```
/* PROCEDURAL MEANS IN SAS: */



/* PART-1:--> */


/* INTRODUCTION: */
/* THE MEANS PROCEDURE DATA SUMMARIZATION TOOLS TO COMPUTE DESCRIPTIVE STATISTICS FOR VARIABLES ACROSS ALL OBSERVATIONS AN
/* EG:- */
/* PROC MEANS, */
/* CALCULATES DESCRIPTIVE STATISTICS BASED ON MOMENTS. */
/* ESTIMATES QUANTILES, WHICH INCLUDES THE MEDIAN. */
/* CALCULATES CONFIDENCE LIMITS FOR THE MEAN. */
/* IDENTIFIES EXTREME VALUES. */
/* PERFORMS A T TEST. */

/* BY DEFAULT, PROC MEANS DISPLAYS OUTPUT.
    YOU CAN ALSO USE THE OUTPUT STATEMENT TO STORE THE STASTISTICS IN A SAS DATASET. */




DATA TEST;
SET SASHELP.BASEBALL;
RUN;

/* IT GIVE ALL INFORMATION ABOUT A TABLE. */
PROC CONTENTS DATA=WORK.TEST;
RUN;



/* a. PROC MEANS FOR DEFAULT OR BASIC INFO : */

PROC MEANS DATA=WORK.TEST;
RUN;




/* PART-2:---> */


/* b. PERFORMING ANALYSIS ON SELECTED VARIABLE:  */

/* IT GIVE INFORMATION ABOUT ONLY ONE VARIABLE THAT IS 'SALARY'. */
PROC MEANS DATA=WORK.TEST;
VAR SALARY;
RUN;

/* IT GIVE INFORMATION ABOUT MORE THAN ONE VARIABLE THAT IS 'SALARY' & 'LOGSALARY'. */
PROC MEANS DATA=WORK.TEST;
VAR SALARY logSalary;
RUN;


/*  IT GIVE THE COLUMN WITH MISSING VALUES AND COUNT OF THESE VARIABLE. */
PROC MEANS DATA=WORK.TEST N NMISS;
RUN;




/* PART-3:---> */


/* c. PERFORMING ONLY SPECIFIC/REQUIRED STATISTICS:  */

PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN;
VAR SALARY logSalary;
RUN;



/* PART-4:--> */
```

```
/* d. LIMITING DECIMAL PLACES FOR STATISTICS:  */


/* 0 DECIMAL PLACES: */
PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=0;
VAR SALARY logSalary;
RUN;


/* 2 DECIMAL PLACES: */
PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY logSalary;
RUN;




/* PART-5:---> */

/* e. STATISTICS BY ONE OR MULTIPLE GROUP OR CATEGORY:  */

PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY logSalary;
CLASS TEAM LEAGUE;
RUN;




/* PART-7:---> */


/* f. SORTING THE CLASSIFICATION ORDER:  */

/* USING ONE VARIABLE: */
PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY logSalary;
CLASS TEAM/DESCEND;
RUN;


/* USING MORE THAN ONE VARIABLE: */
PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY logSalary;
CLASS TEAM LEAGUE/DESCEND;
RUN;


/* USING NUMERICAL VARIABLE: */
PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY logSalary;
CLASS TEAM /ORDER=FREQ;
RUN;


PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY logSalary;
CLASS TEAM LEAGUE/ORDER=FREQ;
RUN;




/* PART-7:---> */


/* g. FILTERING THE STATISTICAL REPORT BY CLASSIFICATION VARIABLES:  */

PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY logSalary;
CLASS TEAM LEAGUE/ORDER=FREQ;
WHERE LEAGUE='American';
