

Learn By Doing – One-Way ANOVA – STEP 2

In each of the following three questions, you'll find two designs for comparing number of credits taken by freshmen versus sophomores versus juniors versus seniors. In each case, one of the designs should not be handled with ANOVA. Your task is to identify which of the two it is.

Question 1:

(i) Survey a random sample of 40 seniors and get them to report the number of credits they took in each of their four years, then compare the four sample mean numbers of credits.

(ii) Survey random samples of 40 seniors, 40 juniors, 40 sophomores, and 40 freshmen, and get them to report the number of credits they are taking.

Which of the above two designs should not be handled with ANOVA?

design (i)

Correct. Indeed, this design should not be handled with ANOVA since the samples are not independent (all four samples have the same 40 students).

design (ii)

Incorrect. Note that in design (i), the samples are not independent (all four samples have the same 40 students), and therefore ANOVA should not be used. In design (ii), the samples are independent (since each of them was chosen randomly) which is required in order to use ANOVA.

Question 2:

(i) Survey random samples of 5 seniors, 5 juniors, 5 sophomores, and 5 freshmen, and get them to report the number of credits they are taking. Part-time students are included, so there are some extreme low outliers in the data.

(ii) Survey random samples of 5 seniors, 5 juniors, 5 sophomores, and 5 freshmen, and get them to report the number of credits they are taking. Only full-time students are being considered, and the distributions do not display any skewness or outliers.

Which of the above two designs should not be handled with ANOVA?**design (i)**

Correct. Indeed, this design should not be handled with ANOVA. The sample sizes are quite low (all of size 5), and there is a violation of the normality assumption in the form of extreme outliers.

design (ii)

Incorrect. Note that in design (i), the sample sizes are quite low (all of size 5), and there is a violation of the normality assumption in the form of extreme outliers. In design (ii), the sample sizes are also small, but the data do not display any violation of the normality assumption, so ANOVA can be used.

Question 3:

(i) Survey random samples of 40 seniors, 40 juniors, 40 sophomores, and 40 freshmen, and get them to report the number of credits they are taking. Sample standard deviations are 3.1, 3.4, 1.9, and 2.7.

(ii) Survey random samples of 20 seniors, 20 juniors, 20 sophomores, and 20 freshmen, and get them to report the number of credits they are taking. Sample standard deviations are 3.1, 5.4, 1.9, and 2.7.

Which of the above two designs should not be handled with ANOVA?**design (i)**

Incorrect. Note that in design (ii) the largest among the sample standard deviations (5.4) is more than twice as large as the smallest one (1.9). Since this rule of the thumb is not satisfied, we cannot assume that condition (iii) of equal population standard deviations is satisfied, and cannot therefore use ANOVA. In design (i), the rule of thumb is satisfied ($3.4 / 1.9$ is less than 2).

design (ii)

Correct. Indeed, the largest among the sample standard deviations (5.4) is more than twice as large as the smallest one (1.9). Since this rule of thumb is not satisfied, we cannot assume that condition (iii) of equal population standard deviations is satisfied, and cannot therefore use ANOVA.