CLASS : 1

SAS Tutorial

PRESENTED BY : SHASHI KUMAR

Syllabus :

Class No.	SAS Syllabus
1	Introduction of SAS, Data Step, and Step Boundary
2	a) SAS Library & related things b) Introduction of SAS Procedure with Proc print and Proc contents
3	a) Backend process : PDV b) Filtrations : Slicing , Dicing , if , where, keep, drop
4	Combining Data : Horizontal and Vertical merging
5	Data Manipulation : SAS functions
6	a) Data Importing : File ref, length statement, Datelines, cards, miss over, trun cover b) Descriptive Statistics : Proc freq, Proc means, Proc summary
7	a) Formatting Data :Format , informant , label, attrib statements, length statement b) Creating Reports : Title, Footnotes, Option, ODS
8	a) Summarization data: Retain, Sum statement b) Loops : do over loop & do, do while & do until loop


python
sas
VS
R

What is SAS & Why SAS ??

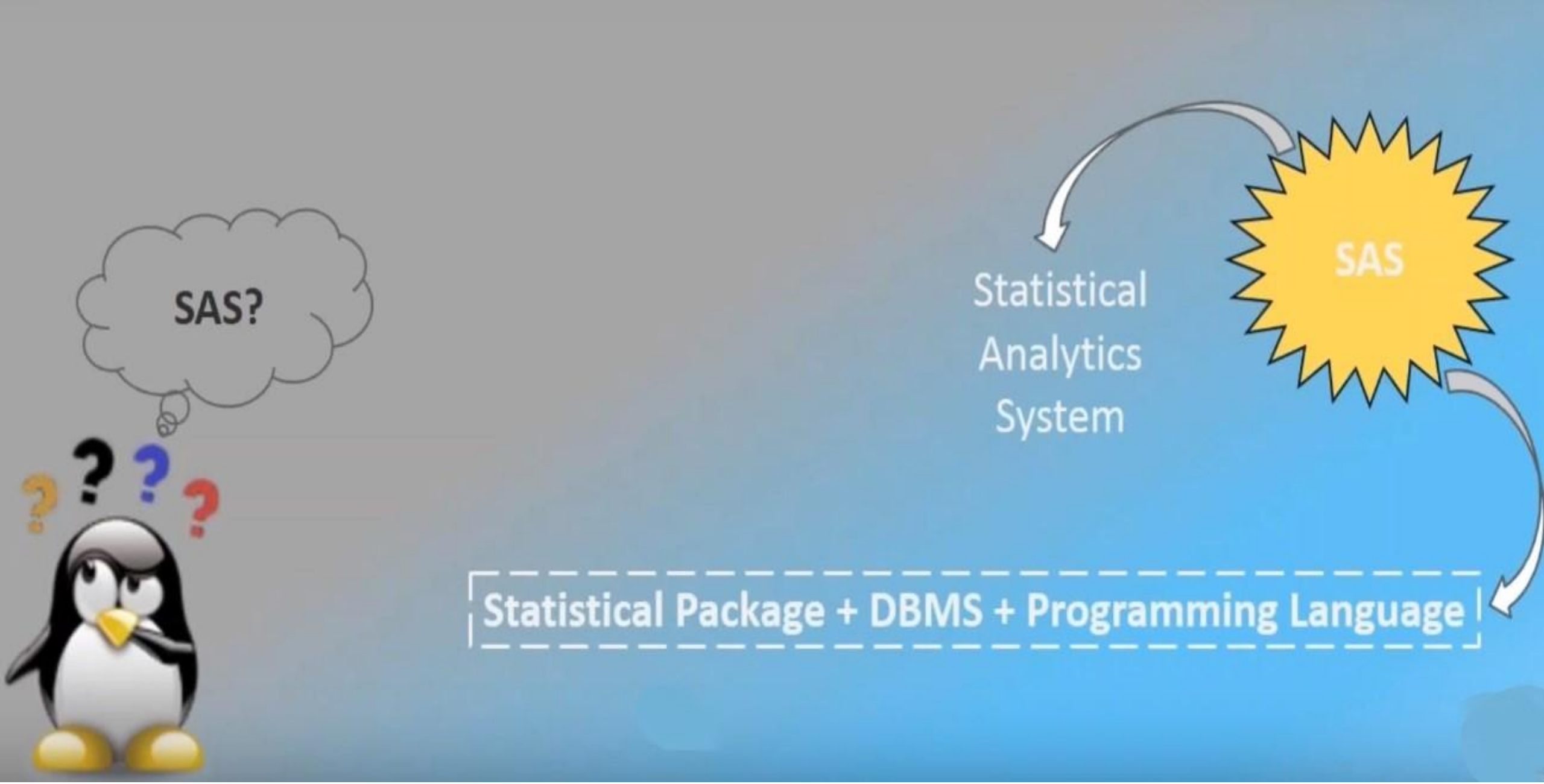
SAS ("Statistical Analysis System") is a software suite developed by SAS Institute for advanced analytics, multivariate analyses, business intelligence, data management, and predictive analytics.

