SAS Industry Orientation Program
Department of Statistics
University of Lucknow

SAS Programming Course

(March 16, 2024 – May 04, 2024)

Department of Statistics
University of Lucknow

Disclaimer:

The topics, thoughts, case studies etc. in the discussions are sole belongs to me and do not belong or represent organization where we work.

Agenda for today:

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

- Concept of Library
- Data Types in SAS
- Concept of Dataset

Concept of Library : Introduction

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

In SAS (Statistical Analysis System), a library is a collection of SAS files stored in one or more directories or folders. These files can include SAS data sets, catalogs, macros, and other types of SAS files. Libraries are used to organize and manage data and other resources within SAS.

Types of Libraries: There are several types of libraries in SAS, including:

1. **SAS Libraries**: These libraries contain SAS data sets and other SAS files. They are created using the LIBNAME statement.
2. **External Databases**: SAS can also access data stored in external databases such as Oracle, SQL Server, and others. These databases can be connected to SAS using LIBNAME statements with appropriate database engine options.
3. **Special Libraries**: SAS also supports special libraries such as WORK and SASUSER. The WORK library is temporary and stores data sets for the duration of a SAS session, while the SASUSER library contains user-specific SAS files.

Concept of Library : Types of SAS Library

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

Work Library

The WORK library in SAS is a special temporary library that is automatically created for each SAS session.
It serves as a scratch space for storing intermediate data sets and other files during the execution of SAS programs.
The data sets stored in the WORK library are typically deleted at the end of the SAS session, making it ideal for temporary storage.

Characteristics of the WORK Library:

Temporary Storage: The WORK library is used to store temporary data sets that are created during the execution of SAS programs. These data sets are typically short-lived and are automatically deleted when the SAS session ends.

Automatic Creation: Unlike user-defined libraries, the WORK library is automatically created by SAS when a new session is initiated. Users do not need to explicitly define or specify the location of the WORK library.

Session Scope: The data sets stored in the WORK library are only accessible within the current SAS session.
They cannot be accessed or referenced outside of the session.

Efficient Memory Usage: Since the WORK library resides in memory, accessing and manipulating data stored in this library is generally faster compared to accessing data stored on disk. This makes it suitable for handling intermediate data processing tasks.

Concept of Library : Types of SAS Library

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

Permanent Library

A permanent library in SAS is a directory or folder on a storage device that contains SAS files, such as data sets, catalogs, macros, and other resources, and it remains accessible across multiple SAS sessions.

Characteristics of the PERMANENT Library:

Persistence: Permanent libraries persist across SAS sessions. Once defined, they remain accessible until explicitly deleted or modified.

Storage: They are typically stored on disk or a network drive, allowing SAS data sets, catalogs, macros, and other files within the library to be retained and accessed even when SAS sessions are closed and reopened.

Accessibility: Permanent libraries are accessible to multiple SAS sessions running concurrently on the same machine or across a network, facilitating collaboration and data sharing among users.

User-Defined Naming: Users can define the names of permanent libraries according to their preferences or organizational conventions, providing flexibility in organizing and managing data and resources.

Concept of Library : Naming SAS Library

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

When naming a SAS library, it's important to follow certain rules and conventions to ensure consistency, readability, and compatibility within your SAS environment. Here are the key rules for naming SAS libraries:

1.Length: The library name can be up to 8 characters long.

2.Characters:

1. Library names can consist of letters, numbers, and underscores (_).
2. The first character must be a letter or underscore.
3. Subsequent characters can be letters, numbers, or underscores.

3.Case Sensitivity:

1. SAS is not case-sensitive for library names; however, it preserves the case of the library name as specified.
2. It's good practice to use consistent capitalization for readability.

4.Special Characters: Avoid using special characters such as spaces, punctuation marks, or symbols in library names. Stick to alphanumeric characters and underscores for simplicity and compatibility.

Concept of Library : Examples

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

❖ Permanent library

```
/* Create a data set in the PERMANENT library */
data mydata.temp_data;
 input id name $ age;
 datalines;
 1 John 30
 2 Alice 25
 3 Bob 35
 ;
 run;
```

❖ Temporary (work) library

```
/* Create a data set in the WORK library */
data work.temp_data;
 input id name $ age;
 datalines;
 1 John 30
 2 Alice 25
 3 Bob 35
 ;
 run;
```

Data Types in SAS

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

In SAS, data types define the kind of values that variables can hold within a dataset. SAS supports several data types, including numeric, character, date/time, and special types like binary and formatted variables. Understanding data types is essential for data manipulation, analysis, and reporting in SAS.

