```
LIBNAME SSN6 '/home/debendra330/BATCH_202402/SESSION_6/A3.SAS_DATASET';
RUN;

DATA PROD_SALES_APAC;
INFILE CARDS DSD DLM='09'X;
INPUT PROD $ YEAR COUNTRY $ UNITS SALES;
CARDS;
APPLE   2020    INDIA   760 17123638
DELL    2020    INDIA   479 5716627
LENOVO  2020    INDIA   182 15120000
HP  2020    INDIA   507 20477520
ACER    2020    INDIA   175 27181250
APPLE   2021    INDIA   705 42244400
DELL    2021    INDIA   871 8536901
LENOVO  2021    INDIA   765 25147808
HP  2021    INDIA   335 14759825
ACER    2021    INDIA   453 12787152
APPLE   2022    INDIA   339 21743625
DELL    2022    INDIA   890 30924453
LENOVO  2022    INDIA   791 7922453
HP  2022    INDIA   609 33623780
ACER    2022    INDIA   879 10523810
APPLE   2020    JAPAN   449 17820306
DELL    2020    JAPAN   720 19908000
LENOVO  2020    JAPAN   838 12082227
HP  2020    JAPAN   398 5978610
ACER    2020    JAPAN   176 14270971
APPLE   2021    JAPAN   769 14358302
DELL    2021    JAPAN   900 39478392
LENOVO  2021    JAPAN   234 25614042
HP  2021    JAPAN   890 6663075
ACER    2021    JAPAN   854 13079043
APPLE   2022    JAPAN   581 27524868
DELL    2022    JAPAN   696 8038506
LENOVO  2022    JAPAN   171 27238806
HP  2022    JAPAN   608 9087780
ACER    2022    JAPAN   801 17522248
APPLE   2020    KOREA   263 13512008
DELL    2020    KOREA   678 33169920
LENOVO  2020    KOREA   569 37174668
HP  2020    KOREA   194 30042600
ACER    2020    KOREA   212 26281038
APPLE   2021    KOREA   794 5761674
DELL    2021    KOREA   448 19125296
LENOVO  2021    KOREA   817 17406485
HP  2021    KOREA   132 8558550
ACER    2021    KOREA   546 23683968
APPLE   2022    KOREA   503 6037295
DELL    2022    KOREA   405 34745466
LENOVO  2022    KOREA   526 31558572
HP  2022    KOREA   835 12083364
ACER    2022    KOREA   411 6609665
;
RUN;

PROC PRINT DATA=PROD_SALES_APAC;
RUN;

/* SUMMARIZING DATA IN BASE SAS */

PROC MEANS DATA=PROD_SALES_APAC;
RUN;

/* BY DEFAULT, ALL NUMERIC VARIABLES WILL BE SUMMARIZED */

/* N */
/* MEAN */
/* STD DEV */
/* MINIMUM */
/* MAXIMUM */

PROC MEANS DATA=PROD_SALES_APAC;
VAR UNITS SALES;
RUN;

PROC MEANS DATA=PROD_SALES_APAC;
CLASS PROD;
VAR UNITS SALES;
RUN;

PROC MEANS DATA=PROD_SALES_APAC;
CLASS PROD COUNTRY;
VAR UNITS SALES;
RUN;

PROC SQL;
SELECT PROD, COUNTRY,
COUNT(UNITS) AS N,
MEAN(UNITS) AS MEAN,
MIN(UNITS) AS MINIMUM,
MAX(UNITS) AS MAXIMUM,
STD(UNITS) AS STD_DEV
FROM PROD_SALES_APAC
GROUP BY PROD, COUNTRY;
```

```
QUIT;

/* EXAMPLE */

OPTIONS VALIDVARNAME=V7;
PROC IMPORT OUT=SSN6.LAPTOP_SALES
            DATAFILE = '/home/debendra330/BATCH_202402/SESSION_6/A1.RAW_DATA/A2.LAPTOP SALES DATA.xlsx'
            DBMS = XLSX REPLACE;
            SHEET = 'WW_SALES';
RUN;

PROC CONTENTS DATA=SSN6.LAPTOP_SALES VARNUM SHORT;
RUN;

/* ORDER_ID LAPTOP COUNTRY PURCHASE_DATE PURCHASE_YEAR PURCHASE_MONTH UNITS PRICE SALES */

PROC MEANS DATA=SSN6.LAPTOP_SALES;
RUN;


PROC MEANS DATA=SSN6.LAPTOP_SALES;
VAR UNITS PRICE SALES;
RUN;

PROC MEANS DATA=SSN6.LAPTOP_SALES;
