

# Exercise 1

Which fields in this file can be read as standard numeric values?

## Partial **sales.csv**

```
120102, Tom, Zhou, M, 108255, Sales Manager, AU, 11AUG1973, 06/01/1993
120103, Wilson, Dawes, M, 87975, Sales Manager, AU, 22JAN1953, 01/01/1978
120121, Irenie, Elvish, F, 26600, Sales Rep. II, AU, 02AUG1948, 01/01/1978
120122, Christina, Ngan, F, 27475, Sales Rep. II, AU, 27JUL1958, 07/01/1982
120123, Kimiko, Hotstone, F, 26190, Sales Rep. I, AU, 28SEP1968, 10/01/1989
```

# Exercise 1 solution

Which fields in this file can be read as standard numeric values?

**The employee ID and salary. The date fields are nonstandard and require special processing.**

## Partial **sales.csv**

```
120102, Tom, Zhou, M, 108255, Sales Manager, AU, 11AUG1973, 06/01/1993
120103, Wilson, Dawes, M, 87975, Sales Manager, AU, 22JAN1953, 01/01/1978
120121, Irenie, Elvish, F, 26600, Sales Rep. II, AU, 02AUG1948, 01/01/1978
120122, Christina, Ngan, F, 27475, Sales Rep. II, AU, 27JUL1958, 07/01/1982
120123, Kimiko, Hotstone, F, 26190, Sales Rep. I, AU, 28SEP1968, 10/01/1989
```



# Exercise 2

Which statement is true?

- a. An input buffer is created only if you are reading data from a raw data file.
- b. The PDV at compile time holds the variable name, type, byte size, and initial value.
- c. The descriptor portion is the first item that is created at compile time.



# Exercise 2 solution

- Which statement is true?

- a. An input buffer is created only if you are reading data from a raw data file.
- b. The PDV at compile time holds the variable name, type, byte size, and initial value.
- c. The descriptor portion is the first item that is created at compile time.

## Exercise 3

Suppose you want the order of the variables to match the order of the fields. You can include the numeric variables in the LENGTH statement. Which of the following produces the correct results?

- a. 

```
length Employee_ID First Name $ 12
        Last_Name $ 18 Gender $ 1
        Salary Job Title $ 25
        Country $ 2;
```

- b. 

```
length Employee_ID 8 First Name $ 12
        Last_Name $ 18 Gender $ 1
        Salary 8 Job Title $ 25
        Country $ 2;
```

## Exercise 3 solution

Suppose you want the order of the variables to match the order of the fields. You can include the numeric variables in the LENGTH statement. Which of the following produces the correct results?

- a. 

```
length Employee_ID First Name $ 12
        Last_Name $ 18 Gender $ 1
        Salary Job Title $ 25
        Country $ 2;
```
- **b.**

```
length Employee_ID 8 First Name $ 12
        Last_Name $ 18 Gender $ 1
        Salary 8 Job Title $ 25
        Country $ 2;
```

## Exercise 4

What problems do you see with the data values for the last two data fields, **Salary** and **Country**?

### Partial **sales3inv.csv**

```
120102,Tom,Zhou,Manager,108255,AU
120103,Wilson,Dawes,Manager,87975,AU
120121,Irenie,Elvish,Rep. II,26600,AU
120122,Christina,Ngan,Rep. II,n/a,AU
120123,Kimiko,Hotstone,Rep. I,26190,AU
120124,Lucian,Daymond,Rep. I,26480,12
120125,Fong,Hofmeister,Rep. IV,32040,AU
```

# Exercise 4 solution

What problems do you see with the data values for the last two data fields, **Salary** and **Country**?

## Partial **sales3inv.csv**

```
120102,Tom,Zhou,Manager,108255,AU
120103,Wilson,Dawes,Manager,87975,AU
120121,Irenie,Elvish,Rep. II,26600,AU
120122,Christina,Ngan,Rep. II,n/a,AU
120123,Kimiko,Hotstone,Rep. I,26190,AU
120124,Lucian,Daymond,Rep. I,26480,12
120125,Fong,Hofmeister,Rep. IV,32040,AU
```



# Exercise 5

Submit program **L6\_E5.sas** and examine the log.

Which statement best describes the reason for the error?

- a. The data in the raw data file is invalid.
- b. The programmer incorrectly read the data.



# Exercise 5 solution

Which statement best describes the reason for the error?

- a. The data in the raw data file is invalid.
- b. The programmer incorrectly read the data.**

## Partial SAS Log

```
404      input Employee_ID First $ Last;  
405  run;
```

**Last** was read as  
numeric but needs  
to be read as character.

**NOTE: Invalid data for Last in line 1 16-17.**

```
RULE:      ----+-----1-----+-----2-----+-----3-----+-----4-----+-----5-----+-----6  
1          120101,Patrick,Lu,M,163040,Director,AU,18AUG1976,01JUL2003 58  
Employee_ID=120101 First=Patrick Last=. _ERROR_=1 _N_=1
```

**NOTE: Invalid data for Last in line 2 15-24.**

```
2          120104,Kareen,Billington,F,46230,Administration Manager,au,1  
61 1MAY1954,01JAN1981 78  
Employee_ID=120104 First=Kareen Last=. _ERROR_=1 _N_=2
```



## Exercise 6

- A *format* is an instruction that tells SAS how to display data values. What formats could you specify to display a SAS date in the styles shown below?

a) 01JAN2000

b) 01/16/2000



## Exercise 6 solution

A *format* is an instruction that tells SAS how to display data values. What formats could you specify to display a SAS date in the styles shown below?

a) 01JAN2000  $\Rightarrow$  **DATE9.**

b) 01/16/2000  $\Rightarrow$  **MMDDYY10.**



# Exercise 7

Use the SAS Help Facility or documentation to investigate the **DATEw.** informat and answer the following questions:

- a) What does the **w** represent?
- b) What is the default width of this informat?

# Exercise 7 solution

Use the SAS Help Facility or documentation to investigate the **DATEw.** informat and answer the following questions:

- a) What does the **w** represent?  
**the width of the input field**
- b) What is the default width of this informat?  
**The default width is 7.**



## Exercise 8

- Submit **L6\_E8.sas** and examine the log and output.
- How many input records were read and how many observations were created?
- Does the output look correct?

```
data work.contacts;  
    length Name $ 20 Phone Mobile $ 14;  
    infile "&path\phone2.csv" dlm=',';  
    input Name $ Phone $ Mobile $;  
run;  
  
proc print data=work.contacts noobs;  
run;
```

# Exercise 8 solution

- Submit **L6\_E8.sas** and examine the log and output.
- How many input records were read and how many observations were created? **five read, three created**
- Does the output look correct? **no**

NOTE: 5 records were read from the infile "S:\workshop\phone2.csv".  
The minimum record length was 31.  
The maximum record length was 44.

NOTE: SAS went to a new line when INPUT statement reached past the end of a line.

NOTE: The data set WORK.CONTACTS has 3 observations and 3 variables.

Name	Phone	Mobile
James Kvarniq	(704) 293-8126	(701) 281-8923
Sandrina Stephano	(919) 871-7830	Cornelia Krah1
Karen Ballinger	(714) 344-4321	Elke Wallstab