

```

/* BEFORE STARTING WITH SAS */

/* ABCDE */

/* TO COMMENT - CTRL + ? */
/* TO UNCOMMENT - CTRL + ? OR CTRL + SHIFT + ? */

/* SAS- STATISTICAL ANALYSIS SYSTEM */

/* SAS COMPONENTS */
/* 1. DBMS PROGRAMMING */
/* 2. BUSINESS INTELLIGENCE AND DATA WAREHOUSING */
/* 3. VISUAL ANALYTICS - SAS VIYA */
/* 4. SAS DATA MINING - ML, DL, AI */

/* SAS JARGONS */

/* 1. IN SQL - WE HAVE A DATABASE. IN SAS WE HAVE A LIBRARY */
/* 2. IN SQL - WE HAVE A TABLE. IN SAS, WE CALL DATASET */
/* 3. IN SQL - WE CALL COLUMNS. IN SAS WE CALL VARIABLES */
/* 4. IN SQL - WE HAVE A ROW. IN SAS, WE CALL AS OBSERVATION */

/* LETS DO SAS PROGRAMMING */

```

---

```

DATA STU_ENQ;
INPUT STU_ID $ STU_NAME $ AGE GENDER $ EDUCATION $ OCCUPATION $ DATE_OF_ENQ COURSE $;
INFORMAT DATE_OF_ENQ MMDDYY8.;
FORMAT DATE_OF_ENQ MMDDYY8.;
CARDS;
S1 DEEPTI 30 FEMALE BTECH IT 01-02-21 MPIDS
S2 SUPRABHA 29 FEMALE BTECH IT 02-02-21 MPIDS
S3 SASMITA 29 FEMALE BTECH IT 03-02-21 MPIDS
S4 SOUVIK 34 MALE BTECH IT 04-02-21 ADA
S5 SUMESH 31 MALE BTECH IT 05-02-21 ADA
S6 PRATYUSH 31 MALE BTECH IT 06-02-21 ADA
S7 PRADEEBA 27 FEMALE BTECH IT 07-02-21 ADA
S8 MANISH 28 MALE MCA IT 08-02-21 DAS
S9 PRITHVI 29 MALE MCA IT 09-02-21 DAS
S10 ARPITA 31 FEMALE MCA IT 10-02-21 DAS
;
RUN;

/* HAPPY BDAY TO SAS DATE - 1ST JAN 1960 */

```

---

```

DATA STU_ENQ;
/* DATA STATEMENT ALLOWS YOU TO CREATE DATASET */
INPUT STU_ID $ STU_NAME $ AGE GENDER $ EDUCATION $ OCCUPATION $ DATE_OF_ENQ COURSE $;
/* INPUT STATEMENT TO CREATE VARIABLE INPUT. IT ALLOWS YOU TO DECLARE THE DATA TYPE */
INFORMAT DATE_OF_ENQ MMDDYY8.;
/* INFORMAT IS USED TO READ DATA IN REQUIRED SAS FORMAT */
FORMAT DATE_OF_ENQ MMDDYY8.;
/* FORMAT IS USED TO WRITE DATA IN REQUIRED SAS FORMAT */
CARDS;
S1 DEEPTI 30 FEMALE BTECH IT 01-02-21 MPIDS
S2 SUPRABHA 29 FEMALE BTECH IT 02-02-21 MPIDS
S3 SASMITA 29 FEMALE BTECH IT 03-02-21 MPIDS
S4 SOUVIK 34 MALE BTECH IT 04-02-21 ADA
S5 SUMESH 31 MALE BTECH IT 05-02-21 ADA
S6 PRATYUSH 31 MALE BTECH IT 06-02-21 ADA
S7 PRADEEBA 27 FEMALE BTECH IT 07-02-21 ADA
S8 MANISH 28 MALE MCA IT 08-02-21 DAS
S9 PRITHVI 29 MALE MCA IT 09-02-21 DAS
S10 ARPITA 31 FEMALE MCA IT 10-02-21 DAS
;
/* CARDS IS USED FOR INPUT OF TEMPORARY DATA */
RUN;
/* RUN STATEMENT IS USED IN PROGRAMME EXECUTION */

/* TO PRINT A DATASET */

```

---

```

PROC PRINT DATA=STU_ENQ;

