Unit 31: One-Way ANOVA

SUMMARY OF VIDEO

A vase filled with coins takes center stage as the video begins. Students will be taking part in an experiment organized by psychology professor John Kelly in which they will guess the amount of money in the vase. As a subterfuge for the real purpose of the experiment, students are told that they are taking part in a study to test the theory of the "Wisdom of the Crowd," which is that the average of all of the guesses will probably be more accurate than most of the individual guesses. However, the real purpose of the study is to see whether holding heavier or lighter clipboards while estimating the amount of money in the jar will have an impact on students' guesses. The idea being tested is that physical experience can influence our thinking in ways we are unaware of – this phenomenon is called embodied cognition.

The sheet on which students will record their monetary guesses is clipped onto a clipboard. For the actual experiment, clipboards, each holding varying amounts of paper, weigh either one pound, two pounds or three pounds. Students are randomly assigned to clipboards and are unaware of any difference in the clipboards. After the data are collected, guesses are entered into a computer program and grouped according to the weights of the clipboards. The mean guess for each group is computed and the output is shown in Table 31.1.

Clipboard Weight	Mean	N	Standard Deviation
1	$106.56	75	$100.62
2	$129.79	75	$204.95
3	$143.29	75	$213.13
Total	$126.55	225	$180.16

Table 31.1. Average guesses by clipboard weight.

Figure 31.1. Mean guess versus clipboard weight.

Looking at the means, the results appear very promising. As clipboard weight goes up, so does the mean of the guesses, and that pattern appears fairly linear (See Figure 31.1.). To test whether or not the apparent differences in means could be due simply to chance, John turns to a technique called a one-way analysis of variance, or ANOVA. The null
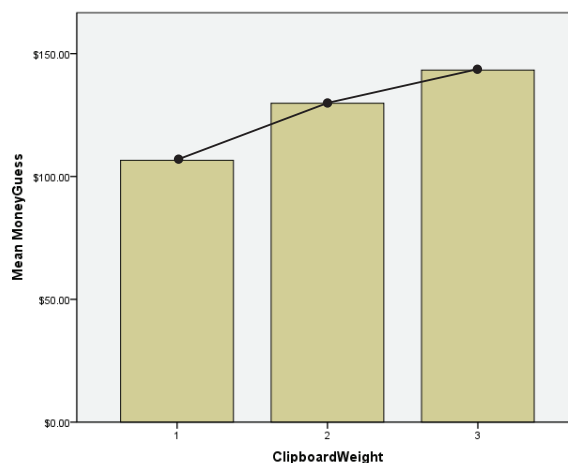
hypothesis for the analysis of variance will be that there is no difference in population means for the three weights of clipboards: $H_0 : \mu_1 = \mu_2 = \mu_3$. He hopes to find sufficient evidence to reject the null hypothesis so that he can conclude that there is a significant difference among the population means. John runs an ANOVA using SPSS statistical software to compute a statistic called F, which is the ratio of two measures of variation:

$$F = \frac{\text{variation among sample means}}{\text{variation within individual observations in the same sample}}$$

In this case, $F = 0.796$ with a p-value of 0.45. That means there is a 45% chance of getting an F value at least this extreme when there is no difference between the population means. So, the data from this experiment do not provide sufficient evidence to reject the null hypothesis.

One of the underlying assumptions of ANOVA is that the data in each group are normally distributed. However, the boxplots in Figure 31.2 indicate that the data are skewed and include some rather extreme outliers. John's students tried some statistical manipulations on the data to make them more normal and reran the ANOVA. However, the conclusion remained the same.
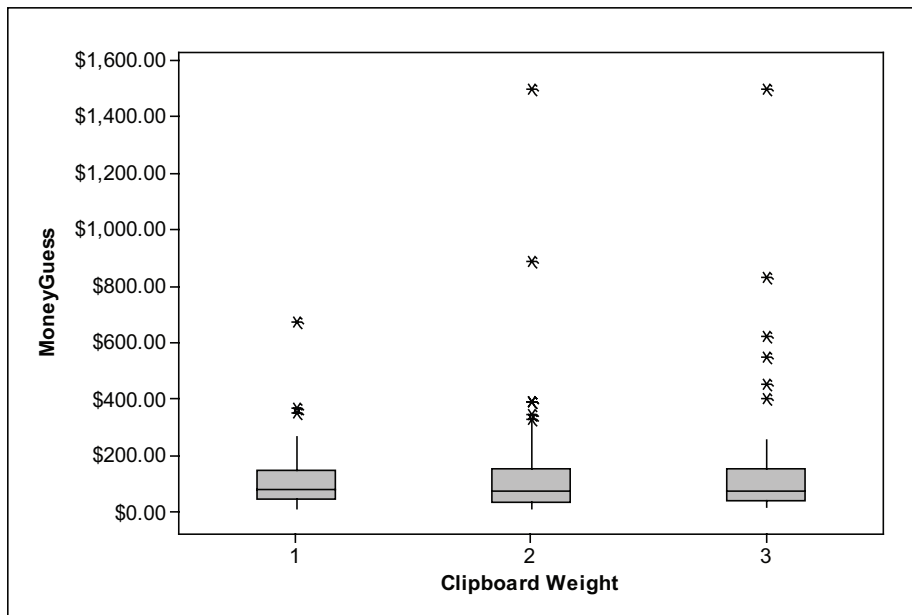

