Exercise 1

a) We want to check if the average score on the OTIS scale for the people in the population from which we selected our sample is equal to 92 units. This problem involves a hypothesis test for the population mean. We will initially check if the following assumptions for the use of this parametric test are satisfied:

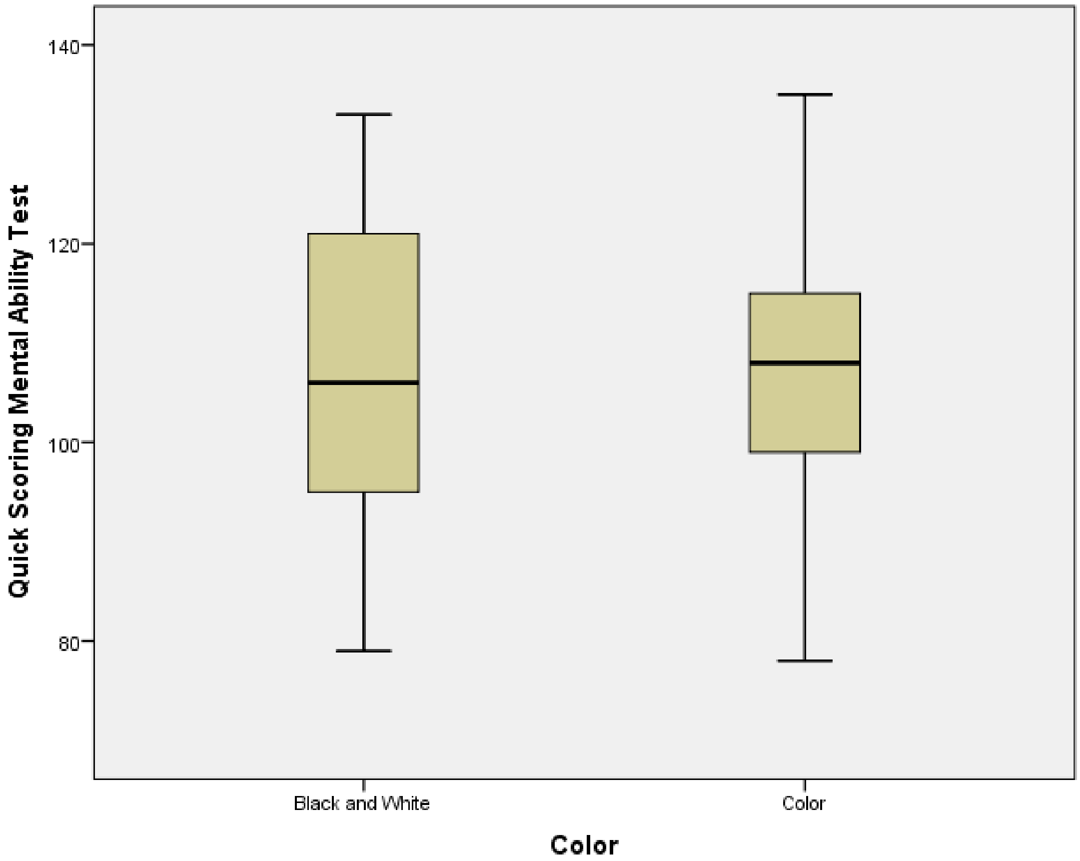
1. Our sample is random.
2. There are no extreme values in our data that exceed 10%.
3. The data follows a normal distribution. The first condition is related to how we selected our sample and is satisfied. The check for extreme values was done using a histogram and showed that there are no extreme values (see histogram 1). Furthermore, the Shapiro-Wilk test indicates that the assumption of a normal distribution for the OTIS score values is not met (test value 0.982, p-value 0.021). Transformations like log, x^2, x^(1/2), 1/x do not resolve the normality issue. Since the number of sample observations is 179 (>30), we proceed with the parametric test of the hypothesis, where the p-value and the confidence interval will be approximate. From the parametric t-Test for testing the hypothesis that the mean score on the OTIS scale for the people in the population is equal to 92 units, it is evident that the average score is statistically significantly different from 92 units (approximate p-value < 0.001). As the sign of the mean difference is positive (15.218), we conclude that the average score on the OTIS scale is statistically significantly higher than 92 units. Finally, an approximate 95% confidence interval for the population mean score on the OTIS scale is (92 + 13.18, 92 + 17.26).

