
ROBERT HOLLAND/ALAMY

CHAPTER 15

Two-Way Analysis of Variance

Introduction

In this chapter, we move from one-way ANOVA, which compares means of several populations, to two-way ANOVA. Two-way ANOVA compares the means of populations that are classified in two ways or the mean responses in two-factor experiments.

- Brand advertising is frequently integrated into video games. Does brand recall depend on the type of controller used by a player or on the location of the ads (central or peripheral) in the game? Researchers investigate this using a racing game with products such as Snickers, Dr. Pepper, and Twix placed on billboards along the race course.[1]
- Coca-Cola marketers are interested in the effects of priming consumers on Powerade. They investigate the effect of this priming on the intent to purchase in both nonthirsty and thirsty consumers.[2]

Many of the key concepts are similar to those of one-way ANOVA, but the presence of more than one factor also introduces some new ideas. We once more assume that the data are approximately Normal and that groups may have different means but have the same standard deviation; we again pool to estimate the variance; and we again use F statistics for significance tests. *The major difference between one-way and two-way ANOVA is in the FIT part of the model.* We carefully study this term, finding much that is both new and useful.

We may, of course, have more than two factors. The statistical analysis is then called **higher-way ANOVA.** Although some details grow more complex, the most important ideas are already present in the two-way setting.

higher-way ANOVA

CHAPTER OUTLINE

15.1 The Two-Way ANOVA Model

We begin with a discussion of the advantages of the two-way ANOVA design and illustrate these with some examples. Then we discuss the model and the conditions that justify two-way ANOVA.

Advantages of two-way ANOVA

factor

In one-way ANOVA, we classify populations according to one categorical variable, or **factor.** In the two-way ANOVA model, there are two factors, each with several levels. When we are interested in the effects of two factors, a two-way design offers great advantages over two single-factor studies. We use several examples to illustrate these advantages.

EXAMPLE 15.1 Experimental Design 1: Best Magazine Layout and Cover

In Example 14.1 (page 712), a magazine publisher wants to compare three different magazine layouts. To do this, she plans to randomly assign the three design layouts equally among 60 supermarkets. The number of magazines sold during a one-week period is the outcome variable.

Now suppose a second experiment is planned for the following week to compare four different covers for the magazine. A similar experimental design will be used, with the four covers randomly assigned equally among the same 60 supermarkets.

Here is the design of the first experiment with the sample sizes:

Layout	n
1	20
2	20
3	20
Total	60

And here is the second experiment:

Cover	n
1	15
2	15
3	15
4	15
Total	60

In the first experiment, 20 stores were assigned to each level of the factor for a total of 60 stores. In the second experiment 15 stores were assigned to each level of the factor for a total of 60 stores. Each experiment lasts one week, so the total amount of time for the two experiments is two weeks.

Each experiment will be analyzed using one-way ANOVA. The factor in the first experiment is magazine layout with three levels, and the factor in the second experiment is magazine cover with four levels. Let's now consider combining the two experiments into one.

EXAMPLE 15.2 Experimental Design 2: Best Magazine Layout and Cover

Suppose we use a two-way approach for the magazine design problem. There are two factors, layout and cover. Because layout has three levels and cover has four levels, this is a 3×4 design. This gives a total of 12 possible combinations of layout and cover. With a total of 60 stores, we could assign each combination of layout and cover to five stores. The number of magazines sold during a one-week period is the outcome variable.

Here is a picture of the two-way design with the sample sizes:

	Cover				
Layout	**1**	**2**	**3**	**4**	**Total**
1	5	5	5	5	20
2	5	5	5	5	20
3	5	5	5	5	20
Total	15	15	15	15	60

cell Each combination of the factors in a two-way design corresponds to a **cell.** The 3×4 ANOVA for the magazine experiment has 12 cells, each corresponding to a particular combination of layout and cover.

With the two-way design for layout and cover, notice that we have 20 stores assigned to each level, the same as we had for the one-way experiment for layout alone. Similarly, there are 15 stores assigned to each level of cover. Thus, the two-way design gives us the *same amount* of information for estimating the sales for each level of each factor as we had with the two one-way designs. The difference is that we can collect all the information in only one experiment. This experiment lasts one week (instead of two weeks) and involves a single observation from each of the 60 stores. By combining the two factors into one experiment, we have increased our efficiency by reducing the amount of data to be collected by half.

EXAMPLE 15.3 Bundling to Introduce a New Product

Bundling is a marketing strategy that involves the sale of two or more separate products in one package. It is frequently used to introduce a new product or brand. For example, a new hair gel might be bundled with a popular shampoo, or a new cell phone might be bundled with a new wireless service contract. While much research has been done in terms of evaluating the quality of a bundle, relatively little has been done on the effects of bundling. Do the characteristics of the partnered product affect a consumer's opinion of the new product? One characteristic of interest would be the partnered product's functional relatedness, or complementarity.[3] If the partnered product is used in conjunction with the new product, does that enhance the consumer's perception of the new product?

To design a study to answer this question, we first need to determine an appropriate target group. This will depend on the new product under consideration. Suppose we're interested in introducing a new surround sound receiver and, therefore, decide to focus on college-age consumers of electronics. For our partnered products, we decide on a six-piece speaker system as the complementary product and a digital camera as the noncomplementary product. Because consumers have been shown to respond differently to products based on the product's brand image, our design should also take that into account. Let's consider a two-way ANOVA for this study.

EXAMPLE 15.4 Two-Way Design for the Bundling Study

The factors for this two-way ANOVA are the partnered product's complementarity and brand image, each with two levels. There are $2 \times 2 = 4$ cells in this study. If 200 consumers were recruited, we would then randomly assign 50 to each cell. The outcome variable will be a measure of the new product's image based on several 7-point scale questions.

Here is a table that summarizes the design:

	Complementarity		
Brand image	**Low**	**High**	**Total**
Low	50	50	100
High	50	50	100
Total	100	100	200

This example illustrates a second reason for using two-way designs. Although we're primarily interested in the effect of the partnered product's complementarity, we included the partnered product's brand image because we thought there might be an effect. Consider an alternative, one-way design where we assign 200 consumers to the two levels of complementarity and ignore brand image. With this design, we have the same number of consumers at each of the complementarity levels; thus, in this way, it is similar to our two-way design. However, suppose that there is, in fact, an effect of brand image. In this case, the one-way ANOVA would assign this variation to the RESIDUAL (within groups) part of the model. In the two-way ANOVA, brand image is included as a factor, and therefore this variation is included in the FIT part of the model. Whenever we can move variation from RESIDUAL to FIT, we reduce the σ of our model and increase the power of our tests.

REMINDER
DATA = FIT + RESIDUAL, p. 489

EXAMPLE 15.5 Disorganized Shelf Display and Product Availability

Despite a retailer's best efforts, products displayed on shelves can become disorganized and scarce. Does this affect the purchase likelihood of a product? Researchers consider the impact of shelf display (organized or disorganized) and product quantity (low, medium, high) on the purchase likelihood of a snack food product.[4]

Suppose we wanted to compare an organized and disorganized shelf display across three different product quantity levels. Instead of three separate one-way ANOVAs (or two-sample t tests), one for each quantity level, we make product quantity a factor in our design.

EXAMPLE 15.6 Two-Way Design for the Shelf Display Study

We use a 2×3 design for our shelf display study. The two factors are shelf display organization and product quantity. For product quantity, we compare the levels low (or scarce), medium, and high. The study that this example is based on used a convenience sample of undergraduate students taking a marketing class. Although aware of the potential hazards with this type of sampling, we obtain 72 student volunteers from our university. Each student is randomly shown one of six pictures of the shelf display. We measure purchase intent through a series of questions that are scored on a 1 to 7 scale.

Here is a table that summarizes the design with the sample sizes:

	Shelf display		
Product quantity	**Organized**	**Disorganized**	**Total**
Low	12	12	24
Medium	12	12	24
High	12	12	24
Total	36	36	72

interaction
main effects

This example illustrates a third reason for using two-way designs. The difference between the two shelf displays may depend on the product quantity. We call this an **interaction** when it occurs. In contrast, the average values for the shelf display effect and the product quantity effect are represented as **main effects.** The two-way model represents FIT as the sum of a main effect for each of the two factors *and* an interaction. One-way designs that vary a single factor and hold other factors fixed cannot discover interactions. We discuss interactions more fully later in this section.

These examples illustrate several reasons why two-way designs are preferable to one-way designs.

Advantages of Two-Way ANOVA

1. It is more efficient to study two factors simultaneously rather than separately.

2. We can reduce the residual variation in a model by including a second factor thought to influence the response.

3. We can investigate interactions between factors.

These considerations also apply to study designs with more than two factors. We are content to explore only the two-way case. The choice of the design for data production (sample or experiment) is fundamental to any statistical study. Factors and levels must be carefully selected by an individual or team who understands both the statistical models and the issues that the study will address.

The two-way ANOVA model

When discussing two-way models in general, we use the labels A and B for the two factors. For particular examples and when using statistical software, it is better to use names for these categorical variables that suggest their meaning. Thus, in Example 15.2, we would say that the factors are layout and cover. The numbers of levels of the factors are often used to describe the model. Again referring to Example 15.2, we would call this a 3×4 ANOVA. Similarly, Example 15.4 illustrates a 2×2 ANOVA. In general, Factor A will have I levels and Factor B will have J levels. Therefore, we call the general two-way problem an $I \times J$ ANOVA.

In a two-way design, every level of A appears in combination with every level of B, so that $I \times J$ groups are compared. The sample size for level i of Factor A and level j of Factor B is n_{ij}. The total number of observations is[5]

$$N = \sum n_{ij}$$

Assumptions for Two-Way ANOVA

We have independent simple random samples (SRSs) of size n_{ij} from each of $I \times J$ Normal populations. The population means μ_{ij} may differ, but all populations have the same standard deviation σ. The μ_{ij} and σ are unknown parameters.

Let x_{ijk} represent the kth observation from the population having Factor A at level i and Factor B at level j. The statistical model is

$$x_{ijk} = \mu_{ij} + \epsilon_{ijk}$$

for $i = 1, \ldots, I$ and $j = 1, \ldots, J$ and $k = 1, \ldots, n_{ij}$. The deviations ϵ_{ijk} are from an $N(0, \sigma)$ distribution.

The FIT part of the model is the means μ_{ij}, and the RESIDUAL part is the deviations ϵ_{ijk} of the individual observations from their group means. To estimate a population mean μ_{ij}, we use the sample mean of the observations from this group:

$$\bar{x}_{ij} = \frac{1}{n_{ij}} \sum_{k} x_{ijk}$$

The k below the Σ means that we sum the n_{ij} observations that belong to the (i, j)th group.

The RESIDUAL part of the model contains the unknown σ. We calculate the sample variances for each SRS and, provided the rule of thumb for equal standard deviations is met, pool these to estimate σ^2:

REMINDER rule for examining standard deviations, p. 720

$$s_p^2 = \frac{\sum (n_{ij} - 1)s_{ij}^2}{\sum (n_{ij} - 1)}$$

Just as in one-way ANOVA, the numerator in this fraction is SSE and the denominator is DFE. Also as in the one-way analysis, DFE is the total number of observations minus the number of groups. That is, DFE $= N - IJ$. The estimator of σ is s_p, the pooled standard error.