Here is a brief description about the 3 ecosystems:

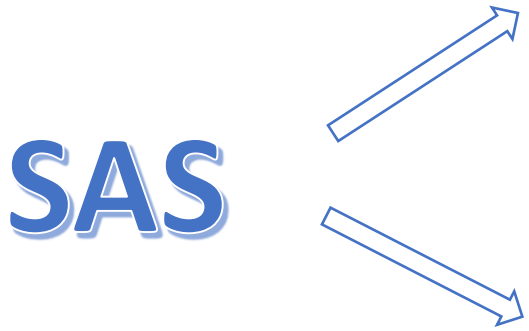
SAS: SAS has been the undisputed market leader in commercial analytics space. The software offers huge array of statistical functions, has good GUI (Enterprise Guide & Miner) for people to learn quickly and provides awesome technical support. However, it ends up being the most expensive option and currently being updated with latest trends like SAS Viya(R &Python).

R: R is the Open source counterpart of SAS, which has traditionally been used in academics and research. Because of its open source nature, latest techniques get released quickly. There is a lot of documentation available over the internet and it is a very cost-effective option.

Python: With origination as an open source scripting language, Python usage has grown over time. Today, it sports libraries (numpy, scipy and matplotlib) and functions for almost any statistical operation / model building you may want to do. Since introduction of pandas, it has become very strong in operations on structured data.



Presented By : Shashi Kumar
YouTube Channel : https://lnkd.in/fNSUTDE



Program

BASE SAS

Advance SAS

Graphical User Interface (GUI)

Data Integration (DI)

Visual Analytics (VA)

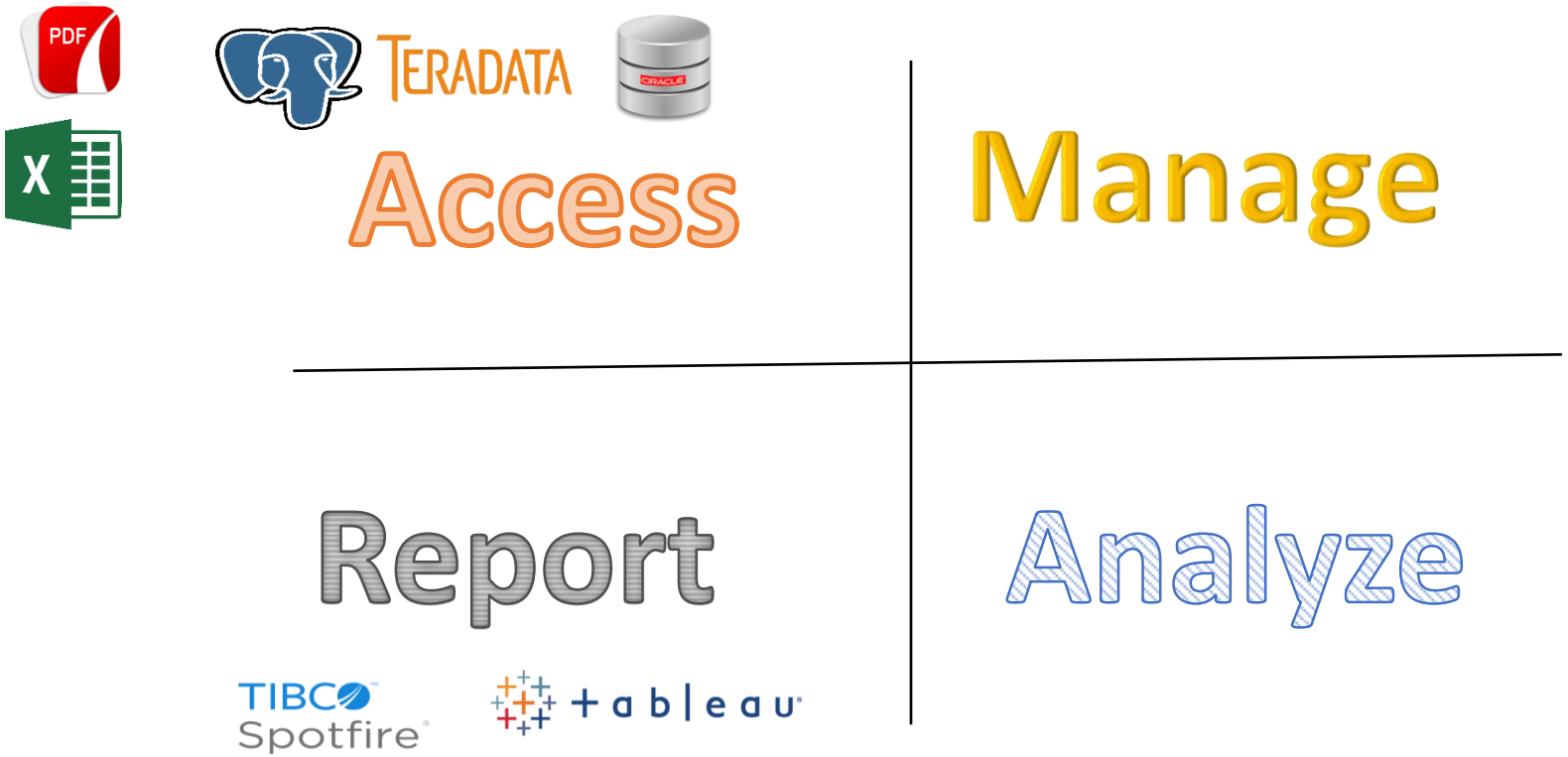
Enterprise guide(EG)

