```
LIBNAME SSN6 '/home/debendra330/BATCH_202404/SESSION_6/A3.SAS_DATASET';
RUN;

/* SAS FUNCTIONS */

/* 1. CHARACTER FUNCTIONS */
/* 2. NUMERIC FUNCTIONS */
/* 3. DATETIME FUNCTIONS */

/* 1. CHARACTER FUNCTIONS -  */
==========================================

DATA BATCH_202301;
INPUT STU_NAME : $20. GENDER $ AGE COURSE $;
CARDS;
ADITYA  MALE 23 MPIDS
aman MALE 23 MPIDS
Akash MALE 25 MPIDS
ankita  FEMALE 23 ADA
PadmaSHREE FEMALE 24 MPIDS
;
RUN;

PROC PRINT DATA=BATCH_202301;
RUN;

DATA BATCH_202301;
SET BATCH_202301;
NAME_CAPITAL = UPCASE(STU_NAME);
NAME_LOWER = LOWCASE(STU_NAME);
NAME_PROPER = PROPCASE(STU_NAME);
RUN;

PROC PRINT;
RUN;

/* SUBSTR */

DATA X;
NAME = 'AYASHKANT';
PULL1 = SUBSTR(NAME, 1, 5);
PULL2 = SUBSTR(NAME, 2, 4);
RUN;

PROC PRINT;
RUN;

DATA GOOGLE_COMMENTS;
INFILE CARDS TRUNCOVER;
INPUT COMMENTS $200.;
CARDS;
IS BAR MODI SARKAR
THIS TIME UPA MAY LEAD THE POSITIONS
MODIJI IS VERY POWERFUL IMPACT OVER THE YOUTH OF INDIA
NDA WILL DEFEAT OTHER PARTIES WITH MODIJI'S PRESENCE
RAHUL GANDHI DID A MISTAKE BY MODIFYING THE HIERARCHY OF CONGRESS
;
RUN;

PROC PRINT DATA=GOOGLE_COMMENTS;
RUN;

/* INDEX, INDEXC, INDEXW */

DATA GOOGLE_COMMENTS;
SET GOOGLE_COMMENTS;
FIND_MODI = INDEX(COMMENTS, 'MODI'); /*SEARCH THE WORD, AND IGNORES ANY OTHER EXTRA CHARACTERS*/
FIND_MODI2 = INDEXW(COMMENTS, 'MODI'); /*SEARCH PARTICULAR WORD ONLY*/
FIND_MODI3 = INDEXC(COMMENTS, 'MODI'); /*SEARCH FOR A PARTICULAR CHARACTER WITHIN THE WORD THAT YOU HAVE PASSED*/
RUN;

PROC PRINT DATA=GOOGLE_COMMENTS;
RUN;

DATA STU_NAME;
INPUT NAME $40.;
CARDS;
```

```
PADMASHREE KULKARNI
ADITYA SWAIN
AKASH SAHOO
AMAN RAMNANI
ANULIKA ATRE
ANKITA DAS
ABDUL SAHID
LESLIN PRADHAN
ILISAYA PRADHAN
PAYAL UDHWANI
;
RUN;

PROC PRINT DATA=STU_NAME;
RUN;

DATA STU_NAME;
SET STU_NAME;
FIRST_NAME = SUBSTR(NAME, 1, INDEX(NAME, " ")-1);
LAST_NAME = SUBSTR(NAME, INDEX(NAME, " ")+1, 20);
RUN;

PROC PRINT;
RUN;

DATA STU_NAME;
INPUT NAME $40.;
CARDS;
PADMASHREE_KULKARNI
ADITYA_SWAIN
AKASH_SAHOO
AMAN_RAMNANI
ANULIKA_ATRE
ANKITA_DAS
ABDUL_SAHID
LESLIN_PRADHAN
ILISAYA_PRADHAN
PAYAL_UDHWANI
;
RUN;

PROC PRINT DATA=STU_NAME;
RUN;

DATA STU_NAME;
SET STU_NAME;
FIRST_NAME = SUBSTR(NAME, 1, INDEX(NAME, "_")-1);
LAST_NAME = SUBSTR(NAME, INDEX(NAME, "_")+1, 20);
RUN;

PROC PRINT;
RUN;

DATA AGDAM_BAGDAM;
INPUT TEXT : $200.;
CARDS;
ADJGJ%^$^$%35RWU328136792749HDASJUDTU
DIASYR75^$^%#%353543763897329231JKHKHKHiyh
FJUDSIFRYEURTWRYERQWYurte%wrw%$^$&^%**JAWUEW8R9
;
RUN;

PROC PRINT DATA=AGDAM_BAGDAM;
RUN;

/* COMPRESS */

DATA AGDAM_BAGDAM;
SET AGDAM_BAGDAM;
ALL_CHARACTERS = COMPRESS(TEXT,,'KA');
ALL_NUMBERS = COMPRESS(TEXT,, 'KD');
ALL_SPECIAL = COMPRESS(TEXT,,'KP');
RUN;

PROC PRINT;
RUN;

DATA PHONE_NUMBER;
```

```
INPUT PHONE $20.;
PHONE2=COMPRESS(PHONE,'-');
