Problem Set #6 - LIFE EXPECTANCY

Life expectancy is often used as an indicator for the well-being of a country. Macroeconomists often speculate that the life expectancy is linked with the economic well-being of a country. To test this hypothesis, the LifeExpectancy.xlsx dataset contains the following information:

Variable Name Description

Country Country name

Group Is the country a member of OECD, Africa, or other?

PPGDP Per person GDP

The Group variable indicates if the country is a member of the Organization for Economic Co-operation and Development, an international think tank charged with promoting policies that will improve global social and economic well-being. Macroeconomists also hypothesize that OECD members will have longer life expectancy. For each of the following questions, assume that your audience are political strategists working for the OECD. Please attach your clearly commented code (R or SAS) as an appendix.

Statistical modeling will be used to help answer the following research questions:

RQ 1 – Is the economic well-being of a country linked with life expectancy? If so, quantify this relationship for political strategists.

RQ2 – Will OECD members have longer life expectancies? If so, quantify this relationship for political strategists.

1. Use the data to assess if a multiple linear regression (MLR) model is suitable to analyze the Life Expectancy data. Provide a plot of the life expectancy vs. ln(PPGDP) where you use different plotting symbols (or colors) for countries of each group type. Also provide boxplots of life expectancy by group. From these plots, provide a discussion on whether or not indicator variables for “Group” should be included in the model, as well as whether or not interactions between ln(PPGDP) and Group should be included in the model.
2. Write out (in mathematical form with Greek letters) an MLR model that would help answer the research questions stated above. Regardless of your answer in #1 above, include an interaction term between log(PPGDP) and Group. Clearly state any assumptions you are using in your model. Provide an interpretation of each mathematical term (variable or parameter) included in your model (be careful to interpret the interaction terms correctly). Using the mathematical form, discuss how your model, after ﬁtting it to the data, will be able to answer the questions in this problem.

Note: You will need two indicator variables for the Group variable. Set up indicator variables for “Africa” and “OCED”. **The category “Other” will be the baseline or control group**.

1. Fit your model in #2 to the Life Expectancy data and summarize the results by displaying the ﬁtted model in equation form (do NOT just provide a screen shot of the R or SAS output). Plot your ﬁtted regression line(s) on the scatterplot of life expectancy vs. log(PPGDP).
2. Remove the covariates from your model in #2 that do not have a significant impact on life expectancy. Use the T-tests to determine which are significant and which are not. Fit this reduced model to the Life Expectancy data and summarize the results by displaying the ﬁtted model in equation form (do NOT just provide a screen shot of the R or SAS output). Plot your ﬁtted regression line(s) on the scatterplot of life expectancy vs. log(PPGDP).
3. Assess the ﬁt of your model and justify your model assumptions using appropriate graphics or summary statistics. Provide discussion of your assessment of the model ﬁt and assumptions on the level of your target audience and interpret your model R2.
4. Regardless of your conclusions regarding the model assumptions in #5, perform an F-test to determine if the removable of the insignificant covariates in #4 was appropriate. Also perform an F-test to determine if any of the remaining predictors (covariates) signiﬁcantly effect life expectancy. Finally, construct 95% conﬁdence intervals for each term in the model and interpret these intervals in the context of life expectancy.
5. Use your analysis to answer both research questions posed by the political strategists.

RQ 1 – Is the economic well-being of a country linked with life expectancy? If so, quantify this relationship for political strategists.

RQ2 – Will OECD members have longer life expectancies? If so, quantify this relationship for political strategists.