Customer Intelligence (CI)

Foundation Of SAS

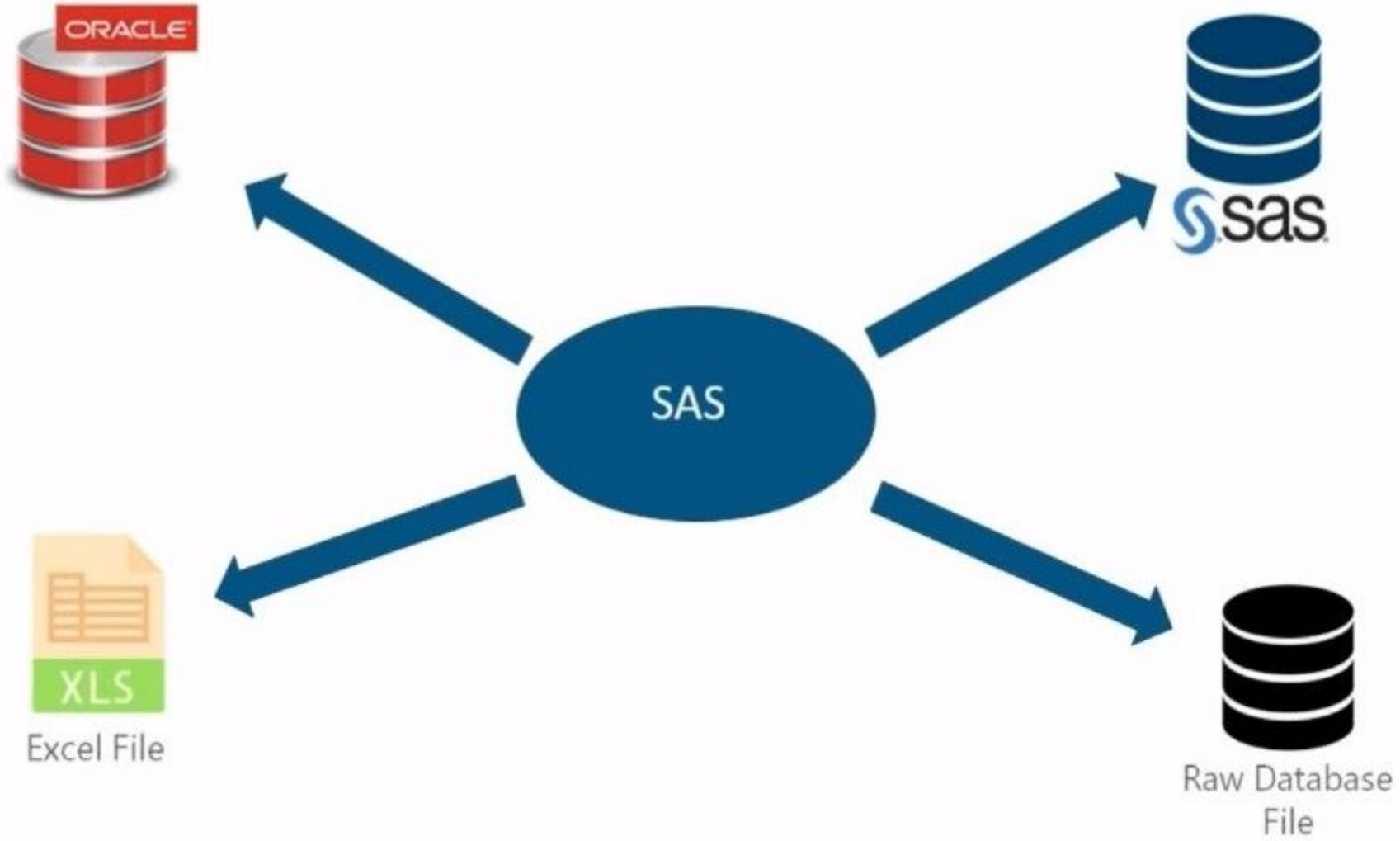
SAS :-

It is a highly flexible and integrated software environment that is used to access, manipulate, manage, analyze and report of data.



