Chapter 9
The Means Procedure

WHILE THE <u>Crosstabs</u> procedure allows you to identify the frequency of certain types of categorical data (Chapter 8), the **Means** command allows you to explore certain characteristics of continuous variables within those categories. By way of comparison, a crosstabulation of **ethnic** by **gender** would indicate that there were 13 White females, 22 White males, 8 Hispanic females, 6 Hispanic males, and so forth. The **Means** command allows you to view certain characteristics of continuous variables (such as total points, GPAs, percents) by groups. Thus if you computed **total** (number of points) for **ethnic** by **gender**, you would find that there were 13 White females who scored an average (mean) of 113.12 points, 22 White males who scored a mean of 115.34 points, 8 Hispanic females who scored a mean of 116.79, 6 Hispanic males with a mean of 113.45, and so forth. This information is, of course, presented in tabular format for ease of reading and interpretation. The utility of the **Means** command for data such as our sample file is several-fold. For a class with more than one section, we might like to see mean scores for each section, or to compare the scores of males with females, or the performance of upper- with lower-division students.

The **Means** command is one of SPSS's simplest procedures. For the selected groups it will list the mean for each group, the standard deviation, and the number of subjects for each category. There is an additional **Options** subcommand with which you may conduct a one-way analysis of variance (ANOVA) based on the means and standard deviations you have just produced. We will include that option in this chapter but will save a detailed explanation of analysis of variance for the **One-Way ANOVA** and **General Linear Models** chapters (Chapters 12–14).

We again make use of the **grades.sav** file ($N = 105$) and the variables **total**, **percent**, **gpa**, **section**, **lowup**, and **gender** to illustrate the **Means** procedure.

9.1 Step by Step
9.1.1 Describing Subpopulation Differences

To access the initial SPSS screen from the Windows display, perform the following sequence of steps:

In Screen	Do This	Step 1
2.1	✳ ⊞ → ✳ All Programs ▷ → ✳ 📁 IBM SPSS Statistics → ✳ ② IBM SPSS Statistics	Front1

Mac users: To access the initial SPSS screen, successively click the following icons:

In Screen	Do This	Step 1a
2.1	✳ 🚀 → ✳ ⬛ → ✳ Σα₊	Front1

After clicking the SPSS program icon, Screen 1 appears on the monitor.

	Step 2

Create and name a data file or edit (if necessary) an already existing file (see Chapter 3).

Screens 1 and 2 (displayed on the inside front cover) allow you to access the data file used in conducting the analysis of interest. The following sequence accesses the **grades.sav** *file for further analyses:*

| In Screen	Do This	Step 3
Front1	✳ <u>F</u>ile �le→ <u>O</u>pen �le→ ✳ D<u>a</u>ta [or ✳ 📁]	
Front2	type `grades.sav` �le→ ✳ <u>O</u>pen [or ✳✳ grades.sav]	Data

Whether first entering SPSS or returning from earlier operations the standard menu of commands across the top is required. As long as it is visible you may perform any analyses. It is not necessary for the data window to be visible.

After completion of Step 3 a screen with the desired menu bar appears. When you click a command (from the menu bar), a series of options will appear (usually) below the selected command. With each new set of options, click the desired item. The sequence to access **<u>Means</u>** *begins at any screen with the menu of commands visible:*

In Screen	Do This	Step 4
Menu	✳ <u>A</u>nalyze �le→ Co<u>m</u>pare Means �le→ ✳ <u>M</u>eans	9.1

At this point a new window opens (Screen 9.1, below) that deals with designating for which variables you wish to compare means. At the top of the window is the **Dependent List** box. This is where the *continuous* variables that you wish to analyze will be placed. For instance, you may compare mean scores for previous GPA, for final exam scores, for total points, or for percentage of total points. You may list several variables in this box; SPSS will calculate a separate set of means for each variable.

