SAS Lesson 02

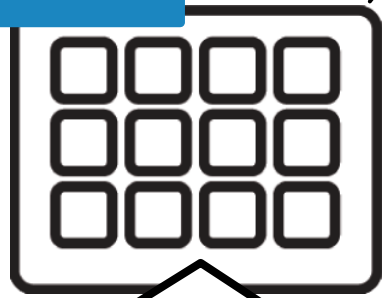
Accessing Data

Understanding SAS Data

Types of Data

Structured data

Has meta Data
→ Data about the rows/columns/size of table.



Unstructured data

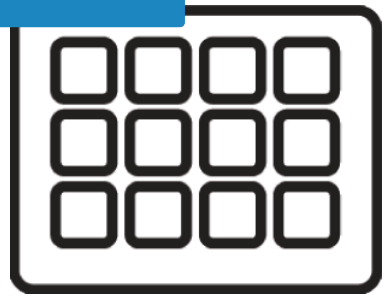
includes defined rows and columns

many types able to be read by SAS

SAS, Oracle, Teradata, Microsoft Excel, Hadoop, Versa tables, and others

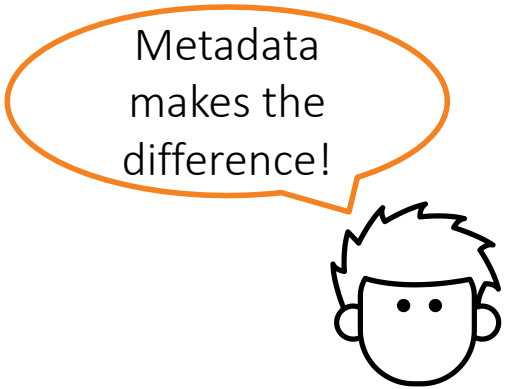
Types of Data

Structured data

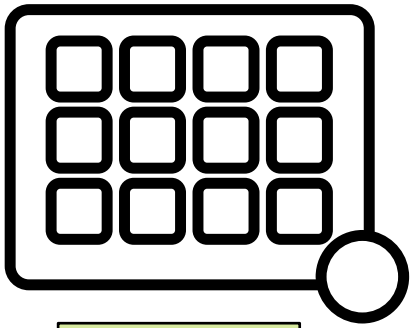


Unstructured data

- includes data, but not defined columns
- must be imported into SAS
- text, delimited, JSON, weblogs, and other files



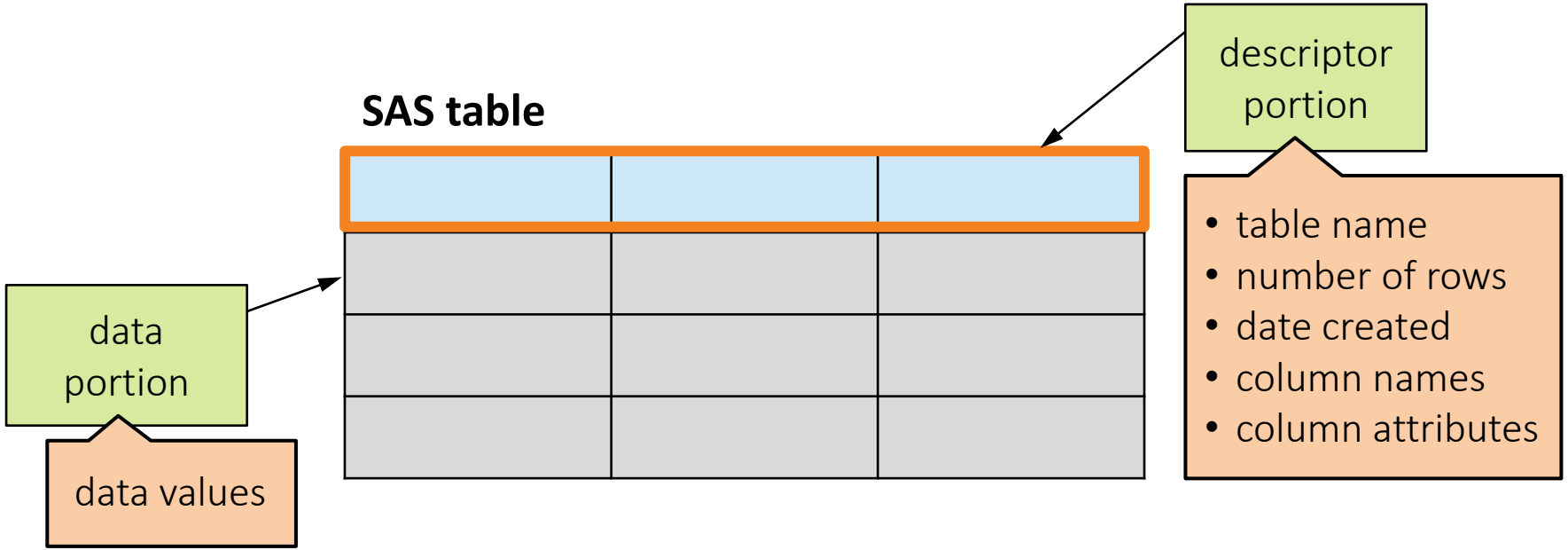
What Is a SAS Table?