Presented By : Shashi Kumar
YouTube Channel : https://lnkd.in/fNSUTDE

1. Access
2. Manage
3. Analyze
4. Present / *Report*



1 Access

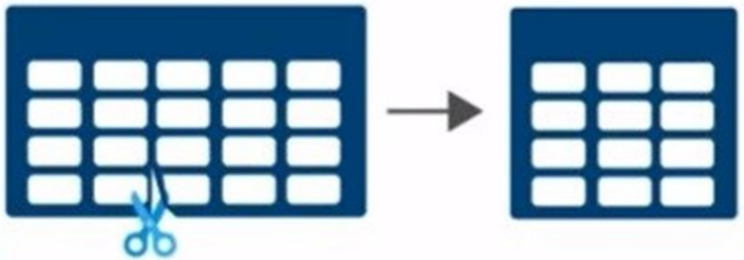
2 **Manage**

3 Analyze

4 Present

SAS gives you excellent **data management** capabilities

1)Subset Data



2)Create Variables

3)Clean & Validate Data

1. Access
2. Manage
3. **Analyze**
4. Present

After Data Management the next step is **data analysis :**

Frequency or Mean calculation

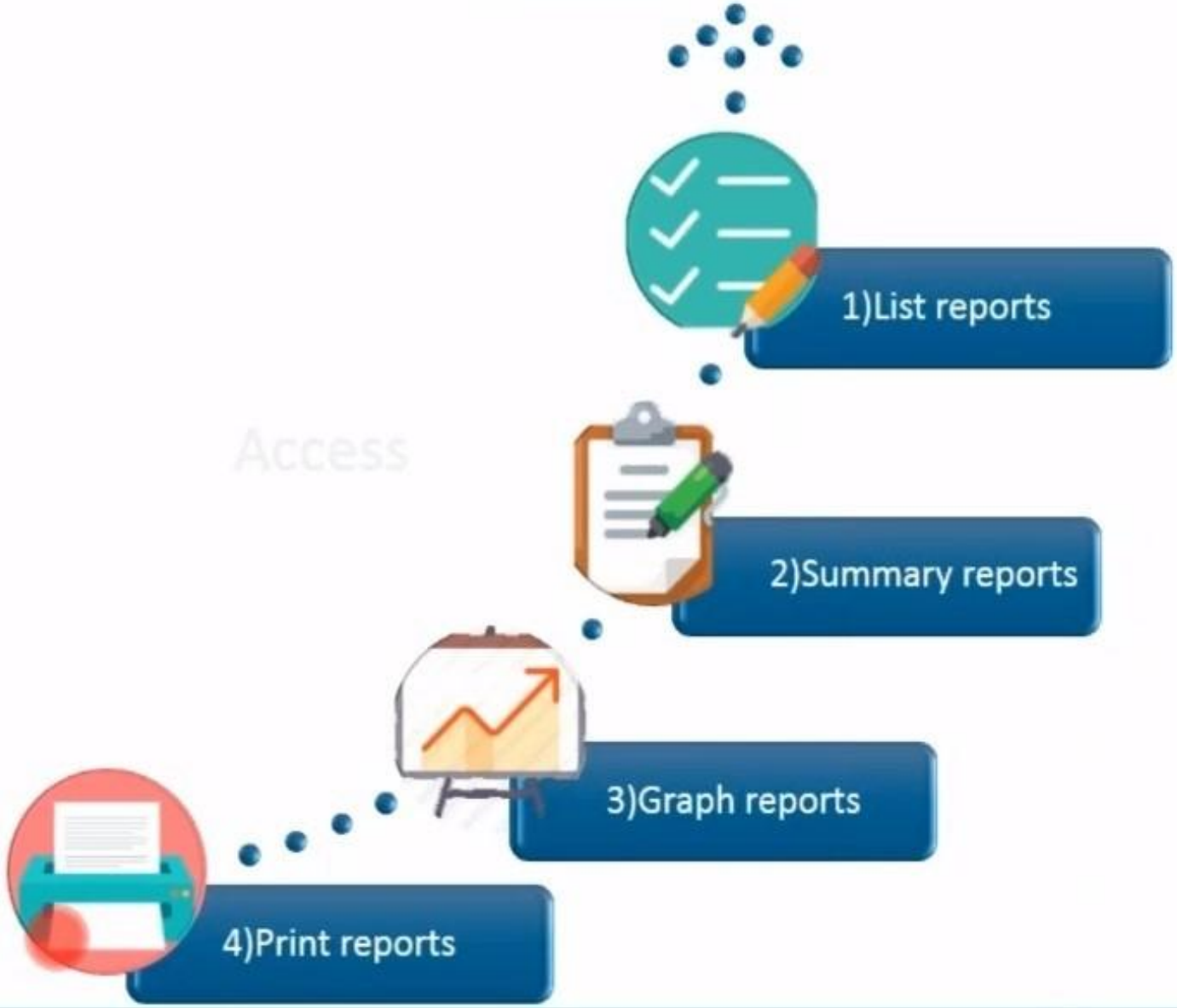
Regression and Forecasting

SAS is the gold standard for statistical analysis.