Figure 31.2. Boxplots of guesses grouped by clipboard weight.

But what if we used the data displayed in Figure 31.3 instead? The sample means are the same, around $107, $130, and $143, but this time the data are less spread out about those means.
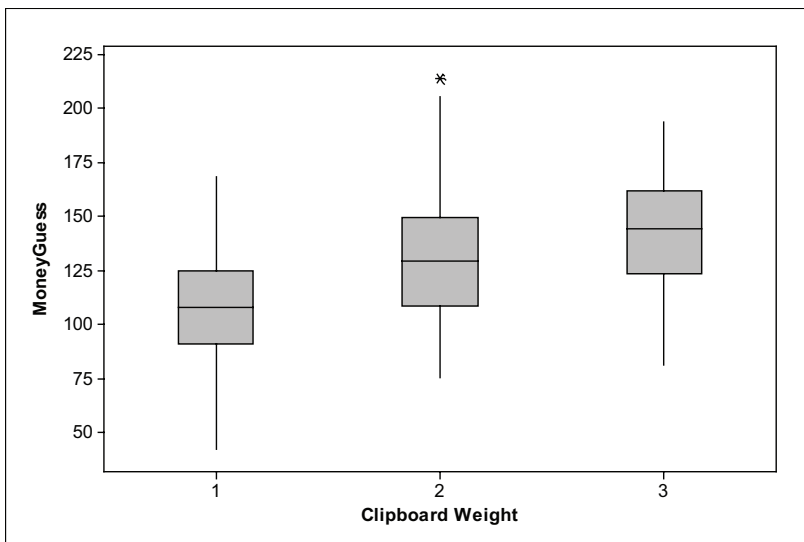
Figure 31.3. Hypothetical guess data.

In this case, after running ANOVA, the result is $F = 33.316$ with a *p*-value that is essentially zero. Our conclusion is to reject the null hypothesis and conclude that the population means are significantly different.

In John's experiment, the harsh reality of a rigorous statistical analysis has shot down the idea that holding something heavy causes people, unconsciously, to make larger estimates, at least in this particular study. But if the real experiment didn't work, what about the cover story – the theory of the Wisdom of the Crowd? The actual amount in the vase is $237.52. Figure 31.4 shows a histogram of all the guesses. The mean of the estimates is $129.22 – more than $100 off, but still better than about three-quarters of the individual guesses. So, the crowd was wiser than the people in it.
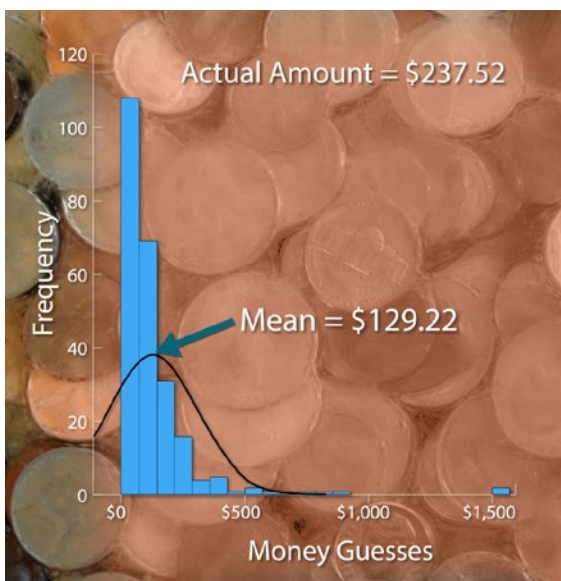

Figure 31.4. Histogram of guess data.

STUDENT LEARNING OBJECTIVES

A. Be able to identify when analysis of variance (ANOVA) should be used and what the null and alternative (research) hypotheses are.

B. Be able to identify the factor(s) and response variable from a description of an experiment.

C. Understand the basic logic of an ANOVA. Be able to describe between-sample variability (measured by mean square for groups (MSG)) and within-sample variability (measured by mean square error (MSE)).