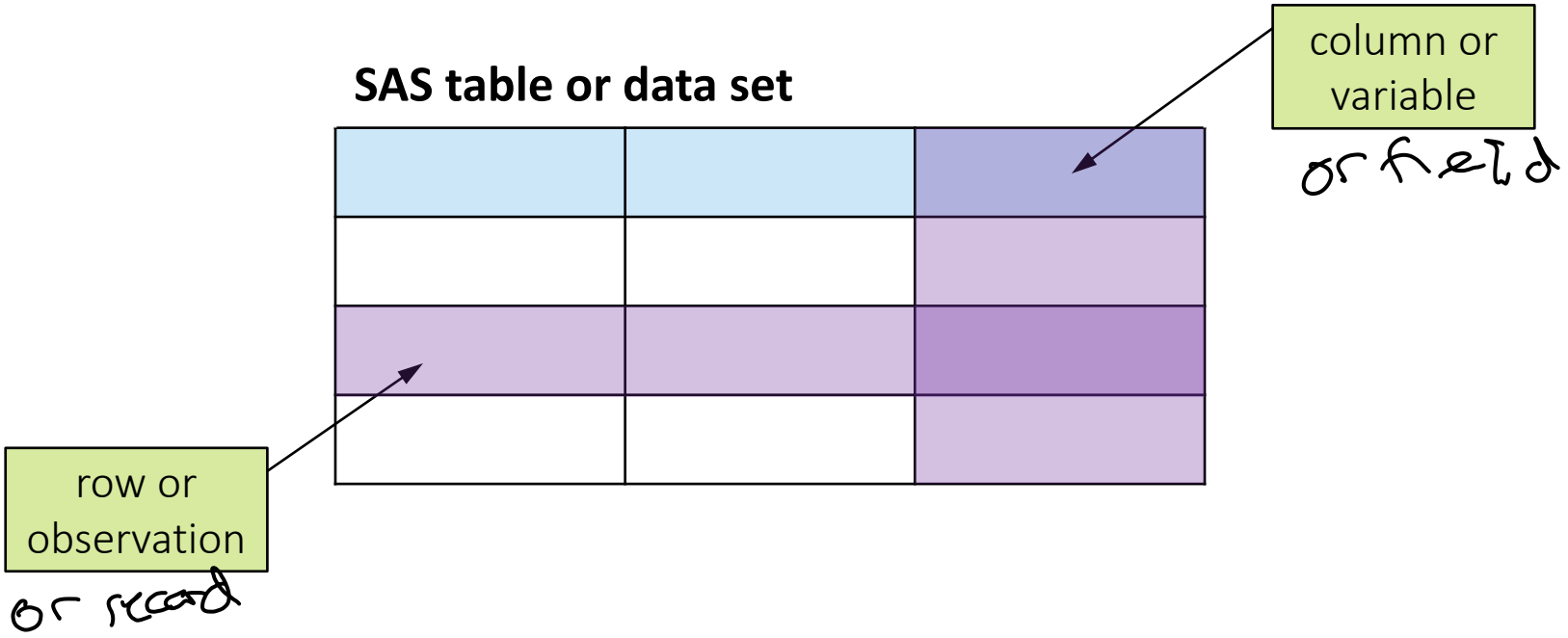
.sas7bdat

structured data

What Is a SAS Table?

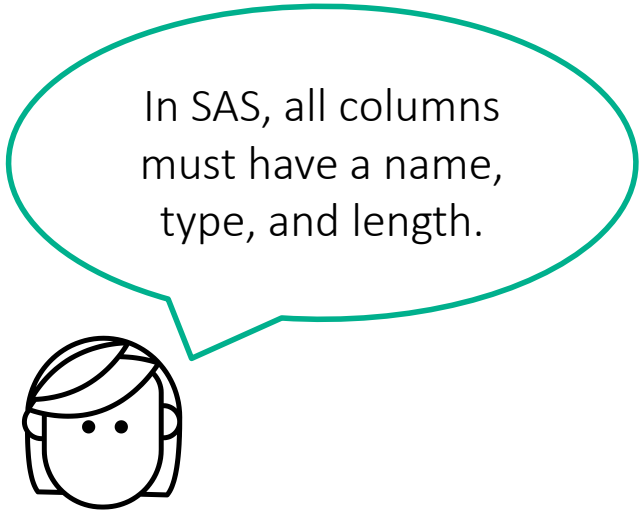


SAS Terminology

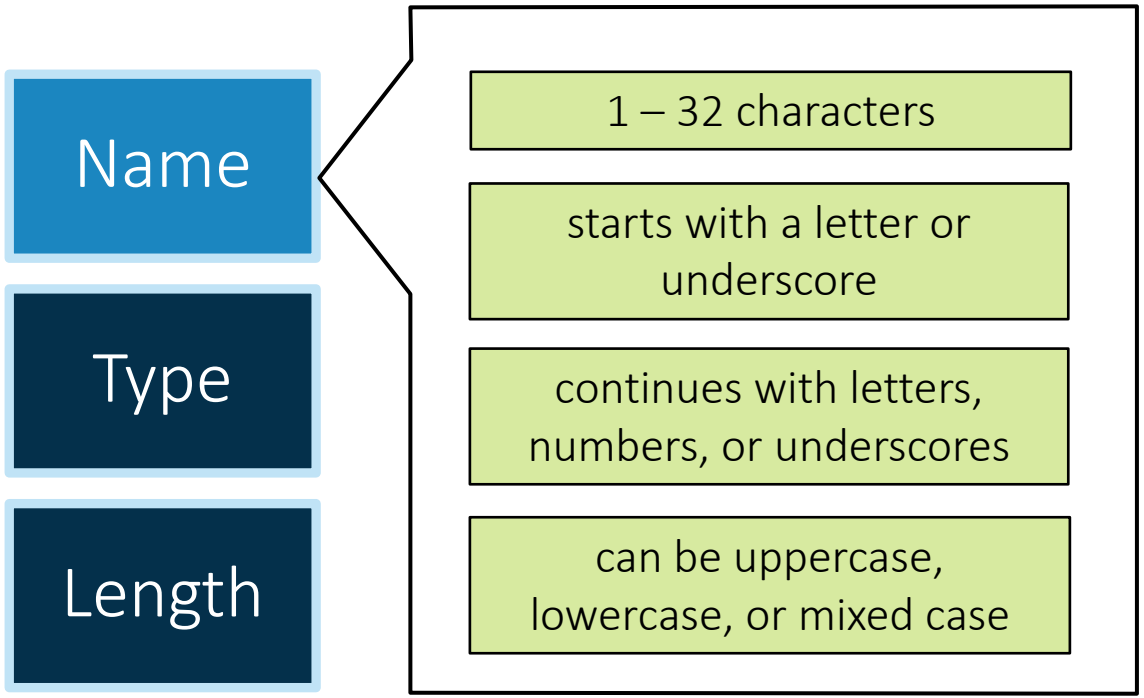


Required Column Attributes for SAS Tables

- Name
- Type
- Length



Required Column Attributes: Name



* no special characters in col names

Multiple Answer Question

Which column names are valid? (Select all that apply.)

