Course > Inference: Relationships C→Q > ANOVA > ANOVA: Example

🔖 **Bookmark this page**

ANOVA: Example

> **Learning Objective: In a given context, carry out the inferential method for comparing groups and draw the appropriate conclusions.**

> **Learning Objective: Specify the null and alternative hypotheses for comparing groups.**

Before we give you hands-on practice in carrying out the ANOVA F-test, let's look at another example:

Example

Do advertisers alter the reading level of their ads based on the target audience of the magazine they advertise in?

In 1981, a study of magazine advertisements was conducted (F.K. Shuptrine and D.D. McVicker, "Readability Levels of Magazine Ads," Journal of Advertising Research, 21:5, October 1981). Researchers selected random samples of advertisements from each of three groups of magazines:

Group 1—highest educational level magazines (such as Scientific American, Fortune, The New Yorker)

Group 2—middle educational level magazines (such as Sports Illustrated, Newsweek, People)

Group 3—lowest educational level magazines (such as National Enquirer, Grit, True Confessions)

The measure that the researchers used to assess the level of the ads was the number of words in the ad. 18 ads were randomly selected from each of the magazine groups, and the number of words per ad were recorded.

The following figure summarizes this problem:
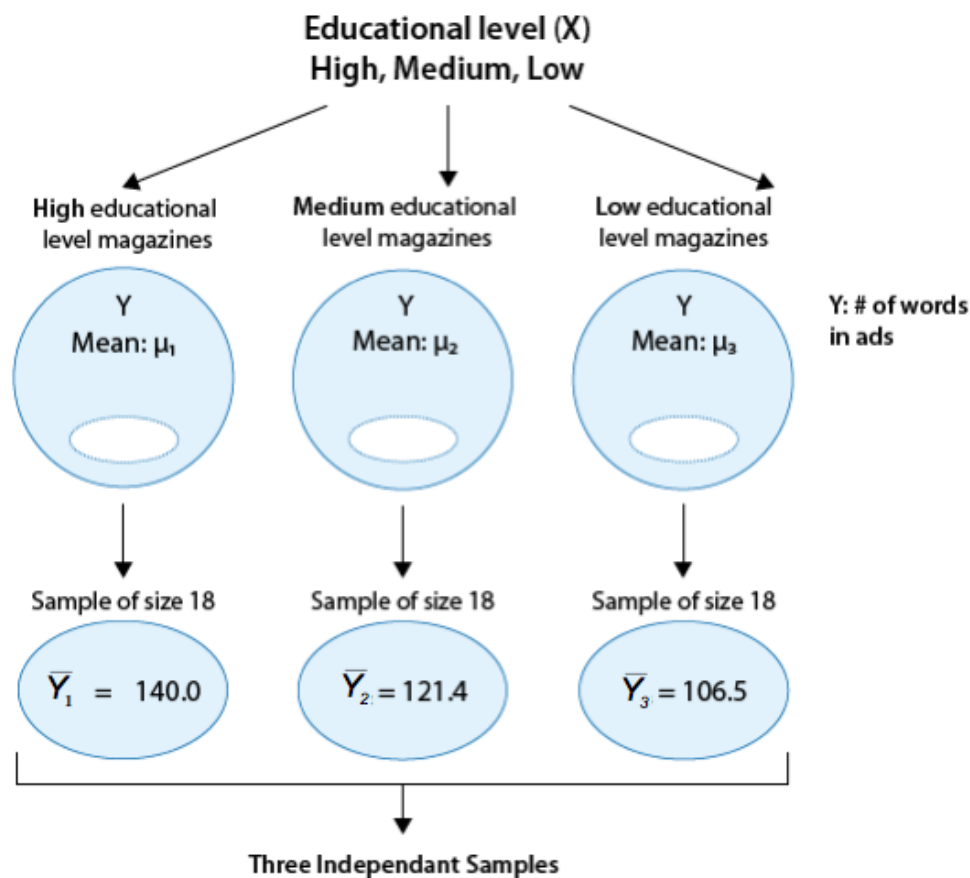
Our question of interest is whether the number of words in ads (Y) is related to the educational level of the magazine (X). To answer this question, we need to compare μ_1, μ_2, μ_3, the mean number of words in ads of the three magazine groups. Note in the figure that the sample means are provided. It seems that what the data suggest makes sense; the magazines in group 1 have the largest number of words per ad (on average) followed by group 2, and then group 3.

The question is whether these differences between the sample means are significant. In other words, are the differences among the observed sample means due to true differences among the µ's or merely due to sampling variability? To answer this question, we need to carry out the ANOVA F-test.

Step 1: Stating the hypotheses.

We are testing:

$$H_0 : \mu_1 = \mu_2 = \mu_3$$
$$H_a : not\ all\ the\ \mu's\ are\ equal$$

Conceptually, the null hypothesis claims that the number of words in ads is not related to the educational level of the magazine, and the alternative hypothesis claims that there is a relationship.

