

## Introduction

Conversion rate is more important than pure page views because restaurants would want action taken to make money and would pay accordingly for higher conversion rates. It could be that the new ad design brings in more clicks but lesser conversions, making it less effective in the eyes of the restaurants, the paying customers of Yelp-style site. Furthermore, page views without impressions does not make sense as the number of impressions could have been different.

We thought of doing a one-way ANOVA test. However, the assumptions of homogeneity of variance and normality for ANOVA was violated according to the Levene test and Anderson-Darling normality test below (p-values of less than 0.01 for both). Given that one-way ANOVA assumptions are violated, we did a Kruskal-Wallis rank sum test, which is the non-parametric alternative of one-way ANOVA. As the p-value is less than 0.01, we can conclude that there are significant differences between the 3 treatment groups for both conversion rates.

## Comparing call and reservation conversion rates between treatment 1 and 2

The reason for doing so is because we want to know whether the new ad design was more effective in terms of conversion rates compared to the current ad design, especially if the new ad design was more relevant than the current ad design such that conversion increases.

We would use the two-samples Wilcoxon rank sum test, which is a non-parametric statistical hypothesis test used to compare independent groups of samples to assess whether their population mean ranks differ, which is an alternative to the t-test since the population is not normally distributed.

Looking at the results of the 2 tests, both with alternative hypothesis that the median of treatment 1 is less than the median of treatment 2, the p-value is less than 0.01. This means we can reject the null hypothesis in favor of the alternative hypothesis at the 1% significance level. This means that the new ad design is more effective than the current ad design for both conversion rates.

We repeated this by differentiating restaurant type. It could be that treatment 2 could only be effective for one type. However, our p-value was still less than 0.01 for all the tests, so we can reject the null hypothesis again in favour of the alternative hypothesis at the 1% significance level. The new ad design is more effective than the current ad design for both conversion rates regardless of restaurant type.

## Comparing call and reservation conversion rates between treatment 0 and 2

Treatment 0 is the one with no advertisements (i.e. purely from organic search). It could be possible that the new ad design was not statistically better than having no advertisements. Then, the ad design would be ineffective and restaurant would be better off not paying the Yelp-style site. We did the two-samples Wilcoxon rank sum test again.

For the results of the 2 tests, both with alternative hypothesis that the median of treatment 0 is less than the median of treatment 2, we saw that the p-value is less than 0.01. This means we can reject the null hypothesis in favour of the alternative hypothesis at the 1% significance level. This means that the new ad design is more effective than the having no ad for both conversion rates.

Tests done for the whole report	p-values
Levene's Test for Homogeneity of Variance	< 0.00000000000000022
Anderson-Darling normality test	< 0.00000000000000022
Kruskal-Wallis rank sum test	< 0.00000000000000022

(p-values for all Kruskal-Wallis rank sum tests done all had the same p-values shown above because it was so extreme)