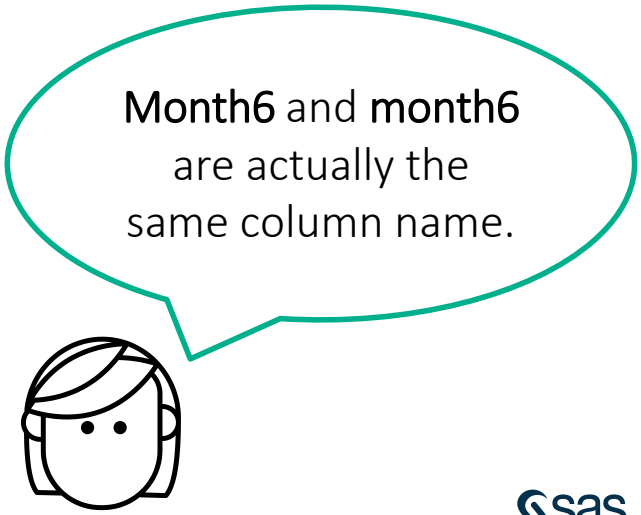
a. month6
b. 6month
c. month#6
d. month 6
e. month_6
f. Month6

Multiple Answer Question – Correct Answers

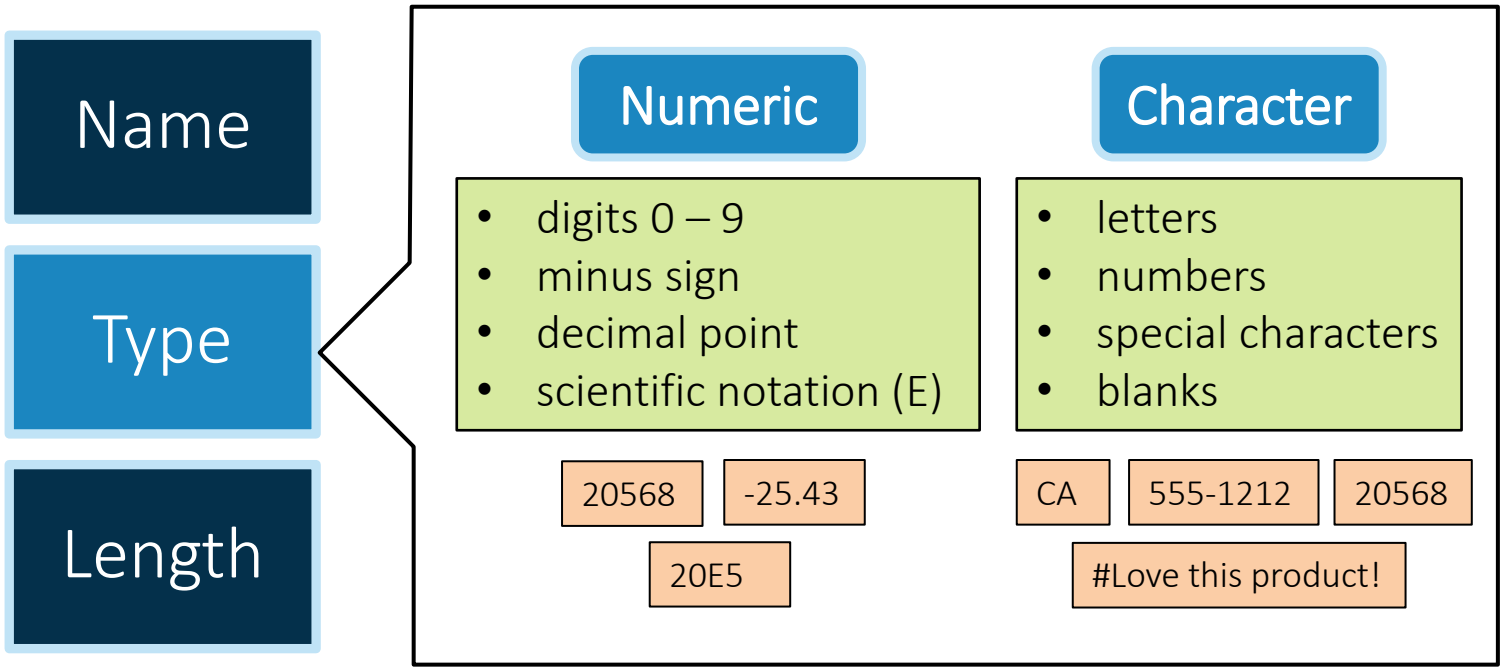
Which column names are valid? (Select all that apply.)

