**Exercise 1**

Open the file **World Happiness 2019.csv**. This dataset has a “happiness score” which was measured by simply asking the people in the country how happy they are on a scale of 0 to 10 and then calculating the average per country. It also contains many other variables which might be related to happiness. The dataset was compiled by the Sustainable Development Solutions Network (2020).

1. Create a scatterplot between the variables happiness score and GDP per capita. What do you see?
2. Calculate the correlation between the variables happiness score and GDP per capita. Does it match with the scatterplot?
3. Create a scatterplot matrix and a correlation heatmap. Are there variables which show no clear relationship with happiness score?

**Exercise 2**

Open the file **video\_games.xls**. This file has data from a large amount of video games sold by blockbuster. More information about this dataset can be found in the article written by Cox (2014). In this exercise you will try to predict the sales of the video games using linear regression.

Create four regression models. All of the should have Sales as the dependent variable but a different independent variable.

What do the coefficients and p-values tell you about the relationship between the sales and the independent variables?

**References**

Cox, J. (2014). What makes a blockbuster video game? An empirical analysis of US sales data. *Managerial and Decision Economics*, *35*(3), 189-198.

Sustainable Development Solutions Network (2020). *World Happiness Report*. Retrieved from: https://www.kaggle.com/unsdsn/world-happiness