Screen 9.1 The Means Window
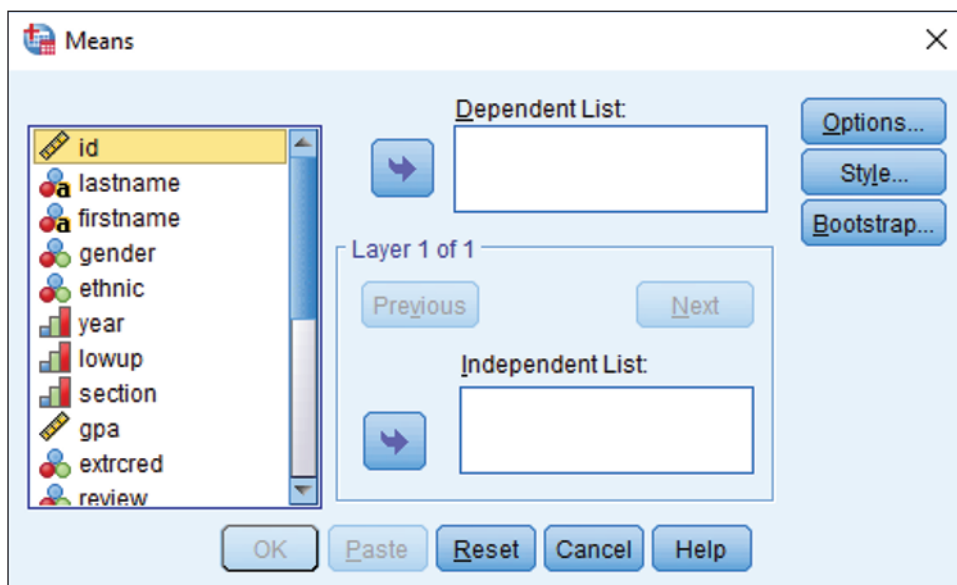
The lower **Independent List** box is where you identify the categorical variables, such as gender, section, grade, or ethnicity. If you include only one variable (such as **gender**) in the lower **Independent List** box, and a single variable in the upper **Dependent List** box (say, **total** points), the **Means** command will indicate the average (or mean) number of total points earned by women and the mean number of total points earned by men. The *N* and standard deviations are also included by default.

More frequently the researcher will desire more than just two categories. A more common operation would be a **gender** by **section** analysis that would give mean total scores for males and females in each of the three sections, or a **gender** by **year** analysis that would yield mean **total** scores for males and females for each year in college. To specify more than one categorical variable, make use of the **Previous** and **Next** to the left and right of **Layer 1 of 1** in the middle of the screen. As you observe the sequences of steps shown below, how these options are used will become clear. Each sequence that follows begins from Screen 9.1.

*To determine the mean number of total points (**total**) in each **section**, perform the following sequence of steps. Screen 9.1 is the starting point.*

In Screen	Do This								Step 5
9.1	⊛ total →	⊛ upper ➡ →	⊛ section →	⊛ lower ➡ →	⊛ OK				Back1

*If you wish to include an additional categorical variable in the analysis (**lowup**— lower or upper division student) in addition to **section**, the appropriate sequence of steps follows:*

In Screen	Do This				Step 5a
9.1	⊛ total →	⊛ upper ➡ →	⊛ section →	⊛ lower ➡ →	
	⊛ **Next** →	⊛ lowup →	⊛ lower ➡ →	⊛ OK	Back1

*You may list more than one dependent variable. SPSS will then produce as many columns in the table as there are dependent variables. In the procedure that follows, we compute means and standard deviations for **gpa**, **total**, and **percent** (three dependent variables) for six categories (three levels of **section** by two levels of **gender**) in the analysis.*

In Screen	Do This				Step 5b
9.1	⊛ gpa →	⊛ upper ➡ →	⊛ total →	⊛ upper ➡	
	→ ⊛ percent →	⊛ upper ➡ →	⊛ section →	⊛ lower ➡	
	→ ⊛ **Next** →	⊛ gender →	⊛ lower ➡ →	⊛ OK	Back1

If you wish to conduct a one-way analysis of variance or to include additional output within each cell, a click on the **Options** button will open a new screen (Screen 9.2, following page).