b) We want to test if there is a difference in the OTIS scale with respect to race. Therefore, it is a test of equality of two means with independent random samples. To use the parametric t-test with independent samples to make a decision, the following assumptions need to be met:

1. Our samples are randomly selected.
2. There are no extreme values in the sample data for each population that exceed 10%.
3. Each population is adequately described by a normal distribution. The first condition is related to how we selected our samples and is satisfied. First, we check for the presence of extreme values in the sample observations for the OTIS score by race. The test for extreme values in the sample of 89 Black & White and 90 Color samples shows that there are no extreme values (see histogram 1). Then, we check the assumption that both samples come from normal populations. It turns out that the assumption that the Black & White sample comes from a normal population is not satisfied (Shapiro-Wilk p-value = 0.012), while the assumption for the Color sample is satisfied (Shapiro-Wilk p-value = 0.230). Transformations like log, x^2, 1/x, x^(1/2) do not resolve the normality issue for the Black & White sample. The number of sample observations for the OTIS score in the Black & White population is 89 (>30), so, using the central limit theorem, we proceed with a parametric test where the p-value and confidence intervals will be approximate. Therefore, we will use the parametric t-test to test the hypothesis of equality of the OTIS scale scores by race for the sample units. From theory, we know that the form of the

t-test to be used is determined by whether or not the population variances are equal. The assumption of equality of population variances is rejected (Levene's test, $F=4.552$, p-value=0.034). There is no difference in the OTIS scale scores with respect to race (p-value = 0.791, approximate), and an approximate confidence interval is (-3.542, 4.642).

Boxplot 1



c) We want to examine if there is a statistically significant difference in the OTIS scale scores with respect to education level. To use parametric statistical methods to make a decision, the following assumptions should be met:

1. Our samples are randomly selected.
2. There are no extreme values in the sample data for each population that exceed 10%.
3. Each population is adequately described by a normal distribution.

The first condition is related to how we selected our samples and is satisfied.

The check for extreme values in the samples of scores for the three education levels (preprofessional, professional, college student) showed that there are no extreme values (see histogram 1).

The test for the assumption that the data for the preprofessional education level follows a normal distribution is not rejected (Shapiro-Wilk test p-value = 0.155).

Similarly, the assumption for the professional education level is not rejected (Shapiro-Wilk test p-value = 0.127), and the assumption for the college student education level is also not rejected (Shapiro-Wilk test p-value = 0.090).

The assumption of equal population variances for the performance of the three education levels is rejected at a 5% significance level (Levene's test p-value = 0.018).

Therefore, to test whether there is a statistically significant difference in OTIS scale scores with respect to education, we will use the Welch test with a 5% significance level.

There is a statistically significant difference in OTIS scale scores with respect to education (Welch test p-value < 0.01). To pinpoint where the statistically significant differences in OTIS scale scores exist among the three education levels, we will use the Tukey multiple comparisons method with a 5% significance level.

The average OTIS scale score for the preprofessional education level differs statistically significantly from both the professional and college student levels and is the lowest, as indicated by the multiple comparisons table (Table 1).