RUN;


PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY logSalary;
CLASS TEAM LEAGUE/ORDER=FREQ;
WHERE DIV IN ('AW','AE');
RUN;
```

```
/* PART-8:--> */

/* h. OUTPUT THE STATISTICAL ANALYSIS IN A SEPARATE DATASET:  */

PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2;
VAR SALARY;
OUTPUT OUT = BASEBALL_SUMM;
RUN;

/* IF YOU DO NOT WANT TO PRINT THE RESULT: */

PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2 NOPRINT;
VAR SALARY;
OUTPUT OUT = BASEBALL_SUMM;
RUN;



/* PART-9:---> */

/* i. ADDING SOME ADDITIONAL STATISTICAL VARIABLES IN OUTPUT DATASET:  */

/* EG-1: */
PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2 NOPRINT;
VAR SALARY;
OUTPUT OUT = BASEBALL_SUMM MEDIAN=MEDIAN_REPORT;
RUN;

/* EG-2: */
PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2 NOPRINT;
VAR SALARY;
OUTPUT OUT = BASEBALL_SUMM MEDIAN=MEDIAN_REPORT SUM=SUM_REPORT;
RUN;



/* PART-10:---> */

/* j. AUTOMATICALLY NAMING THE VARIABLES IN OUTPUT DATASET:  */

/* EG-1: */
PROC MEANS DATA=WORK.TEST N NMISS SUM STD P75 MEDIAN MAXDEC=2 NOPRINT;
VAR SALARY;
OUTPUT OUT = BASEBALL_SUMM MEDIAN= SUM= T=/AUTONAME;
RUN;



/* PART-11:---> */

/* k. PERFORMING DIFFERENT STATISTICS ON DIFFERENT VARIABLE IN THE SAME OUTPUT DATASET:  */

/* EG-1: */
PROC MEANS DATA=WORK.TEST  MAXDEC=2 NOPRINT;
OUTPUT OUT = BASEBALL_SUMM
MEDIAN(logSalary)=SUM(SALARY)=T(CrRuns)=/AUTONAME;
RUN;

PROC PRINT DATA=BASEBALL_SUMM;
RUN;



/* PART-12:---> */
```

```
/* j. _TYPE_ VARIABLE IN OUTPUT DATASET:  */

/* EG-1: */
PROC MEANS DATA=WORK.TEST;
CLASS TEAM;
OUTPUT OUT = TESTING_BASE1
MEAN(CrRuns)=STD(SALARY)=/AUTONAME;
RUN;

PROC PRINT DATA=TESTING_BASE1;
RUN;

/* EG-2: */
PROC MEANS DATA=WORK.TEST;
CLASS TEAM DIV;
OUTPUT OUT = TESTING_BASE2
MEAN(CrRuns)=STD(SALARY)=/AUTONAME;
RUN;

PROC PRINT DATA=TESTING_BASE2;
RUN;

/* EG-3: */
PROC MEANS DATA=WORK.TEST;
CLASS TEAM DIV LEAGUE;
OUTPUT OUT = TESTING_BASE
MEAN(CrRuns)=STD(SALARY)=/AUTONAME;
RUN;

PROC PRINT DATA=TESTING_BASE;
RUN;




/* PART-12:---> */


/* l. USING NWAY OPTION TO OUTPUT THE REPORT:  */

/* IT FILTER ONLY MAXIMUM NUMBER OF 'TYPE' VARIABLE. */

/* EG-1: */
PROC MEANS DATA=WORK.TEST NWAY;
CLASS TEAM DIV LEAGUE;
OUTPUT OUT = TESTING_BASE4
MEAN(CrRuns)=STD(SALARY)=/AUTONAME;
RUN;

PROC PRINT DATA=TESTING_BASE4;
RUN;


/* EG-2: */
PROC MEANS DATA=WORK.TEST NWAY;
CLASS TEAM DIV;
OUTPUT OUT = TESTING_BASE5
MEAN(CrRuns)=STD(SALARY)=/AUTONAME;
RUN;

PROC PRINT DATA=TESTING_BASE5;
RUN;


/* EG-3: */
PROC MEANS DATA=WORK.TEST NWAY;
CLASS TEAM;
OUTPUT OUT = TESTING_BASE6
MEAN(CrRuns)=STD(SALARY)=/AUTONAME;
RUN;

PROC PRINT DATA=TESTING_BASE6;
RUN;
```