Screen 9.2 The Means: Options Window
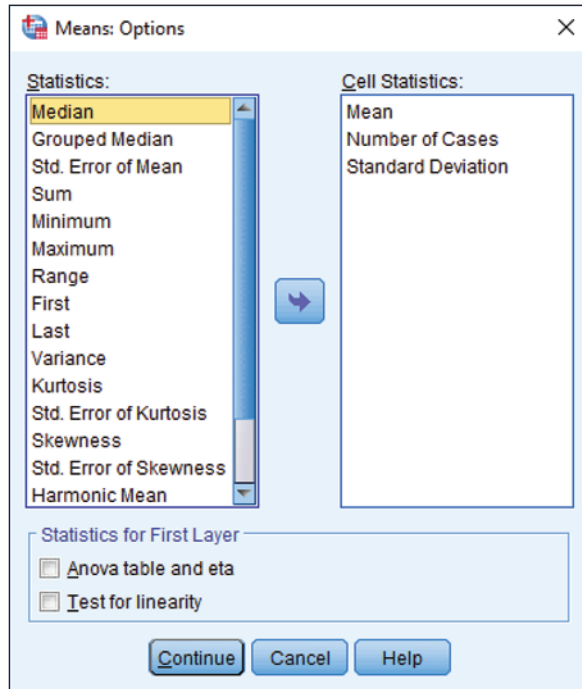
This screen allows the researcher to include additional output options for each analysis. For instance, in addition to the default (**Mean**, **Standard Deviation**, and **Number of Cases**), you may, by clicking over items from the box to the left, also include a number of other options. Using the **Means** command, it is possible to conduct a simple one-way analysis of variance (ANOVA). We will save detailed description of this procedure until Chapter 12, but will show how to access that option here. Under the **Statistics for the First Layer** box, click the **Anova table and eta** option. If you conduct this analysis with a "first layer" grouping variable of **section** and a dependent variable of **total**, then an analysis will be conducted that compares the mean total points for each of the three sections.

*The following sequence of steps will produce means on **total** points for the six categories of **gender** by **section**. It will also conduct a one-way analysis of variance for the first-layer variable (**section**) on **total** points. We begin at Screen 9.1. Do whichever of Steps 1–4 (pages 132–133) are necessary to arrive at this screen. You may need to click the **Reset** button.*

In Screen	Do This	Step 5c
9.1	✳ **total** ⟶ ✳ *upper* ➡ ⟶ ✳ **section** ⟶ ✳ *lower* ➡ ⟶ ✳ **Next**	
	⟶ ✳ **gender** ⟶ ✳ *lower* ➡ ⟶ ✳ **Options**	
9.2	✳ **Anova table and eta** ⟶ ✳ **Continue**	
9.1	✳ **OK**	Back1

Upon completion of Step 5, 5a, 5b, or 5c, the output screen will appear (Screen 1, inside back cover). All results from the just-completed analysis are included in the Output Navigator. Make use of the scroll bar arrows (▲▼▶◀) to view the

results. Even when viewing output, the standard menu of commands is still listed across the top of the window. Further analyses may be conducted without returning to the data screen.

9.2 Printing Results

Results of the analysis (or analyses) that have just been conducted require a window that displays the standard commands (**File Edit Data Transform Analyze . . .**) across the top. A typical print procedure is shown below beginning with the standard output screen (Screen 1, inside back cover).

To print results, from the Output screen perform the following sequence of steps:

In Screen	Do This	Step 6
Back1	*Select desired output or edit (see pages 16–22)* ⟶ ✳ **File** ⟶ ✳ **Print**	
2.9	*Consider and select desired print options offered, then* ⟶ ✳ **OK**	

To exit you may begin from any screen that shows the **File** *command at the top*.

In Screen	Do This	Step 7
Menu	✳ **File** ⟶ ✳ **Exit**	2.1

Note: After clicking **Exit**, there will frequently be small windows that appear asking if you wish to save or change anything. Simply click each appropriate response.

9.3 Output

9.3.1 Describing Subpopulation Differences

What follows is complete output from sequence Step 5c, page 135.