CARDS;
895-110-8662
789-562-5656
902-456-9090
;
RUN;

PROC PRINT;
RUN;

DATA MOBILE_NUMBER;
INPUT MOBILE $30.;
MOBILE2=COMPRESS(MOBILE);
MOBILE3=COMPBL(MOBILE);
CARDS;
895      11     08662
895   11    08663
891    11    08664
;
RUN;

PROC PRINT DATA=MOBILE_NUMBER;
RUN;

/* INTERVIEW QUESTION, DIFFERENCE BETWEEN COMPRESS AND COMPBL */


/* INTERVIEW QUESTION - CAT, CATX, CATS, CATT */

DATA STU_NAME;
INPUT FIRST_NAME :$20. LAST_NAME :$20.;
FULL_NAME1 = CAT(FIRST_NAME, LAST_NAME);
FULL_NAME2 = CATX('_', FIRST_NAME, LAST_NAME);
FULL_NAME3 = CATS(FIRST_NAME, LAST_NAME);
FULL_NAME4 = CATT(FIRST_NAME, LAST_NAME);
CARDS;
ARPIT AGARWAL
AGNIJA MOHANTY
PRAGNYA ROUT
DIPENDAR RAY
;
RUN;


PROC PRINT;
RUN;

/* CAT FUNCTION WILL CONCATENATE BY GIVING A SPACE IN BETWEEN */
/* CATX FUNCTION WILL CONCATENATE WITH A GIVEN DELIMITER */
/* CATS FUNCTION WILL CONCATENATE WITHOUT ANY GIVEN SPACES */
/* CATT FUNCTION WILL CONCATENATE BY TRIMMING ALL POSITIONAL VALUES */

/* TRANWRD */

DATA PLAYER_NAME;
NAME = 'SACHIN ENDULKAR';
FIND_NAME = TRANWRD(NAME, 'ENDULKAR', 'TENDULKAR');
RUN;

PROC PRINT;
RUN;

/* FIND */
DATA PLAYER_NAME;
NAME = 'SACHIN TENDULKAR';
FIND_NAME = FIND(NAME, 'TENDULKAR');
RUN;

PROC PRINT;
RUN;

/* SCAN */

DATA STU_FULL_DETAILS;
DETAIL = 'AYASHKANT,PANDA,27,BTECH,ODISHA,MPIDS';
FIRST_NAME = SCAN(DETAIL, 1, ',');
```

```
LAST_NAME = SCAN(DETAIL, 2, ',');
AGE = SCAN(DETAIL, 3, ',');
QUALIFICATION = SCAN(DETAIL, 4, ',');
STATE = SCAN(DETAIL, 5, ',');
COURSE = SCAN(DETAIL, 6, ',');
RUN;

PROC PRINT;
RUN;

DATA STU_FULL_DETAILS;
DETAIL = 'ANKITA:DAS;24-MSC|ODISHA<MPIDS';
FIRST_NAME = SCAN(DETAIL, 1, ':');
LAST_NAME = SCAN(DETAIL, 2, ':;');
AGE = SCAN(DETAIL, 2, ';-');
QUALIFICATION = SCAN(DETAIL, 2, '-|');
STATE = SCAN(DETAIL, 2, '|<');
COURSE = SCAN(DETAIL, 2, '<');
RUN;

PROC PRINT;
RUN;

/* WE USE SCAN FUNCTION TO RETREIVE DATA BASED ON DELIMITERS */

/* TRANSLATE */

DATA X;
ACCOUNT_NUMBER = 33059237584;
ACCOUNT_MASKED = TRANSLATE(ACCOUNT_NUMBER, 'ZEQWTRYUI', '0123456789');
RUN;

DATA BANK_ACCOUNT_NUMBER;
INPUT ACCOUNT_NUMBER;
MASK=TRANSLATE(ACCOUNT_NUMBER,'XQWOPFKELA','0398162457');
CARDS;
7927092
3735735
7864864
4794797
5897590
;
RUN;

PROC PRINT;
RUN;

/* INPUT AND PUT */

DATA PROD_SALES;
INPUT PROD $ YEAR SALES $;
CARDS;IN
APPLE 2020 90000
DELL 2020 23000
HP 2020 95000
ACER 2020 45000
;
RUN;

PROC PRINT DATA=PROD_SALES;
RUN;

DATA PROD_SALES_V1;
SET PROD_SALES;
NEW_YEAR = PUT(YEAR, $10.);
NEW_SALES = INPUT(SALES, 10.);
RUN;

PROC PRINT;
RUN;

/* INPUT IS USED TO CONVERT CHARACTER TO NUMERIC */
/* PUT IS USED TO CONVERT NUMERIC TO CHARACTER */

/* PNC */
/* ICN */

/* 2. NUMERIC FUNCTIONS */
```

```
==================================


DATA PROD_SALES;
INPUT PROD $ JAN FEB MAR APR MAY JUN;
CARDS;
