Print

Quiz, Lesson 9: Combining SAS Data Sets

Your Score: 100%

Congratulations! Your score of 100% indicates that you've mastered the topics in this lesson. If you'd like, you can review the feedback for each question.

When you're ready to start the next lesson, exit this lesson and begin the next one.

1. Which statement is true concerning match-merging?

   a. The MERGE statement must refer to permanent data sets.
   b. The variables in the BY statement can be in only one of the data sets.
   c. Only two data sets can be specified in the MERGE statement.
   d. When you use the MERGE statement with the BY statement, the data must be sorted or indexed on the BY variable.

   Your answer: d
   Correct answer: d

   With match-merging, each input data set must first be sorted on the values of the BY variable(s), or have an appropriate index.

   Review: Match-Merging SAS Data Sets

---

2. Which of the following data set options can be added to the MERGE statement to help identify data set contributors (that is, identify the matches)?

   a. IN=
   b. OBS=
   c. RENAME=
   d. MATCHES=

   Your answer: a
   Correct answer: a

   When you combine two data sets, you can use IN= data set options to track which of the original data sets contributes to each observation in the new data set.

   Review: Identifying Data Set Contributors

---

3. The following program will execute without errors:

```
data bonuses;
   merge managers(in=M)
         staff(in=S);
   by EmpID;
   if M=0 and S=1;
run;;
```

   a. True
   b. False

   Your answer: a

Correct answer: a

If **M**=0 and **S**=1, then only observations from the **staff** file that do not have a match in the **managers** file are selected.

Review: Identifying Data Set Contributors

---

4. Which subsetting IF statement selects observations for subsequent processing only if all three input data sets contribute to the current observation?

```
data merged.flowers;
   merge spring.roses(rename=(red=BrickRed) in=yellow)
         spring.lilacs(in=purple)
         spring.petunias(in=pink);
   by ID;
   _____________________________
run;
```

a. `if yellow=1 and purple=1 or pink=1;`
b. `if yellow=0 or purple=0 or pink=0;`
c. `if yellow=1 and purple=1 and pink=1;`
d. `if yellow=1 or purple=1 and pink=0;`

Your answer: c
Correct answer: c

The subsetting IF statement specifies all three IN= variables in the IF condition. The AND operator joins expressions in the condition. You can also write the IF statement as:

```
if yellow and purple and pink;
```

Review: Identifying Data Set Contributors

---

5. The following program will execute without errors:

```
data work.merged;
   merge blood.donors1 blood.donors2;
   by ID;
run;
```

blood.donors1

ID	Type	Units
2304	O	16
1129	A	48
1129	A	50
1129	A	57
2486	B	63

blood.donors2

ID	Code	Units
6488	65	27
1129	63	32
5438	62	39
2304	61	45
1387	64	67

a. True
b. False

Your answer: b

Correct answer: b

The two input data sets are not sorted by the values of the BY variables, so the DATA step produces errors and stops processing.

Review: Match-Merging SAS Data Sets

---

6. The following tables list the variables in three SAS data sets: **a**, **b**, and **c**. Can these three data sets be merged in one DATA step?

a

Hair
Makeup
Nails
Jewelry
Shoes

b

Bows
Hairspray
Hair
Rubber Bands
Brush

c

Hose
Socks
Anklet
Boots
Shoes

a. Yes
b. No

Your answer: b
Correct answer: b

Match-merging requires that the BY variable be present in all data sets being merged. There is no common variable to each of these data sets.

Review: Match-Merging Multiple Data Sets

---

7. Which of the following LIBNAME statements correctly assigns the libref **mymusic** to the file **AllMusic.xls**, which is stored in the **Entertainment** directory of the **C**: drive on the Windows operating environment?

a.
```
libname mymusic 'Allmusic.xls';
```
b.
```
libname mymusic 'C:\Entertainment\Allmusic.xls';
```
c.
```
libname mymusic 'C:\Entertainment\Allmusic';
```
d.
```
libname mymusic 'C:\Entertainment\Allmusic' filetype='excel';
```

Your answer: b
Correct answer: b

Specify the libref **MyMusic** and then the filename **AllMusic.xls**. Include the full path enclosed in quotation marks. Include the file extension when you specify the filename.

Review: Accessing an Excel Workbook in SAS

---

8. The data sets **ensemble.spring** and **ensemble.summer** both contain a variable named **Blue**, and **Blue** is not the BY variable. Which program prevents the values of the variable **Blue** from being overwritten when you merge the two data sets?

a.
```
data ensemble.merged;
   merge ensemble.spring(keep=blue)
         ensemble.summer;
```

```
   by fabric;
run;
```

b.

```
data ensemble.merged;
   merge ensemble.spring(blue=navy)
         ensemble.summer;
   by fabric;
run;
```

c.

```
data ensemble.merged;
   merge ensemble.spring(rename=(blue=navy))
         ensemble.summer;
   by fabric;
run;
```

Your answer: c
Correct answer: c

Match-merging overwrites same-named variables in the first data set with same-named variables in subsequent data sets. To prevent overwriting, rename variables using the RENAME= data set options in the MERGE statement.

Review: Match-Merging a SAS Data Set and an Excel Worksheet

---

9. The variable **Location** appears in the three data sets represented below. Which value appears in the output data set when the three data sets are merged in the order shown?

```
merge dataset1 dataset2 dataset3;
```

dataset1

Location
Florida

dataset2

Location
Canada

dataset3

Location
New York

a. Florida
b. Canada
c. New York

Your answer: c
Correct answer: c

Because DATA step match-merging overwrites values of same-named variables, the value *New York* from the last data set merged (**Dataset3**) appears in the output data set.

Review: Match-Merging a SAS Data Set and an Excel Worksheet

---

10. Suppose the **empinfo.bonuses** data set contains the variables **ID**, **Name**, **Office**, **Manager**, **Location**, and **Amount**. Specify a data set option in the MERGE statement below to use only the variables **ID**, **Name**, and **Amount** in the data set.

```
data mergedata.emppay;
   merge sales.reps(rename=(office=OfficeNumber))
         empinfo.sales
         empinfo.bonuses ____________________;
   by ID;
run;
```

a. `(drop=ID Name Amount)`
b. `(output ID Name Amount)`
c. `(keep=ID Name Amount)`
d. `(in=ID Name Amount)`

Your answer: c
Correct answer: c

The KEEP= data set option specifies the variables to keep in the **empinfo.bonuses** data set.

Review: Controlling Match-Merge Output

Close