Once you have analyzed data you can present it better with SAS

1. Access
2. Manage
3. Analyze
4. **Present**

/ Report



SAS Applications

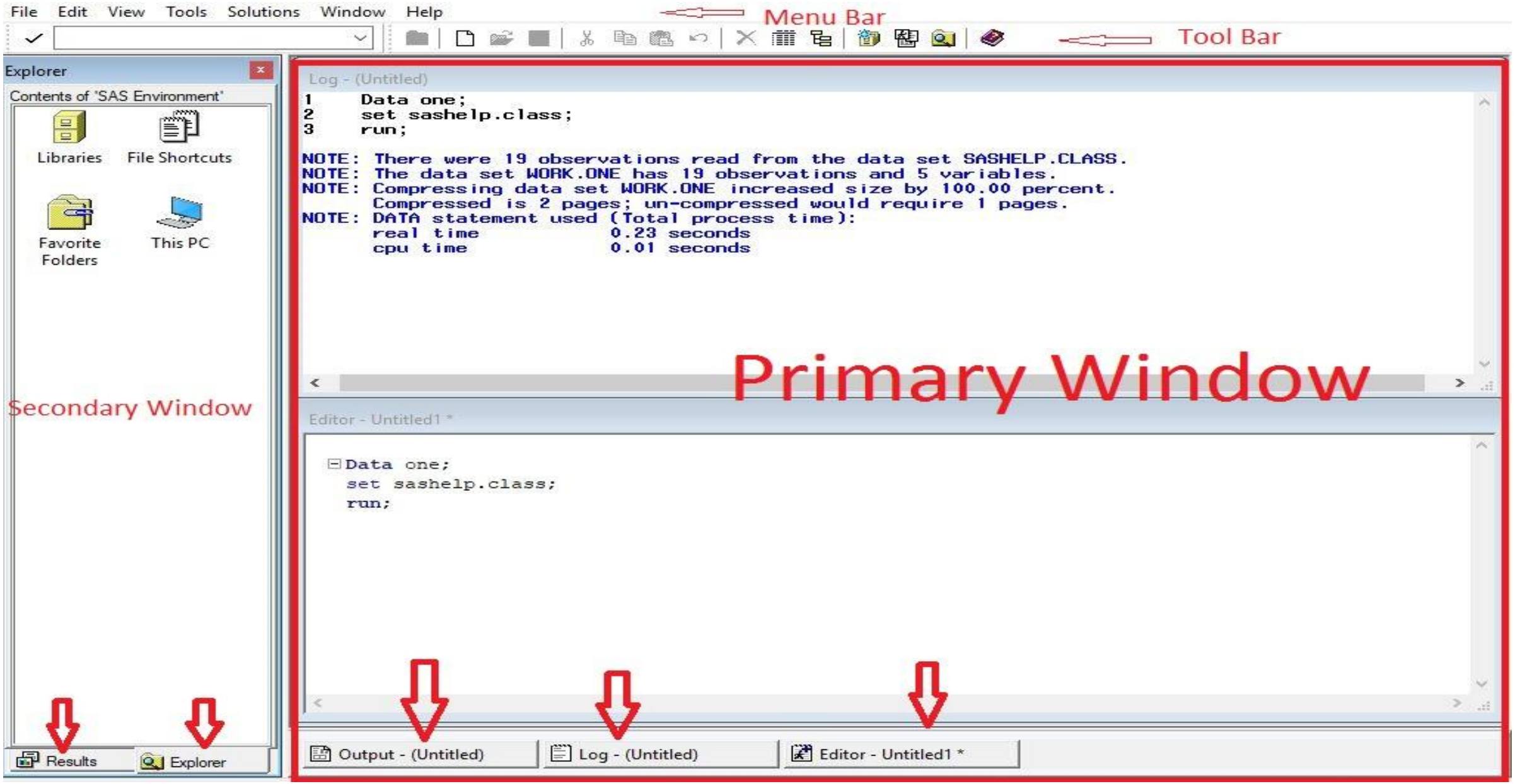
1)Stock Prediction

2)Create Safe Drugs

3)Fight Fraud

4)Optimize Workflow

SAS Session



Presented By : Shashi Kumar
YouTube Channel : https://lnkd.in/fNSUTDE

SAS Session

Primary Window

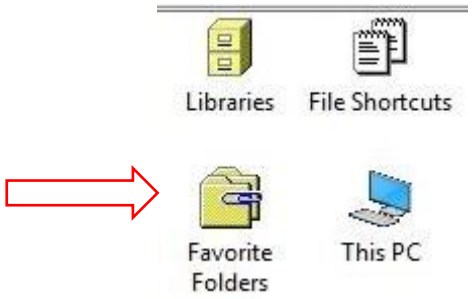
1. Outputs: Contain the report of procedure that have submitted and executed
2. Log: Provide Information about SAS program execution.

   a. Note : Blue color
   Numbers of observation
   Data set names
   b. Warning: Green Color
   Execution Continue
   c. Error: Red Color
   Depend on error it will stop or continue the execution