```

```
RUN;
```

```
/* WE HAVE USED A FEW SAS WINDOWS */
```

```
/* 1. CODE - WRITING SAS CODES */
```

```
/* 2. LOG - ALLOWS YOU TO SEE ERRORS, WARNING, NOTES, OR SUCCESS MESSAGES */
```

```
/* 3. RESULTS - TO PRINT THE OUTPUT OF A DATASET */
```

```
/* 4. OUTPUT - IT IS USED TO SHOW THE DATASET. IT IS NOT ALWAYS AVAILABLE IN SAS WINDOW */
```

```
/* 5. EXPLORER - IT IS USED TO NAVIGATE BETWEEN YOUR FILES AND DATASET */
```

```
/* 1. SAS DATA STEP - */
```

```
DATA STU_ENQ;
```

```
INPUT STU_ID $ STU_NAME $ AGE GENDER $ EDUCATION $ OCCUPATION $ DATE_OF_ENQ COURSE $;
```

```
INFORMAT DATE_OF_ENQ MMDDYY8.;
```

```
FORMAT DATE_OF_ENQ MMDDYY8.;
```

```
CARDS;
```

```
S1 DEEPTI 30 FEMALE BTECH IT 01-02-21 MPIDS
```

```
S2 SUPRABHA 29 FEMALE BTECH IT 02-02-21 MPIDS
```

```
S3 SASMITA 29 FEMALE BTECH IT 03-02-21 MPIDS
```

```
S4 SOUVIK 34 MALE BTECH IT 04-02-21 ADA
```

```
S5 SUMESH 31 MALE BTECH IT 05-02-21 ADA
```

```
S6 PRATYUSH 31 MALE BTECH IT 06-02-21 ADA
```

```
S7 PRADEEBA 27 FEMALE BTECH IT 07-02-21 ADA
```

```
S8 MANISH 28 MALE MCA IT 08-02-21 DAS
```

```
S9 PRITHVI 29 MALE MCA IT 09-02-21 DAS
```

```
S10 ARPITA 31 FEMALE MCA IT 10-02-21 DAS
```

```
;
```

```
RUN;
```

```
/* 2. PROC STEP */
```

```
PROC PRINT DATA=STU_ENQ;
```

```
RUN;
```

```
/* LETS CREATE A DATASET FROM EXTERNAL FIXED FORMAT DATA FILE */
```

```
DATA STAFF;
```

```
INFILE '/home/debendra330/BATCH 202404/SESSION_1/A1.RAW_DATA/a1.EMPLIST FM.txt';
```

```
/* INFILE STATEMENT IS USED TO GET EXTERNAL FILE PATH TO READ DATA */
```

```
INPUT LAST_NAME $ 1-20 FIRST_NAME $ 21-30 EMPLOYEE_CODE $ 31-35 JOBTITLE $ 36-43 DOJ SALARY 54-59;
```

```
INFORMAT DOJ MMDDYY8.;
```

```
FORMAT DOJ MMDDYY8.;
```

```
RUN;
```

```
PROC PRINT DATA=STAFF;
```

```
RUN;
```

```
/* LIBRARY - */
```

```
LIBNAME LIBREF <PATH>;
```

```
RUN;
```

```
LIBNAME SSN1 '/home/debendra330/BATCH 202404/SESSION_1/A2.SAS_DATASET';
```

```
RUN;
```

```
DATA SSN1.STAFF;
```

```
INFILE '/home/debendra330/BATCH 202404/SESSION_1/A1.RAW_DATA/a1.EMPLIST FM.txt';
```

```
/* INFILE STATEMENT IS USED TO GET EXTERNAL FILE PATH TO READ DATA */
```

```
INPUT LAST_NAME $ 1-20 FIRST_NAME $ 21-30 EMPLOYEE_CODE $ 31-35 JOBTITLE $ 36-43 DOJ SALARY 54-59;
```

```
INFORMAT DOJ MMDDYY8.;
```

```
FORMAT DOJ MMDDYY8.;
```

```
RUN;
```

```
PROC PRINT DATA=SSN1.STAFF;
```

```
RUN;
```

```
/* SAS7BDAT - IT IS THE EXTENSION OF DATASET */
```

```
/* .SAS - IS THE EXTENSION FOR SAS CIODES */
```

```
/* LIBRARY NAME CRITERIA */
```

```
=====
```

```
/* 1. LIBNAME SHOULD NOT BE MORE THAN 8 CHARACTERS */
/* 2. IT SHOULD START WITH AN ALPHABET OR UNDERSCORE */
/* 3. IT MUST BE FOLLOWED WITH ALPHABETS, NUMBERS AND UNDERSCORE */

/* TEST@1 --- NO */
/* SSN 1 --- NO */
/* SESSION_1 --- NO */
/* _SSN1 --- YES */
/* 1_SSN1 --- NO */

/* WHILE CREATING A DATASET OR VARIABLE NAME, DO WE NEED TO FOLLOW ANY CRITERIA */

/* YES, WE NEED TO */

/* CRITERIA FOR DATASET AND VARIABLE NAME */
=====

/* 1. DATASET OR VARIABLE NAME SHOULD NOT BE MORE THAN 32 CHARACTERS */
/* 2. IT SHOULD START WITH AN ALPHABET OR UNDERSCORE */
/* 3. IT MUST BE FOLLOWED WITH ALPHABETS, NUMBERS AND UNDERSCORE */

/* STU_ENQ_BANGALORE - YES */
/* CUSTOMER ID - NO */

/* VARIABLES AND DATA TYPES */
=====

/* 1. CHARACTER - */
/*          DEFAULT LENGHT - 8 */
/*          MINIMUM LENGHT - 1 */
/*          MAXIMUM - 32,767 */

/* 2. NUMERIC - */
/*          DEFAULT - 8 BYTES WHICH CAN STORE UPTO 16-17 DIGITS */
```

