Weekly Homework II
Survey Methodology I
Due date: **December 20, 2023**

1. Design effects are used to evaluate the precision of statistics for different sample designs. Based on the readings assigned to this class (Sections 4.4, Survey Methodology book by Groves et al.):

    a. Is the design effect of a clustered element sample likely to be larger or smaller than one?
    b. Is the design effect of a stratified element sample likely to be bigger or smaller than one?
    c. In a single-stage clustered sample, if within a cluster, a variable has nearly the same value for all elements within the cluster, what value will the intraclass correlation be close to?
    d. For a single-staged clustered sample, the intraclass correlation for a key variable is 0.021, and the cluster size is 11. Calculate the design effect for the mean of that key variable.
    e. What does the design effect in part d mean?

2. You have received a dataset that includes two variables: gender (female and male) and age group (18-30, 31-50, 51 or more). The survey file, named "survey_data.csv," can be found in the assignment folder. Furthermore, you have access to the gender and age composition data from the Census:

	Male	Female	Total
18-30	15.89%	16.82%	32.71%
31-50	21.19%	23.10%	44.29%
51+	10.94%	12.06%	23.00%
Total	48.02%	51.98%	100.00%

    a. Create adjustment weights to align the sample with the provided table. Calculate the design effect associated with the estimated weights and estimate the proportion of drug use, along with confidence intervals.
    b. Generate raking weights to adjust only the marginal distributions. Calculate the design effect associated with the estimated weights and estimate the proportion of substance use, along with confidence intervals.
    c. Based on the findings, explain and justify which solution is better.