3. Editor: The Place where SAS program is written, edited, submitted the program for execution.

Secondary Window

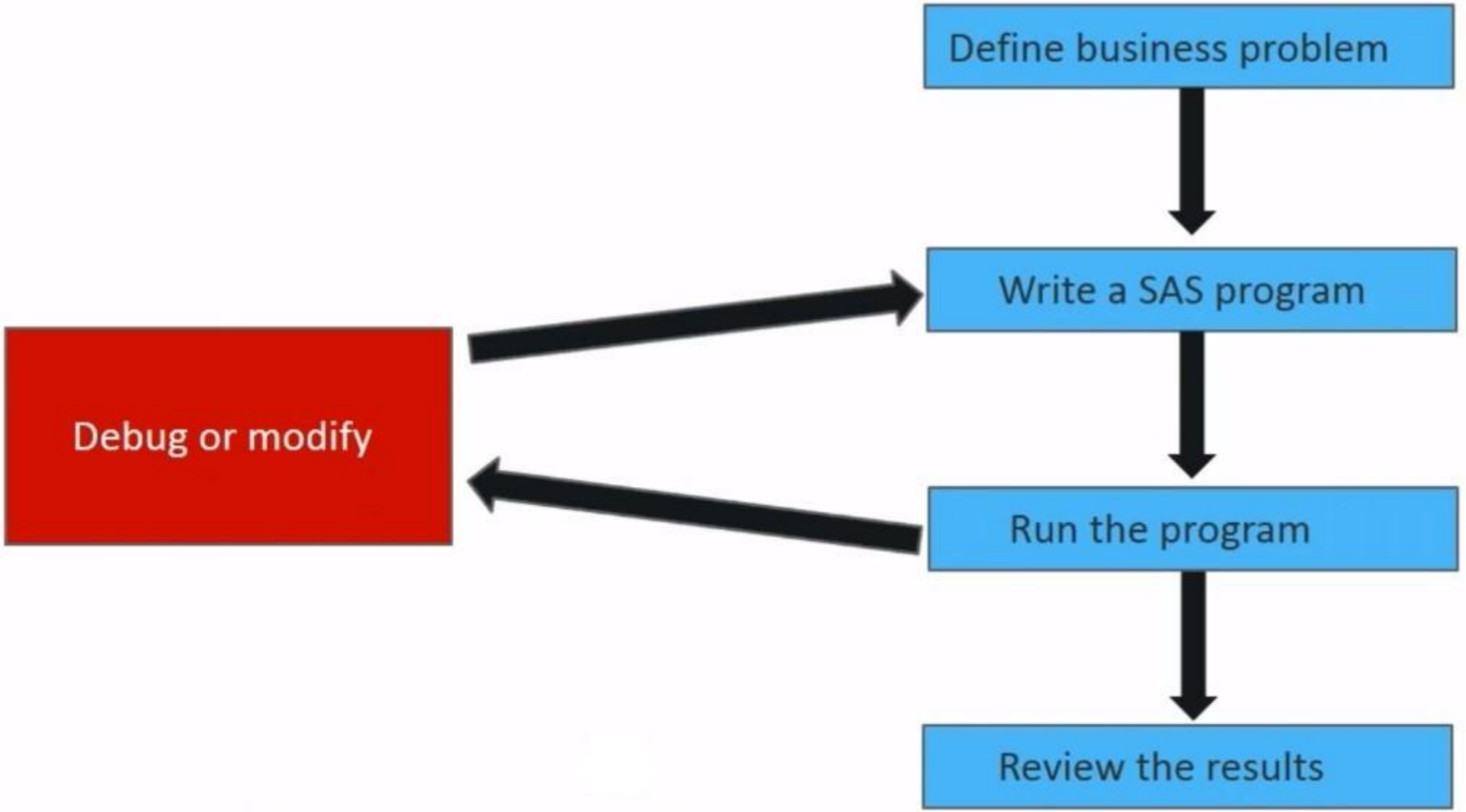
1. Result : It contains list of procedure which are submitted and executed successfully.
2. Explorer : Provide easy navigation to SAS library icon ,window system, my computer etc.