---

```
DATA STU_DETAILS;
INPUT STU_NAME $ AGE EDUCATION $ COURSE $;
CARDS;
ABHISEKH 25 BTECH MPIDS
PRIYADARSHINI 26 BTECH ADA
AYASHKANT 26 MBA MPIDS
;
RUN;
```

---

```
PROC PRINT DATA=STU_DETAILS;
RUN;
```

---

```
DATA STU_DETAILS;
INPUT STU_NAME : $20. AGE EDUCATION $ COURSE $;
CARDS;
ABHISEKH 25 BTECH MPIDS
PRIYADARSHINI 26 BTECH ADA
AYASHKANT 26 MBA MPIDS
;
RUN;
```

---

```
PROC PRINT DATA=STU_DETAILS;
RUN;
```

---

```
/* IMPORTING AND EXPORTING DATA */
=====

/* IMPORTING DATA FROM TEXT FILE */
OPTIONS VALIDVARNAME = V7;
/* TO READ DATA WITH VALID SAS NAME CRITERIA */
```

---

```
PROC IMPORT OUT=SSN1.MED_2024_TXT
            DATAFILE='/home/debendra330/BATCH 202404/SESSION_1/A1.RAW_DATA/MED_2023.txt'
            DBMS=DLM REPLACE;
            DELIMITER='|';

RUN;
```

---

```
PROC PRINT DATA=SSN1.MED_2024_TXT;
```

```
RUN;

/* EXPORTING DATA TO A TEXT FILE */

PROC EXPORT DATA=SSN1.MED_2024_TXT
    OUTFILE='/home/debendra330/BATCH 202404/SESSION_1/A4.SAS_OUTPUT/MED_2024_TXT_EXPORT.txt'
    DBMS = DLM REPLACE;
    DELIMITER = '/';
RUN;

/* IMPORTING DATA FROM CSV FILE */

OPTIONS VALIDVARNAME = V7;

PROC IMPORT OUT = SSN1.MED_2024_CSV
    DATAFILE = '/home/debendra330/BATCH 202404/SESSION_1/A1.RAW_DATA/MED_2023.csv'
    DBMS = CSV REPLACE;
RUN;

/* EXPORTING DATASET TO A CSV */

PROC EXPORT DATA=SSN1.MED_2024_CSV
    OUTFILE='/home/debendra330/BATCH 202404/SESSION_1/A4.SAS_OUTPUT/MED_2024_CSV_EXPORT.csv'
    DBMS=CSV REPLACE;
RUN;

/* IMPORTING DATA FROM EXCEL */

OPTIONS VALIDVARNAME=V7;

PROC IMPORT OUT=SSN1.MED_2024_EXCEL
    DATAFILE='/home/debendra330/BATCH 202404/SESSION_1/A1.RAW_DATA/MED_2023.xlsx'
    DBMS = XLSX REPLACE;
    SHEET='MED_AUSTRALIA';
RUN;

/* EXPORTING DATA TO EXCEL */

PROC EXPORT DATA=SSN1.MED_2024_EXCEL
    OUTFILE = '/home/debendra330/BATCH 202404/SESSION_1/A4.SAS_OUTPUT/MED_2024_EXCEL_EXPORT.xlsx'
    DBMS= XLSX REPLACE;
    SHEET='MED_2024_AUS';
RUN;

/* HOW TO DESCRIBE A DATASET AND LIBRARY */
=====

PROC PRINT DATA=SSN1.STAFF;
RUN;

PROC CONTENTS DATA=SSN1.STAFF;
RUN;

/* WE USE PROC CONTENTS TO DESCRIBE OUR DATASETS INTO 3 PARTS */
/* 1. DATASET INFORMATION */
/* 2. HOST ENGINE INFORMATION */
/* 3. VARIABLES, DATATYPES AND FORMATS */

PROC CONTENTS DATA=SSN1.STAFF VARNUM SHORT;
RUN;

/* LAST_NAME FIRST_NAME EMPLOYEE_CODE JOBTITLE DOJ SALARY */

PROC CONTENTS DATA=SSN1.STAFF VARNUM;
RUN;

/* VARNUM - IT IS USED TO GET THE VARIABLES IN THE SAME ORDER AS DATASET */
/* SHORT - IT IS USED TO ONLY GET VARIABLE LIST */

/* HOW TO DESCRIBE THE ENTIRE LIBRARY TOGETHER */
=====
```

---

```
PROC CONTENTS DATA=SSN1._ALL_  
RUN;
```

---

```
PROC CONTENTS DATA=SSN1._ALL_ NODS;  
RUN;
```

```
/* WE USE NODS AS AN OPTION TO AVOID DESCRIBING EACH DATASET IN LIBRARY */
```

```
/* HOW TO CREATE A NEW DATASET FROM EXISTING DATASET */
```

```
=====
```

---

```
DATA UPDATED_STAFF;  
SET SSN1.STAFF;  
RUN;
```

```
/* WE USE SET STATEMENT TO READ DATA FROM ONE DATASET TO ANOTHER */
```