D. Know how to compute the F statistic and determine its degrees of freedom given the following summary statistics: sample sizes, sample means and sample standard deviations. Be able to use technology to compute the p-value for F.

E. Be able to use technology to produce an ANOVA table.

F. Recognize that statistically significant differences among population means depend on the size of the differences among the sample means, the amount of variation within the samples, and the sample sizes.

G. Recognize when underlying assumptions for ANOVA are reasonably met so that it is appropriate to run an ANOVA.

H. Be able to create appropriate graphic displays to support conclusions drawn from ANOVA output.

CONTENT OVERVIEW

In Unit 27, Comparing Two Means, we compared two population means, the mean total energy expenditure for Hadza and Western women, and used a two-sample t-test to test whether or not the means were equal. But what if you wanted to compare three population means? In that case, you could use a statistical procedure called **Analysis of Variance** or **ANOVA**, which was developed by Ronald Fisher in 1918.

For example, suppose a statistics class wanted to test whether or not the amount of caffeine consumed affected memory. The variable caffeine is called a **factor** and students wanted to study how three **levels** of that factor affected the **response variable**, memory. Twelve students were recruited to take part in the study. The participants were divided into three groups of four and randomly assigned to one of the following drinks:

A. Coca-Cola Classic (34 mg caffeine)

B. McDonald's coffee (100 mg caffeine)

C. Jolt Energy (160 mg caffeine).

After drinking the caffeinated beverage, the participants were given a memory test (words remembered from a list). The results are given in Table 31.2.

Group A (34 mg)	Group B (100 mg)	Group C (160 mg)
7	11	14
8	14	12
10	14	10
12	12	16
7	10	13

Table 31.2. Number of words recalled in memory test.

For an ANOVA, the null hypothesis is that the population means among the groups are the same. In this case, $H_0 : \mu_A = \mu_B = \mu_C$, where μ_A is the population mean number of words recalled after people drink Coca Cola and similarly for μ_B and μ_C. The alternative or research hypothesis is that there is some inequality among the three means. Notice that there is a lot of variation in the number of words remembered by the participants. We break that variation into two components:

(1) variation in the number of words recalled among the three groups also called **between-groups variation**

(2) variation in number of words among participants within each group also called **within-groups variation**.

To measure each of these components, we'll compute two different variances, the **mean square for groups (MSG)** and the **mean square error (MSE)**. The basic idea in gathering evidence to reject the null hypothesis is to show that the between-groups variation is substantially larger than the within-groups variation and we do that by forming the ratio, which we call F:

$$F = \frac{\text{between-groups variation}}{\text{within-groups variation}} = \frac{MSG}{MSE}$$

In the caffeine example, we have three groups. More generally, suppose there were k different groups (each assigned to consume varying amounts of caffeine) with sample sizes n_1, n_2, … n_k. Then the null hypothesis is $H_0: \mu_1 = \mu_2 = \ldots = \mu_k$ and the alternative hypothesis is that at least two of the population means differ. The formulas for computing the between-groups variation and within-groups variation are given below:

$$MSG = \frac{n_1(\overline{x}_1 - \overline{x})^2 + n_2(\overline{x}_2 - \overline{x})^2 + \ldots + n_k(\overline{x}_k - \overline{x})^2}{k-1}$$

where $\overline{x}$ is the mean of all the observations and $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k$ are the sample means for each group.

$$MSE = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots, + (n_k - 1)s_k^2}{N-k}$$

where $N = n_1 + n_2 + \ldots + n_k$ and $s_1, s_2, \ldots, s_k$ are the sample standard deviations for each group.

When H_0 is true, then $F = MSG/MSE$ has the F distribution $k-1$ and $N-k$ degrees of freedom. We use the F distribution to calculate the p-value for the F-test statistic.

We return to our three-group caffeine experiment to see how this works. To begin, we calculate the sample means and standard deviations (See Table 31.3.).

Group	Sample Mean	Sample Standard Deviation
A	8.8	2.168
B	12.2	1.789
C	13.0	2.236

Table 31.3. Group means and standard deviations.

Before calculating the mean square for groups we also need the grand mean of all the data values: $\bar{x} \approx 11.3$. Now we are ready to calculate *MSG*, *MSE*, and *F*:

$$MSG = \frac{5(8.8 - 11.33)^2 + 5(12.2 - 11.33)^2 + 5(13.0 - 11.33)^2}{3 - 1} = \frac{49.73}{2} = 24.87$$

$$MSE = \frac{(5 - 1)(2.168)^2 + (5 - 1)(1.789)^2 + (5 - 1)(2.236)^2}{15 - 3} = \frac{51.602}{12} \approx 4.30$$

$$F = \frac{24.87}{3.40} \approx 5.78$$

All that is left is to find the *p*-value. If the null hypothesis is true, then the *F*-statistic has the *F* distribution with 2 and 12 degrees of freedom. We use software to see how likely it would be to get an *F* value at least as extreme as 5.78. Figure 31.5 shows the result giving a *p*-value of around 0.017. Since $p < 0.05$, we conclude that the amount of caffeine consumed affected the mean memory score.