APPLY YOUR KNOWLEDGE

15.1 What's wrong? For each of the following, explain what is wrong and why.

(a) A two-way ANOVA is used when there are two outcome variables.
(b) In a 2×3 ANOVA, each level of Factor A appears with only two levels of Factor B.
(c) The FIT part of the model in a two-way ANOVA represents the variation that is sometimes called error or residual.
(d) You can perform a two-way ANOVA only when the samples sizes are the same in each cell.

15.2 Are some colors more attractive to impulsive shoppers? A marketing experiment compares four different colors of for-sale tags at an outlet mall. Each color tag is used for one week. Shoppers are classified as impulse buyers or not through a survey instrument. The total dollar amount each of the 138 shoppers spent on sale items is recorded. Identify the response variable, both factors, and state the number of levels for each factor (I and J) and the total number of observations (N).

15.3 Compare employee training programs. A company wants to compare three different training programs for its new employees. Each of these programs takes four hours to complete. The training can be given for four hours on one day or for two hours on two consecutive days. The next 90 employees hired by the company will be the subjects for this study. After the training is completed, the employees are asked to evaluate the effectiveness of the program on a 7-point scale. Describe the two factors, and give the number of levels of each and the total number of observations.

Main effects and interactions

Because we have independent samples from each of $I \times J$ groups, we can first think of the two-way ANOVA as a one-way ANOVA with IJ groups. Each population mean μ_{ij} is estimated by the corresponding sample mean $\bar{x}_{ij}$, and we can calculate sums of squares and degrees of freedom as in one-way ANOVA. Thus, SSG is the group sum of squares constructed from deviations of the form $\bar{x}_{ij} - \bar{x}$, where $\bar{x}$ is the average of all the observations and $\bar{x}_{ij}$ is the mean of the (i, j)th group. Similarly, DFG is simply $IJ - 1$.

The first important distinction between one-way and two-way ANOVA is that in two-way ANOVA we break down the FIT part of the model (the population means μ_{ij}) in a way that reflects the presence of two factors. This means the terms SSG and DFG are broken down into terms corresponding to a main effect for A, a main effect for B, and an AB interaction. Each of SSG and DFG is then a sum of terms:

$$\text{SSG} = \text{SSA} + \text{SSB} + \text{SSAB}$$

and

$$\text{DFG} = \text{DFA} + \text{DFB} + \text{DFAB}$$

The term SSA represents variation among the means for the different levels of Factor A. Because there are I such means, DFA $= I - 1$ degrees of freedom. Similarly, SSB represents variation among the means for the different levels of Factor B, with DFB $= J - 1$.

Interactions are a bit more involved. We can see that SSAB, which is SSG $-$ SSA $-$ SSB, represents the variation in the group means that is not accounted for by the main effects. By subtraction we see that its degrees of freedom are

$$\text{DFAB} = (IJ - 1) - (I - 1) - (J - 1) = (I - 1)(J - 1)$$

There are many kinds of interactions. The easiest way to study them is through some examples.

EXAMPLE 15.7 Per Capita Income

The American Community Survey provides annual per capita income for various subpopulations of the United States. Here are the per capita incomes for individuals of two ethnicities in two regions of the United States.[6] (The data include people who earned nothing.)

Region	White	Asian	Mean
Midwest	\$28,528	\$29,166	\$28,847
Northeast	\$35,192	\$32,295	\$33,744
Mean	\$31,860	\$30,731	\$31,295

The table also includes averages of the means in the rows and columns (rounded to the nearest dollar). For example, the second entry in the far-right margin is the average of the per capita income for individuals who claim to be of white or Asian descent in the Northeast:

$$\frac{35{,}192 + 32{,}295}{2} = 33{,}744$$

Similarly, the average per capita income of a white individual in the two regions is

$$\frac{28{,}528 + 35{,}192}{2} = 31{,}860$$

marginal means

These averages are called **marginal means** because of their location at the margins of the table. The grand mean (31,295 in this case) can be obtained by averaging either set of marginal means.

It is clear from the marginal means that white individuals have a higher per capita income than Asian individuals and that individuals in the Northeast have a higher per capita income than those in the Midwest. These are *main effects* for the two factors. We can describe the main effects by the differences between the marginal means. On average, a person of white descent has a per capita income $1129 higher than someone of Asian descent and a person in the Northeast has a per capita income that is $4897 more than an individual in the Midwest.

What about the interaction between region and ethnicity? An *interaction* is present if the main effects provide an incomplete description of the data. That is, if the ethnicity earnings gap is different in the two regions, then ethnicity and region interact. In this survey, the earnings gap is much larger in the Northeast:

	Midwest	Northeast
White-Asian difference	−$638	$2897

Figure 15.1(a) is a plot of the four group means. Because the difference between the per capita incomes is different in the two regions, the gap between the lines increases from left to right. That is, the white and Asian lines are not parallel.

How would the plot look if there were no interaction? No interaction says that the regional earnings gap is the same in both ethnicities. That is, the effect

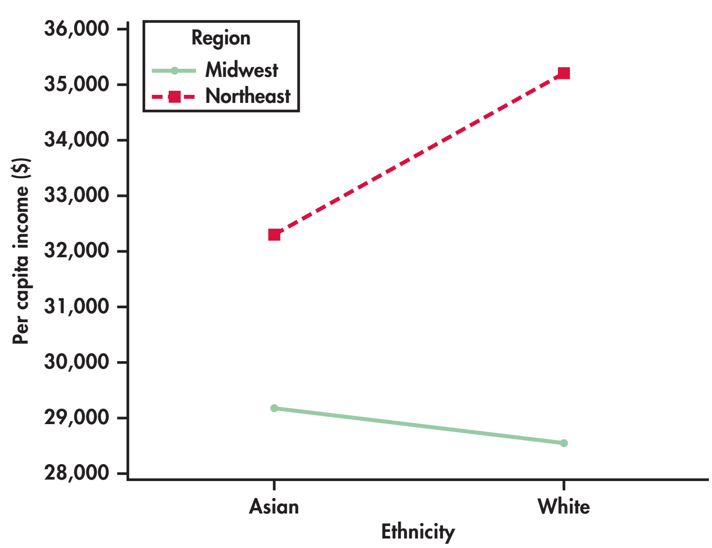
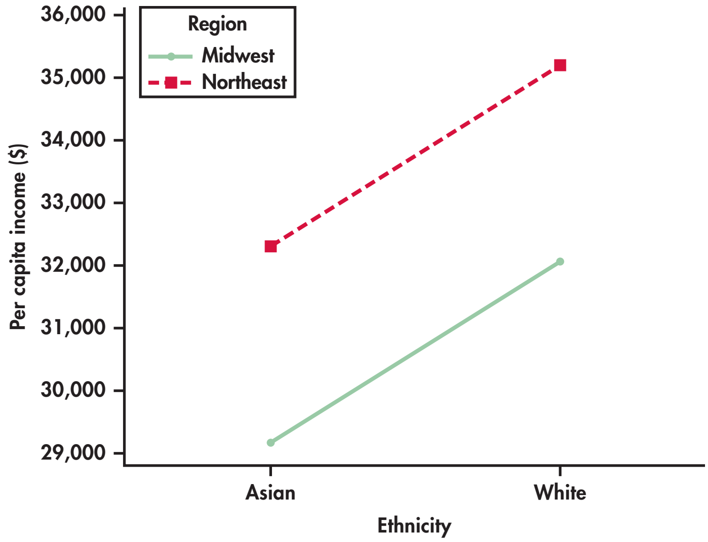


FIGURE 15.1 (a) Plot of per capita incomes for individuals of white or Asian descent in two U.S. regions, Example 15.7. Interaction between ethnicity and region is visible in the lack of parallelism of the lines.

FIGURE 15.1 (*Continued*) (b) The plot as it would appear if there were no interaction. The ethnicity lines are now parallel.



of ethnicity does not depend on region. Suppose that the gap were $2897 in both regions. Figure 15.1(b) plots the means. The white and Asian lines are now parallel. *Interaction between the factors is visible as lack of parallelism in a plot of the group means.*

To examine two-way ANOVA data for a possible interaction, always construct a plot similar to Figure 15.1. Profiles that are roughly parallel imply that there is *no* clear interaction between the two factors. When no interaction is present, the marginal means provide a reasonable description of the two-way table of means.

In this case, it is clear that the two profiles (the collections of marginal means for a given region) are not parallel. *When there is an interaction, the marginal means do not tell the whole story.* For example, with these data, the marginal mean difference between regions is $4897. This is larger than the difference for Asian individuals ($3129) and smaller than the difference for white individuals ($6664).

EXAMPLE 15.8 Per Capita Income, Continued

The American Community Survey in fact reports per capita income for four regions of the United States. Here are the data:

Region	White	Asian	Mean
Midwest	$28,528	$29,166	$28,847
Northeast	$35,192	$32,295	$33,744
South	$28,455	$30,246	$29,351
West	$30,264	$31,176	$30,720
Mean	$30,610	$30,721	$30,665

Including the additional regions changes the marginal means for ethnicity and the overall mean of all groups. Figure 15.2 is a plot of the group means. There is a clear main effect for region: the per capita income of both white and Asian individuals is highest in the Northeast and West. The plot does not show a clear main effect for ethnicity because of the Northeast region. In all other regions, a person of Asian descent has a larger per capita income, but in the Northeast, the Asian per capita income is lower. As a result, the two lines are not parallel, indicating that an interaction is present. *When there is interaction, main effects can be meaningful and important, but this is not always the case.* For example, the main effect of region is meaningful despite the interaction, but the main effect of ethnicity is not as clear because of the interaction.

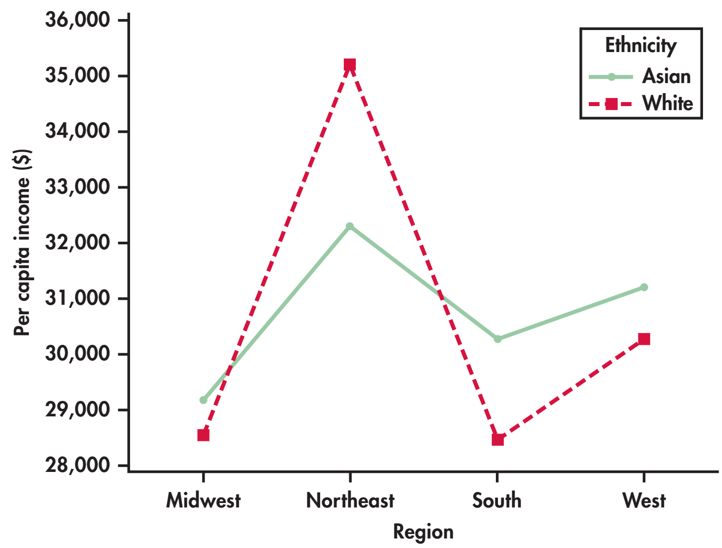


FIGURE 15.2 Plot of per capita income for individuals of white or Asian descent in four U.S. regions, Example 15.8. The plot shows a clear main effect for region as well as interaction between the two factors.

APPLY YOUR KNOWLEDGE

15.4 Marginal means. Verify the marginal mean for Asian given in Example 15.8. Then verify that the overall mean at the lower-right of the table is the average of the two ethnicity means and also the average of the four region means.

