Block 13: Analysis of variance and covariance

(Activity solutions can be found at the end of the document.)

Analysis of variance (ANOVA) is a straightforward way to examine the differences between groups of responses which are measured on interval or ratio scales. If the set of independent variables consists of both categorical and metric variables, the technique is called analysis of covariance (ANCOVA). In this case, the categorical independent variables are still referred to as factors, whereas the metric independent variables are referred to as covariates.

Learning Objectives

- Discuss the scope of the ANOVA technique and its relationship to the *t* test and regression
- Describe one-way ANOVA, including decomposition of the total variation, measurement of effects significance testing and interpretation of results
- Describe nn-way ANOVA and the testing of the significance of the overall effect, the interaction effect and the main effect of each factor
- Describe ANCOVA and show how it accounts for the influence of uncontrolled independent variables
- Explain key factors pertaining to the interpretation of results with emphasis on interactions and the relative importance of factors.

Reading List

Malhotra, N.K., D. Nunan and D.F. Birks. Marketing Research: An Applied Approach. (Pearson, 2017) 5th edition [ISBN 9781292103129] Chapter 21 (excluding from 'Multiple comparisons' on page 622).

13.1 Analysis of variance and covariance

For each section of *Analysis of variance and covariance*, use the LSE ELearning resources to test your knowledge with the Key terms and concepts flip cards.

Overview

Analysis of variance (ANOVA) is used as a test of means for two or more populations. The null hypothesis, typically, is that all means are equal. ANOVA must have a dependent variable which is metric (measured using an interval or ratio scale). There must also be one or more independent variables which are all categorical (non-metric).

Categorical independent variables are also called factors. A particular combination of factor levels, or categories, is called a treatment.

One-way ANOVA involves only one categorical variable, or a single factor. In one-way ANOVA, a treatment is the same as a factor level. If two or more factors are involved, the analysis is termed n-way analysis of variance.

Figure 21.1 of the textbook shows the relationship between the t test, ANOVA, ANCOVA and regression.

Activity 13.1

Discuss the similarities and differences between analysis of variance and analysis of covariance.

Activity 13.2

What is the relationship between analysis of variance and the *t* test?

One-way analysis of variance

Market researchers are often interested in examining the differences in the mean values of the dependent variable for several categories of a single independent variable or factor. The following are examples.

- Do the various segments differ in terms of their volume of product consumption?
- Do the brand evaluations of groups exposed to different commercials vary?
- What is the effect of consumers' familiarity with the car dealership (measured as high, medium and low) on preference for the car dealership?

The one-way ANOVA statistics are as follows:

- **Eta squared, η2** - the strength of the effect of X (the independent variable or factor) on Y (the dependent variable) is measured by η2. The value of η2 varies between 0 and 1.
- **F statistic** - the null hypothesis that the category means are equal in the population is tested by an F statistic based on the ratio of mean square related to X and mean square related to error.
- **Mean square** - this is the sum of squares divided by the appropriate degrees of freedom.
- **$SS_{Between}$=SSX** - this is the variation in Y related to the variation in the means of the categories of X. This represents variation between the categories of X or the portion of the sum of squares in Y related to X.
- **SS_{Within}=SS_{Error}** - this is the variation in Y due to the variation within each of the categories of X. This variation is not accounted for by X.
- **SS_Y** - this is the total variation in Y.

The process to conduct one-way ANOVA is as follows:

Identify the dependent and independent variables

⇓

Decompose the total variation

⇓

Measure the effects

⇓

Test significance

⇓

Interpret the results

Conducting one-way ANOVA

Conducting one-way ANOVA

The total variation in Y, denoted by SS_Y, can be decomposed into two components:

$$SS_Y = SS_{Between} + SS_{Within}$$

where the subscripts 'Between' and 'Within' refer to the categories of X.

$SS_{Between}$ is the variation in Y related to the variation in the means of the categories of X. For this reason, $SS_{Between}$ is also denoted as SS_X.

SS_{Within} is the variation in Y related to the variation within each category of X. SS_{Within} is not accounted for by X. Therefore, it is also referred to as SS_{Error}.

The total variation in Y may be decomposed as:

$$SS_Y = SS_X + SS_{Error}$$

where:

$$SS_Y = \sum_{j=l}^{c}\sum_{i=l}^{nj}(Y_{ij} - \bar{Y})^2, SS_X = \sum_{j=l}^{c} n_j \left(\bar{Y}_j - \bar{Y}\right)^2 \ and \ SS_{Error} = \sum_{j=l}^{c}\sum_{i=l}^{nj}(Y_{ij} - \bar{Y}_j)^2$$

where:

- Y_{ij} = the ith observation in the jth category
- $\bar{Y}$ = the mean over the whole sample, i.e. the overall mean
- $\bar{Y}_j$ = the mean for category j.
- nj = the sample size for category j.

