UCD School of Mathematics and Statistics

STAT40840: Data programming with SAS
Laura Kirwan

Exercises 7

www.ucd.ie/mathstat

Exercise 1

Which of the following statements creates a numeric variable, **Bonus**, with a value

a. Bonus=$500;
b. Bonus=500;
c. label Bonus='500';
d. format Bonus 500.;

Partial **work.comp**

Bonus	Compensation	Bonus Month
500	108755	6
500	88475	1
500	27100	1

Exercise 1 solution

Which of the following statements creates a numeric variable, **Bonus**, with a value

a. Bonus=$500;

b. Bonus=500;

c. label Bonus='500';

d. format Bonus 500.;

Partial **work.comp**

Bonus	Compensation	Bonus Month
500	108755	6
500	88475	1
500	27100	1

You use an assignment statement to set the value of the variable, Bonus, equal to 500. Numeric constants do not include commas or currency symbols.

Exercise 2

A DROP statement was added to this DATA step. Can the program calculate **Compensation** and **BonusMonth** correctly?

```
data work.comp;
   set orion.sales;
   drop Gender Salary Job_Title Country
        Birth_Date Hire_Date;
   Bonus=500;
   Compensation=sum(Salary,Bonus);
   BonusMonth=month(Hire_Date);
run;
```

L7_E2.sas

Exercise 2 solution

A DROP statement was added to this DATA step. Can the program calculate **Compensation** and **BonusMonth** correctly?

```
23   data work.comp;
24      set orion.sales;
25      drop Gender Salary Job_Title Country
26                Birth_Date Hire_Date;
27      Bonus=500;
28      Compensation=sum(Salary,Bonus);
29      BonusMonth=month(Hire_Date);
30   run;
NOTE: There were 165 observations read from the data set ORION.SALES.
NOTE: The data set WORK.COMP has 165 observations and 6 variables.
```

Yes. A drop flag is set for the dropped variables, but the variables are in the PDV and therefore available for processing. DROP is a compile-time only statement.

Exercise 3

In the program L7_D2.sas, is it possible for more than one condition to be true for a single observation?

a. Yes, more than one condition can be true.

b. No, the conditions are mutually exclusive, so only one condition can be true.

Exercise 3 solution

In the program L2_D2.sas, is it possible for more than one condition to be true for a single observation?

a. Yes, more than one condition can be true.

b. No, the conditions are mutually exclusive, so only one condition can be true.

For each observation, there is only one value for Job_Title. If that value matches one of the conditions, then it cannot match any other condition.

Exercise 4

Program L7_E4.sas reads **orion.nonsales**, a non-validated data set. Open and submit the program and review the results. Why is **Bonus** set to 300 in observations 125, 197, and 200?

```
data work.bonus;
   set orion.nonsales;
   if Country='US' then Bonus=500;
   else Bonus=300;
run;
```

Exercise 4 solution

Program **L7_E4.sas** reads **orion.nonsales**, a non-validated data set. Open and submit the program and review the results. Why is **Bonus** set to 300 in observations 125, 197, and 200?

```
data work.bonus;
   set orion.nonsales;
   if Country='US' then Bonus=500;
   else Bonus=300;
run;
```

The Country variable has some mixed case values in orion.nonsales. Observations with a country value of *US* are assigned 500. All others are assigned 300, including *us*.

Exercise 5

How would you prevent **Freq** from being truncated?

Exercise 5 solution

How would you prevent **Freq** from being truncated?

Possible solutions:

- **Pad the first occurrence of the Freq value with blanks to be the length of the longest possible value.**
- **Switch conditional statements to place the longest value of Freq in the first conditional statement.**
- **Add a LENGTH statement to declare the byte size of the variable up front.**

Exercise 6

How many variables will be in **empsall2** after concatenating **empscn** and **empsjp**?

empscn

First	Gender	Country
Chang	M	China
Li	M	China
Ming	F	China

empsjp

First	Gender	Region
Cho	F	Japan
Tomi	M	Japan

```
data empsall2;
   set empscn empsjp;
run;
```

Exercise 6 - solution

How many variables will be in **empsall2** after concatenating **empscn** and **empsjp**?

empscn

First	Gender	Country
Chang	M	China
Li	M	China
Ming	F	China

empsjp

First	Gender	Region
Cho	F	Japan
Tomi	M	Japan

Four variables: First, Gender, Country, and Region

Exercise 7

Which statement has correct syntax?

a.
```
set empscn(rename(Country=Location))
    empsjp(rename(Region=Location));
```

b.
```
set empscn(rename=(Country=Location))
    empsjp(rename=(Region=Location));
```

c.
```
set empscn rename=(Country=Location)
    empsjp rename=(Region=Location);
```

Exercise 7 solution

Which statement has correct syntax?

a.
```
set empscn(rename(Country=Location))
    empsjp(rename(Region=Location));
```

b.
```
set empscn(rename=(Country=Location))
    empsjp(rename=(Region=Location));
```

c.
```
set empscn rename=(Country=Location)
    empsjp rename=(Region=Location);
```

Exercise 8

Which of the following BY statements correctly sorts by descending **salary** within **gender**?

a. by descending salary within gender;
b. by descending salary gender;
c. by gender descending salary;
d. by gender salary descending;

Exercise 8 solution

Which of the following BY statements correctly sorts by descending **salary** within **gender**?

a. by descending salary within gender;
b. by descending salary gender;
c. by gender descending salary;
d. by gender salary descending;

The keyword "descending" is placed before the variable to which it applies.

Exercise 9

- Complete program **L7_E9.sas t**o match-merge the sorted SAS data sets referenced in the PROC SORT steps.
- Submit the program. Correct and resubmit it, if necessary.
- What are the modified, completed statements?

Exercise 9 solution

```
proc sort data=orion.employee_payroll
          out=work.payroll;
   by Employee_ID;
run;

proc sort data=orion.employee_addresses
          out=work.addresses;
   by Employee_ID;
run;

data work.payadd;
   merge work.payroll work.addresses;
   by Employee_ID;
run;
```

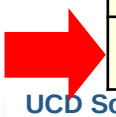
Exercise 10

Consider the data set **empsauc** created by the program in the previous example. Which input data sets contributed information to the last observation?

a. empsau
b. phonec
c. both empsau and phonec
d. There is insufficient information.

empsauc

First	Gender	EmpID	Phone
Togar	M	121150	+61(2)5555-1795
Kylie	F	121151	
Birin	M	121152	+61(2)5555-1667
		121153	+61(2)5555-1348



Exercise 10 solution

Consider the data set **empsauc** created by the program in the previous example. Which input data sets contributed information to the last observation?

a. empsau
b. phonec
c. both empsau and phonec
d. There is insufficient information.

empsauc

First	Gender	EmpID	Phone
Togar	M	121150	+61(2)5555-1795
Kylie	F	121151	
Birin	M	121152	+61(2)5555-1667
		121153	+61(2)5555-1348