Step 2: Checking conditions and summarizing the data.

(i) The ads were selected at random from each magazine group, so the three samples are independent.

In order to check the next two conditions, we'll need to look at the data (condition ii), and calculate the sample standard deviations of the three samples (condition iii). Here are the side-by-side boxplots of the data, followed by the standard deviations:


```
Group 1 StDev: 74.0
Group 2 StDev: 64.3
Group 3 StDev: 57.6
```
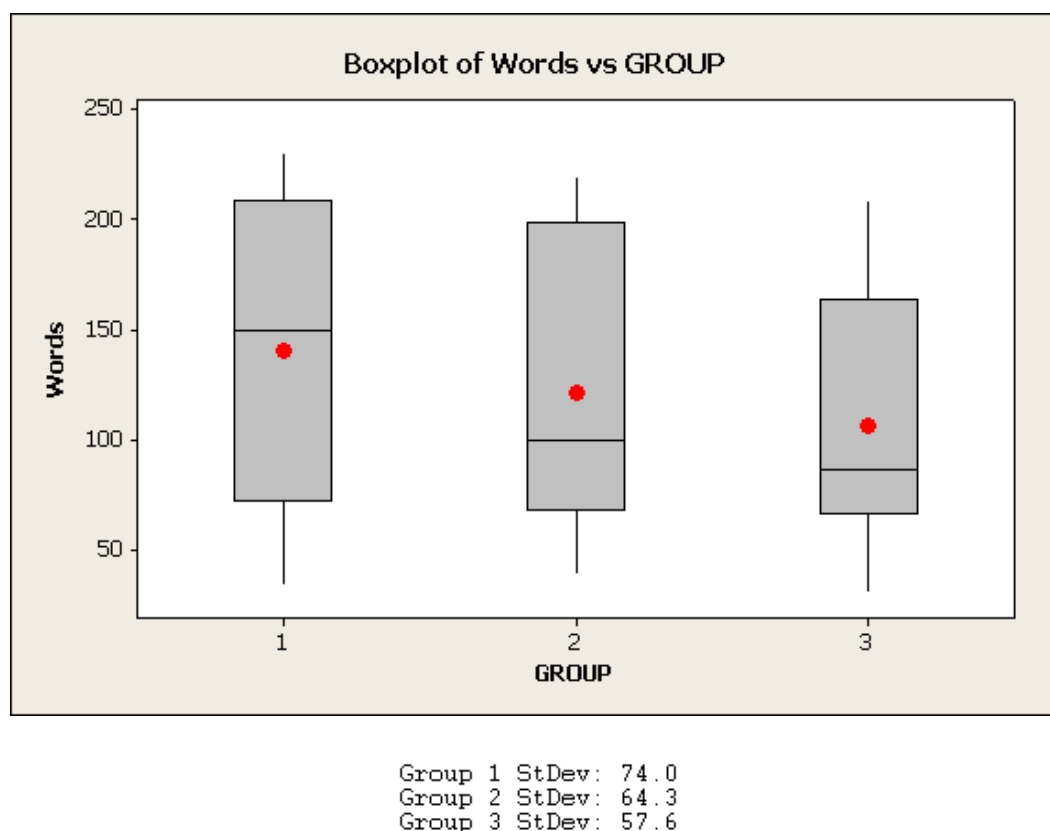
(ii) The graph does not display any alarming violations of the normality assumption. It seems like there is some skewness in groups 2 and 3, but not extremely so, and there are no outliers in the data.

(iii) We can assume that the equal standard deviation assumption is met since the rule of thumb is satisfied: the largest sample standard deviation of the three is 74 (group 1), the smallest one is 57.6 (group 3), and 74/57.6 < 2.

Before we move on, let's look again at the graph. It is easy to see the trend of the sample means (indicated by red circles). However, there is so much variation within each of the groups that there is almost a complete overlap between the three boxplots, and the differences between the means are over-shadowed and seem like something that could have happened just by chance. Let's move on and see whether the ANOVA F-test will support this observation.

Using statistical software to conduct the ANOVA F-test, we find that the F statistic is 1.18, which is not very large. We also find that the p-value is 0.317.

Step 3. Finding the p-value.

The p-value is 0.317, which tells us that getting data like those observed is not very surprising assuming that there are no differences between the three magazine groups with respect to the mean number of words in ads (which is what H_o claims).

In other words, the large p-value tells us that it is quite reasonable that the differences between the observed sample means could have happened just by chance (i.e., due to sampling variability) and not because of true differences between the means.

Step 4: Making conclusions in context.

The large p-value indicates that the results are not significant, and that we cannot reject H_o.

We therefore conclude that the study does not provide evidence that the mean number of words in ads is related to the educational level of the magazine. In other words, the study does not provide evidence that advertisers alter the reading level of their ads (as measured by the number of words) based on the educational level of the target audience of the magazine.