Table 21.1 of the textbook shows the decomposition of the total variation in one-way ANOVA.

In ANOVA, we estimate two measures of variation: within groups (SS_{Within}) and between groups ($SS_{Between}$). Therefore, by comparing the Y variance estimates based on between-groups and within-groups variation, we can test the null hypothesis.

The **strength of the effect** of X on Y is measured as follows:

$$\eta 2 = \frac{SSX}{SSY} = SSY - \frac{SSError}{SSY}.$$

The value of $\eta 2$ varies between 0 and 1, with larger values indicating stronger effects.

In one-way ANOVA, the interest lies in testing the null hypothesis that the *category means are equal in the population*, given by:

$$H_0: \mu 1 = \mu 2 = \mu 3 = \cdots = \mu c.$$

Under the null hypothesis, SSXSSX and SSErrorSSError come from the same source of variation. In other words, the estimate of the population variance of YY is:

S2Y=SSXc−1=Mean square due to X=MSXSY2=SSXc−1=Mean square due to X=MSX or:

$$S_Y^2 = \frac{SSError}{n - c} = Mean\ square\ due\ to\ error = MSError$$

where *n* is the overall sample size.

The null hypothesis may be tested by the *F* statistic based on the *ratio* between these two estimates:

$$F = \frac{SSX/(c-1)}{SSError/(n-c)} = \frac{MS_X}{MSError} \sim F_{C-1,\ n-c.}$$

This statistic follows an *F* distribution, with c−1 and n−c degrees of freedom (df), in the numerator and denominator, respectively. If the null hypothesis of equal category means is *not rejected*, then the independent variable *does not have a significant effect* on the dependent variable. On the other hand, if the null hypothesis is *rejected*, then the effect of the independent variable is *statistically significant*. A *comparison of the category mean values* will indicate the nature of the effect of the independent variable.

Activity 13.3

What is total variation? How is it decomposed in a one-way analysis of variance?

Activity 13.4

What is the null hypothesis in one-way ANOVA? Which basic statistic is used to test the null hypothesis in one-way ANOVA? How is this statistic computed?

Illustrative application of one-way ANOVA

Mercedes is attempting to determine the effect of dealership promotion, *X*, on sales, *Y*. For the purpose of illustrating hand calculations, the data are transformed to show the dealership sales, Y*ij*, for each level of promotion. The null hypothesis is that the category means are equal:

$$H_0 : \mu1 = \mu2 = \mu3.$$

Table 21.2 of the textbook provides the data on direct mail offer, dealership promotion, sales of new cars and clientele rating (the data can be downloaded in the file *Mercedes.sav* or can be seen in the table below).

Dealership number	Dealership offer	Dealership promotion	Sales	Clientele Rating
1	1	1	10	9
2	1	1	9	10
3	1	1	10	8
4	1	1	8	4
5	1	1	9	6
6	1	2	8	8
7	1	2	8	4
8	1	2	7	10
9	1	2	9	6
10	1	2	6	9
11	1	3	5	8
12	1	3	7	9
13	1	3	6	6

14	1	3	4	10
15	1	3	5	4
16	2	1	8	10
17	2	1	9	6
18	2	1	7	8
19	2	1	7	4
20	2	1	6	9
21	2	2	4	6
22	2	2	5	8
23	2	2	5	10
24	2	2	6	4
25	2	2	4	9
26	2	3	2	4
27	2	3	3	6
28	2	3	2	10
29	2	3	1	9
30	2	3	2	8

Table 21.3 of the textbook shows the effect of dealership promotion on sales of new cars.

To test the null hypothesis, the various sums of squares are computed as follows:

$$SS_Y = (10 - 6.067)^2 + (9 - 6.067)^2 + \cdots + (2 - 6.067)^2 = 185.867$$

$$SS_X = 10(8.3 - 6.067)^2 + 10(6.2 - 6.067)^2 + 10(3.7 - 6.067)^2 = 106.067$$

$$SS_{Error} = (10 - 8.3)^2 + \cdots + (8 - 6.2)^2 + \cdots + (2 - 3.7)^2 = 79.80.$$

It can be verified that:

$$SS_Y = SS_X + SS_{Error} \Rightarrow 185.867 = 106.067 + 79.80.$$

The strength of the effect of X on Y is measured as follows:

$$\eta 2 = \frac{SS_X}{SS_Y} = \frac{106.067}{185.867} = 0.571.$$

In other words, 57.1% of the variation in sales, YY, is accounted for by dealership promotion, XX, indicating a modest effect. The null hypothesis may now be tested:

$$F = \frac{(SSX/(c-1)}{SSError/n-c} = \frac{MS_X}{MSError} = \frac{106.067/(3-1)}{79.80/30-3} = 17.944.$$

Using statistical tables, $F_{0.05,2,27} = 3.35 < 17.944$, we reject H_0 and conclude that dealership promotion affects sales. The *p*-value, to three decimal places, is 0.000 which is less than any reasonable α when testing at the 100α% significance level (for α∈(0,1)), hence we also reject the null hypothesis when using the pp-value approach to hypothesis testing.