15.5 How do the differences depend on region? One way to describe the interaction between region and ethnicity in Example 15.8 is to give the differences between the per capita income of individuals of white descent and individuals of Asian descent in the four regions. Plot the differences versus region, and write a short summary of what you conclude from the plot.

15.6 Lack of interaction. Suppose that the difference between the per capita income of white and Asian individuals remained fixed at $2897 for all four regions in Example 15.8 and that the per capita income for whites in each region is as given in the table. Find the per capita income for individuals of Asian descent in each region, and make a plot of the eight group means. In what important way does your plot differ from Figure 15.2?

Interactions come in many forms. When we find them, a careful examination of the means is needed to properly interpret the data. Simply stating that interactions are significant tells us little. Plots of the group means are very helpful.

EXAMPLE 15.9 Bundling as a Product Introduction Strategy

In Examples 15.3 and 15.4, we discussed a two-way design to investigate the effects of the partnered product's brand image and complementarity on the new product's perceived quality. Here are the means from a study that performed an experiment like this. The Sony brand name was used for the strong brand image, and the Haier brand name was used for the weak brand image.[7] Figure 15.3 is the plot of the group means.

	Complementarity		
Brand image	**Low**	**High**	**Total**
Low	4.06	4.40	4.23
High	4.16	5.06	4.61
Total	4.11	4.73	4.42

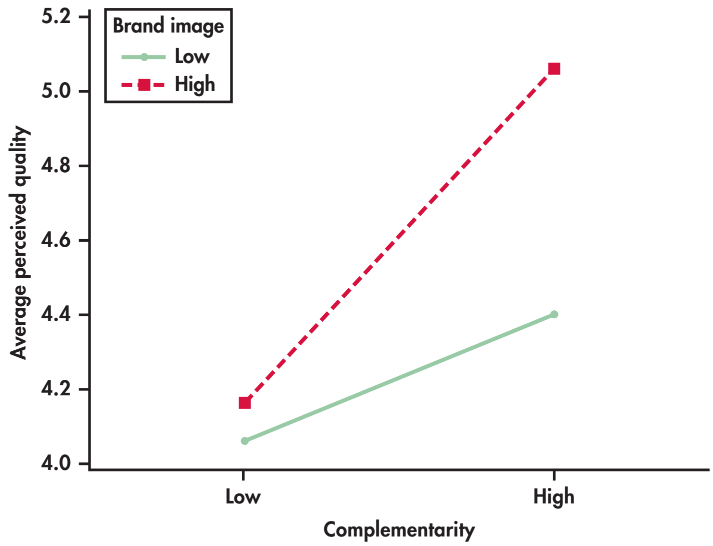


FIGURE 15.3 Plot of average perceived quality by complementarity and brand image, Example 15.9.

When the partnered product's brand image is low, perceived quality of the surround sound receiver is about the same. When the brand image is high, the perceived quality increases regardless of complementarity. However, the perceived quality increases more when the partnered product's complementarity is high (speaker system) than when it is low (camera).

In a statistical analysis, the pattern of means shown in Figure 15.3 produced significant main effects for brand image and complementarity in addition to a brand image–by–complementarity interaction. The main effects record that perceived quality is higher when the partnered product has high brand image and when the partnered product complements the new product. This clearly does not tell the whole story. We need to discuss the complementarity effect in each of the brand image levels to fully understand how perceived quality is affected.

A different kind of interaction is present in the next example. Here, we must be very cautious in our interpretation of the main effects as one of them can lead to a distorted conclusion.

EXAMPLE 15.10 Right Shade of Green

It is commonly thought that environmentally friendly consumption is driven by a desire to protect the environment and not to enhance one's self-image. Is that the case in all settings? To study this, a two-factor experiment was run, in which subjects looked at an advertisement for a fuel-efficient vehicle and then reported their purchase intentions. One factor was whether the advertisement had a self-benefit or other-benefit appeal. The other factor was the type of accountability (public or private).[8] Here are the mean purchase intentions:

	Accountability		
Appeal	**Public**	**Private**	**Mean**
Self-benefit	3.26	4.58	3.92
Other-benefit	4.84	3.37	4.11
Mean	4.05	3.98	4.01

The means are plotted in Figure 15.4. In the analysis of this experiment, only the interaction is statistically significant. How are we to interpret these results?

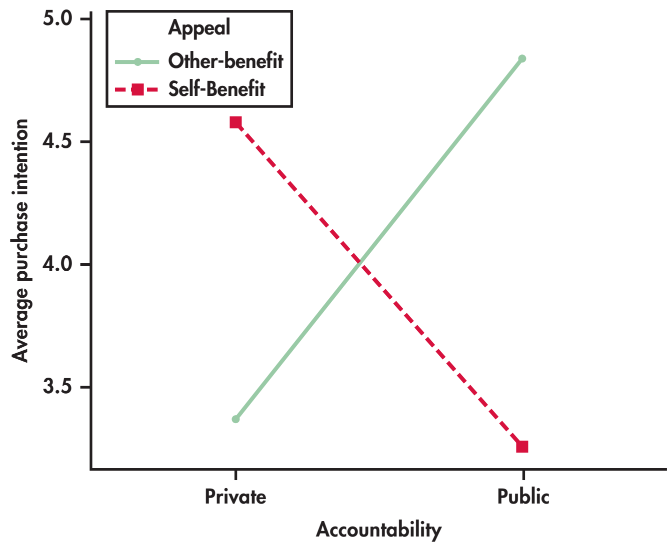


FIGURE 15.4 Plot of average purchase intent by accountability and appeal, Example 15.10.

What catches our eye in the plot is that the lines cross. The purchase intention is much higher for the self-benefit advertisement in the private accountability condition, but the purchase intention was much higher for the other-benefit advertisement in the public accountability condition. This interaction between accountability and benefit is the important result. *The main effects for accountability and appeal are not practically meaningful because of this interaction*. Both factors have effects, but we must know which type of accountability a person is under in order to say which type of appeal results in higher purchase intent.

CAUTION

APPLY YOUR KNOWLEDGE

15.7 Is there an interaction? Each of the following tables gives means for a two-way ANOVA. Make a plot of the means with the levels of Factor A on the x axis. State whether or not there is an interaction, and if there is, describe it.

(a)

	Factor A		
Factor B	**1**	**2**	**3**
1	12	18	24
2	5	8	11

(b)

	Factor A		
Factor B	**1**	**2**	**3**
1	10	15	20
2	30	35	40

(c)

	Factor A		
Factor B	**1**	**2**	**3**
1	10	5	15
2	20	25	15

(d)

	Factor A		
Factor B	**1**	**2**	**3**
1	20	5	20
2	10	25	10

SECTION 15.1 Summary

- **Two-way analysis of variance** (ANOVA) is used to compare population means when populations are classified according to two factors. ANOVA assumes that the populations are Normal and that independent SRSs are drawn from each population. The populations may have different means, but they all have the same standard deviation.
- **Marginal means** are calculated by taking averages of the group means, when organized in a two-way table, either across rows or down columns. These means can be used in an interaction plot to aid in the interpretation of the results.
- ANOVA separates the total variation into parts for the **model** and **error.** Pooling is used to estimate the error, or within-group variance. The model variation is separated into parts for each of the **main effects** and the **interaction.** We describe main effects by the differences between marginal means. An interaction is present if the main effects provide an incomplete description of the group means.

15.2 Inference for Two-Way ANOVA

Because two-way ANOVA breaks the FIT part of the model into three parts, corresponding to the two main effects and the interaction, inference for two-way ANOVA includes an F statistic for each of these effects. As with one-way ANOVA, the calculations are organized in an ANOVA table.

The ANOVA table for two-way ANOVA

The results of a two-way ANOVA are summarized in an ANOVA table based on splitting the total variation SST and the total degrees of freedom DFT among the two main effects and the interaction. *When the sample size is the same for all groups,* both the sums of squares (which measure variation) and the degrees of freedom add:

$$\text{SST} = \text{SSA} + \text{SSB} + \text{SSAB} + \text{SSE}$$

$$\text{DFT} = \text{DFA} + \text{DFB} + \text{DFAB} + \text{DFE}$$

CAUTION

When the n_{ij} are not all equal, there are several ways to decompose SST, and the sums of squares may not add. Whenever possible, design studies with equal sample sizes to avoid these complications. We consider inference only for the equal-sample-size case.

The sums of squares are always calculated in practice by statistical software. From each sum of squares and its degrees of freedom, we find the mean square in the usual way:

$$\text{mean square} = \frac{\text{sum of squares}}{\text{degrees of freedom}}$$

The significance of each of the main effects and the interaction is assessed by an F statistic that compares the variation due to the effect of interest with the within-group variation. Each F statistic is the mean square for the source of interest divided by MSE. Here is the general form of the two-way ANOVA table:

Source	**Degrees of freedom**	**Sum of squares**	**Mean square**	***F***
A	$I - 1$	SSA	SSA/DFA	MSA/MSE
B	$J - 1$	SSB	SSB/DFB	MSB/MSE
AB	$(I - 1)(J - 1)$	SSAB	SSAB/DFAB	MSAB/MSE
Error	$N - IJ$	SSE	SSE/DFE	
Total	$N - 1$	SST	SST/DFT	

There are three null hypotheses in two-way ANOVA, with an F test for each. We can test for significance of the main effect of A, the main effect of B, and the AB interaction. It is generally good practice to examine the test for interaction first because the presence of a strong interaction may influence the interpretation of the main effects. *Be sure to plot the means as an aid to interpreting the results of the significance tests.*

Significance Tests in Two-Way ANOVA

To test the main effect of A, use the F statistic

$$F_{\mathrm{A}} = \frac{\mathrm{MSA}}{\mathrm{MSE}}$$

To test the main effect of B, use the F statistic

$$F_{\mathrm{B}} = \frac{\mathrm{MSB}}{\mathrm{MSE}}$$

To test the interaction of A and B, use the F statistic

$$F_{\mathrm{AB}} = \frac{\mathrm{MSAB}}{\mathrm{MSE}}$$

If the effect being tested is zero, the calculated F statistic has an F distribution with numerator degrees of freedom corresponding to the effect and denominator degrees of freedom equal to DFE. Large values of the F statistic lead to rejection of the null hypothesis. The P-value is the probability that a random variable having the corresponding F distribution is greater than or equal to the calculated value.

APPLY YOUR KNOWLEDGE

15.8 What's wrong? For each of the following, explain what is wrong and why.

(a) You should reject the null hypothesis that there is no interaction in a two-way ANOVA when the test statistic is small.
(b) Sums of squares are equal to mean squares divided by degrees of freedom.
(c) The significance tests for the main effects in a two-way ANOVA have a chi-square distribution when the null hypothesis is true.
(d) The estimate s_p^2 is obtained by pooling the marginal sample variances.

15.9 Customers' preferences for packaging. Exercise 15.2 (page 15-6) describes the setting for a two-way ANOVA design that compares different types of buyers (impulse or not) and the color of sales tags. Give the degrees of freedom for each of the F statistics that are used to test the main effects and the interaction for this problem.

15.10 Comparing employee training programs. Exercise 15.3 (pages 15-6 to 15-7) describes the setting for a two-way ANOVA design that compares employee training programs. Give the degrees of freedom for each of the F statistics that are used to test the main effects and the interaction for this problem.