Means and One-Way ANOVA Results

Total				
Section	**Gender**	**Mean**	**N**	**Std. Deviation**
1	Female	103.95	20	18.135
	Male	106.85	13	13.005
	Total	105.09	33	16.148
2	Female	100.00	26	12.306
	Male	98.46	13	11.822
	Total	99.49	39	12.013
3	Female	102.83	18	10.678
	Male	90.73	15	21.235
	Total	97.33	33	17.184
Total	Female	102.03	64	13.896
	Male	98.29	41	17.196
	Total	100.57	105	15.299

ANOVA Table

		Sum of Squares	df	Mean Square	F	Sig.
TOTAL * SECTION	Between Groups (Combined)	1065.910	2	532.955	2.335	.102
	Within Groups	23277.804	102	228.214		
	Total	24343.714	104			

Measures of Association

	Eta	Eta Squared
TOTAL * SECTION	.209	.044

Note that the first portion of the output (table, previous page) is simply the means and frequencies for the entire group and for each of the selected categories. SPSS lists the mean, standard deviation, and number of cases for the first of the two categorical variables (**section**) in the upper portion of the chart, and similar information for gender in the lower portion of the chart. Definitions of terms in the ANOVA output are listed below. With means of 105.1, 99.5, and 97.3, the test statistics yield a p value of .102. This finding suggests that the sections may have a trend toward a significant influence on the total scores. Pairwise comparisons are not possible with this very simple ANOVA procedure, but visual inspection reveals that the greatest difference is between Section 1 ($M = 105.1$) and Section 3 ($M = 97.3$). See Chapter 12 for a more complete description of one-way ANOVAs. We conclude with definitions of output terms and the exercises.

Term	Definition/Description
Within-Group Sum of Squares	The sum of the squared deviations between the mean for each group and the observed values of each subject within that group.
Between-Group Sum of Squares	The sum of squared deviations between the grand mean and each group mean, weighted (multiplied) by the number of subjects in each group.
Between-Groups Degrees of Freedom	Number of groups minus one (3 − 1 = 2).
Within-Groups Degrees of Freedom	Number of subjects minus number of groups (105 − 3 = 102).
Mean Square	Sum of squares divided by degrees of freedom.
F ratio	Between-groups mean square divided by within-groups mean square.
Significance	The probability of the observed values happening by chance. This probability answers the question, "Are you fairly certain that the difference tested by the F statistic is real instead of random?" In this case, the p value indicated here ($p = .10$) indicates that a marginally significant difference between means may exist among at least one of the three pairings of the three sections. We aren't fairly certain, but it's probably worth exploring further.
Eta	This effect size measures the strength of the relation between the two variables, and answers the question "How big is the difference tested by the F statistic?"
Eta Squared	The proportion of the variance in the dependent variable accounted for by the independent variable. For instance, an eta squared of .044 indicates that 4.4% of the variance in the **total** scores is due to membership in one of the three **sections**. Whether you want to use eta or eta squared to answer the question, "How big is the difference tested by the F statistic?" will depend on what those in your particular field use; mathematically either is fine as you can convert from one to the other by squaring or taking the square root. Eta squared is easier to compare with other analyses, but eta gives a bigger number so some people prefer it for that reason.

Exercises

Answers to selected exercises are downloadable at **www.spss-step-by-step.net**.

1. Using the **grades.sav** file use the Means procedure to explore the influence of **ethnic** and **section** on **total**. Print output, fit on one page, and, in general terms, describe what the value in each cell means.

2. Using the **grades.sav** file use the Means procedure to explore the influence of **year** and **section** on **final**. Print output, fit on one page, and, in general terms, describe what the value in each cell means.

3. Using the **divorce.sav** file use the Means procedure to explore the influence of gender (**sex**) and marital status (**status**) on **spiritua** (spirituality—high score is spiritual). Print output and, in general terms, describe what the value in each cell means.

4. Using the **grades.sav** file and the Means procedure, examine the difference in **total** points between students who did or did not do the extra credit (**extrcredit**) project. Can you confidently say that doing the extra credit project helped student grades?

5. Using the **grades.sav** file and the Means procedure, examine the difference in **total** points between students who did or did not attend **review** sessions. Can you confidently say attending review sessions helped student grades?

6. Dr. Toob believes that people who meditate regularly are less jealous. She has ten participants meditate for a week, and ten participants not meditate for a week. After she measures level of jealousy for her participants, she uses the means procedure to determine whether her hypothesis is supported. People who meditated scored 45, 38, 34, 40, 30, 41, 37, 32, 28, and 29. People who did not meditate scored 27, 23, 27, 25, 27, 13, 13, 20, 37, and 23. Perform this analysis. Is her hypothesis supported? How confident can you be?