Table 21.4 of the textbook shows the statistical results of a one-way ANOVA investigating the effect of dealership promotion on the sale of new cars

Assumptions in ANOVA

The salient assumptions in ANOVA can be summarised as follows.

Ordinarily, the categories of the independent variable are assumed to be fixed. Inferences are made only to the specific categories considered. This is referred to as the *fixed-effects model*.

The error term is *normally distributed*, with a zero mean and a constant variance. The error is not related to any of the categories of *X*.

The error terms are *uncorrelated*. If the error terms are correlated (i.e. the observations are not independent), the *F* ratio can be seriously distorted.

n-way ANOVA

In market research, one is often concerned with the effects of *more than one factor simultaneously*. The following are examples.

- How do advertising levels (high, medium and low) interact with price levels (high, medium and low) to influence a brand's sales?
- Do educational levels (less than high school, high school graduate, some college and college graduate) and age group (under 35, 35-55, over 55) affect consumption of a brand
- What is the effect of consumers' familiarity with a car dealership (high, medium and low) and dealership image (positive, neutral and negative) on preference for the dealer?

Consider the simple case of two factors X1X1 and X2X2 having c1c1 and c2c2 categories, respectively. The total variation in this case is partitioned as follows:

$$SSTotal = SS\ due\ to\ X1 + SS\ due\ to\ X2 + SS\ due\ to\ interaction\ of\ X1\ and\ X2 + SSWithin$$

or:

$$SSY = SSX1 + SSX2 + SSX1X2 + SSError.$$

The strength of the joint effect of two factors, called the **overall effect**, or multiple eta-squared (η2), is measured as follows:

$$\text{multiple } \eta 2 = \frac{\text{SS}_{X1} + \text{SS}_{X2} + \text{SS}_{X1X2}}{SS_Y}.$$

If the overall effect is significant, the next step is to examine the **significance of the interaction effect**. Under the null hypothesis of no interaction, the appropriate F test is:

$$F = \frac{SS_{X1X2}/df_n}{SSError/df_d} = \frac{MS_{X1X2}}{MSError}$$

where:

- $df_n = (\text{c}1-1)(\text{c}2-1)$
- $df_d = \text{n}-\text{c}1\text{c}2$

The **significance of the overall effect** may be tested by an *F* test, as follows:

$$F = \frac{(SSX1 + SSX2 + SSX1X2)/dfn}{SSError/dfd} = \frac{SSX1, X2, X1X2/dfn}{SSError/dfd} = \frac{MS_{X1X2,X1X2}}{MSError}$$

where:

- dfn = degrees of freedom for the numerator, given by:

$$(c1-1)+(c2-1)+(c1-1)(c2-1)=c1c2-1$$

- dfd = degrees of freedom for the denominator = n−c1c2n−c1c2
- MS = mean square.

The **significance of the main effect of each factor** may be tested as follows for X_1:

$$F = \frac{SS_{X1}/df_n}{SSError/df_d} = \frac{MS_{x_1}}{MSError}$$

where:

- dfn = c1 − 1
- dfd = n − c1c2

Similarly, we may test for X_2 using:

$$F = \frac{SS_{X2}/df_n}{SSError/df_d} = \frac{MS_{x_2}}{MSError}$$

where:

- dfn=c2−1dfn=c2−1
- dfd=n−c1c2dfd=n−c1c2.

Returning to the Mercedes example, Table 21.5 of the textbook shows the statistical results of a two-way ANOVA.

Activity 13.5

How does n-way analysis of variance differ from the one-way procedure?

Activity 13.6

How is the total variation decomposed in n-way analysis of variance?

Analysis of covariance

When examining the differences in the mean values of the dependent variable related to the effect of the controlled independent variables, it is often necessary to take into account the influence of uncontrolled independent variables. For example:

In determining how different groups exposed to different commercials evaluate a brand, it may be necessary to control for prior knowledge

In determining how different price levels will affect a household's cereal consumption, it may be essential to take household size into account.

Suppose that we wanted to determine the effect of dealership promotion and direct mail offers on sales while controlling for the effect of clientele ratings.

Returning to the Mercedes example, Table 21.6 of the textbook shows the statistical results of an ANCOVA.

Activity 13.7

What is the most common use of the covariate in ANCOVA?

Issues in interpretation

Important issues involved in the interpretation of ANOVA results include interactions, the relative importance of factors and multiple comparisons.