SAS Programming Language

Presented By : Shashi Kumar
YouTube Channel : https://lnkd.in/fNSUTDE

SAS Programming Process



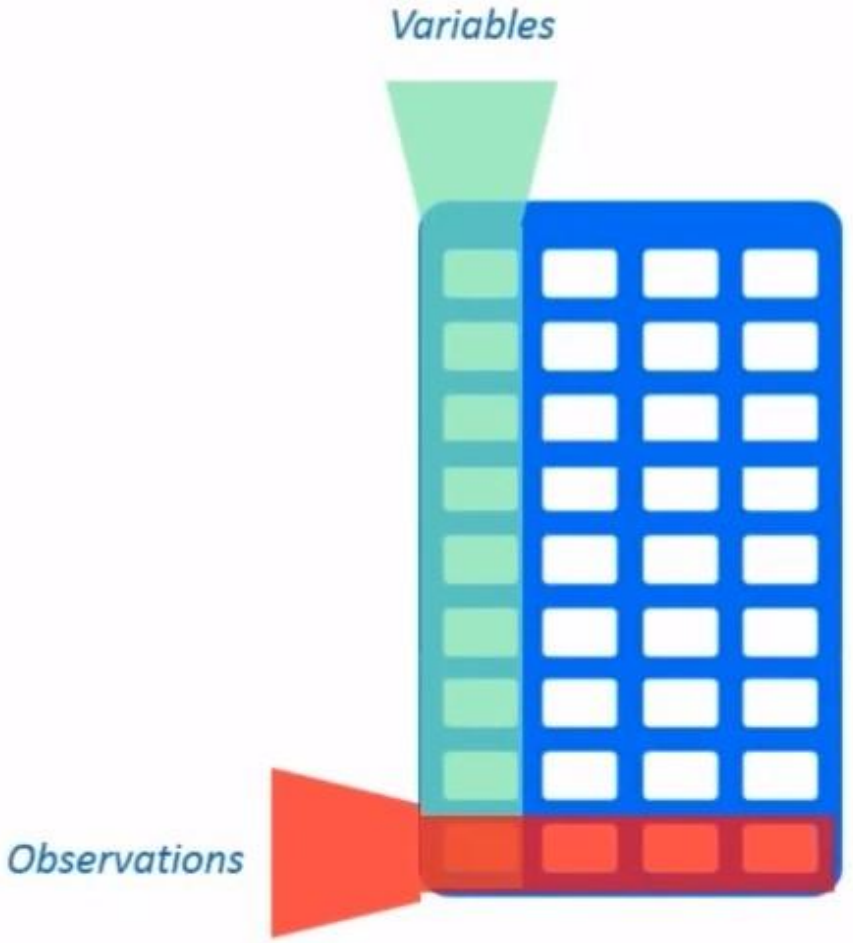
SAS Data

Data is central to every data set.

- In SAS Data is in tabular form
- Variables occupy the columns
- Observations occupy the rows

Data types:

- Numeric
- Characters



SAS Program Structure

SAS programming is based on two building blocks:

1)DATA Steps

DATA steps create or modify SAS data sets. Using DATA steps you can:

- Add data to a data set
- Compute values of variables
- Create new data sets (by sub-setting, merging)

SAS Program Structure

SAS programming is based on two building blocks:

2)PROC Steps

PROC steps analyse and process SAS data sets. Using PROC steps you can:

- Print a report
- Produce descriptive analysis
- Create a tabular report
- Produce plots and charts

Topic: Data Step , Boundary, Statement

```
*****************************************
/* What is Statement  */


/* Statement beigns with reserved keyword and ends with ; */;

data a;  * this is also a statement ;
set zzz; * this is also  a stement ;
run;     * this is also a statemnent;

*************************************;
```

```
/*Multiple SAS statements is written in single row */
Data one;set sashelp.class;run;
/*A single SAS statement can be written in multiple row*/
Data
one;
set
sashelp.class;run;
```

The **DATA step** consists of a group of **SAS** statements that begins with a DATA statement. The DATA statement begins the process of building a **SAS** data set and names the data set. The statements that make up the **DATA step** are compiled, and the syntax is checked.