Distribution Plot
F, df1=2, df2=12
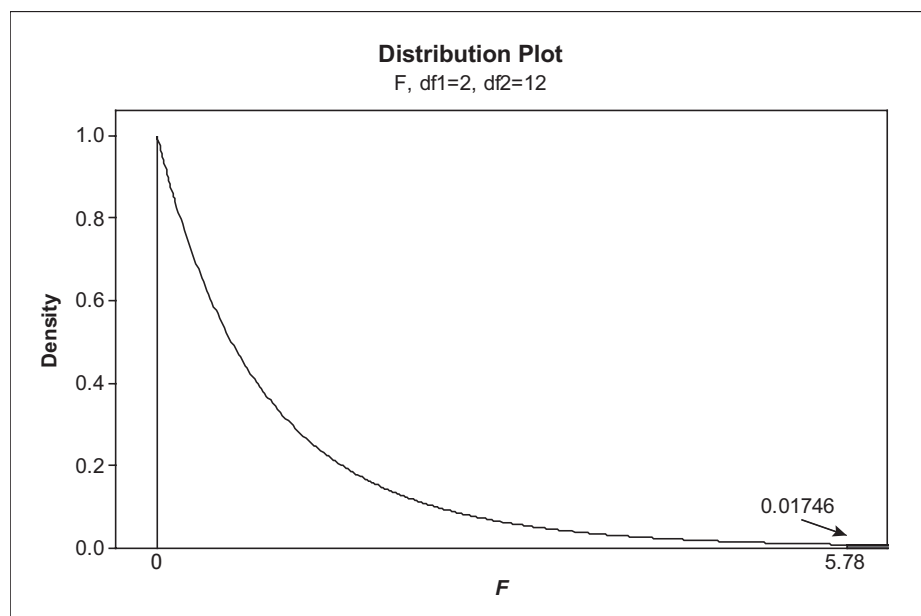
Figure 31.5. Finding the p-value from the F-distribution.

It takes a lot of work to compute *F* and find the *p*-value. Here's where technology can help. Statistical software such as Minitab, spreadsheet software such as Excel, and even graphing calculators can calculate ANOVA tables. Table 31.4 shows output from Minitab. Now, match the calculations above with the values in Table 31.4. Check out where you can find the values for *MSG*, *MSE*, *F*, the degrees of freedom for *F*, and the *p*-value directly from the output of ANOVA. That will be a time saver!

One-way ANOVA: Group A (34 mg), Group B (100 mg), Group C (160 mg)

```
Source  DF       SS      MS      F       P
Factor   2    49.73   24.87   5.78   0.017
Error   12    51.60    4.30
Total   14   101.33
```

Table 31.4. ANOVA output from Minitab.

It is important to understand that ANOVA does not tell you which population means differ, only that at least two of the means differ. We would have to use other tests to help us decide which of the three population means are significantly different from each other. However, we can also get a clue by plotting the data. Figure 31.6 shows comparative dotplots for the number of words for each group. The sample means are marked with triangles. Notice that the biggest difference in sample means is between groups A (34 mg caffeine) and C (160 mg of caffeine). The sample means for groups B and C are quite close together. So, it looks as if consuming Coca Cola doesn't give the memory boost you could expect from consuming coffee or Jolt Energy.
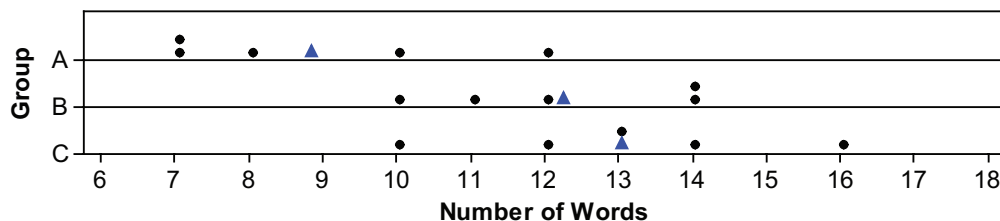
Figure 31.6. Comparative dotplots.

There is one last detail before jumping into running an ANOVA – there are some underlying assumptions that need to be checked in order for the results of the analysis to be valid. What we should have done <u>first</u> with our caffeine experiment, we will do last. Here are the three things to check.

1. Each group's data need to be an independent random sample from that population. In the case of an experiment, the subjects need to be randomly assigned to the levels of the factor.

