1. From the data set "data" extract information you need and save it as "gapdata":

- you need only data of 2007 and all the continents except "Oceania" (instead of the including "==" operator that can help to select continents you need, you can choose the excluding"!=" operator that is opposite, when you don't want to include "Oceania" only).

 - you need only columns "continent" and "lifeExp".

2. Formulate your research question and null hypothesis, what would you like to test there?

3. Check the size of your sample. Is it large enough?

4. Check normality of the numerical variable. Use both analytical and graphical ways.

Do you meet the normality assumption?

5. Check normality of each group in the categorical variable.

Do you meet the normality assumption?

6. We failed the normality assumption. Which test should you perform instead of ANOVA?

Which assumption is needed to be met? Check this assumption.

7. Perform the test. What conclusion can you make? Do you reject or accept the null hypothesis?

8. You want to check between which groups this difference is significant using a Post-hoc test. Which test should you perform to check it?

Perform the test. Build the box plot. Does the box plot demonstrate the Dunn's test output?

9. If you have time, perform all the steps but instead of life expectancy choose the population ("pop") variable.

Check list

Null and alternative hypothesis (H0 and H1)	
Assumptions for ANOVA/ Kruskal-Wallis Test:	
Independence of observations	yes
Sample size	
Normality of the whole numerical variable	
Normality of each group	
Similar variance between the groups (for ANOVA)	
Similar distributions between the groups (for Kruskal-Wallis)	
Test to perform:	
Post-hoc test to perform:	
Conclusion:	