1. **SAS Statement** : A SAS Statement begin with SAS identifying Keyword and end with semi column (**;**).
   Properties:-
   - 1.1 A Single SAS statement can be written in multiple row.
   - 1.2 Multiple SAS statement is written in single row.
   - 1.3 One or more blank separated the word.
2. **Step Boundary** : Program ends with SAS identifying Keywords e.g. **run**, **quit** and **begin of new SAS program.**
3. **SAS Step** : It is a combination of SAS statements ,the SAS step begin or start with identifying keyword i.e. **data** or **proc** and end with step boundary.

Topic: Data Step , Boundary, Statement

```
/* Sample Code */
data a;
set sashelp.class;
run;

*******************************************;

/*Data step  begins with a  keyword "data" and this step ends with either of 3 ways (run; or beginning of new pgm or quit; )   */
 /* First method  */

 data a;
 set sashelp.class;
 run;


 ******************************;

 /* Second Method; */

 data b;
 set sashelp.class; *this pgm works without run statement because after this there is a new beginning of pgm;

 data c;
 set sashelp.class;
 run;

 /* Third Method  */

 /* Quit statement is used with procs only */

 proc print data=a;
 quit;
```

Presented By : Shashi Kumar
YouTube Channel : https://lnkd.in/fNSUTDE

Exercise :

```
*%%%%%%%%%% Create SAS dataset using Set statement %%%%%;
Data  one;
Set  sashelp.class;
Run;

*%%%%%%%%%% Create SAS dataset using datalines %%%%%;
Data two;
Infile datalines dlm=" ";
Input ID NAMES $;
Datalines;
1 Shashi
2 Ravi
3 Mohan
;
run;
```

Thank You

Presented By : Shashi Kumar