Carrying out a two-way ANOVA

The following case illustrates how to do a two-way ANOVA. As with the one-way ANOVA, we focus our attention on interpretation of the computer output.

CASE 15.1

DATA FREQD1

Discounts and Expected Prices Does the frequency with which a supermarket product is offered at a discount affect the price that customers expect to pay for the product? Does the percent reduction also affect this expectation? These questions were examined by researchers in a study conducted on students enrolled in an introductory management course at a large midwestern university. For 10 weeks, 160 subjects received information about the products. The treatment conditions corresponded to the number of promotions (one, three, five, and seven) during this 10-week period, and the percent that the product was discounted (10%, 20%, 30%, and 40%). Ten students were randomly assigned to each of the $4 \times 4 = 16$ treatments.[9] For our case study, we examine the data for two levels of promotions (1 and 3) and two levels of discount (40% and 20%). Thus, we have a two-way ANOVA with each of the factors having two levels and 10 observations in each of the four treatment combinations. Here are the data:

Number of promotions	Percent discount	Expected price ($)									
1	40	4.10	4.50	4.47	4.42	4.56	4.69	4.42	4.17	4.31	4.59
1	20	4.94	4.59	4.58	4.48	4.55	4.53	4.59	4.66	4.73	5.24
3	40	4.07	4.13	4.25	4.23	4.57	4.33	4.17	4.47	4.60	4.02
3	20	4.88	4.80	4.46	4.73	3.96	4.42	4.30	4.68	4.45	4.56

As usual we start our statistical analysis with a careful examination of the data.

EXAMPLE 15.11 Plotting the Data

CASE 15.1 With 10 observations per treatment, we can plot the individual observations. To do this, we created an additional variable, "Comb," that has four distinct values corresponding to the particular combination of the number of promotions and the discount. The value "d20-p1" corresponds to 20% discount with one promotion, and the values "d20-p3," "d40-p1," and "d40-p3" have similar interpretations. The data are plotted in Figure 15.5. The lines in the figure connect the four group means.

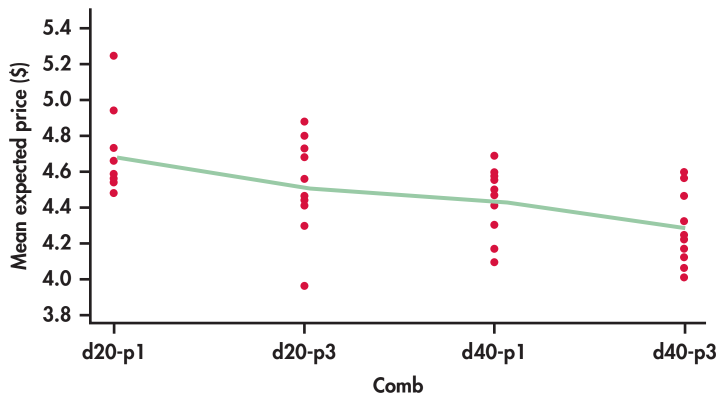


FIGURE 15.5 Plot of the data for the promotions and discount study, Example 15.11.

The spreads of the data within the groups are similar, and there are no outliers or other unusual patterns. That is, the conditions for ANOVA inference appear to be satisfied. The treatment means appear to differ.

DATA FREQD2

CASE 15.2

Expected Prices, Continued Our second case study is a variation on the first. We use data from the experiment described in Case 15.1 but with different treatment combinations. Here are the data for the factor promotions at levels 1 and 5, and the factor discount at levels 30% and 10%:

Number of promotions	Percent discount	Expected price ($)									
1	30	3.57	3.77	3.90	4.49	4.00	4.66	4.48	4.64	4.31	4.43
1	10	5.19	4.88	4.78	4.89	4.69	4.96	5.00	4.93	5.10	4.78
5	30	3.90	3.77	3.86	4.10	4.10	3.81	3.97	3.67	4.05	3.67
5	10	4.31	4.36	4.75	4.62	3.74	4.34	4.52	4.37	4.40	4.52

APPLY YOUR KNOWLEDGE

DATA FREQD2

CASE 15.2 **15.11 Plot the data for Case 15.2.** Make a plot similar to the one given in Figure 15.5 for the levels of the factors given in Case 15.2. Connect the means with lines. Do the conditions for ANOVA inference appear to be met? Describe the pattern of the group means.

After looking at the data for Case 15.1 graphically, we proceed with numerical summaries.

EXAMPLE 15.12 Means and Standard Deviations

FREQD1

CASE 15.1 The software output in Figure 15.6 gives descriptive statistics for the data of Case 15.1. In the row with 1 under the heading "Promo" and 20 under the heading "Discount," the mean of the 10 observations in this treatment combination is given as 4.689. The standard deviation is 0.2331. We would report these as 4.69 and 0.23. The next row gives results for one promotion and a 40% discount. The marginal results for all 20 students assigned to one promotion appear in the following "Total" row. The marginal standard deviation 0.2460 is not useful because it ignores the fact that the 10 observations for the 20% discount and the 10 observations for the 40% discount come from different populations. The overall mean for all 40 observations appears in the last row of the table. The standard deviations for the four groups are quite similar, and we have no reason to suspect a serious violation of the condition that the population standard deviations must all be the same.

FIGURE 15.6 Descriptive statistics from SPSS for the promotions and discount study, Example 15.12.

IBM SPSS Statistics Viewer

Descriptive Statistics

Dependent Variable: Eprice

Promo	Discount	Mean	Std. Deviation	N
1	20	4.689	.2331	10
	40	4.423	.1848	10
	Total	4.556	.2460	20
3	20	4.524	.2707	10
	40	4.284	.2040	10
	Total	4.404	.2638	20
Total	20	4.607	.2600	20
	40	4.353	.2024	20
	Total	4.480	.2633	40

Often, we display the means in a table similar to the following:

	Discount		
Promotions	**20%**	**40%**	**Total**
1	4.69	4.42	4.56
3	4.52	4.28	4.40
Total	4.61	4.35	4.48

In this table, the marginal means give us information about the main effects. When promotions are increased from one to three, the expected price drops from \$4.56 to \$4.40. Furthermore, when the discount is increased from 20% to 40%, the expected price drops from \$4.61 to \$4.35.

Numerical summaries with marginal means enable us to describe the main effects in a two-way ANOVA. For interactions, however, graphs are much better.

EXAMPLE 15.13 Plotting the Group Means

CASE 15.1 The means for the promotions and discount data of Case 15.1 are plotted in Figure 15.7. We have chosen to put the two values of promotion on the x axis. We see that the mean expected price for the 40% discount condition is consistently lower than the mean expected price for the 20% discount condition. Similarly, the means for three promotions are consistently less than the means for one promotion. The two lines are approximately parallel, suggesting that there is little interaction between promotion and discount in this example.

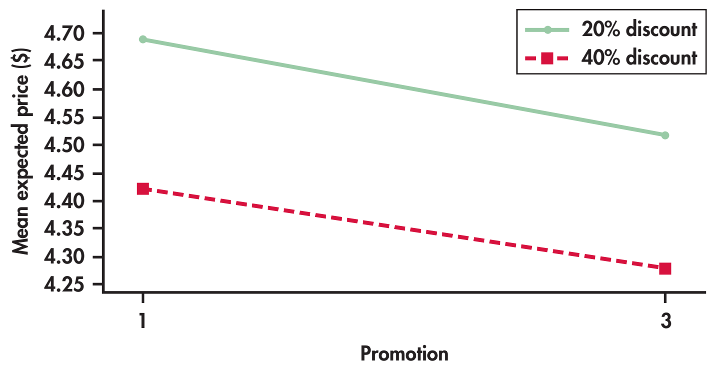


FIGURE 15.7 Plot of the means for the promotions and discount study, Example 15.13.

APPLY YOUR KNOWLEDGE

CASE 15.1 **15.12 Group means in Excel output.** The first part of the Excel output in Figure 15.8 gives the group means and the marginal means for the data in Case 15.1. Find these means and display them in a table. Report them with the digits exactly as given in the output. Does Excel agree with the SPSS output in Figure 15.6?

DATA FREQD2

CASE 15.2 **15.13 Numerical summaries for Case 15.2.** Find the means and standard deviations for each of the promotion-by-discount treatment combinations for the data in Case 15.2. Display the means in a table that also includes the marginal means. Plot the means and describe the main effects and the interaction. Do the standard deviations suggest that it is reasonable to pool the group standard deviations to get MSE?

Having examined the data carefully using numerical and graphical summaries, we are now ready to proceed with the statistical examination of the data using the two-way ANOVA model.

EXAMPLE 15.14 ANOVA Software Output

DATA FREQD1

CASE 15.1 Figure 15.8 gives the two-way ANOVA output from Minitab, Excel, SPSS, and SAS. Look first at the ANOVA table in the Minitab output. The form of the table is very similar to the general form of the two-way ANOVA table given on page 15-14. In place of A and B as the generic factors, the output gives the labels that we specified when we entered the data. We have main effects for “Discount” and “Promo.” The interaction between these two factors is labeled “Promo*Discount,” and the last two rows are "Error" and "Total." The results of the significance tests are in the last two columns, labeled "F-Value" and "P-Value." As expected, the interaction is not statistically significant ($F = 0.03$, df = 1 and 36, $P = 0.856$). On the other hand, the main effects of discount ($F = 12.59$, df = 1 and 36, $P = 0.001$) and promotion ($F = 4.54$, df = 1 and 36, $P = 0.040$) are significant.

Minitab

General Linear Model: Eprice versus Promo, Discount

```
Analysis of Variance

Source           DF   Adj SS    Adj MS  F-Value  P-Value
  Promo           1  0.23104  0.231040     4.54    0.040
  Discount        1  0.64009  0.640090    12.59    0.001
  Promo*Discount  1  0.00169  0.001690     0.03    0.856
Error            36  1.83038  0.050844
Total            39  2.70320


Model Summary
       S    R-sq  R-sq(adj)  R-sq(pred)
0.225486  32.29%     26.65%      16.41%
```

Excel

	A	B	C	D	E	F
1	Anova: Two-Factor With Replication					
2						
3	SUMMARY	Promo 1	Promo 3	Total		
4	*Discount 40*					
5	Count	10	10	20		
6	Sum	44.23	42.84	87.07		
7	Average	4.423	4.284	4.3535		
8	Variance	0.034134	0.041627	0.040971		
9						
10	*Discount 20*					
11	Count	10	10	20		
12	Sum	46.89	45.24	92.13		
13	Average	4.689	4.524	4.6065		
14	Variance	0.054321	0.073293	0.067613		
15						
16	*Total*					
17	Count	20	20			
18	Sum	91.12	88.08			
19	Average	4.556	4.404			
20	Variance	0.06052	0.069594			
21						
22						
23	ANOVA					
24	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*
25	Sample	0.64009	1	0.64009	12.58932	0.0011
26	Columns	0.23104	1	0.23104	4.544106	0.039928
27	Interaction	0.00169	1	0.00169	0.033239	0.856358
28	Within	1.83038	36	0.050844		
29						
30	Total	2.7032	39			

FIGURE 15.8 Two-way ANOVA output from Minitab, Excel, SPSS, and SAS for the promotions and discount example, Exercise 15.12 and Example 15.14.