Interactions:

- Figure 21.3 of the textbook provides a classification of possible interaction effects which could arise when conducting ANOVA on two or more factors. Figure 21.4 of the textbook shows examples of different patterns of interactions.

Relative importance of factors:

- Experimental designs are usually balanced, in that each cell contains the same number of participants. This results in an orthogonal design in which the factors are uncorrelated. Hence it is possible to determine unambiguously the relative importance of each factor in explaining the variation in the dependent variable.

The most commonly used measure in ANOVA is **omega squared**, ω^2. This measure indicates what proportion of the variation in the dependent variable is related to a particular independent variable or factor.

The relative contribution of a factor XX is calculated as follows:

$$\omega_X^2 = \frac{SS_X - (df_X \times MS_{Error})}{SS_{Total} + MS_{Error}}.$$

Normally, ω^2 is interpreted only for statistically significant effects. ω^2 associated with the level of dealership promotion is calculated as follows:

$$\omega_P^2 = \frac{106.067 - (2 \times 0.967)}{185.867 + 0.967} = \frac{104.133}{186.834} = 0.557$$

Note that:

$$SS_{Total} = 106.067 + 53.333 + 3.267 + 23.2 = 185.867.$$

Likewise, the ω^2 associated with direct mail offers is:

$$\omega_D^2 = \frac{53.333 - (1 \times 0.967)}{185.867 + 0.967} = \frac{52.366}{186.834} = 0.280.$$

As a guide to interpreting ω^2, a large experimental effect produces an index of 0.15 or greater, a medium effect produces an index of around 0.06, and a small effect produces an index of 0.01.

While the effect of promotion and direct mail offers are both large, the effect of promotion is much larger.

Activity 13.8

What is the difference between ordinal and disordinal interaction?

Activity 13.9

How is the relative importance of factors measured in a balanced design?

Discussion forum and Activities

To access the solutions to these questions and case study, click here to access the printable Word document or click here to go to LSE's Elearning resources.

Activitys on the block's topics

1. A market researcher wants to test the hypothesis that there is no difference in the importance attached to shopping by consumers living in Belgium, France, Germany and the Netherlands. A study is conducted and analysis of variance is used to analyse the data. The results obtained are presented in the following table.

Source	**Sum of squares**	**df**	**Mean squares**	ff**squares**	pp**-value**
Between groups	70.212	3	23.404	1.12	0.3
Within groups	20812.416	996	20.896		

2.

   a. Is there sufficient evidence to reject the null hypothesis?

   b. What conclusion can be drawn from the table?

   c. If the average importance was computed for each group, would you expect the sample means to be similar or different?

   d. What was the total sample size in this study?

3. An experiment tested the effects of package design and shelf display on the likelihood of buying a breakfast cereal. Package design and shelf display were varied at two levels each, resulting in a 2×22×2design. Purchase likelihood was measured on a seven-point Likert scale. The results are partially described in the following table.

Source of variation	**Sum of squares**	**df**	**Mean square**	F	p-**value**	ω^2
Package design	68.76	1				
Self display	320.19	1				
Two-way interaction	55.05	1				
Residual error	176.00	40				

4.
   a. Complete the table by calculating the mean square, F, p-value and ω^2 values.
   b. How should the main effects be interpreted?

5. In a pilot study examining the effectiveness of three commercials (A, B and C), 10 consumers were assigned to view each commercial and rate it on a 9-point Likert scale. The data obtained from the 30 participants are given in the data file *Commercial.sav*. (An Excel version of the dataset is *Commercial.xlsx*.)
   a. Are the three commercials equally effective on average?

Use **Analyze > Compare Means > One-Way ANOVA**……. Select ‘Effectiveness rating’ as the dependent variable, and ‘Commercial’ as the factor.

Video walkthrough of activity 3a.

   b. Produce some descriptive statistics, carry out a test for the homogeneity of variances, and produce a plot of means for the data. Interpret your results.

Use **Analyze > Compare Means > One-Way ANOVA**……. Select ‘Effectiveness rating’ as the dependent variable, and ‘Commercial’ as the factor. Select the ‘Options……’ box and check ‘Descriptive’, ‘Homogeneity of variance test’ and ‘Means plot’. Click ‘Continue’, then ‘OK’.

Video walkthrough of activity 3b.

6. In an experiment designed to measure the effect of gender and frequency of travel on preference for long-haul holidays, a 2 (gender) $\times$ 3 (frequency of travel) between-subjects design was adopted. Five participants were assigned to each cell for a total sample size of 30. Preference for long-haul holidays was measured on a 9-point Likert scale (1 = no preference, 9 = strong preference). Gender was coded as ‘1 = male’ and ‘2 = female’. Frequency of travel was coded as ‘1 = light’, ‘2 = medium’ and ‘3 = heavy’. The data obtained are given in the data file *Travel.sav*. (An Excel version of the dataset is *Travel.xlsx*.)
   a. Do males and females differ on average in their preference for long-haul travel? Hint: Perform an independent samples t test.