Boxplot 1

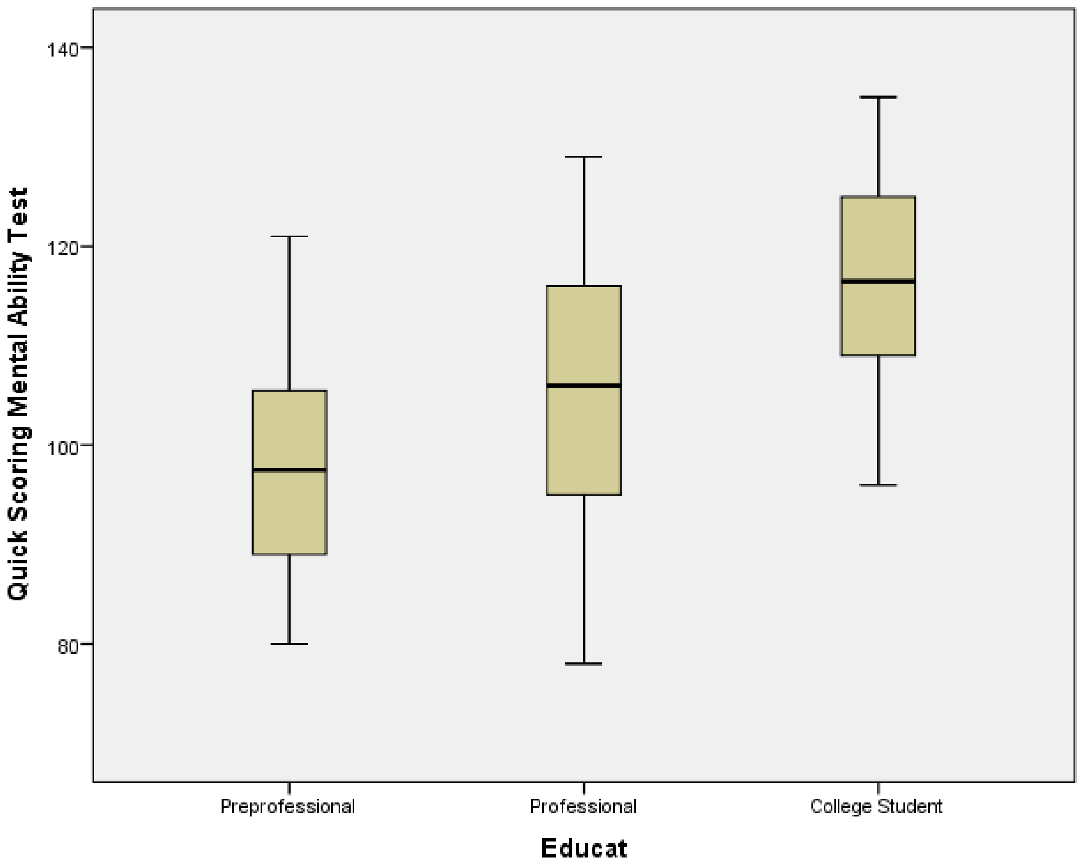


Table 1

	(I) Educat	(J) Educat	Mean Difference (I-J)
Tamhane	Preprofessional	Professional	-7,066*
		College Student	-18,632*
	Professional	Preprofessional	7,066*
		College Student	-11,566*
	College Student	Preprofessional	18,632*
		Professional	11,566*

δ) To examine if there is a statistically significant difference in the OTIS scale scores with respect to the region, you have set three assumptions for parametric statistical methods:

1. Samples should be randomly selected.
2. There should be no extreme values in the sample data for each population that exceed 10%.
3. Each population should be adequately described by a normal distribution.

The first condition is satisfied as it pertains to the way you selected your samples.

The check for extreme values in the samples of scores for the three regions (Hospital A, Hospital B, Hospital C, Student Dormitory) showed that there are no extreme values (see histogram 1).

The assumption that the data for Hospital A follows a normal distribution is not rejected (Shapiro-Wilk test p-value = 0.283). Similarly, the assumption for Hospital B is not rejected (Shapiro-Wilk test p-value = 0.752). However, the assumption for Hospital C is rejected (Shapiro-Wilk test p-value = 0.023), and the assumption for Student Dormitory is also rejected (Shapiro-Wilk test p-value = 0.090).