IBM SPSS Statistics Viewer

Tests of Between-Subjects Effects

Dependent Variable: Eprice

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.873[a]	3	.291	5.722	.003
Intercept	802.816	1	802.816	15789.823	.000
Promo	.231	1	.231	4.544	.040
Discount	.640	1	.640	12.589	.001
Promo * Discount	.002	1	.002	.033	.856
Error	1.830	36	.051		
Total	805.519	40			
Corrected Total	2.703	39			

a. R Squared = .323 (Adjusted R Squared = .266)

SAS

The SAS System

The GLM Procedure

Dependent Variable: Eprice

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.87282000	0.29094000	5.72	0.0026
Error	36	1.83038000	0.05084389		
Corrected Total	39	2.70320000			

R-Square	Coeff Var	Root MSE	Eprice Mean
0.322884	5.033167	0.225486	4.480000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Promo	1	0.23104000	0.23104000	4.54	0.0399
Discount	1	0.64009000	0.64009000	12.59	0.0011
Promo*Discount	1	0.00169000	0.00169000	0.03	0.8564

FIGURE 15.8 *(Continued)*

The statistical significance tests assure us that the differences that we saw in the graphical and numerical summaries can be distinguished from chance variation. We summarize as follows: When promotions are increased from one to three, the expected price drops from \$4.56 to \$4.40 ($F = 4.54$, df $= 1$ and 36, $P = 0.040$). Furthermore, when the discount is increased from 20% to 40%, the expected price decreases from \$4.61 to \$4.35 ($F = 12.59$, df $= 1$ and 36, $P = 0.001$).

Some software does not explicitly give s, the estimate of the parameter σ of our model. To find this, we take the square root of the mean square error, $s = \sqrt{0.0508} = 0.225$.

APPLY YOUR KNOWLEDGE

CASE 15.1 **15.14 Verify the ANOVA calculations.** Use the output in Figure 15.8 to verify that the four mean squares are obtained by dividing the corresponding sums of squares by the degrees of freedom. Similarly, show how each F statistic is obtained by dividing two of the mean squares.

CASE 15.1 **15.15 Compare software outputs.** Examine the outputs of Figure 15.8 from Minitab, Excel, SPSS, and SAS carefully. Write a short evaluation comparing the formats. Indicate which you prefer and why.

DATA FREQD2

CASE 15.2 **15.16 Run the ANOVA for Case 15.2** Analyze the data for Case 15.2 using the two-way ANOVA model and summarize the results.

SECTION 15.2 Summary

- As with one-way ANOVA, the model assumptions should be assessed. Preliminary analysis includes examination of means, standard deviations, scatterplots, histograms, and Normal quantile plots.
- The calculations for two-way ANOVA are organized in an **ANOVA table.**
- F statistics and P-values are used to test hypotheses about the main effects and the interaction.
- Careful inspection of the means is necessary to interpret significant main effects and interactions. Plots are a useful aid.

CHAPTER 15 Review Exercises

For Exercises 15.1 to 15.3, see pages 15-6 to 15-7; for 15.4 to 15.6, see page 15-10; for 15.7, see pages 15-12 to 15-13; for 15.8 to 15.10, see page 15-15; for 15.11, see page 15-16; for 15.12 and 15.13, see page 15-18; and for 15.14 to 15.16, see pages 15-20 to 15-21.

15.17 Describing a two-way ANOVA. A 4×2 ANOVA was run with six observations per cell.
(a) Give the degrees of freedom for the F statistic that is used to test for interaction in this analysis and the entries from Table E that correspond to this distribution.
(b) Sketch a picture of this distribution with the information from Table E included.
(c) The calculated value of the F statistic is 2.68. What is the P-value?
(d) Would you expect a plot of the means to look parallel? Explain your answer.

15.18 How large does the statistic need to be? For each of the following situations, state how large the F statistic needs to be for rejection of the null hypothesis at the 5% level. Sketch each distribution and indicate the region where you would reject.
(a) The main effect for the first factor in a 5×3 ANOVA with three observations per cell.
(b) The interaction in a 2×3 ANOVA with five observations per cell.
(c) The interaction in a 6×6 ANOVA with six observations per cell.

15.19 Describe the design. Each of the following situations is a two-way study design. For each case, identify the response variable and both factors, and state the number of levels for each factor (I and J) and the total number of observations (N).
(a) A video game developer wants to see if haptic feedback (forces and vibrations applied through a joystick) enhances a player's excitement level. He considers two types of games (racing and shooter) and four different levels of haptic feedback (none, force only, vibration only, and force and vibration). He plans to assign five players to each combination. A wrist band will monitor the player's skin temperature.
(b) A restaurant chain is interested in whether calorie-posted menus lead to lower-calorie choices. Two hundred participants were recruited on Amazon Mechanical Turk and asked to order their lunch. Each participant was presented an identical food item menu but the menu varied in terms of type (traditional, calories also posted, and calories posted with food items organized by calories) and price pattern (prices positively or negatively correlated with calories). The total calories of the lunch ordered was recorded.
(c) The strength of concrete depends upon the formula used to prepare it. An experiment compares six different mixtures. Nine specimens of concrete are poured from each mixture. Three of these specimens are subjected to 0 cycles of freezing and thawing, three are subjected to

100 cycles, and three are subjected to 500 cycles. The strength of each specimen is then measured.

15.20 Outline the ANOVA table. For each part of the previous exercise, outline the ANOVA table, giving the sources of variation and the degrees of freedom. (Do not compute the numerical values for the sums of squares and mean squares.)

15.21 What can you conclude? Analysis of data for a 2 × 3 ANOVA with five observations per cell gave the F statistics in the following table:

Effect	F
A	1.87
B	3.94
AB	2.04

What can you conclude from the information given?

15.22 What additional information is needed? A study reported the following results for data analyzed using the methods that we studied in this chapter:

Effect	F	P-value
A	4.75	0.009
B	14.26	0.001
AB	5.14	0.007

(a) What can you conclude from the information given?
(b) What additional information would you need to write a summary of the results for this study?

15.23 Where are your eyes? The objectifying gaze, often referred to as "ogling" or "checking out," can have many adverse consequences. A group of researchers used eye-tracking technology to better understand the nature and causes for this gaze. They asked 29 women and 36 men to look at images of college-aged women. Each woman had the same clothes and neutral expression but varied in body shape (ideal, average, and below average). Prior to looking at the images, each participant was told to focus on either the appearances or personalities of the women. Here is a summary of the amount of time (in milliseconds) the eyes focused on the chest of the women.

	Gender			
	Male		Female	
Focus	$\bar{x}$	SE	$\bar{x}$	SE
Appearance	448.25	35.98	463.22	48.09
Personality	338.78	54.25	276.48	46.06

(a) Plot the means. Do you think there is an interaction? Explain your answer.
(b) Do you think the marginal means would be useful for understanding the results of this study? Explain why or why not.
(c) The researchers broke these results down further using body shape as a third factor. Describe why the inclusion of this factor complicates the analysis. In other words, why is this not a standard three-factor experiment.

15.24 Acceptance of functional foods. Functional foods are foods that are fortified with health-promoting supplements, like calcium-enriched orange juice or vitamin-enriched cereal. Although the number of functional foods is growing in the marketplace, very little is known about how the next generation of consumers views these foods. Because of this, a questionnaire was given to college students from the United States, Canada, and France.[10] This questionnaire measured the students' attitudes and beliefs about general food and functional food. One of the response variables collected concerned cooking enjoyment. This variable was the average of numerous items, each measured on a 10-point scale, where 1 = most negative value and 10 = most positive value. Here are the means:

	Culture		
Gender	Canada	United States	France
Female	7.70	7.36	6.38
Male	6.39	6.43	5.69

(a) Make a plot of the means and describe the patterns that you see.
(b) Does the plot suggest that there is an interaction between culture and gender? If your answer is Yes, describe the interaction.

15.25 Estimating the within-group variance. Refer to the previous exercise. Here are the cell standard deviations and sample sizes for cooking enjoyment:

	Culture					
	Canada		United States		France	
Gender	s	n	s	n	s	n
Female	1.668	238	1.736	178	2.024	82
Male	1.909	125	1.601	101	1.875	87

Find the pooled estimate of the standard deviation for these data. Use the rule for examining standard deviations in ANOVA from Chapter 14 (page 720) to determine if it is reasonable to use a pooled standard deviation for the analysis of these data.

15.26 Comparing the groups. *Challenge* Refer to Exercises 15.24 and 15.25. The researchers presented a table of means with different superscripts indicating pairs of means that differed at the 0.05 significance level, using the Bonferroni method.
(a) What denominator degrees of freedom would be used here?
(b) How many pairwise comparisons are there for this problem?
(c) Perform these comparisons using $t^{**} = 2.94$ and summarize your results.

15.27 More on acceptance of functional foods. Refer to Exercise 15.24. The means for four of the response variables associated with functional foods are as follows.

	General attitude			Product benefits		
	Culture			Culture		
Gender	Canada	United States	France	Canada	United States	France
Female	4.93	4.69	4.10	4.59	4.37	3.91
Male	4.50	4.43	4.02	4.20	4.09	3.87
	Credibility of information			**Purchase intention**		
	Culture			**Culture**		
Gender	**Canada**	**United States**	**France**	**Canada**	**United States**	**France**
Female	4.54	4.50	3.76	4.29	4.39	3.30
Male	4.23	3.99	3.83	4.11	3.86	3.41

For each of the four response variables, give a graphical summary of the means. Use this summary to discuss any interactions that are evident. Write a short report summarizing any differences in culture and gender with respect to the response variables measured.

15.28 Interpreting the results. The goal of the study in the previous exercise was to understand cultural and gender differences in functional food attitudes and behaviors among young adults, the next generation of food consumers. The researchers used a sample of undergraduate students and had each participant fill out the survey during class time. How reasonable is it to generalize these results to the young adult population in these countries? Explain your answer.

15.29 Smart shopping carts. In Example 7.10 (pages 381–382) we compared spending by shoppers on a budget using a shopping cart equipped with or without real-time feedback. In Exercise 7.67 (page 397) we compared spending by shoppers not on a budget using a shopping cart equipped with or without real-time feedback. Let's now perform a two-factor ANOVA of these data. SMART2

(a) Construct a plot of the means and describe the main features of the plot.
(b) Analyze the data using a two-way ANOVA. Report the F statistics, degrees of freedom, and P-values. Because the n_{ij} are not equal, different software may give slightly different F statistics and P-values.
(c) Write a short summary of your findings.
(d) Describe the benefits of this analysis compared to the two t tests previously performed.

15.30 Fuzzy fish? Drugs used to treat anxiety persist in wastewater effluent, resulting in relatively high concentrations of these drugs in our rivers and streams. A regional commercial fishing business wants to better understand the effects of these drugs on fish. They hire researchers who expose fish to various levels of an anxiety drug in a laboratory setting and observe their behavior. In one study, researchers considered the effects of three doses of oxazepam on the behavior of the European perch.[11] Twenty-five perch were each assigned to doses of 0, 1.8, or 910 micrograms per liter of water (μg/l). Each fish was first observed prior to treatment and then observed seven days after treatment. The following table summarizes the results for activity (number of swimming bouts greater than 0.25 cm during 10 minutes).