   Check: The subjects in the caffeine-memory experiment were divided into groups. Groups were then randomly assigned to the level of caffeine.

2. Next, each population has a Normal distribution. The results from ANOVA will be approximately correct as long as the sample group data are roughly normal. Problems can arise if the data are highly skewed or there are extreme outliers.

    Check: The normal quantile plots of Words Recalled for each group are shown in Figure 31.7. Based on these plots, it seems reasonable to assume these data are from a Normal distribution.
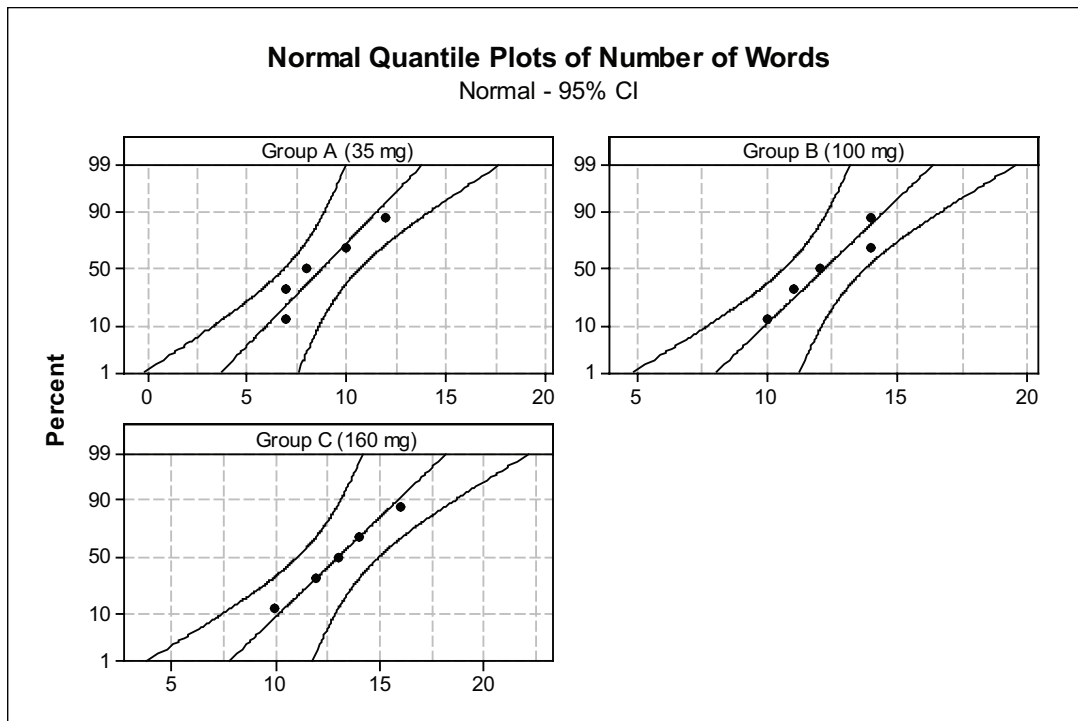

Figure 31.7. Checking the normality assumption.

3. All populations have the same standard deviation. The results from ANOVA will be approximately correct as long as the ratio of the largest standard deviation to the smallest standard deviation is less than 2.

    Check: The ratio of the largest to the smallest standard deviation is 2.236/1.789 or around 1.25, which is less than 2.

KEY TERMS

A **factor** is a variable that can be used to differentiate population groups. The **levels** of a factor are the possible values or settings that a factor can assume. The variable of interest is the **response variable**, which may be related to one or more factors.

An **analysis of variance** or **ANOVA** is a method of inference used to test whether or not three or more population means are equal. In a **one-way ANOVA** there is one factor that is thought to be related to the response variable.

An analysis of variances tests the equality of means by comparing two types of variation, **between-groups variation** and **within-groups variation**. Between-groups variation deals with the spread of the group sample means about the **grand mean**, the mean of all the observations. It is measured by the **mean square for groups**, *MSG*. Within-groups variation deals with the spread of individual data values within a group about the group mean. It is measured by the **mean square error**, *MSE*.

The *F*-**statistic** is the ratio *MSG*/MSE.

THE VIDEO

Take out a piece of paper and be ready to write down answers to these questions as you watch the video.

1. What is the theory called the Wisdom of the Crowd?

2. What was different about the clipboards that students were holding?

3. State the null hypothesis for the one-way ANOVA.

4. What is the name of the test statistic that results from ANOVA?

5. Was the professor able to conclude from the F-statistic that the population means differed depending on the weight of the clipboard? Explain.