You addressed the non-normality issue in Hospital B by transforming the logarithm of the data. However, it introduced an extreme value in Hospital B with serial number 65 (see histogram 2). You continued your analysis by removing this specific observation from the sample units. You observed that the normality issue is corrected with the logarithmic transformation (Table 1).

Continuing, you will test if the mean of the logarithm of the OTIS scale scores differs significantly with respect to the region.

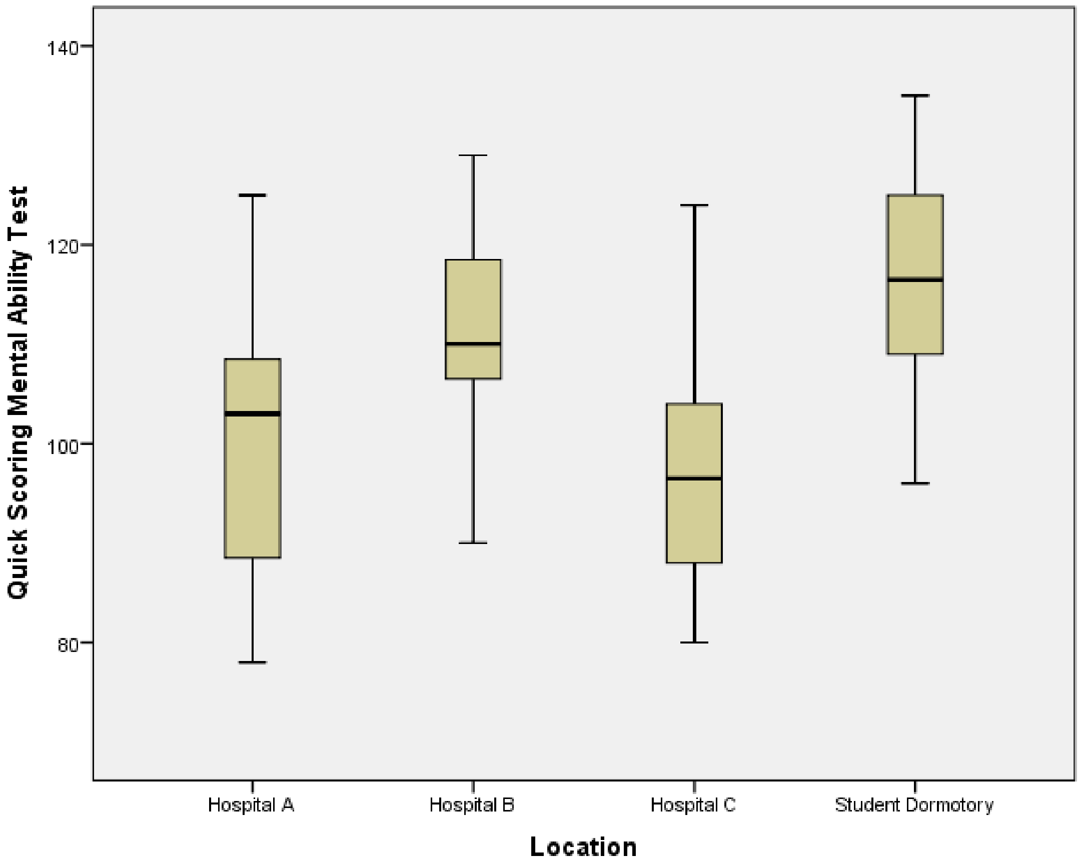
The assumption of equal population variances for the performance of the four available regions is rejected at a 5% significance level.

Therefore, to test whether there is a statistically significant difference in the logarithm of OTIS scale scores with respect to the region, you will use the Welch test with a 5% significance level.

There is a statistically significant difference in the logarithm of the OTIS scale scores among the four regions (Welch test p-value < 0.001). To pinpoint where the statistically significant differences in the logarithm of OTIS scale scores exist among the four regions, you will use the Tukey multiple comparisons method with a 5% significance level.

The logarithm of the OTIS scale scores in Hospital C differs statistically significantly from those in Hospital A, B, and Student Dormitory and is the worst, as indicated by the multiple comparisons table (Table 2).