	Number of movements			
Dose	Pretreatment		Posttreatment	
(μg/1)	$\bar{x}$	s	$\bar{x}$	s
0	3.92	2.38	3.68	1.80
1.8	3.76	1.94	6.32	2.01
910	4.08	1.58	8.68	3.05

(a) The response is the number of movements in 10 minutes, so this variable takes only integer values. Should we be concerned about violating the assumption of Normality? Explain your answer.
(b) Often with this type of count, one considers taking the square root of the count and performing ANOVA on the transformed response. Explain why a transformation might be used here.
(c) Construct an interaction plot and comment on the main effects of dose and time and their interaction.

15.31 The influences of transaction history and a thank you. A service failure is defined as any service-related problem (real or perceived) that transpires during a customer's experience with a firm. In the hotel industry, there is a high human component, so these sorts of failures commonly occur regardless of extensive training and established policies. As a result, hotel firms must learn to effectively react to these failures. One study investigated the relationship between a consumer's transaction history (levels: long and short) and an employee's statement of thanks (levels: yes and no) on a consumer's repurchase intent.[12] Each subject was randomly assigned to one of the four treatment groups and asked to read some service failure/resolution scenarios and respond accordingly. Repurchase intent was measured using a 9-point scale. Here is a summary of the means:

	Thank you	
History	**No**	**Yes**
Short	5.69	6.80
Long	7.53	7.37

(a) Plot the means. Do you think there is an interaction? If yes, describe the interaction in terms of the two factors.
(b) Find the marginal means. Are they useful for understanding the results of this study? Explain your answer.

15.32 Transaction history and a thank you. Refer to the previous exercise. The numbers of subjects in each cell were not equal, so the researchers used linear regression to analyze the data. This was done by creating an indicator variable for each factor and the interaction. Here is a partial ANOVA table. Complete it and state your conclusions regarding the main effects and interaction described in the previous exercise.

Source	DF	SS	MS	*F*	*P*-value
Transaction history		61.445			
Thank you		21.810			
Interaction		15.404			
Error	160	759.904			

15.33 The effect of humor. In advertising, humor is often used to overcome sales resistance and stimulate customer purchase behavior. One experiment looked at the use of humor as an approach to offset the negative feelings often associated with website encounters.[13] The setting of their experiment was an online travel agency, and they used a three-factor design, each factor with two levels. They were humor (used, not used), process (favorable, unfavorable), and outcome (favorable, unfavorable). For the humor condition, cartoons and jokes about skiing were used on the site. For the no-humor condition, standard pictures of ski sites were used. Two hundred and forty-one business students from a large Dutch university participated in the experiment. Each was randomly assigned to one of the eight treatment conditions. The students were asked to book a skiing holiday and then rate their perceived enjoyment and satisfaction with the process. All responses were measured on a 7-point scale. The following is a summary of the results for satisfaction.

Treatment	n	$\bar{x}$	s
No humor–favorable process–unfavorable outcome	27	3.04	0.79
No humor–favorable process–favorable outcome	29	5.36	0.47
No humor–unfavorable process–unfavorable outcome	26	2.84	0.59
No humor–unfavorable process–favorable outcome	31	3.08	0.59
Humor–favorable process–unfavorable outcome	32	5.06	0.59
Humor–favorable process–favorable outcome	30	5.55	0.65
Humor–unfavorable process–unfavorable outcome	36	1.95	0.52
Humor–unfavorable process–favorable outcome	30	3.27	0.71

(a) Plot the means of the four treatments that did not use humor. Do you think there is an interaction? If yes, describe the interaction in terms of the process and outcome factors.
(b) Plot the means of the four treatments that used humor. Do you think there is an interaction? If yes, describe the interaction in terms of the process and outcome factors.
(c) The three-factor interaction can be assessed by looking at the two interaction plots created in parts (a) and (b). If the relationship between process and outcome is different across the two humor conditions, there is evidence of an interaction among all three factors. Do you think there is a three-factor interaction? Explain your answer.

15.34 Pooling the standard deviations. Refer to the previous exercise. Find the pooled estimate of the standard deviation for these data. What are its degrees of freedom? Using the rule from Chapter 14 (page 720), is it reasonable to use a pooled standard deviation for the analysis? Explain your answer.

15.35 Describing the effects. Refer to Exercise 15.33. The *P*-values for all main and two-factor interactions are significant at the 0.05 level. Using the table, find the marginal means and use them to describe these effects.

15.36 Trust of individuals and groups. Trust is an essential element in any exchange of goods or services. The following trust game is often used to study trust experimentally:

> A *sender* starts with \$X and can transfer any amount $x \leq X$ to a *responder*. The responder then gets \$$3x$ and can transfer any amount $y \leq 3x$ back to the sender. The game ends with final amounts $X - x + y$ and $3x - y$ for the sender and responder, respectively.

The value x is taken as a measure of the sender's trust, and the value $y/3x$ indicates the responder's trustworthiness. A study used this game to study the dynamics between individuals and groups of three.[14] The following table summarizes the average amount x sent by senders.

Sender	Responder	n	$\bar{x}$	s
Individual	Individual	32	65.5	36.4
Individual	Group	25	76.3	31.2
Group	Individual	25	54.0	41.6
Group	Group	27	43.7	42.4

(a) Find the pooled estimate of the standard deviation for this study and its degrees of freedom.
(b) Is it reasonable to use a pooled standard deviation for the analysis? Explain your answer.
(c) Compute the marginal means.
(d) Plot the means. Do you think there is an interaction? If yes, then describe it.
(e) The F statistics for sender, responder, and interaction are 9.05, 0.001, and 2.08 respectively. Compute the P-values and state your conclusions.

15.37 Trustworthiness of individuals and groups. Refer to the previous exercise. Here is a summary of the percent returned to the sender.

Sender	Responder	n	$\bar{x}$	s
Individual	Individual	32	25.1	19.5
Individual	Group	25	25.1	17.5
Group	Individual	25	23.2	22.1
Group	Group	27	16.7	18.7

Repeat parts (a)–(d) of Exercise 15.36 using these results. In this case, none of the effects is statistically significant.

15.38 Switching between work tasks. In many businesses, employees are expected to handle multiple projects simultaneously. To work efficiently, this requires the employee to reduce or eliminate cognitions about one task and fully focus on another. Is this easy to do? A study looked at how having to transition between sequential tasks affects worker ability to dedicate their full attention to a given task.[15] Each participant was asked to work on two tasks (each five minutes in length). The first task was a word puzzle, and the second involved reading four résumés and selecting the best candidate. The 78 participants were randomly assigned within a 2 (first task: finished/unfinished) $\times$ 2 (time pressure: high/low) design and were evaluated on how many characteristics of the résumés they recalled from Task 2. The following table summarizes the means.

	Time Pressure	
Task 1	**Low**	**High**
Finished	46.97	64.57
Unfinished	44.26	45.10

Plot the means and describe the primary features of the data in terms of main effects and interaction.

15.39 Repeating an advertising message. Does repetition of an advertising message increase its effectiveness? One theory suggests that there are two phases in the process. In the first phase, called "wearin," negative or unfamiliar views are transformed into positive views. In the second phase, called "wearout," the effectiveness of the ad is decreased because of boredom or other causes. One study designed to investigate this theory examined two factors. The first was familiarity of the ad, with two levels, familiar and unfamiliar; the second was repetition, with three levels, 1, 2, and 3.[16] One of the response variables collected was attitude toward the ad. This variable was the average of four items, each measured on a 7-point scale, anchored by bad–good, low quality–high quality, unappealing–appealing, and unpleasant–pleasant. Here are the means for attitude:

	Repetition		
Familiarity	**1**	**2**	**3**
Familiar	4.56	4.73	5.24
Unfamiliar	4.14	5.26	4.41

(a) Make a plot of the means and describe the patterns that you see.
(b) Does the plot suggest that there is an interaction between familiarity and repetition? If yes, describe the interaction.

15.40 Other response variables. Refer to the previous exercise. In settings such as this, researchers collect data for several response variables. For this study, they also constructed variables that were called attitude toward the brand, total thoughts, support arguments, and counterarguments. Here are the means:

	Attitude to brand			Total		
	Repetition			Repetition		
Familiarity	**1**	**2**	**3**	**1**	**2**	**3**
Familiar	4.67	4.65	5.06	1.33	1.93	2.55
Unfamiliar	3.94	4.79	4.26	1.52	3.06	3.17
	Support			**Counter**		
	Repetition			**Repetition**		
Familiarity	**1**	**2**	**3**	**1**	**2**	**3**
Familiar	0.63	0.67	0.98	0.54	0.70	0.49
Unfamiliar	0.76	1.40	0.64	0.52	0.75	1.14

For each of the four response variables, give a graphical summary of the means. Use this summary to discuss any interactions that are evident. Write a short report summarizing the effect of repetition on the response variables measured, using the data in this exercise and the previous one.

15.41 Pooling the standard deviations. Refer to the previous exercise. Here are the standard deviations for attitude toward brand:

	Repetition		
Familiarity	**1**	**2**	**3**
Familiar	1.16	1.46	1.16
Unfamiliar	1.39	1.22	1.42

Assuming that the cell sizes are equal, find the pooled estimate of the standard deviation for these data. Use the rule for examining standard deviations in ANOVA from Chapter 14 (page 720) to determine if it is reasonable to use a pooled standard deviation for the analysis of these data.

15.42 More pooling. Refer to Exercise 15.40. Here are the standard deviations for total thoughts:

	Repetition		
Familiarity	**1**	**2**	**3**
Familiar	1.63	1.42	1.52
Unfamiliar	1.64	2.16	1.59

Assuming that the cell sizes are equal, find the pooled estimate of the standard deviation for these data. Use the rule for examining standard deviations in ANOVA from Chapter 14 (page 720) to determine if it is reasonable to use a pooled standard deviation for the analysis of these data.

15.43 Interpret the results. Refer to Exercises 15.39 and 15.40. The subjects were 94 adult staff members at a West Coast university. They were evenly split into familiar and unfamiliar groups. Each subject watched a half-hour local news show from a different state that included ads at all the repetition levels. The selected ads were judged to be "good" by some experts and had been shown in regions other than where the study was conducted. The real names of the products were replaced by either familiar or unfamiliar brand names by a professional video editor. The ads were pretested, and no one in the pretest sample suggested that the ads were not real. Discuss each of these facts in terms of how you would interpret the results of this study.

15.44 Interpret the results. Refer to Exercises 15.39 and 15.40. The ratings for this study were each measured on a 7-point scale, anchored by bad–good, low quality–high quality, unappealing–appealing, and unpleasant–pleasant. The results presented were averaged over three ads for different products: a bank, women's clothing, and a health care plan. Write a short report summarizing the Normality assumption for two-way ANOVA and the extent to which it is reasonable for the analysis of these data.