Video walkthrough of activity 4a.

   b. Do the light, medium and heavy travellers differ on average in their preference for long-haul travel?

Use **Analyze > Compare Means > One-Way ANOVA**……. Under 'Options……', select 'Descriptive'.

Video walkthrough of activity 4b.

c. Conduct a 2×32×3 analysis of variance with preference for long-haul travel as the dependent variable, and gender and frequency of travel as the independent variables. Interpret the results.

Use **Analyze > General Linear Model > Univariate**……. Choose 'Preference' as the 'Dependent Variable', and 'Gender' and 'Travel group' as the 'Fixed Factor(s)'. Under 'Options……', select 'Descriptive statistics'.

Video walkthrough of activity 4c.

Learning outcomes checklist

Use this to assess your own understanding of the chapter. You can always go back and amend the checklist when it comes to revision!

- Discuss the scope of the ANOVA technique and its relationship to the t test and regression
- Describe one-way ANOVA, including decomposition of the total variation, measurement of effects significance testing and interpretation of results
- Describe n -way ANOVA and the testing of the significance of the overall effect, the interaction effect and the main effect of each factor
- Describe ANCOVA and show how it accounts for the influence of uncontrolled independent variables
- Explain key factors pertaining to the interpretation of results with emphasis on interactions and the relative importance of factors.

Block 13: Analysis of variance and covariance

Solution to Exercise 13.1

Essentially, both analysis of variance (ANOVA) and analysis of covariance (ANCOVA) are used to test the significance of differences among two or more sample means, or equivalently, to test the hypothesis that all population means are equal. Both ANOVA and ANCOVA must have a dependent variable which is metric. They may have more than one independent variable. However, ANOVA requires that all the independent variables be categorical (non-metric). These categorical independent variables are also called factors. In ANCOVA, the independent variables may consist of both categorical and metric variables. Here the categorical independent variables are called factors, but the metric independent variables are referred to as covariates.

Solution to Exercise 13.2

Both techniques involve a metric dependent variable and are used to examine the significance of differences in sample means. However, the *t* test involves a single independent variable which is binary and is generally used in cases involving two sample means. However, analysis of variance may have more than one independent categorical variable and is used in cases involving two or more sample means.

Solution to Exercise 13.3

The total variation is defined as the sum of the squared deviations of each measurement from the overall mean and is given by:

$$SS_Y = SS_X + SS_{Error}$$

where:

$$SS_Y = \sum_{j=l}^{c} \sum_{i=l}^{nj} (Y_{ij} - \bar{Y})^2, SS_X = \sum_{j=l}^{c} n_j \left(\bar{Y}_j - \bar{Y}\right)^2 \; and \; SS_{Error} = \sum_{j=l}^{c} \sum_{i=l}^{nj} (Y_{ij} - \bar{Y}_j)^2$$

and:

$$SS_{Error} = \sum_{j=l}^{c} \sum_{i=l}^{nj} (Y_{ij} - \bar{Y}_j)^2$$

where:

- Y*ij* = the *i*th observation in the *j*th category
- $\bar{Y}$ = the mean over the whole sample, i.e. the overall mean
- $\bar{Y}_j$ = the mean for category *j*
- *Nj* = the sample size for category *j*.

$SS_{Between}$ represents the variation among the means of Y in the categories of X and involves the squares of the deviations of various category means from the overall mean.

SS_{Within} represents the variation in Y due to the variation within each category of X and involves the squares of the deviations of each measurement of Y from the corresponding category mean.

$SS_{Between}$ is often referred to as SS_X, and SS_{Within} is often referred to as SS_{Error}.

Solution to Exercise 13.4

The null hypothesis here is that all the category means are equal in the population, i.e. we have:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_c$$

It is called a joint hypothesis because several independent hypotheses are being assumed, namely:

$$H_0: \mu_1 = \mu_2, \quad H_0: \mu_1 = \mu_3, \quad H_0: \mu_2 = \mu_3 \text{ etc.}$$

The basic statistic used to test the null hypothesis in one-way ANOVA is the FF statistic. Therefore, the estimate of the population variance of YY can be based on either between-category variation or within-category variation, i.e. we have:

$$S_Y^2 = \frac{SS_X}{c-1} = \text{mean square due to } X = MS_X$$

or:

$$S_Y^2 = \frac{SS_{Error}}{n-c} = \text{mean square due to error} = MS_{Error}$$

where c is the number of categories and n is the total number of observations. Therefore, the null hypothesis may be tested by the F statistic based on the ratio between these estimates:

$$F = \frac{SS_X/(c-1)}{SS_{Error}/(n-c)} = \frac{MS_X}{MS_{Error}} \sim F_{c-1,n-c}$$

which follows an F distribution with c−1 and n−c degrees of freedom in the numerator and denominator, respectively.