Boxplot 1



Boxplot 2

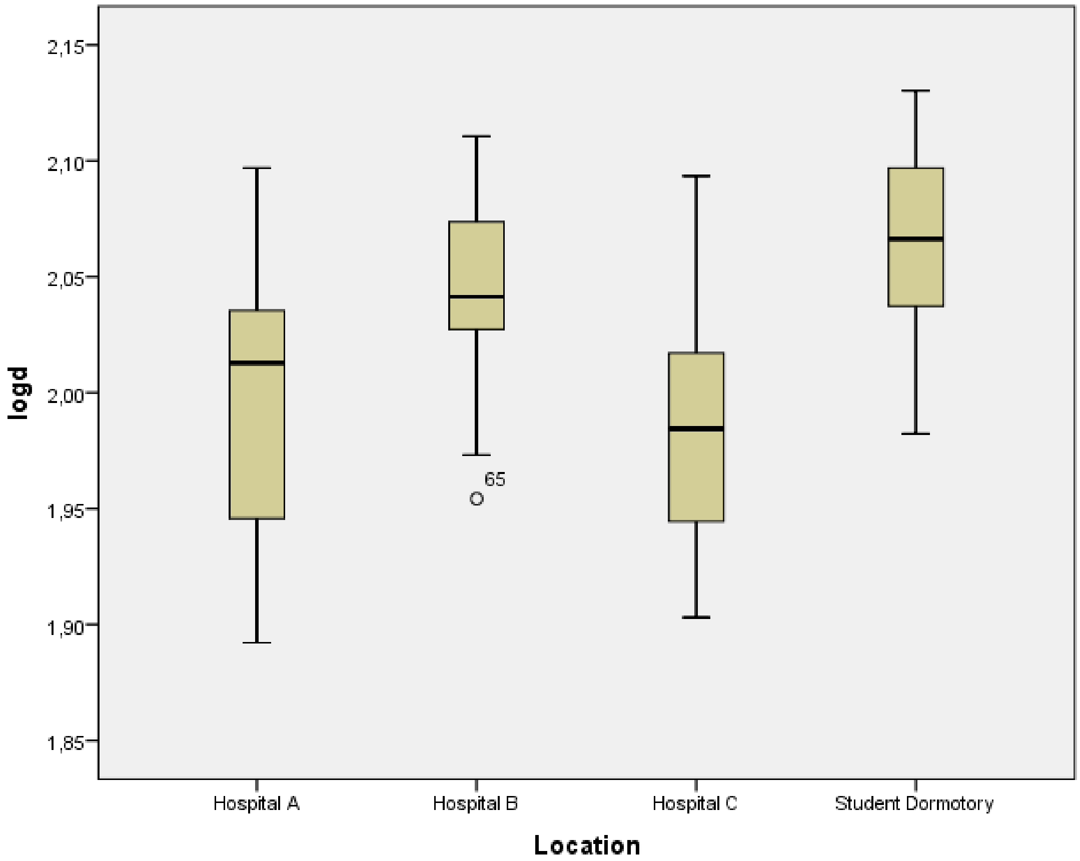


Table 1

	Location	Shapiro-Wilk		
		Statistic	df	Sig.
logd	Hospital A	,918	15	,181
	Hospital B	,972	27	,649
	Hospital C	,966	66	,067
	Student Dormotory	,968	70	,069

Table 2

	(I) Location	(J) Location	Mean Difference (I-J)
Tamhane	Hospital A	Hospital B	-,05186
		Hospital C	,00910
		Student Dormotory	-,06924*
	Hospital B	Hospital A	,05186
		Hospital C	,06096*
		Student Dormotory	-,01737
	Hospital C	Hospital A	-,00910
		Hospital B	-,06096*
		Student Dormotory	-,07833*
	Student Dormotory	Hospital A	,06924*
		Hospital B	,01737
		Hospital C	,07833*