6. Did the crowd prove to be wiser than the individual students?

UNIT ACTIVITY:
CONTROLLING WAFER THICKNESS

You will use the Wafer Thickness tool to collect data for this activity. There are three control settings that affect wafer thickness during the manufacture of polished wafers used in the production of microchips.

1. Leave Controls 2 and 3 set at level 2. Your first task will be to perform an experiment to collect data and determine whether settings for Control 1 affect the mean thickness of polished wafers.

a. Open the Wafer Thickness tool. Set Control 1 to level 1, and Controls 2 and 3 to level 2 (the middle setting). In Real Time mode, collect data from 10 polished wafers. Store the data in a statistical package or Excel spreadsheet or in a calculator list. Make a sketch of the histogram produced by the interactive tool.

b. Set Control 1 to level 2. Leave Controls 2 and 3 set at level 2. Repeat (a). Sketch the second histogram using the same scales as was used on the first. Store the data in your spreadsheet or a calculator list.

c. Set Control 1 to 3. Leave Controls 2 and 3 set at level 2. Repeat (a). Sketch your third histogram, again using the same scales as were used on the first histogram. Store the data in your spreadsheet or a calculator list.

d. Calculate the means and standard deviations for each of your three samples. Based on the sample means and on your histograms, do you think that there is sufficient evidence that changing the level of Control 1 changes the mean thickness of the polished wafers produced? Or might these sample-mean differences be due simply to chance variation? Explain your thoughts.

e. Use technology to run an ANOVA. State the null hypothesis being tested, the value of F, the p-value, and your conclusion.

2. Your next task will be to perform an experiment to collect data and determine whether settings for Control 2 affect the mean thickness of polished wafers.

a. Leave Controls 1 and 3 set at level 2. Adapt the process used in question 1(a – c) to collect the data on Control 2.

---

b. Compute the standard deviations for the three samples. Is the underlying assumption of equal standard deviations reasonably satisfied? Explain.

c. Provided you answered yes to (b), use technology to run an ANOVA. State the null hypothesis being tested, the value of F, the p-value, and your conclusion. (If you answered no to (b), skip this part.)

3. Your final task will be to perform an experiment to collect data and determine whether settings for Control 3 affect the mean thickness of polished wafers.

a. Leave Controls 1 and 2 set at level 2. Adapt the process used in question 1(a – c) to collect the data on Control 3.

b. Compute the standard deviations for the three samples. Is the underlying assumption of equal standard deviations reasonably satisfied? Explain.

c. Provided you answered yes to (b), use technology to run an ANOVA. State the null hypothesis being tested, the value of F, the p-value, and your conclusion. (If you answered no to (b), skip this part.)

EXERCISES

1. A professor predicts that students will learn better if they study to white noise (similar to a fan) compared to music or no sound. She randomly divides 27 students into three groups and sends them to three different rooms. In the first room, students hear white noise, in the second, music from a local radio station, and in the third, the door is closed to help block out normal sound from the hall. Each group is given 30 minutes to study a section of their text after which they take a 10-question multiple-choice exam. Table 31.5 contains the results.

White Noise	Music	No Sound
8	5	4
5	7	7
5	3	6
7	5	2
8	5	4
7	8	3
8	4	5
3	5	5
10	7	4

Table 31.5. Test results.

a. Calculate the mean test score for each group. Calculate the standard deviation of the test scores for each group.

b. Make comparative dotplots for the test results of the three groups. Do you think that the dotplots give sufficient evidence that there is a difference in population mean test results depending on the type of noise? Explain.

c. Run an ANOVA. State the hypotheses you are testing. Show the calculations for the F-statistic. What are the degrees of freedom associated with this F-statistic?

d. What is the p-value of the test statistic? What is your conclusion?

2. Not all hotdogs have the same calories. Table 31.6 contains calorie data on a random sample of Beef, Poultry, and Veggie dogs. (One extreme outlier for Veggie dogs was omitted from the data.) Does the mean calorie count differ depending on the type of hotdog? You first encountered this topic in Unit 5, Boxplots.

Calories		
Beef	Poultry	Veggie
110	60	40
110	60	45
130	60	45
130	70	45
140	70	50
150	70	50
160	80	55
160	90	57
170	90	60
170	100	60
175	100	70
180	100	80
180	110	80
180	110	81
190	110	90
190	120	95
190	120	100
200	130	100
210	140	110
230	150	

Table 31.6. Calorie content of hotdogs.