a. month6
b. 6month
c. month#6
d. month 6
e. month_6
f. Month6

same col name sas isn't case sensitive



Required Column Attributes: Type

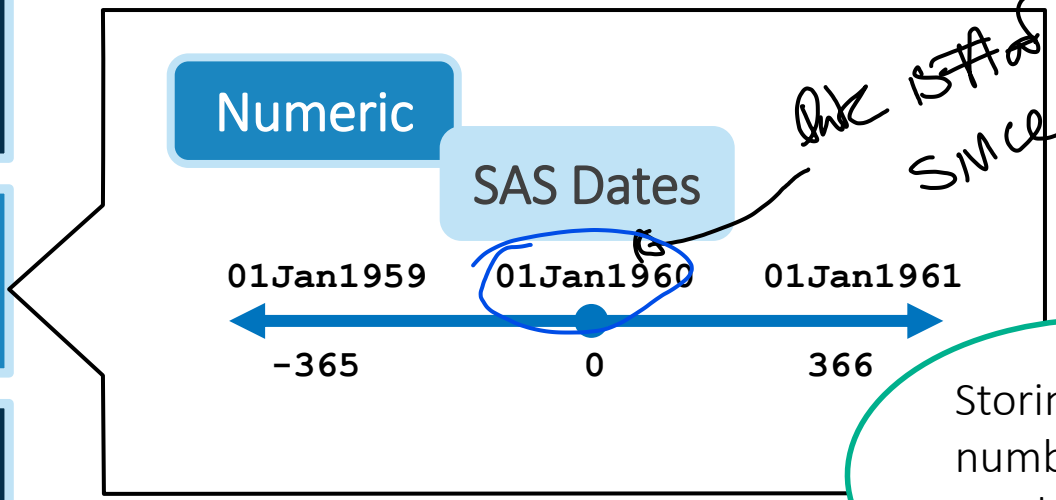


Required Column Attributes: Type

Name

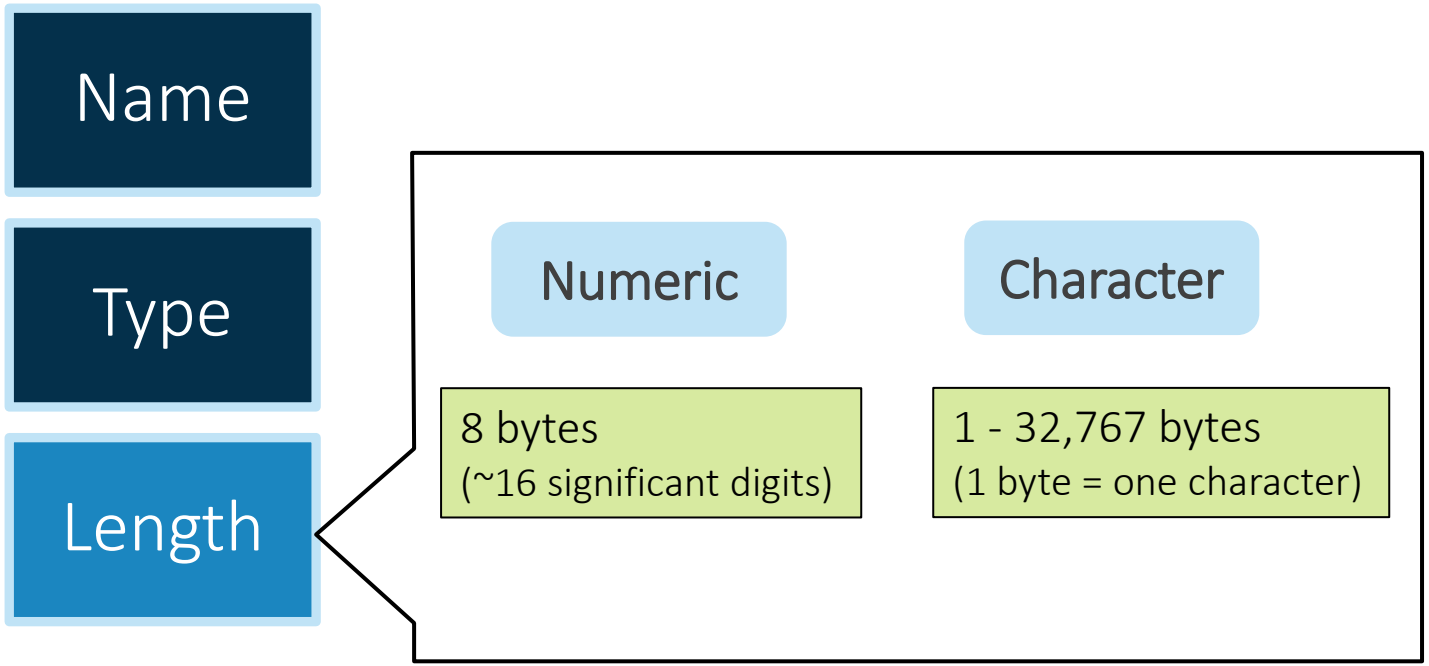
Type

Length



Storing dates as numbers makes calculations easy!

Required Column Attributes: Length



Activity

How are missing character and numeric values represented in the data shown below?

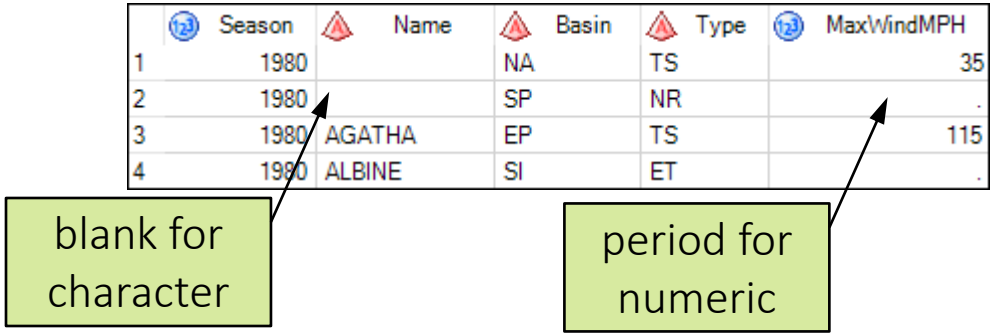
blank for char
period for numeric

VIEWTABLE: _EXP0_.storm_summary

	Season	Name	Basin	Type	MaxWindMPH	MinPressure
1	1980		na	TS	35	.
2	1980		SP	NR	.	998
3	1980	AGATHA	EP	TS	115	.
4	1980	ALBINE	SI	ET	.	.
5	1980	ALEX	WP	TS	40	998
6	1980	ALLEN	NA	TS	190	899
7	1980	AMY	SI	NR	132	915
8	1980	BERENICE	SI	TS	.	.
9	1980	BETTY	WP	ET	115	925
10	1980	BLAS	EP	TS	58	.
11	1980	BONNIE	NA	ET	98	975
12	1980	BRIAN	SI	NR	115	930
13	1980	CARMEN	WP	TS	69	985
14	1980	CARY	WP	TS	52	996
15	1980	CELIA	EP	TS	75	.
16	1980	CHARLEY	NA	TS	81	989

Activity – Correct Answer

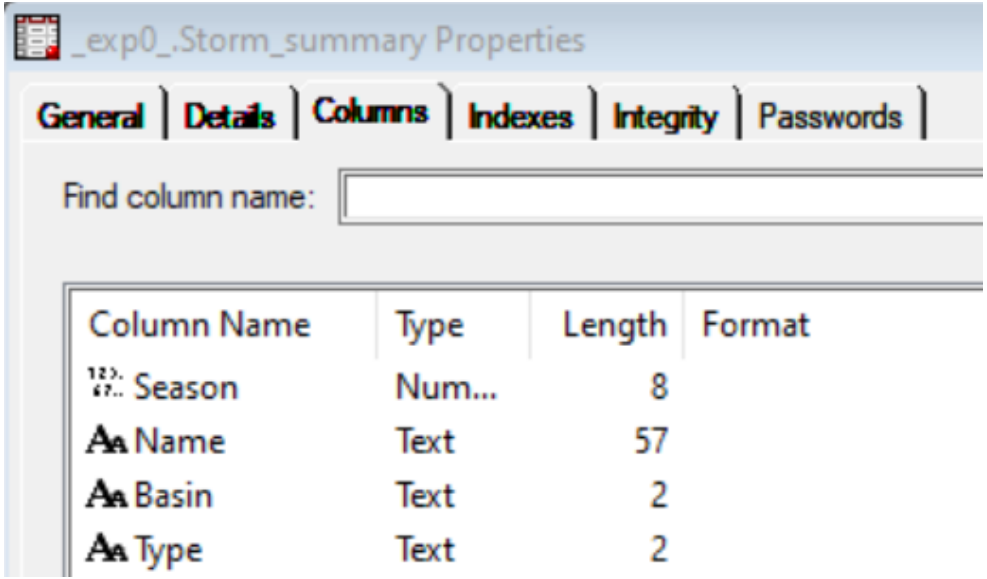
How are missing character and numeric values represented in the data?



