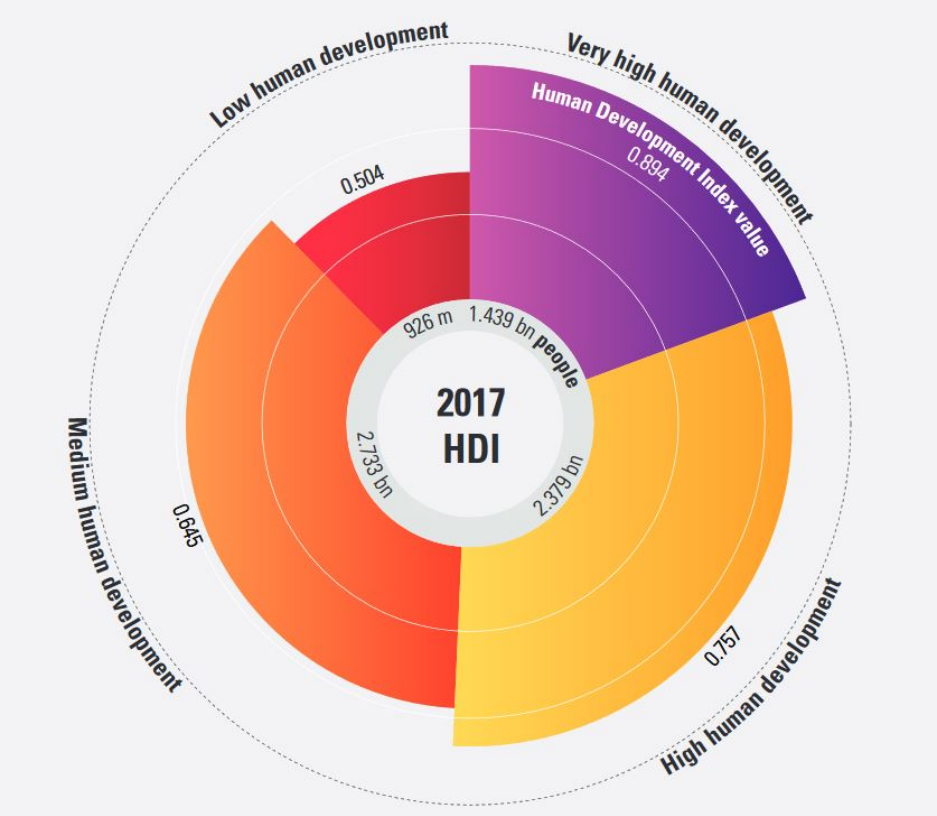
Human Development Index is a measure of average achievement in three key dimensions of human development: length of life, education level and standard of living. The HDI is the geometric mean of normalised indices for each of these three dimensions. HDI values are between 0 and 1; a value closer to 1 indicates that an individual from that country has more opportunities to develop to his or her full potential.

The health dimension is assessed by life expectancy at birth. The education dimension is measured by mean years of schooling for adults aged 25 years and more, and expected years of schooling for children. The standard of living is measured by national income per capita. The scores for the three dimensions are then aggregated into a composite index using the geometric mean.

The data in `Human Development Index (HDI).xlsx` contains the HDI ratings for 189 countries from 1990 until 2017. Some values are missing.

Question 1 (10 marks)

Here is a visualisation of HDI, taken from the internet.

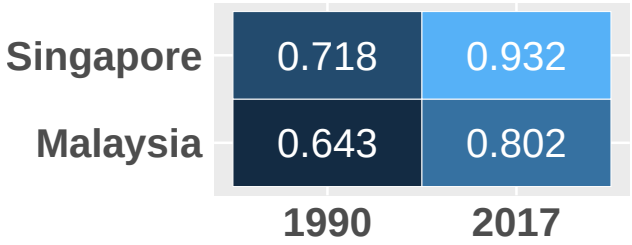


In the text box below, comment on the visualisation. Remember to address the following points:

1. What's good and what's bad about this graphic?
2. What do you understand/conclude from this graphic?
3. What alternative geoms could have been used?
4. What questions/visualisations would you follow up with?

Question 2 (10 marks)

Re-create the following plot, which depicts how HDI has changed from 1990 until 2017 for two countries in South-East Asia. Note that the fill aesthetic of the tile geom has been mapped to HDI.



Question 3 (10 marks)

Write a function `get_highest_rank()`, taking arguments country and data, that will compute and print the highest yearly rank that a country has reached, and the year in which it first achieved that rank. The function should return NULL. For instance, here is the output for Colombia.

Just to clarify, rank number 1 is the best - it corresponds to the country with the highest HDI for that year. Use the min_rank() function from within dplyr to compute the ranks.

```
get_highest_rank("Colombia", hdi_tidy)
```

```
## The highest rank achieved by Colombia was 76.
## This was first achieved in 1997.
```

Question 4 (20 marks)

Suppose we wish to invesigate the change in HDI over the years. Create **two plots** that visualise this data, and summarize your findings/investigations in words. There is an additional dataset (`country_conti.csv`) that groups the countries into broad geographical regions, in case you wish to use it.

Remember that you will be graded on your ability to follow up on an initial hypothesis, so the two plots should ideally be related, with the second one uncovering or explaining more about the data than the first.