First-Year Cumulative Grade Point Average

High Rating	Medium Rating	Low Rating
3.37	2.92	3.19
3.28	2.11	1.57
1.73	3.92	0.99
3.64	2.83	3.58
3.04	3.26	3.87
2.80	3.18	1.43
3.83	2.28	2.50
3.22	3.13	3.45
3.55	2.90	2.63
2.28	3.41	2.59
2.51	2.64	3.50
1.74	3.71	1.95
2.88	2.52	2.21
2.86	2.24	2.50
4.00	0.98	3.35
2.67	3.65	2.88
3.75	2.87	1.27
2.30	2.21	3.10
3.24	3.02	3.29
2.43	3.66	1.06

Table 31.7. First-year college GPA by high school rating.

a. Verify that the standard deviations allow the use of ANOVA to compare population means.

b. Use technology to run an ANOVA. State the value of the *F*-statistic, the degrees of freedom for *F*, the *p*-value of the test, and your conclusion.

c. Make boxplots that compare the calorie data for each type of hot dog. Add a dot to each boxplot to mark the sample means. Do your plots help confirm your conclusion in (b)?

3. Many states rate their high schools using factors such as students' performance, teachers' educational backgrounds, and socioeconomic conditions. High school ratings for one state have been boiled down into three categories: high, medium, and low. The question for one of the state universities is whether or not college grade performance differs depending on high school rating. Table 31.7 contains random samples of students from each high school rating level and their first-year cumulative college grade point averages (GPA).

a. Calculate the sample means for the GPAs in each group. Based on the sample means alone, does high school rating appear to have an impact on mean college GPA? Explain.

b. Check to see that underlying assumptions for ANOVA are reasonably satisfied.

c. Run an ANOVA to test whether there is a significant difference among the population mean GPAs for the three groups. Report the value of the *F*-statistic, the *p*-value, and your conclusion.

4. Researchers in a nursing school of a large university conducted a study to determine if differences exist in levels of active and collaborative learning (ACL) among nursing students, other health professional students, and students majoring in education. A random sample of 1,000 students from each of these three majors was selected from students who completed the National Survey of Student Engagement (NSSE).

a. The sample mean ACL scores for nursing, other health professional students, and education majors were 46.44, 45.58, and 48.59, respectively. Do these sample means provide sufficient evidence to conclude that there was some difference in population mean ACL scores among these three majors? Explain.

b. A one-way analysis of variance was run to determine if there was a difference among the three groups on mean ACL scores. Assuming that all students answered the NSSE questions related to ACL, what were the degrees of freedom of the *F*-test?

c. The results from the ANOVA gave *F* = 8.382. Determine the *p*-value. What can you conclude?

REVIEW QUESTIONS

1. Random samples of three types of candy were given to children. The three types of candy were chocolate bars (A), hard candy (B), and chewy candy (C). A group of 15 children rated samples of each type of candy on a scale from 1 (lowest) to 10 (highest). Two hypothetical data sets are given in Tables 31.8 and 31.9.

Ratings for A	Ratings for B	Ratings for C
8	4	6
10	5	5
7	7	7
8	8	5
6	7	6
7	8	5
4	6	6
7	5	5
6	5	4
8	6	2
6	6	3
5	7	4
6	3	5
7	8	5
8	5	6

Table 31.8. Data set #1.

Ratings for A	Ratings for B	Ratings for C
8	4	6
10	5	5
7	6	8
8	9	2
3	6	7
7	8	3
4	6	7
9	5	5
6	4	4
8	7	2
5	6	3
4	9	4
6	3	3
8	9	7
10	3	8

Table 31.9. Data set #2.

a. Find the sample means of each candy type based on the ratings in Table 31.8. Then do the same for the ratings in Table 31.9. Based on these results, can you tell if there is a significant difference in population mean ratings among the different types of candies? Explain.

b. Make comparative boxplots for the data in Table 31.8. Then do the same for the data in Table 31.9. For both sets of plots, mark the mean with a dot on each boxplot. For which data set is it more likely that the results from a one-way ANOVA will be significant? Explain.

c. Run an ANOVA based on Data Set #1. Report the value of the F-statistic, the p-value, and your conclusion. Then do the same for Data Set #2. Explain why you should not be surprised by the results.

2. The data in Table 31.10 were part of a study to investigate online questionnaire design. The researcher was interested in the effect that type of answer entry and type of question-to-question navigation would have on the time it would take to complete online surveys. Twenty-

seven volunteers participated in this study. Each completed two questionnaires, one dealing with credit and the other focusing on vacations. Each questionnaire had 14 questions, and participants could select only one answer.