15.45 A manufacturing problem. One step in the manufacture of large engines requires that holes of very precise dimensions be drilled. The tools that do the drilling are regularly examined and are adjusted to ensure that the holes meet the required specifications. Part of the examination involves measurement of the diameter of the drilling tool. A team studying the variation in the sizes of the drilled holes selected this measurement procedure as a possible cause of variation in the drilled holes. They decided to use a designed experiment as one part of this examination. Some of the data are given in Table 15.1. The diameters in millimeters (mm) of five tools were measured by the same operator at three times (8:00 A.M., 11:00 A.M., and 3:00 P.M.). Three measurements were taken on each tool at each time. The person taking the measurements could not tell which tool was being measured, and the measurements were taken in random order.[17] DRILL

TABLE 15.1 Tool diameter data

Tool	Time	Diameter		
1	1	25.030	25.030	25.032
1	2	25.028	25.028	25.028
1	3	25.026	25.026	25.026
2	1	25.016	25.018	25.016
2	2	25.022	25.020	25.018
2	3	25.016	25.016	25.016
3	1	25.005	25.008	25.006
3	2	25.012	25.012	25.014
3	3	25.010	25.010	25.008
4	1	25.012	25.012	25.012
4	2	25.018	25.020	25.020
4	3	25.010	25.014	25.018
5	1	24.996	24.998	24.998
5	2	25.006	25.006	25.006
5	3	25.000	25.002	24.999

(a) Make a table of means and standard deviations for each of the 5 × 3 combinations of the two factors.
(b) Plot the means and describe how the means vary with tool and time. Note that we expect the tools to have slightly different diameters. These will be adjusted as needed. It is the process of measuring the diameters that is important.
(c) Use a two-way ANOVA to analyze these data. Report the test statistics, degrees of freedom, and *P*-values for the significance tests.
(d) Write a short report summarizing your results.

15.46 Convert from millimeters to inches. Refer to the previous exercise. Multiply each measurement by 0.04 to convert from millimeters to inches. Redo the plots and rerun the ANOVA using the transformed measurements. Summarize what parts of the analysis have changed and what parts have remained the same.

CASE 15.1 **15.47 Discounts and expected prices.** Case 15.1 (page 15-15) describes a study designed to determine how the frequency that a supermarket product is promoted at a discount and the size of the discount affect the price that customers expect to pay for the product. In the exercises that followed, we examined the data for two levels of each factor. Table 15.2 gives the complete set of data. FREQD
(a) Summarize the means and standard deviations in a table and plot the means. Summarize the main features of the plot.
(b) Analyze the data with a two-way ANOVA. Report the results of this analysis.
(c) Using your plot and the ANOVA results, prepare a short report explaining how the expected price depends on the number of promotions and the percent of the discount.

TABLE 15.2 Expected price data

Number of promotions	Percent discount	Expected price ($)									
1	40	4.10	4.50	4.47	4.42	4.56	4.69	4.42	4.17	4.31	4.59
1	30	3.57	3.77	3.90	4.49	4.00	4.66	4.48	4.64	4.31	4.43
1	20	4.94	4.59	4.58	4.48	4.55	4.53	4.59	4.66	4.73	5.24
1	10	5.19	4.88	4.78	4.89	4.69	4.96	5.00	4.93	5.10	4.78
3	40	4.07	4.13	4.25	4.23	4.57	4.33	4.17	4.47	4.60	4.02
3	30	4.20	3.94	4.20	3.88	4.35	3.99	4.01	4.22	3.70	4.48
3	20	4.88	4.80	4.46	4.73	3.96	4.42	4.30	4.68	4.45	4.56
3	10	4.90	5.15	4.68	4.98	4.66	4.46	4.70	4.37	4.69	4.97
5	40	3.89	4.18	3.82	4.09	3.94	4.41	4.14	4.15	4.06	3.90
5	30	3.90	3.77	3.86	4.10	4.10	3.81	3.97	3.67	4.05	3.67
5	20	4.11	4.35	4.17	4.11	4.02	4.41	4.48	3.76	4.66	4.44
5	10	4.31	4.36	4.75	4.62	3.74	4.34	4.52	4.37	4.40	4.52
7	40	3.56	3.91	4.05	3.91	4.11	3.61	3.72	3.69	3.79	3.45
7	30	3.45	4.06	3.35	3.67	3.74	3.80	3.90	4.08	3.52	4.03
7	20	3.89	4.45	3.80	4.15	4.41	3.75	3.98	4.07	4.21	4.23
7	10	4.04	4.22	4.39	3.89	4.26	4.41	4.39	4.52	3.87	4.70

CASE 15.1 **15.48 Rerun the data as a one-way ANOVA.** Refer to the previous exercise. Rerun the analysis as a one-way ANOVA with $4 \times 4 = 16$ treatments. Summarize the results of this analysis. Use the Bonferroni multiple-comparisons procedure to describe combinations of number of promotions and percent discounts that are similar or different.

15.49 Consumer-generated ads. More and more companies involve consumers in the process of developing advertisements. Is it beneficial to let consumers know this? In one study, 125 undergraduate students were randomly assigned to a 2×2 design.[18] Each student watched one of two consumer-generated Doritos ads. Half of the students were told that the ad was consumer-generated and the other half were not. After viewing the ad, the students provided their reactions to the ad and the advertised brand with higher scores reflecting a more favorable opinion. Here is part of the ANOVA table for their reactions to the ad:

Source	Degrees of freedom	Sum of squares	Mean square	F
A (Ad)		3.054		
B (Informed)		7.813		
AB		1.876		
Error		146.807		
Total				

(a) Fill in the missing values in the ANOVA table.
(b) What is the value of the F statistic to test the null hypothesis that there is no interaction? What is its distribution when the null hypothesis is true? Using Table E, find an approximate P-value for this test.
(c) Answer the questions in part (b) for the main effect of advertisement and the main effect of being informed that the ad was consumer-generated.
(d) What is s_p^2, the within-group variance? What is s_p?
(e) The mean score when disclosing that the ads were consumer-generated was 4.52. The mean score when this was not disclosed was 5.23. Using these results and your answers to parts (a)–(d), summarize the results.

15.50 Use of animated agents in a multimedia environment. Multimedia learning environments are designed to enhance learning by providing a more hands-on and exploratory investigation of a topic. Often, animated agents (human-like characters) are used with the hope of enhancing social interaction with the software and thus improving learning. One group of researchers decided to investigate whether the presence of an agent and the type of verbal feedback provided were actually helpful.[19] To do this, they recruited 135 college students and randomly divided them among four groups: agent/simple feedback, agent/elaborate feedback, no agent/simple feedback, and no agent/elaborate feedback. The topic of the software was thermodynamics. The change in score on a 20-question test taken before and after using the software was the response. AGENT
(a) Make a table giving the sample size, mean, and standard deviation for each group.
(b) Use these means to construct an interaction plot. Describe the main effects for agent presence and for feedback type as well as their interaction.
(c) Analyze the change in score using analysis of variance. Report the test statistics, degrees of freedom, and P-values.
(d) Use the residuals to check model assumptions. Are there any concerns? Explain your answer.
(e) Based on parts (b) and (c), write a short paragraph summarizing your findings.

NOTES AND DATA SOURCES

1. This example is based on Laura Herrewijn and Karolien Poels, "Recall and recognition of in-game advertising: The role of game control," *Frontiers in Psychology* 4 (2014), pp. 1–14.

2. This example is based on Laura Smarandescu and Terence A. Shimp, "Drink Coca-Cola, eat popcorn, and choose Powerade: Testing the limits of subliminal persuasion," doi:10.1007/s11002-014-9294-1 (2014).

3. This example is based on Shibin Sheng and Yue Pan, "Bundling as a new product introduction strategy: The role of brand image and bundle features," *Journal of Retailing and Consumer Services* 16 (2009), pp. 367–376.

4. This example is based on Iana A. Castro et al., "The influence of disorganized shelf displays and limited product quantity on consumer purchase," *Journal of Marketing* 77 (2013), pp. 118–133.

5. We present the two-way ANOVA model and analysis for the general case in which the sample sizes may be unequal. If the sample sizes vary a great deal, serious complications can arise. There is no longer a single standard ANOVA analysis. Most computer packages offer several options for the computation of the ANOVA table when group counts are unequal. When the counts are approximately equal, all methods give essentially the same results.

6. U.S. Census Bureau, *American Community Survey,* 2012 American Community Survey 1-Year Estimates.

7. See Note 3.

8. Example 15.10 is based on a study described in Todd Green and John Peloza, "Finding the right shade of green: The effect of advertising appeal type on environmentally friendly consumption," *Journal of Advertising* 43, no. 2 (2014), pp. 128–141.

9. Based on M. U. Kalwani and C. K. Yim, "Consumer price and promotion expectations: An experimental study," *Journal of Marketing Research* 29 (1992), pp. 90–100.

10. Jane Kolodinsky et al., "Sex and cultural differences in the acceptance of functional foods: A comparison of American, Canadian, and French college students," *Journal of American College Health* 57 (2008), pp. 143–149.

11. Tomas Brodin et al., "Dilute concentrations of a psychiatric drug alter behavior of fish from natural populations," *Science* 339 (2013), pp. 814–815.

12. Vincent P. Magnini and Kiran Karande, "The influences of transaction history and thank you statements in service recovery," *International Journal of Hospitality Management* 28 (2009), pp. 540–546.

13. Willemijn M. van Dolen, Ko de Ruyter, and Sandra Streukens. "The effect of humor in electronic service encounters," *Journal of Economic Psychology* 29 (2008), pp. 160–179.

14. Tamar Kugler et al., "Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy," *Journal of Economic Psychology* 28 (2007), pp. 646–657.

15. S. Leroy, "Why is it so hard to do my work? The challenge of attention residue when switching between work tasks," *Organizational Behavior and Human Decision Processes* 109 (2009), pp. 168–181.

16. Margaret C. Campbell and Kevin Lane Keller, "Brand familiarity and advertising repetition effects," *Journal of Consumer Research* 30 (2003), pp. 292–304.

17. Based on a problem from Renée A. Jones and Regina P. Becker, Department of Statistics, Purdue University.

18. Debora V. Thompson and Prashant Malaviya, "Consumer-generated ads: Does awareness of advertising co-creation help or hurt persuasion?" *Journal of Marketing* 77 (2013), pp. 33–47.

19. Lijia Lin et al., "Animated agents and learning: Does the type of verbal feedback they provide matter?" *Computers and Education,* 2013, doi: 10.1016/j.compedu.2013.04.017.

ANSWERS TO ODD-NUMBERED EXERCISES

15.1 **(a)** A two-way ANOVA is used when there are two factors (explanatory variables) not outcomes. **(b)** Each level of Factor A appears with each level of Factor B. **(c)** This is true for the RESIDUAL part of the model, not the FIT. **(d)** The sample sizes in each cell can be different.

15.3 Factors: Training program has 3 levels (the 3 different programs); Method, or how the program is administered, has 2 levels (one 4-hour session, or two 2-hour sessions). $N = 90$.

15.5 In the Northeast, the per capita income for white individuals is substantially higher than for Asian individuals; in the other 3 regions, the per capita income for Asian individuals is higher than for white individuals.

15.7 **(a)** There is a slight interaction. Factor B level 1 is higher than Factor B level 2, but this difference increases as the level of Factor A increases, hence the interaction. **(b)** There is no interaction, Factor B level 2 is consistently higher than level 1 regardless of Factor A. Additionally, the means increase as Factor A level increases. **(c)** There is an interaction. For Factor A level 1, Factor B level 2 is higher than level 1; this difference increases for Factor A level 2. Finally there is no difference between level 2 and 1 for Factor B when Factor A is level 3. **(d)** There is an interaction, for Factor A levels 1 and 3, Factor B level 1 is higher than level 2; however, this reverses for Factor A level 2, where Factor B level 1 is now much higher than level 2.

15.9 For the tag color main effect, DF = 3, 130. For the impulse main effect, DF = 1, 130. For the interaction effect, DF = 3, 130.