Question

Examine the length of the **Basin** column. Could *East Pacific* be properly stored as a data value in the **Basin** column?

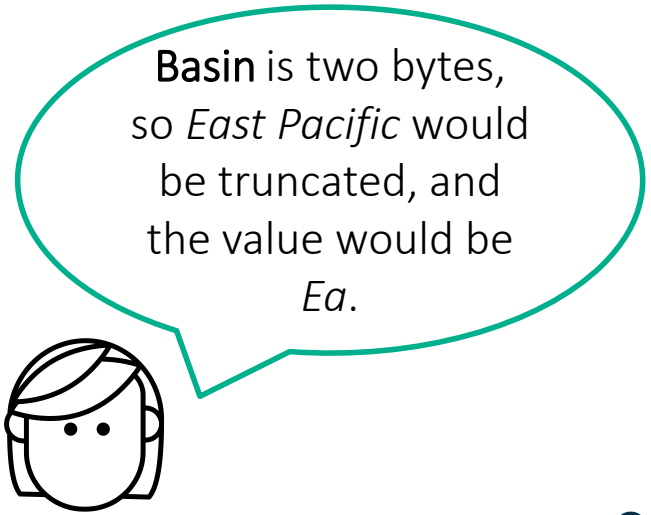
- Yes
- No



Question – Correct Answer

Examine the length of the **Basin** column. Could *East Pacific* be properly stored as a data value in the **Basin** column?

- ○ Yes
- ◉ No



Accessing Data

Accessing Data through Libraries – Chapter 3

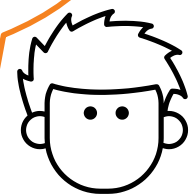
Using a Library to Read SAS Files

create the library*

LIBNAME *libref engine* "*path*";

- name of library
  - eight-character maximum
  - starts with a letter or underscore
  - continues with letters, numbers, or underscores
- what type of data it is
- where the data is located

*By default, LIBNAME does not actually create the data location.



Using a Library to Read SAS Files

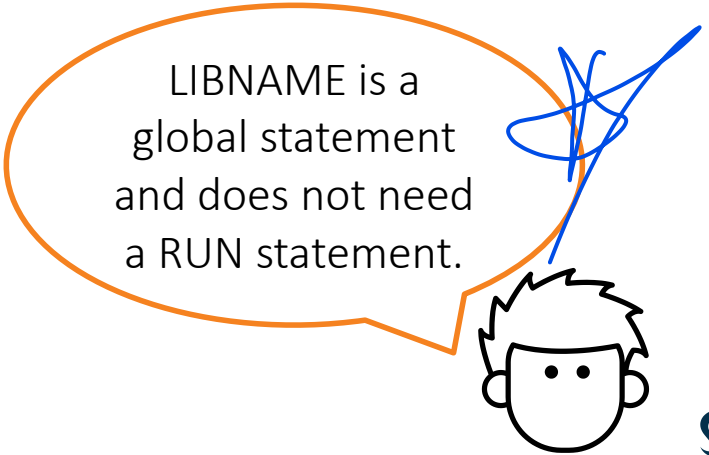
create the library

```
libname cert base "c:\users\userid\cert";
```

library name **cert**

Base SAS tables (optional)

data located in cert folder

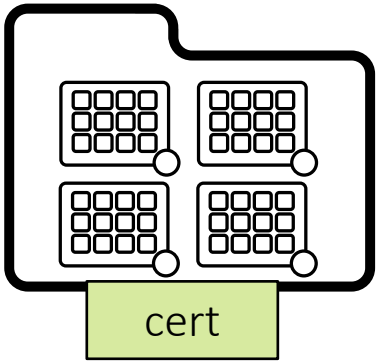


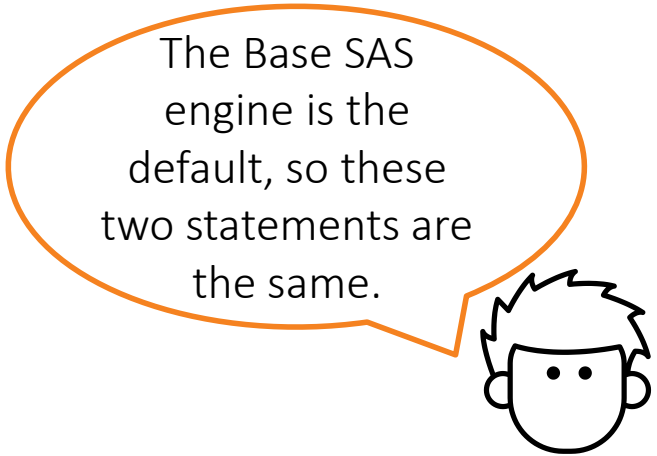
Using a Library to Read SAS Files

create the library

```
libname cert base "c:\users\userid\cert ";
```

```
libname cert "c:\users\userid\cert";
```



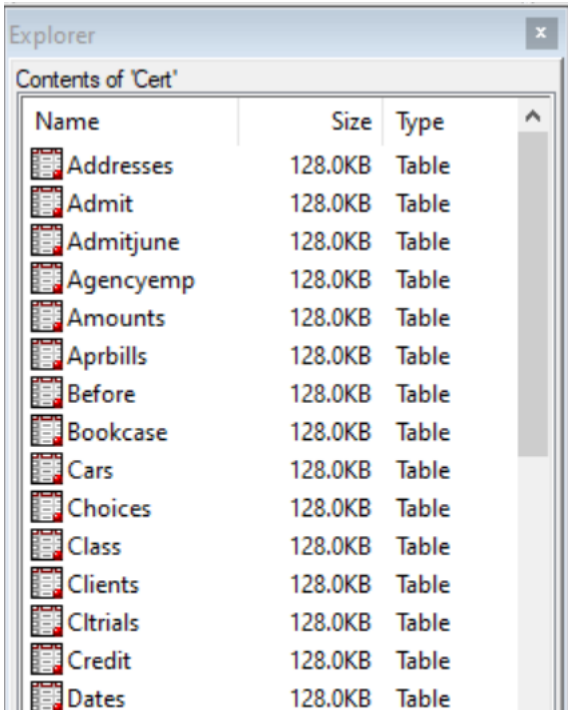


Assigning a SAS Libref

This demonstration/activity will assign a libref to the practice data that accompanies the certification prep guide.