Solution to Exercise 13.5

If there is only one categorical variable (or a single factor) we have one-way analysis of variance. However, if two or more factors are involved, the analysis is termed an n-way ANOVA and this allows the researcher to examine the interactions between the factors.

Solution to Exercise 13.6

Suppose Y is the dependent variable and X_1 and X_2 are two factors having c_1 and $c_2$categories, respectively. The total variation in this case is decomposed as:

$$SS_Y = SS_{X1} + SS_{X2} + SS_{X_1X_2} + SS_{Error}$$

Solution to Exercise 13.7

The covariate, an independent variable which is metric, is commonly used to remove the extraneous variations from the dependent variable by using regression procedures.

Solution to 13.8

In the ordinal case, the rank order of the effects due to one factor does not change across the levels of the second factor. In the disordinal case, however, the rank order changes. Besides, disordinal interaction is stronger than ordinal interaction.

Solution to 13.9

Usually, if an experimental design is balanced (i.e. each cell contains the same number of participants), then it results in an orthogonal design in which the factors are uncorrelated. Therefore, it becomes easier to determine accurately the relative importance of each factor in explaining the variation in the dependent variable. In ANOVA, the relative contribution of a factor X is calculated as:

$$\omega_X^2 = \frac{SS_X - (d_{f_X} - MS_{Error})}{SS_{Total} + MS_{Error}}$$

ω^2 is the most commonly-used measure in ANOVA for this purpose.

Discussion forum and exercises

Solutions to exercises on the block’s topics

1.

   a. No, with a p-value of 0.3 we cannot reject the null hypothesis.

   b. No difference in shopping behaviour exists between the countries.

   c. They should be similar.

   d. n=1000.

2.

   a. The complete table is:

Source of variation	**Sum of squares**	**df**	**Mean square**	F	p-**value**	$\boldsymbol{\omega^2}$
Package design	68.76	1	68.76	15.63	\lt 0.0001	0.103
Self display	320.19	1	320.19	72.77	\lt 0.0001	0.506
Two-way interaction	55.05	1	55.05	12.51	\lt 0.0001	0.081
Residual error	176.00	40	4.4			

   b. As the interaction effect is significant, the effect of one factor should be interpreted for each level of the other factor.

3.

a. We test the null hypothesis that all commercials are equally effective, against the alternative that they are not all equally effective. Specifically:

$H_0 : \mu_1 = \mu_2 = \mu_3$ vs. H_1 : Not all μis are equal.

Here we perform one-way ANOVA as there is one categorical independent variable (the commercial) to explain the continuous dependent variable (the effectiveness rating). Under H0H0, the test statistic is:

$$F = \frac{SS_X/(c-1)}{SS_{Error}/(n-c)} \sim F_{c-1,n-c}$$

The test statistic value, noting that c=3 and n=30, is:

$$f = \frac{46.667/(3-1)}{26.000/(30-3)} = \frac{23.333}{0.963} = 24.231,$$

For a 1% significance level, say, the critical value using statistical tables is $F_{0.01,2,27}$=5.49. Since 5.49 < 24.231 we reject the null hypothesis and conclude that there is strong evidence that the mean effectiveness ratings are not all equal across the commercials (i.e. that the commercials are not equally effective). Indeed, we could have instead consulted the p-value of the F statistic, which is 0.000, and hence we would have arrived at the same conclusion.

SPSS output is:

ANOVA

Effectiveness rating

	Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Between Groups	46.667	2	23.333	24.231	.000
Within Groups	26.000	27	.963		
Total	72.667	29			

b. Looking at the descriptive statistics output, we see the sample means for commercials A, B and C are $\bar{x}_A$ = 4, $\bar{x}_B$ = 5 and $\bar{x}_C$ = 7, respectively. These are the point estimates of the true group means, i.e. μ_A, μ_B, and μ_C, respectively.
The overall sample mean is 5.33 and this is the point estimate of the common group mean under the null hypothesis in (a). The one-way ANOVA *F* test was highly significant (see above) which means the test detected some significant difference(s) between the group sample means. However, does this mean all the group means are different, or is one different from the other two (which are the same)?

To answer this question we need to consider the 95% confidence intervals for each group mean and see whether these overlap.