1- Numeric Data Type

2- Character Data Type

3- Date/Time Data Type

Data Types in SAS

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

1- Numeric Data Type:

Numeric variables in SAS hold numerical values such as integers or decimals. Numeric variables can be used for mathematical calculations and statistical analyses.

```
/* Numeric variable declaration */

 data mydata;
 input ID Age Height Weight;
 datalines;
 1 30 175.5 70.2
 2 25 162.3 55.7
3 35 180.0 80.0
;
run;
```

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

2- Character Data Type:

Character variables in SAS hold alphanumeric values such as names, labels, or textual data. Character variables are enclosed in single or double quotation marks.

```
/* Character variable declaration */

data mydata;
  input ID $ Name $ Gender $;
  datalines;
1 John M
2 Alice F
3 Bob M
;
run;
```

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

3- Date/Time Data Type:

Date and time variables in SAS represent specific dates or times. SAS provides date and time formats for storing and manipulating date/time values.

```
/* Date/time variable declaration */

data mydata;
  input ID DateOfBirth :mmddyy10. AppointmentDate :datetime.;
  format DateOfBirth mmddyy10. AppointmentDate datetime.;
  datalines;
1 05/20/1985 01JAN2024:09:00:00
2 12/15/1990 15MAR2024:10:30:00
3 09/30/1980 25APR2024:13:45:00
;
run;
```

Concept of Data

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

In SAS, datasets are structured collections of data organized into rows and columns. Each row represents an observation, while each column represents a variable. Datasets are fundamental to SAS programming and analysis, as they serve as the primary means of storing, manipulating, and analyzing data within the SAS environment.

Characteristics of SAS Datasets:

Structured Format: SAS datasets are stored in a structured format that allows for efficient storage and retrieval of data.
Each dataset consists of two main components: the data portion (observations) and the descriptor portion (metadata).

Variables and Observations: Variables represent the characteristics or attributes being measured, while observations represent individual instances or records. Variables can be of different types, including numeric, character, and date/time.

Data Types: SAS supports various data types, including numeric, character, date/time, and special types like binary and formatted variables.

Attributes: Each variable in a SAS dataset can have associated attributes, such as variable labels, formats, and informat.

Metadata: SAS datasets contain metadata, which includes information about the dataset structure, variable attributes, and other properties.

Concept of Data : Naming a SAS Dataset

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

Naming conventions for SAS datasets help maintain consistency, clarity, and organization within your SAS environment. Here are some rules and best practices for naming SAS datasets:

Length:
- Dataset names can be up to 32 characters long.

Characters:
- Dataset names can consist of letters, numbers, and underscores (_).
- The first character must be a letter or underscore.
- Subsequent characters can be letters, numbers, or underscores.

Case Sensitivity:
- SAS dataset names are case-insensitive by default. However, SAS preserves the case of the dataset name as specified.
- It's good practice to use consistent capitalization for readability.

Special Characters:
- Avoid using special characters such as spaces, punctuation marks, or symbols in dataset names.
- Stick to alphanumeric characters and underscores for simplicity and compatibility.

Concept of Data : Naming a SAS Dataset

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

Reserved Words:

- Avoid using SAS reserved words as dataset names. Using reserved words as dataset names can lead to confusion and errors in your SAS programs. You can find a list of reserved words in the SAS documentation.

Meaningful and Descriptive:

- Choose dataset names that are meaningful and descriptive of the data they contain or the purpose they serve.
- Use clear and concise names that reflect the content or purpose of the dataset.

Avoid Confusion:

- Avoid using dataset names that are too similar to existing datasets in your SAS environment to prevent confusion.
- Ensure that dataset names are unique and easily distinguishable from other datasets.

Consistency:

- Maintain consistency in dataset naming conventions across your SAS projects and programs.
- Follow organizational or team-specific naming conventions if applicable.

Concept of Data : Development of SAS Data

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

SAS datasets can be created in several ways, including:

Using DATA Step: The DATA step is a fundamental component of SAS programming used for data manipulation and dataset creation.

/* Create a SAS dataset using the DATA step */

```
data mydata;
  input ID Name $ Age;
  datalines;
1 John 30
2 Alice 25
3 Bob 35
;
run;
```

Concept of Data : Development of SAS Data

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

Importing External Data: SAS can import data from various external sources, such as CSV files, Excel spreadsheets, and databases, to create SAS datasets.

/* Import data from a CSV file */

```
proc import datafile="path_to_csv_file.csv" out=mydata dbms=csv replace;
  getnames=yes;
run;
```

Concept of Data : Development of SAS Data

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

Using PROC SQL: The SQL procedure in SAS can be used to create datasets by querying existing datasets or external databases.

/* Create a dataset using PROC SQL */

```
proc sql;
  create table mydata as
  select ID, Name, Age
  from existing_dataset;
quit;
```

SAS Industry Orientation Program
Department of Statistics
University of Lucknow

Thank You.

Have a Question:

Anant Prakash Awasthi
M: +91-88846-52929
E: anant.awasthi@outlook.com

Dhananjay Srivastava
M: +91-809014236
E: Srivastava.dhananjay@outlook.com