**1. Problem Formulation and Structure of the Assignment**

In this exam, you must formulate your own research question and analyze it based on the Danish part of the European Social Survey 2018. Your research question must meet the following requirements: **It must address an ethnic background-related gap in an optional attitudinal outcome and attempt to explain this difference with confounding and/or mediating variables. Make sure that you choose to explain an ethnic background-related gap that is statistically significant!**

Examples of possible attitudes are: environmental attitudes, justice attitudes, attitudes towards redistribution and the welfare state, or gender role attitudes. You can be creative and choose the difference you find most interesting.

Examples of research questions:

1. Why do immigrants and their descendants have less trust in the police?
2. Is there an ethnic difference in attitudes towards redistribution, and if so, can this difference be explained by differences in education and income?

You must present a methodological design you will use to identify and then explain the gap. Note that this part has two subsections. (i) First, you must propose a design that you will implement (see below), and which is based on the statistical tools you have learned. In this part, you should explain the statistical tools, but also introduce the dataset and your operationalization of variables. (ii) Second, you must discuss which causal assumptions in your design may be problematic, and discuss an ideal design that would allow you to test the causal effects you are theorizing about. What could an RCT, an IV, or an RDD look like that would allow you to identify the causal effect of interest, and what are the crucial assumptions in this ideal setup?

In the third and final part, you must empirically implement your research design and test whether the factors you have identified really help explain the gap of interest, and to what extent.

**2. Data and Variables**

The data comes from round 9 of the Danish part of the European Social Survey (ESS), which was collected in 2018. The ESS interviews a representative sample of the adult population. You can read more about the ESS here: <https://www.europeansocialsurvey.org/>.

**2.1 Dependent Variables**

The dataset contains many possible dependent variables. You are free to choose one of these, as long as it is an attitude variable. If several variables appear to measure an overarching larger concept and have been measured on the same kind of scale (e.g., 1-10), such as attitudes towards immigration, you can combine them into one additive scale by adding the values of the variables you want to combine and dividing by the number of variables you have combined: x = (x1 + x2 + x3 + x4 + … Xk) / k.

**2.2 The Gap**

The gap must be identified by a simple binary dummy variable that distinguishes between immigrants and their descendants and native Danes (you can also focus on immigrants and their descendants outside the EU or any other subgroup of immigrant minorities you are specifically interested in, as long as you have more than 100 cases).

**2.3 Independent Variables**

The dataset contains a number of possible explanatory factors, from material position indicators such as education level, personal experiences such as discrimination, to subjective attitudes and beliefs. You are free to identify any explanatory factor of your choice. However, be sure to also consider important control variables. You are not expected to identify any factor whose causal effect you can perfectly identify. Nevertheless, try to consider the problem of causal identification when choosing your explanatory factors and the control variables or instruments you could potentially use.

**2.4 Weight**

Be sure to use the post-stratification weights in all your analyses.

**Requirements for the Empirical Analysis:**

* You must estimate a series of OLS regression models of your dependent variable.
* An initial model must identify and quantify the size of the gap you are interested in by regressing your outcome on the dummy variable that identifies the difference in gender, ethnicity, or parental education you are interested in.
* In a series of subsequent models, you must then try to explain the difference based on confounding and mediating factors. Express how much your explanatory factors can explain the difference.
* Note that the goal is not to add as many predictors until the difference is fully explained. The goal is to develop a theoretically sound and convincing conceptual framework that identifies between one and two factors, identifies them as confounding or mediating factors, and tests them.
* You must also test for at least one non-linear relationship or one interaction effect and plot these using model predictions – and interpret the interaction effect if it is statistically significant.
* You must use the design weight in all analyses.
* You must consider any limitations in your empirical design in relation to being able to make causal interpretations. Could there be conditions that challenge a causal interpretation?

**4. Practical Information**

* The scope of the assignment is a maximum of 10 pages of 2400 characters if it is done by one student. 5 pages are added per student if the assignment is done in groups. A reasonable (but in no way binding) disposition of an assignment of 10 pages could be the following: introduction (1 page), conceptual framework (2 pages), methodology/data/variables (2.5 pages), empirical analysis (4 pages), and conclusion (0.5 page).
* You must not write your name in the assignment – all assessments are anonymous.
* You must add your declaration of the use of AI as explained in the Study Regulations!