15.11 The d30-p1 group seems more spread out than the other 3 groups; also, the d10-p5 group has a low outlier. These concerns likely violate the conditions for ANOVA. The mean prices do seem different for the different promotion and discount combinations.

15.13 We cannot pool the standard deviations because the largest s is more than twice the smallest s, $0.38561 > 2(0.15202) = 0.30404$. There doesn't appear to be an interaction. The expected price is higher for a 10% discount than for a 30% discount; the expected price is also higher with only 1 promotion compared with 5 promotions.

Promotions	Discount	N	Mean	Standard Deviation
.	.	40	4.357	0.45211
.	10	20	4.6565	0.34379
.	30	20	4.0575	0.33546
1	.	20	4.5725	0.45661
5	.	20	4.1415	0.33661
1	10	10	4.92	0.15202
1	30	10	4.225	0.38561
5	10	10	4.393	0.26854
5	30	10	3.89	0.16289

15.17 **(a)** DF = 3, 40. **(c)** $0.05 < P\text{-value} < 0.10$.

P-value	*F*
0.100	2.23
0.050	2.84
0.025	3.46
0.010	4.31
0.001	6.59

(d) The plot would be roughly parallel because the interaction is only marginally significant.

15.19 **(a)** Response: skin temperature. Factors: haptic feedback (4 levels) and type of game (2 levels). $N = 40$. **(b)** Response: total calories. Factors: menu type (3 levels) and price pattern (2 levels). $N = 200$. **(c)** Response: strength. Factors: mixture (6 levels) and cycles of freezing (3 levels). $N = 54$.

15.21 Main effect A is not significant, $P\text{-value} > 0.100$ (DF = 1, 24). Main effect B is significant, $0.025 < P\text{-value} < 0.050$ (DF = 2, 24). The interaction is not significant, $P\text{-value} > 0.100$ (DF = 2, 24).

15.23 **(a)** There appears to be an interaction effect; the lines are not parallel. **(b)** There appears to be a significant Focus main effect, so the marginal means for Focus would be useful in explaining this difference. **(c)** If each participant looked at a picture of each body type, then his or her responses likely would be related to each other, which violates the independence assumption.

15.25 $s_p^2 = 3.1493$, so $s_p = 1.7746$. It is reasonable to pool the standard deviations because the largest s is less than twice the smallest s, $2.024 < 2(1.601) = 3.202$.

15.27 For general attitude, there may be a slight interaction effect. Differences in gender are largest for Canada, then the United States, and smallest for France, with females higher than males in all cases. Additionally, Canada attitudes are generally the highest, followed by the United States, with France having the worst attitudes among the three.

For product benefits, there may be a slight interaction effect. Differences in gender are largest for Canada, then the United States, and very little for France, with females higher than males in all cases. Additionally, Canada benefit scores are generally the highest, followed by the United States, with France having the lowest scores among the three.

For credibility of information, there seems to be an interaction effect. Differences in gender are largest for the United States, then Canada, with females higher in both, but in France, males had a slightly higher score for credibility than females. Additionally, Canada credibility scores are generally the highest, followed by the United States, with France having the lowest scores among the three.

For purchase intention, there seems to be a small interaction effect. Differences in gender are largest for the United States with females higher, and for Canada and France, there are small differences between genders. Canada and the United States seem to have much higher purchase intent scores than France.

15.29 **(a)** For those with real-time feedback, the average total cost for those not informed was only slightly larger than the average total cost of those informed, while for those without feedback, the not informed average was much larger than for those who were informed. Additionally, we can see an interaction effect. For those not informed, the lack of real-time feedback increased their spending, while for those who were informed, the lack of real-time feedback decreased their spending. **(b)** $F = 18.18$, DF = 3, 190, P-value < 0.0001. **(c)** The interaction term is significant ($F = 16.18$, P-value < 0.0001) as is the Informed term ($F = 37.06$, P-value < 0.0001). The Smartcart term is not significant ($F = 1.16$, P-value $= 0.2828$). Effects are interpreted as in part (a). **(d)** The two-factor ANOVA is better than two t tests because not only do we see the difference with and without feedback, we can also see the interaction effect where the lack of feedback increases spending for those not informed and decreases spending for those informed.

15.31 **(a)** There appears to be an interaction between history and statement of thanks. For consumers with a short history, the employee's statement of thanks drastically increases the repurchase intent. For consumers with a long history, there is not much difference in repurchase intent whether or not a thanks was given. **(b)** Short: 6.245. Long: 7.45. No: 6.61. Yes: 7.085. The marginal mean for history is somewhat meaningful, as generally the long history consumers are more likely to repurchase. But the marginal mean for thank you is misleading, suggesting that generally "no thanks" is lower than "yes thanks," but that is only true for the short history group.

15.33 **(a)** There appears to be an interaction effect. For the unfavorable outcome, there is no difference in satisfaction between the favorable and unfavorable process; both scored equally low. For the favorable outcome, those with a favorable process scored substantially higher than those with an unfavorable process. **(b)** There doesn't appear to be an interaction effect. Overall, the favorable outcome means were only slightly higher than the unfavorable outcomes means. But both favorable process means were generally quite high, much higher than both unfavorable process means, regardless of whether or not the outcome was favorable or not. **(c)** Yes, there appears to be a three-factor interaction effect. Particularly for those with an unfavorable outcome and a favorable process, the humor group scored much higher than those with no humor. Also, for those with unfavorable outcome and unfavorable process, the humor group scored worse than those with no humor.

15.35 A favorable process seems to have the greatest impact on satisfaction, followed by a favorable outcome. Humor doesn't seem to add much to satisfaction. For humor and process, humor added to satisfaction when the process was favorable but hurt satisfaction when the process was unfavorable. For humor and outcome, humor helped regardless of whether the outcome was favorable or not. For process and outcome, a favorable process had a bigger impact than a favorable outcome on raising the satisfaction and of course the combination of the two was the highest overall satisfaction.

15.37 **(a)** $s_p = 19.51$. **(b)** Yes, the largest s is less than twice the smallest s, $22.1 < 2(17.5) = 35$. **(c)** Sender Individual: 25.1, Sender Group: 19.825, Responder Individual: 24.27, Responder Group: 20.74. **(d)** There appears to be an interaction effect. When the sender was an individual, there was no difference in the amount returned, but when the sender was a group, they returned more money to an individual than they did to a group.

15.39 **(a)** For those familiar with the ad, as repetition increased, attitude also generally increased. For those unfamiliar with the ad, we see the "wearin" and "wearout" effects. At level 1 repetition, the attitude was the lowest of any combination, but once we moved to repetition level 2, the "wearin" effect drastically improved the attitude. Lastly, after moving to repetition level 3, the "wearout" effect drastically decreased the attitude, although it still scored higher than the original attitude at repetition level 1. **(b)** Yes, there appears to be an interaction effect, as described in part (a).

15.41 Assuming equal cell sizes, $s_p = 1.30765$. It is reasonable to pool the standard deviations because the largest s is less than twice the smallest s, $1.46 < 2(1.16) = 2.32$.

15.43 Because the experiment was performed at a West Coast university, it does limit somewhat how much generalizing we can do as there could be aspects of West Coast students that make them more or less willing to accept familiar or unfamiliar brands. Putting the ads into a news program was extremely good because it doesn't place the focus on the ads themselves, which could eliminate some biases.

The facts about the ads being judged to be "good" and edited by professionals is good and makes our results more likely to apply in other professional settings, especially as they were also pretested and thought to be real. Overall, the experiment was set up quite well, though it might be good to test the ads in other regions.

15.45 **(a)**

Level of Tool	Level of Time	N	Diameter Mean	Diameter Std Dev
1	1	3	25.030667	0.0011547
1	2	3	25.028	0
1	3	3	25.026	0
2	1	3	25.016667	0.0011547
2	2	3	25.02	0.002
2	3	3	25.016	0
3	1	3	25.006333	0.0015275
3	2	3	25.012667	0.0011547
3	3	3	25.009333	0.0011547
4	1	3	25.012	0
4	2	3	25.019333	0.0011547
4	3	3	25.014	0.004
5	1	3	24.997333	0.0011547
5	2	3	25.006	0
5	3	3	25.000333	0.0015275

(b) There are differences in diameter among the five different tools as expected. There are also differences in diameter in the different shifts, although not as large as the tool differences. There also appears to be some interaction between tool and time. Particularly, tools 3, 4, and 5 are fairly consistent across time differences, with time 2 having the largest diameters and time 1 the smallest diameters. For tool 2, however, time 1 diameters get larger than time 3s, and for tool 1, time 1 diameters get the largest, bigger than both time 2 and time 3 diameters. **(c)** The Tool effect is extremely significant, $F = 412.94$, DF = 4, 30, P-value < 0.0001. The Time effect is also very significant, $F = 43.60$, DF = 2, 30, P-value < 0.0001. The interaction effect is significant, $F = 7.65$, DF = 8, 30, P-value < 0.0001. **(d)** Because all terms are significant we would interpret effects as stated in part (b).

15.47 **(a)** As promotions increase, expected price goes down. Expected price is also different for different discounts. The expected price from highest to lowest is 10%, 20%, 40%, and 30%. For some reason, the 30% discount gives lower expected prices than the 40% discount. There doesn't appear to be an interaction effect; only with 7 promotions at the 40% level do we see something unusual happen, where it is much lower than where we might expect it to be.

Level of Promotions	Level of Discount	N	Price Mean	Price Std Dev
1	10	10	4.92	0.1520234
1	20	10	4.689	0.2330689
1	30	10	4.225	0.3856092
1	40	10	4.423	0.1847551
3	10	10	4.756	0.2429083
3	20	10	4.524	0.2707274
3	30	10	4.097	0.2346179
3	40	10	4.284	0.2040261
5	10	10	4.393	0.2685372
5	20	10	4.251	0.2648459
5	30	10	3.89	0.1628906
5	40	10	4.058	0.1759924
7	10	10	4.269	0.2699156
7	20	10	4.094	0.2407488
7	30	10	3.76	0.2617887
7	40	10	3.78	0.2143725

(b) Promotions is very significant, $F = 47.73$, P-value < 0.0001. Discount is also very significant, $F = 47.42$, P-value < 0.0001. The interaction effect is not significant, $F = 0.44$, P-value $= 0.9121$. **(c)** As promotions increase, expected price goes down. Expected price is also different for different discounts. The expected price from highest to lowest is 10%, 20%, 40%, and 30%. For some reason the 30% discount gives lower expected prices than the 40% discount.

15.49 **(a)**

Source	DF	SS	MS	F
A (Ad)	1	3.054	3.054	2.52
B (Informed)	1	7.813	7.813	6.44
AB	1	1.876	1.876	1.55
Error	121	146.807	1.2133	
Total	124	159.55		

(b) $F = 1.55$, the distribution is $F(1, 121)$, P-value > 0.100. **(c)** Ad: $F = 2.52$, the distribution is $F(1, 121)$, P-value > 0.100. Informed: $F = 6.44$, the distribution is $F(1, 121)$, $0.01 < P\text{-value} < 0.025$. **(d)** $s_p^2 = 1.2133$, $s_p = 1.10149$. **(e)** There is no significant interaction, as well as no significant Ad effect. Informed was significant, those who were informed that the ad was consumer-generated had a higher average opinion, 5.23, than those that weren't informed, 4.52.