Activity Question

Why are the Excel and text files in the **cert** folder not included in the library?



The **CERT** library uses the BASE engine, so it reads only Base SAS tables.

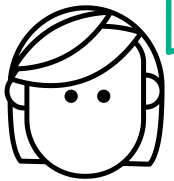
Viewing Table and Column Attributes

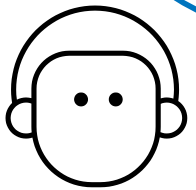
PROC CONTENTS DATA=*data-set* <*varnum*>;
RUN;

```
proc contents data=cert.class;
run;
```

PROC CONTENTS creates a report about the descriptor portion of the data.

The **varnum** option lists variables in the order created.



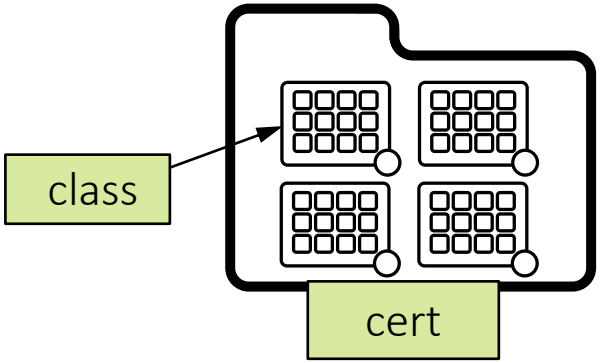


Using a Library with PROC CONTENTS

use the library

libref.table-name

```
proc contents data=cert.class;
run;
```



cert indicates the type of data and the location of the **class** table.

Browsing a SAS Data Library

The CONTENTS procedure with the _ALL_ keyword produces a list of all the SAS files in the data library.

```
PROC CONTENTS DATA=libref._ALL_ NODS;
RUN;
```

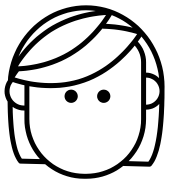
- The NODS option suppresses the descriptor portions of the data sets.
- NODS is only used in conjunction with the keyword _ALL_.

Viewing Table Data

PROC PRINT DATA=*data-set <options>*;
RUN;

```
proc print data=cert.class;
run;
```

PROC PRINT creates a report of all rows and columns by default.



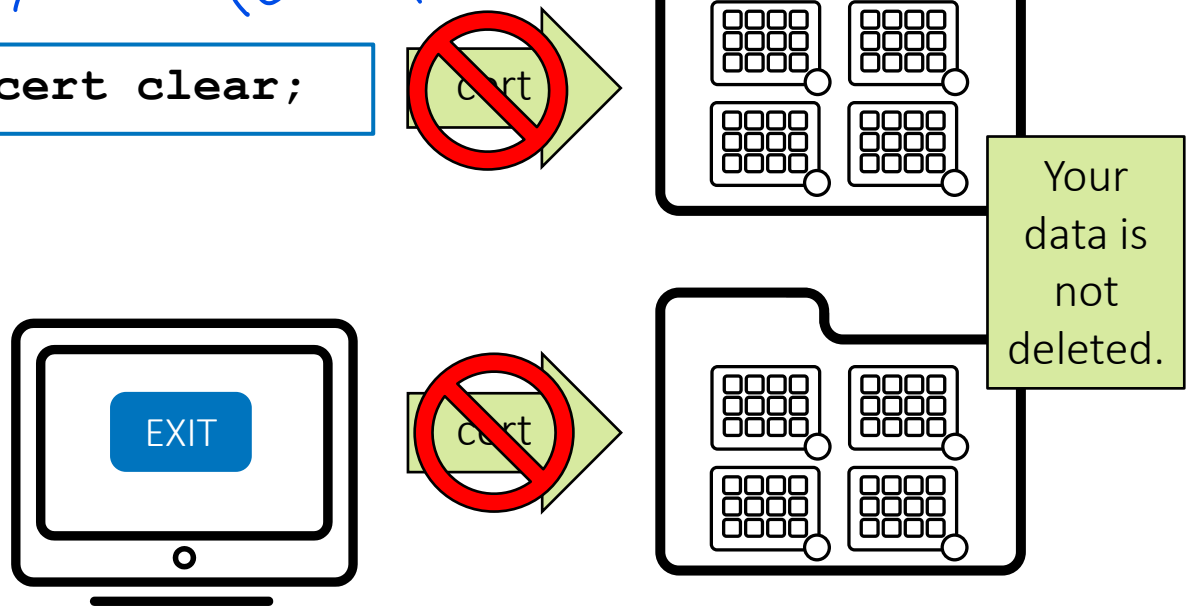
Exploring Automatic SAS Libraries

This demonstration illustrates using the **Contents Procedure** to explore automatic and user defined libraries.

Using a Library to Read SAS Files

delete library reference

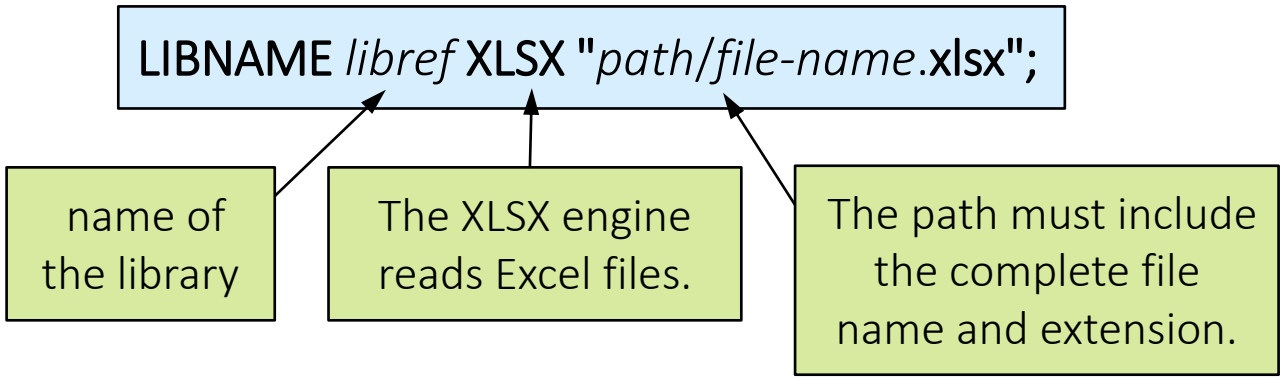
```
libname cert clear;
```



Referencing SAS Files

SAS ACCESS Libraries – Prep Guide Page 49

Using a Library to Read Excel Files




```
libname certxl xlsx "c:\users\userid\cert\exercise.xlsx";
```

The XLSX engine requires a license for SAS/ACCESS Interface to PC Files.