Descriptives

Effectiveness rating

	N	Mean	Std. Deviation	Std.Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Commercial A	10	4.00	.816	0.258	3.42	4.58	3	5
Commercial B	10	5.00	1.054	0.333	4.25	5.75	4	7
Commercial C	10	7.00	1.054	0.333	6.25	7.75	5	*8*
Total	30	5.33	1.583	0.289	4.74	5.92	3	*8*

Noting that $t_{0.025,9}$=2.262, using statistical tables, from the output we see that 95% confidence intervals for the three groups are, respectively:

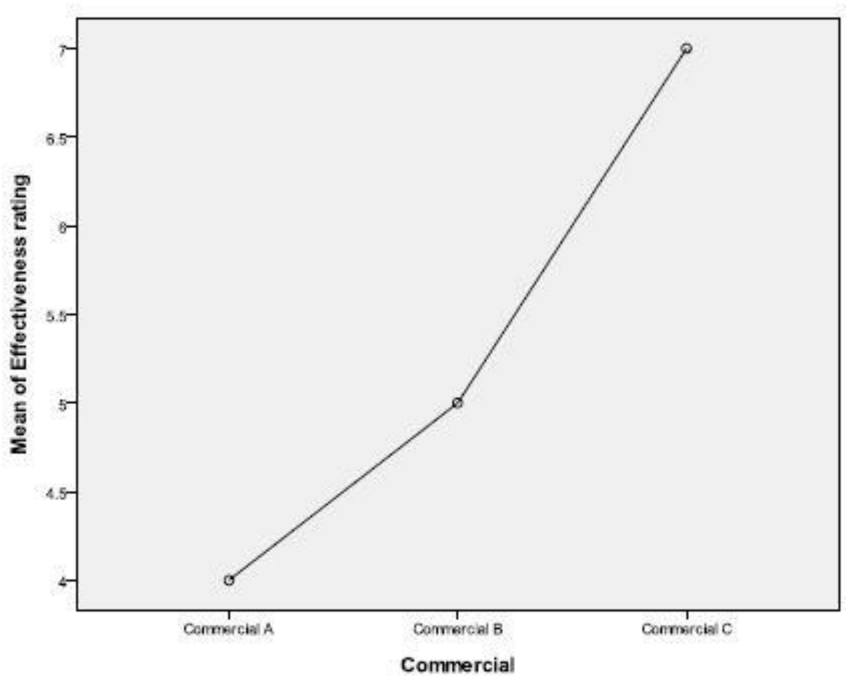
$$\bar{x}_A \pm t_{0.025}, n_{A-1} \times \frac{s_A}{\sqrt{n_A}} \Rightarrow 4 \pm 2.262 \times \frac{0.816}{\sqrt{10}} = 4 \pm 2.262 \times 0.258 \Rightarrow (3.42,4.58)$$

$$\bar{x}_B \pm t_{0.025}, n_{B-1} \times \frac{s_B}{\sqrt{n_B}} \Rightarrow 5 \pm 2.262 \times \frac{1.054}{\sqrt{10}} = 5 \pm 2.262 \times 0.333 \Rightarrow (4.25, 5.75)$$

$$\bar{x}_C \pm t_{0.025}, n_{C-1} \times \frac{s_C}{\sqrt{n_C}} \Rightarrow 7 \pm 2.262 \times \frac{1.054}{\sqrt{10}} = 7 \pm 2.262 \times 0.333 \Rightarrow (6.25, 5.75)$$

The first two confidence intervals overlap (the upper endpoint of the first (4.58) exceeds the lower endpoint of the second (4.25)), indicating no significant difference between the mean effectiveness ratings for commercials A and B. This is equivalent to saying we would not the reject the null hypothesis of $H_0 : \mu_A = \mu_B$ at the 5% significance level. However, the 95% confidence interval for commercial C is above the others, indicating that commercial C is more effective, on average, than commercials A and B.

The means plot illustrates this difference:



So we could rank the commercials from the least effective to the most effective as A, B and C, although there is no significant difference between the effectiveness of commercials A and B.

The one-way ANOVA test assumes the same variance for each group, i.e. the variance of effectiveness ratings should be a constant value, σ^2, for each group, so it would be advisable to test this assumption.

Test of Homogeneity of Variances

Effectiveness rating

Levene Statistic	df1	df2	Sig.
.375	2	27	.691

The output above is for the test of homogeneous (i.e. equal) variances, i.e. we test $H_0: \sigma_A^2 = \sigma_B^2 = \sigma_C^2$ against the alternative hypothesis that the variances are not all equal. We use the 'Levene statistic' (omitting the details here) and note the p-value is 0.691. Therefore, we fail to reject H_0 and are content that our assumption of homogeneous variances holds. In which case we can now improve on the earlier 95% confidence intervals for the individual group means. We can now effectively pool all 30 observations to derive a more precise estimate of the common variance, σ2σ2. The estimator is $S^2 = SS_{Error}$ / (n − c), which gives us a point estimate here of 0.963 (the within groups mean sum of squares in the one-way ANOVA table produced in (a)).