APPLE 800 400 300 200 500 800
DELL 200 300 800 900 200 500
HP 200 800 900 300 500 200
;
RUN;

PROC PRINT DATA=PROD_SALES;
RUN;


DATA PROD_SALES_V1;
SET PROD_SALES;
TOTAL_SALES1 = JAN + FEB + MAR + APR + MAY + JUN;
TOTAL_SALES2 = SUM(JAN--JUN);
TOTAL_SALES3 = SUM(OF JAN--JUN);
TOTAL_SALES4 = SUM(JAN,FEB,MAR,APR,MAY,JUN);
MEAN_SALES = MEAN(OF JAN--JUN);
MAX_SALES = MAX(OF JAN--JUN);
MIN_SALES = MIN(OF JAN--JUN);
STD_SALES = STD(OF JAN--JUN);
RUN;

PROC PRINT DATA=PROD_SALES_V1;
RUN;

DATA ALL_NUMBER;
X=123.999;
A=ROUND(X);
B=INT(X);
C=CEIL(X);
D=FLOOR(X);
RUN;

PROC PRINT;
RUN;

/* 3. DATETIME */
============================

/* CURRENT DATE AND TIME */

DATA X;
CURRENT_DATE = DATE();
FORMAT CURRENT_DATE MMDDYY10.;
RUN;

PROC PRINT;
RUN;

DATA X;
CURRENT_DATE = TODAY();
FORMAT CURRENT_DATE MMDDYY10.;
RUN;

PROC PRINT;
RUN;

DATA X;
CURRENT_DATE = DATETIME();
FORMAT CURRENT_DATE DATETIME40.;
RUN;

PROC PRINT;
RUN;

/* HOW TO GET MONTH, DAY, YEAR, WEEK */

DATA X;
CURRENT_DATE = TODAY();
FORMAT CURRENT_DATE MMDDYY10.;
DATE = DAY(CURRENT_DATE);
RUN;
```

```
PROC PRINT;
RUN;


DATA X;
CURRENT_DATE = TODAY();
FORMAT CURRENT_DATE MMDDYY10.;
MONTH = MONTH(CURRENT_DATE);
RUN;

PROC PRINT;
RUN;


DATA X;
CURRENT_DATE = TODAY();
FORMAT CURRENT_DATE MMDDYY10.;
YEAR = YEAR(CURRENT_DATE);
RUN;

PROC PRINT;
RUN;


DATA X;
CURRENT_DATE = TODAY();
FORMAT CURRENT_DATE MMDDYY10.;
WEEK = WEEK(CURRENT_DATE);
RUN;

PROC PRINT;
RUN;

/* MDY */

DATA X;
TODAY = DATE();
MONTH = MONTH(TODAY);
DAY = DAY(TODAY);
YEAR = YEAR(TODAY);
NEW_DATE = MDY(MONTH, DAY, YEAR);
FORMAT TODAY DATE9. NEW_DATE MMDDYY10.;
RUN;

PROC PRINT;
RUN;

/* HOW TO GET HARCODED DATE VALUE */

DATA X ;
TODAY = '08JAN2020'D;
FORMAT TODAY MMDDYY10.;
RUN;

PROC PRINT;
RUN;

/* HOW TO GET DATE DIFFERENCE */
========================================

DATA AGE_DIFFERENCE;
DOB = '08JAN1997'D;
YEAR_OLD = INTCK('YEAR', DOB, TODAY());
MONTHS_OLD = INTCK('MONTH', DOB, TODAY());
DAYS_OLD = INTCK('DAY', DOB, TODAY());
QTR_OLD = INTCK('QTR', DOB, TODAY());
WEEKS_OLD = INTCK('WEEK', DOB, TODAY());
RUN;

PROC PRINT;
RUN;

/* HOW TO ADD DAYS, MONTHS, YEARS IN SAS */
================================

DATA ADD;
CURRENT_DATE = DATE();
```

```
ADD_5D = INTNX('DAY', CURRENT_DATE, 5);
ADD_5M = INTNX('MONTH', CURRENT_DATE, 5);
ADD_5M_V2 = INTNX('MONTH', CURRENT_DATE, 5, 'S');
ADD_5Y = INTNX('YEAR', CURRENT_DATE, 5);
ADD_5Y_V2 = INTNX('YEAR', CURRENT_DATE, 5, 'S');
ADD_5Y_V3 = INTNX('YEAR', CURRENT_DATE, 5, 'E');
ADD_5Y_V4 = INTNX('YEAR', CURRENT_DATE, 5, 'B');
FORMAT CURRENT_DATE ADD_5D ADD_5M ADD_5M_V2 ADD_5Y ADD_5Y_V2 ADD_5Y_V3 ADD_5Y_V4 MMDDYY10.;
RUN;

PROC PRINT;
RUN;

/* SIGNIFICANCE  */

/* S - SAME DATE AS TODAY */
/* E - END OF THE TIME PERIOD */
/* B - BEGINNING OF TIME PERIOD */
```