Using a Library to Read Excel Files

OPTIONS VALIDVARNAME=V7;

forces table and column names to follow SAS naming conventions

	A	B	C
1	First Name	Last Name	Days Employed
2	Brad	Majors	136
3	Janet	Weiss	
4	Everette	Scott	
5	Frank	Furter	

	First_Name	Last_Name	Days_Employed
1	Brad	Majors	136
2	Janet	Weiss	136
3	Everette	Scott	89
4	Frank	Furter	160

LIBNAME *libref* CLEAR;

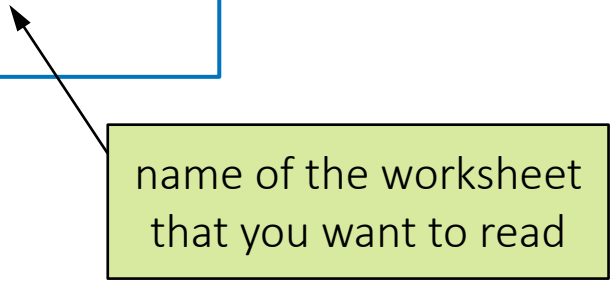
clears the connection to the Excel file

Using a Library to Read Excel Files

```
libname certxl xlsx "c:\users\userid\cert\exercise.xlsx";
```

```
proc contents data=certxl.ActivityLevels;
run;
```

```
libname certxl clear;
```



Using a Library to Read Excel Files

This demonstration illustrates creating a library to connect to an Excel workbook.

The SAS/ACCESS LIBNAME Statement

The *SAS/ACCESS LIBNAME statement* assigns a library reference name (libref) to a relational database.

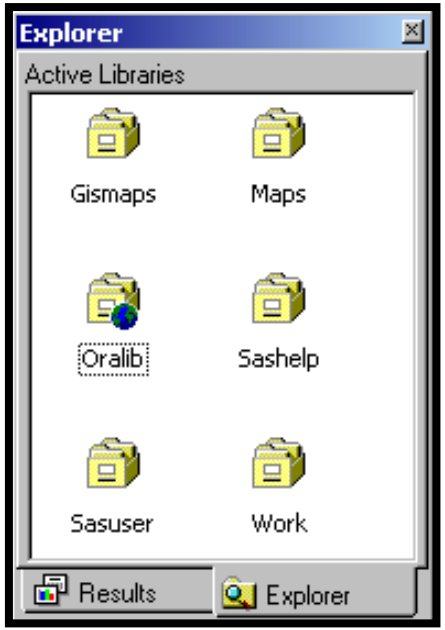
General form of the SAS/ACCESS LIBNAME statement:

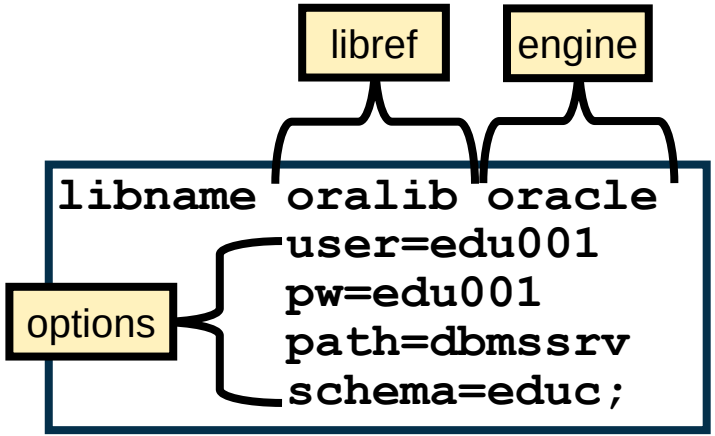
LIBNAME *libref engine-name <SAS/ACCESS-options>*;

After a database is associated with a libref, you can use a SAS two-level name to specify any table in the database and then work with the table as you would with a SAS data set.

Oracle Example

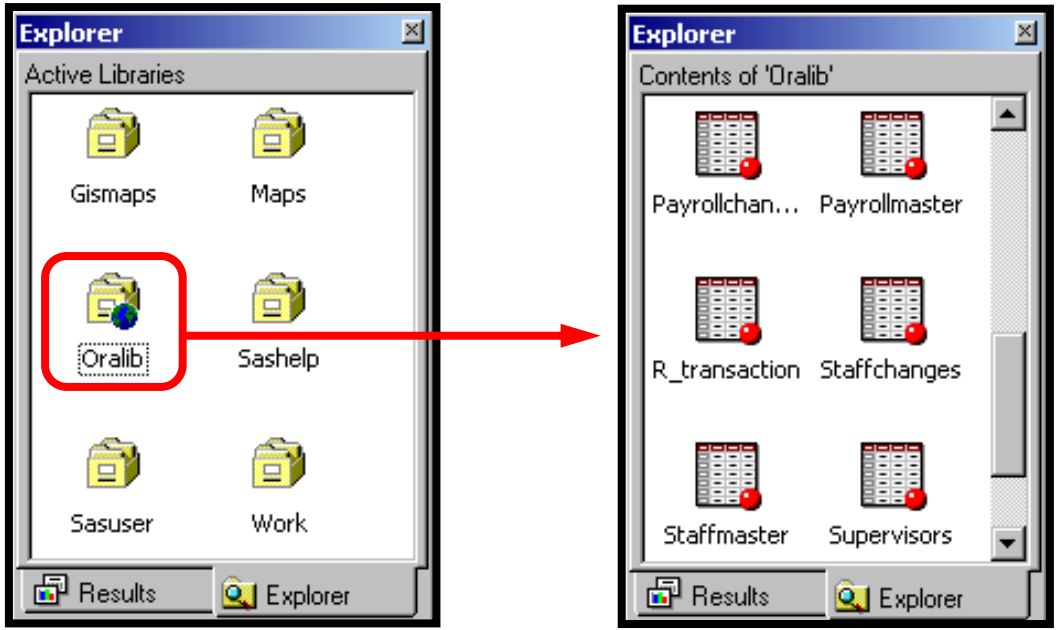
This example uses the LIBNAME statement as supported in the SAS/ACCESS interface to Oracle.