There were three display types for answers:

(1) radio button

(2) drop down list

(3) list box

There were three navigational methods:

(1) questions were on a single page

(2) click the Next/Prev button

(3) press Tab

Display Type	Navigation Type	Time (sec)	Display Type	Navigation Type	Time (sec)
1	1	97	1	3	117
3	3	83	2	2	74
1	1	102	1	3	66
3	3	85	3	1	62
1	1	92	1	2	93
3	3	71	3	1	62
1	2	105	1	2	64
3	3	92	3	1	48
1	2	67	1	2	57
3	3	71	3	1	96
1	2	54	2	3	68
3	3	66	3	1	90
1	3	63	2	3	71
2	1	61	3	1	74
1	3	101	2	3	74
2	1	117	3	2	78
1	3	124	2	3	92
2	1	97	3	2	71
2	1	126	2	3	80
3	2	83	3	2	49
2	1	107	2	3	67
3	2	88	1	1	101
2	1	88	2	2	111
3	2	62	1	1	103
2	2	55	2	2	80
1	3	73	1	1	103
2	2	126	2	2	111

Table 31.10. Time to complete on-line questionnaires.

a. Enter the data in Table 31.10 into a statistical software, Excel, or graphing calculator spreadsheet. The display types and navigation types have been coded as numbers to facilitate data entry. Once the data are entered, you can replace the coded values with their categorical values. (For example, in Display Type replace 1 with radio button, 2 with drop down list, and 3 with list box).

Display Type and Time

b. Make comparative boxplots of the times for each level of Display Type. Mark the location of the means on your boxplot. Do you see anything unusual in the data that might make it not appropriate to use ANOVA? If so, follow up with normal quantile plots to check the assumption of normality.

c. Run an ANOVA using Display Type as the factor. State the null hypothesis you are testing. Report the value of the F-statistic, the p-value, and your conclusion.

d. Make comparative boxplots of the times for each level of Navigation Type. Mark the location of the means on your boxplot. Do you see anything unusual in these data that might make it not appropriate to use ANOVA? If so, follow up with normal quantile plots to check the assumption of normality.

e. Run an ANOVA using Navigation Type as the factor. Report the value of the F-statistic, the degrees of freedom of the F-statistic, the p-value, and your conclusion.

f. Based on this study, what recommendations would you make to online questionnaire designers?

3. A group researching wage discrepancies among the four regions of the U.S. focused on full-time, hourly-wage workers between the ages of 20 and 40. Researchers randomly selected 200 workers meeting the age criteria from the northeast, midwest, south and west and recorded their hourly pay rates. The mean hourly rate for the combined regions was $15.467. A summary of the data are given in Table 31.11. The researchers ran an ANOVA on these data.

Region	Sample size	Sample Mean	Standard Deviation
Northeast	200	16.560	9.164
Midwest	200	15.154	6.381
South	200	13.931	6.933
West	200	16.223	9.289

Table 31.11. Summary of hourly rate data.

a. Is the homogeneous standard deviations assumption for ANOVA reasonably satisfied? Explain.

b. State the researchers' null hypothesis.

c. Calculate the value of the F-statistic and give its degrees of freedom. Show calculations.

d. Determine the p-value.

e. Based on the evidence in Table 31.11 and your answers to (a – d), what conclusions can the researchers make?

4. A study focusing on women's wages was investigating whether there was a significant difference in salaries in four occupations commonly (but not exclusively) held by women – cashier, customer service representative, receptionist, and secretary/administrative assistant. Weekly wages from 50 women working in each occupation are recorded in Table 31.12.

Cashier		Customer Service		Receptionist		Secretary/Adm. Asst.	
385	320	440	473	673	807	640	850
427	450	540	400	232	2038	600	580
480	333	222	673	523	529	420	840
520	300	1000	84	380	520	554	522
540	1200	625	769	715	600	577	528
690	240	540	1428	596	383	712	382
680	288	600	289	402	400	447	800
690	315	738	1885	787	945	380	440
364	548	344	429	238	400	1384	92
113	369	800	555	538	620	705	877
360	340	481	480	360	969	600	923
360	350	340	788	500	450	225	1154
2885	387	720	1480	430	577	440	481
720	345	885	596	540	440	1058	812
360	350	680	600	528	650	673	560
297	540	760	920	568	672	769	918
300	548	700	624	561	400	692	600
340	400	560	340	603	500	826	227
360	439	800	769	510	420	900	865
511	320	431	400	400	345	1900	640
508	315	500	390	500	919	769	692
321	302	1162	615	520	400	746	654
400	665	400	481	1280	480	1384	543
290	220	400	640	560	188	360	320
729	331	1270	430	386	680	415	323

Table 31.12. Weekly wages of women in four occupations.
Data from 2012 March Supplement, Current Population Survey.

a. Calculate the means and standard deviations for the weekly wages in each occupation category.

b. Is the homogeneous standard deviations assumption for ANOVA reasonably satisfied? Explain.

c. Run an ANOVA. Record the ANOVA table and highlight the value of F, and the p-value.

d. Based on your answers to (a – c), what are your conclusions?