Noting that $t_{0.025,27} = 2.052$, using statistical tables, revised 95% confidence intervals are:

$$\bar{x}_A \pm t_{0.025,n-k} \times \frac{S}{\sqrt{n_A}} \Rightarrow 4 \pm 2.052 \times \frac{\sqrt{0.963}}{\sqrt{10}} \Rightarrow (3.36,4.64)$$

$$\bar{x}_B \pm t_{0.025,n-k} \times \frac{S}{\sqrt{n_B}} \Rightarrow 5 \pm 2.052 \times \frac{\sqrt{0.963}}{\sqrt{10}} \Rightarrow (4.36,5.64)$$

$$\bar{x}_C \pm t_{0.025,n-k} \times \frac{S}{\sqrt{n_C}} \Rightarrow 7 \pm 2.052 \times \frac{\sqrt{0.963}}{\sqrt{10}} \Rightarrow (6.36,7.64)$$

Note the confidence interval for μ_A is now slightly wider (due to $S_A^2 < S^2$, i.e. $(0.816)^2 = 0.666 < 0.963$, but the other intervals are now narrower.

One-way ANOVA also assumes the dependent variable (here the effectiveness rating) is normally distributed, which again should be tested. The popular Kolmogorov-Smirnov test would be fine, although we omit the details here for brevity.

4.

a. Given we are comparing two groups (males and females), it is appropriate to conduct an independent samples tt test. We test: $H_0 : \mu_1 = \mu_2$ vs . $H_1 : \mu_1 \neq \mu_2$

The SPSS output is:

Group statistics

	Gender	N	Mean	Std. Devidation	Std. Error Mean
Preferences	**Male**	15	4.93	2.314	.597
	Female	15	5.87	1.345	.347

Independent Samples Test

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Preference	Male	15	4.93	2.314	.597
	Female	15	5.87	1.345	.347

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Preference	Equal variances assumed	5.700	.024	-1.061	28	.298	-.733	.691	-2.149	.682
	Equal variances not assumed			-1.061	22.495	.300	-.733	.691	-2.165	.698

The FF test for the equality of variances, i.e.$H_0 : \sigma_1^2 = \sigma_2^2$, is significant (the test statistic value of 5.700 has a p-value of 0.024), so we reject the null hypothesis of equal variances, hence equal variances are *not* assumed. This means we look at the second row of the *t* test box of output. The *t* test is not significant (t = −1.061 and the p-value = 0.300). Therefore, the null hypothesis of no difference *cannot* be rejected.

b. The three usage groups differ in their preference for long-haul travel (the *F* statistic = 15.294 and the *p*-value = 0.000). Comparing means and 95% confidence intervals for each group, it is clear that heavy users exhibit the greatest preference.

Descriptives

Preference

	N	Meam	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
Lower Bound					**Upper Bound**			
Light	10	4.60	1.897	0.600	3.24	5.96	2	7
Medium	10	4.10	.876	.277	3.47	4.73	3	5
Heavy	10	7.20	1.033	.327	6.46	7.94	6	9
Total	30	5.30	1.896	.346	4.59	6.01	2	9

ANOVA

Preference

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	55.400	2	27.700	15.294	0.000
Within Groups	48.900	27	1.811		
Total	104.300	29			

c. The overall *F* test is significant (the *F* statistic = 20.743 and the *p*-value = 0.000). As individual factors, both gender and frequency of travel are significant, as is the interaction between gender and frequency of travel. For males, the heavy travellers exhibit the greatest preference and the light travellers the lowest preference. For females, the medium travellers exhibit the lowest preference.

Descriptive Statistics

Dependent Variable: Preference

Gender	Travel group	Mean	Std. Deviation	N
Male	Light	3.00	1.0000	5
	Medium	4.00	1.0000	5
	Heavy	7.80	.8370	5
	Total	4.93	2.3140	15

Female	Light	6.20	.8370	5
	Medium	4.20	.8370	5
	Heavy	6.60	.8940	5
	Total	5.67	1.3450	15
Total	Light	4.60	1.8970	10
	Medium	4.10	.8760	10
	Heavy	7.20	1.0330	10
	Total	5.30	1.8960	30

Test of Between-Subjects Effects

Dependent Variable: Preference

Source	Type Ill Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	84.700[a]	5	16.940	20.743	.000
Intercept	842.700	1	842.700	1031.878	.000
Gender	4.033	1	4.033	4.939	.036
Travel	55.400	2	27.700	33.918	.000
Gender • Travel	25.267	2	12.633	15.469	.000
Error	19.600	24	.817		
Total	947.000	30			
Corrected Total	104.300	29			

a. R Squared = .812 (Adjusted R Squared = .773)