Oracle Example

Any table in this Oracle database can be referenced using a SAS two-level name.



Oracle Example

```
libname oralib oracle
        user=edu001 pw=edu001
        path=dbmssrv schema=educ;

proc print data=oralib.supervisors;
run;

data work.staffpay;
   merge oralib.staffmaster
         oralib.payrollmaster;
   by empid;
run;

libname oralib clear;
```

SQL Server Example

```
libname sqllib oledb
            init_string="Provider=SQLOLEDB;
            password=edu01;
            Persist Security Info=True;
            initial catalog=mydata;
            User ID=dal;
            data source=edserver"
            schema=dbo
            IGNORE_READ_ONLY_COLUMNS=YES;
```

Analyzing and Reporting on Data

Enhancing Reports with Titles and Footnotes
Prep Guide Pages 95-100

Using Titles and Footnotes

TITLE<*n*> "*title-text*";

FOOTNOTE<*n*> "*footnote-text*";

```
title1 "Heart Rates for Patients with";
title3 "Increased Stress Tolerance";
footnote1 "Data from Treadmill Tests";
footnote3 "1st Quarter Admissions";

proc print data=cert.stress;
   var resthr maxhr rechr;
   where tolerance="I";
run;
```

Heart Rates for Patients with

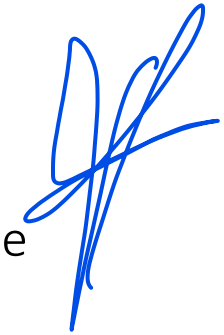
Increased Stress Tolerance

Obs	RestHR	MaxHR	RecHR
2	68	171	133
3	78	177	139
8	70	167	122
11	65	181	141
14	74	152	113
15	75	158	108
20	78	189	138

Data from Treadmill Tests

1st Quarter Admissions

Changing Titles and Footnotes



TITLE*n* or FOOTNOTE*n*

- replaces value of a previous title or footnote with the same number
- cancels all titles or footnotes with higher(larger) numbers.

Clearing Titles and Footnotes

TITLE;
FOOTNOTE;

clears titles and footnotes

ODS NOPROCTITLE;

turns off procedure titles

```
title;footnote;
ods noproctitle;
proc means data=sashelp.heart;
     var height weight;
run;
```

Creating Titles and Footnotes

This demonstration the use of titles and footnotes.

Differing Behavior

Did the PROC MEANS have titles?

- Yes (PC SAS)
- No (SAS Studio)

It depends. SAS Studio automatically clears existing titles before a new program is submitted. PC SAS does not.

Some procedures automatically add a procedure title.

SAS Studio

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
RestHR	7	72.5714286	5.0284903	65.0000000	78.0000000
MaxHR	7	170.7142857	12.9449383	152.0000000	189.0000000

PC SAS

Heart Rates for Patients with

Increased Stress Tolerance

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
RestHR	7	72.5714286	5.0284903	65.0000000	78.0000000
MaxHR	7	170.7142857	12.9449383	152.0000000	189.0000000

Lesson Quiz

1. In this PROC CONTENTS output, what is the default length of the **Birth_Date** column?

a. 4 bytes

b. 8 bytes

c. 32,767 bytes

d. It does not have a default length.

#	Variable	Type
4	Birth_Date	Num
3	Customer_Address	Char
1	Customer_ID	Num
2	Customer_Name	Char

1. In this PROC CONTENTS output, what is the default length of the **Birth_Date** column?

a. 4 bytes

b. 8 bytes

c. 32,767 bytes

d. It does not have a default length.

#	Variable	Type
4	Birth_Date	Num
3	Customer_Address	Char
1	Customer_ID	Num
2	Customer_Name	Char

2. Which LIBNAME statement has the correct syntax?

a. `libname reports "filepath/workshop";`
b. `libname orion filepath/workshop;`
c. `libname 3456a "filepath/workshop";`

2. Which LIBNAME statement has the correct syntax?

a. `libname reports "filepath/workshop";`

b. `libname orion filepath/workshop;`

c. `libname 3456a "filepath/workshop";`

3. Which of the following tables is available at the beginning of a new SAS session?

a. sales
b. work.newsalesemps
c. sashelp.class

3. Which of the following tables is available at the beginning of a new SAS session?

a. sales
b. work.newsalesemps
c. sashelp.class

4. In this table, what type of column is **Employee_ID**?

a. character
b. numeric
c. temporary
d. missing

Obs	Employee_ID	Last	Salary
1	.	Ralston	29250
2	120101	Lu	163040
3	120104	Billington	46230
4	120105	Povey	27110
5	120106	Hornsey	.

4. In this table, what type of column is **Employee_ID**?

a. character
b. numeric
c. temporary
d. missing

Obs	Employee_ID	Last	Salary
1	.	Ralston	29250
2	120101	Lu	163040
3	120104	Billington	46230
4	120105	Povey	27110
5	120106	Hornsey	.

5. Which statement about SAS dates is false?

a. A SAS date is one of three of SAS column types: numeric, character, and date.
b. SAS dates represent the number of days from January 1, 1960.
c. SAS date values can be positive or negative.
d. SAS date values can be used in calculations.

5. Which statement about SAS dates is false?

a. A SAS date is one of three of SAS column types: numeric, character, and date.

b. SAS dates represent the number of days from January 1, 1960.

c. SAS date values can be positive or negative.

d. SAS date values can be used in calculations.

7. Which library name (libref) is valid?

a. 2010Car
b. car/2010
c. car2010
d. cars_2010

7. Which library name (libref) is valid?

a. 2010Car
b. car/2010
c. car2010
d. cars_2010

8. To disassociate a libref that you previously assigned, you can use the UNASSIGN option in the LIBNAME statement.

a. True
b. False

8. To disassociate a libref that you previously assigned, you can use the UNASSIGN option in the LIBNAME statement.

a. True

b. False

10. In which portion of a SAS table are the following found?
    - name of the table
    - type of the column **Salary**
    - creation date of the table

a. descriptor portion

b. data portion

10. In which portion of a SAS table are the following found?
    - name of the table
    - type of the column **Salary**
    - creation date of the table

a. descriptor portion
b. data portion