CLASS LAPTOP;
VAR UNITS PRICE SALES;
RUN;

PROC MEANS DATA=SSN6.LAPTOP_SALES;
CLASS LAPTOP COUNTRY;
VAR UNITS PRICE SALES;
RUN;

PROC MEANS DATA=SSN6.LAPTOP_SALES;
CLASS LAPTOP COUNTRY;
VAR UNITS PRICE SALES;
OUTPUT OUT=SUMMARY_TABLE1;
RUN;

PROC EXPORT DATA=SUMMARY_TABLE1
            OUTFILE='/home/debendra330/BATCH_202402/SESSION_6/A4.SAS_OUTPUT/SUMMARY_TABLE1.xlsx'
            DBMS= XLSX REPLACE;
RUN;

PROC MEANS DATA=SSN6.LAPTOP_SALES NWAY;
CLASS LAPTOP COUNTRY;
VAR UNITS PRICE SALES;
OUTPUT OUT=SUMMARY_TABLE2 (DROP=_TYPE_ _FREQ_);
RUN;

PROC EXPORT DATA=SUMMARY_TABLE2
            OUTFILE='/home/debendra330/BATCH_202402/SESSION_6/A4.SAS_OUTPUT/SUMMARY_TABLE2.xlsx'
            DBMS= XLSX REPLACE;
RUN;

PROC MEANS DATA=SSN6.LAPTOP_SALES NWAY;
CLASS LAPTOP COUNTRY;
VAR UNITS PRICE SALES;
OUTPUT OUT=SUMMARY_TABLE3 (DROP=_TYPE_ _FREQ_);
RUN;

/* ONE MORE REALTIME EXAMPLE */

OPTIONS VALIDVARNAME=V7;
PROC IMPORT OUT = SSN6.MED_2023
            DATAFILE='/home/debendra330/BATCH_202402/SESSION_6/A1.RAW_DATA/A3.MED_2020.xlsx'
            DBMS = XLSX REPLACE;
            SHEET='MED_NEW_2016';
RUN;

/* SAS SQL */

PROC SQL NUMBER;
SELECT STATE_CODE, COMPANY,
COUNT(CUSTOMER_ID) AS SUBS,
MEAN(AGE) AS AVG_AGE,
SUM(NO_OF_TRIPS) AS VISITS,
SUM(SPENT_AMOUNT) AS TOTAL_SPENT
FROM SSN6.MED_2023
GROUP BY 1,2
ORDER BY 1,2;
QUIT;

PROC MEANS DATA=SSN6.MED_2023 NWAY;
CLASS STATE_CODE COMPANY;
VAR CUSTOMER_ID AGE NO_OF_TRIPS SPENT_AMOUNT;
OUTPUT OUT=MED_SUMMARY (DROP = _TYPE_ _FREQ_) N(CUSTOMER_ID)=SUBS MEAN(AGE)=AVG_AGE SUM(NO_OF_TRIPS)=VISITS SUM(SPENT_AMOUNT)=TOTAL_SPENT;
RUN;

/* PROC SUMMARY */

PROC SUMMARY DATA=SSN6.MED_2023 PRINT;
RUN;
```

```
PROC SUMMARY DATA=SSN6.MED_2023 PRINT;
VAR SPENT_AMOUNT;
RUN;

PROC SUMMARY DATA=SSN6.MED_2023 PRINT;
CLASS COMPANY;
VAR SPENT_AMOUNT;
RUN;

PROC SUMMARY DATA=SSN6.MED_2023 PRINT;
CLASS STATE_CODE COMPANY;
VAR SPENT_AMOUNT;
RUN;

PROC SUMMARY DATA=SSN6.MED_2023 NWAY PRINT;
CLASS STATE_CODE COMPANY;
VAR CUSTOMER_ID AGE NO_OF_TRIPS SPENT_AMOUNT;
OUTPUT OUT=MED_SUMMARY (DROP = _TYPE_ _FREQ_) N(CUSTOMER_ID)=SUBS MEAN(AGE)=AVG_AGE SUM(NO_OF_TRIPS)=VISITS SUM(SPENT_AMOUNT)=TOTAL_SPENT;
RUN;

/* DIFFERENCE BETWEEN PROC MEANS AND SUMMARY */

/* 1. PROC MEANS DOES NOT REQUIRE PRINT AS AN OPTION, PROC SUMMARY REQUIRES PRINT TO EXECUTE */

/* PROC UNIVARIATE STATEMENT */
PROC PRINT DATA=PROD_SALES_APAC;
RUN;

PROC UNIVARIATE DATA=PROD_SALES_APAC;
RUN;

PROC UNIVARIATE DATA=PROD_SALES_APAC;
VAR UNITS;
RUN;

PROC UNIVARIATE DATA=PROD_SALES_APAC;
CLASS PROD;
VAR UNITS;
RUN;

/* WHAT IS THE DIFFERENCE BETWEEN PROC MEANS AND PROC UNIVARIATE */

/* 1. PROC MEANS SHOWS COUNT, AVERAGE, STD_DEV, MIN, MAX */
/* 2. APART FROM ALL THESE STATS, PROC UNIVARIATE ALSO SHOWS YOU VARAINCE, SKEWNESS, QUANTILE, KUTOSIS */

/* PROC FREQ */

PROC FREQ DATA=PROD_SALES_APAC;
RUN;

PROC FREQ DATA=PROD_SALES_APAC;
TABLES PROD;
RUN;

PROC FREQ DATA=PROD_SALES_APAC;
TABLES PROD*COUNTRY;
RUN;

PROC FREQ DATA=SSN6.MED_2023;
TABLES COMPANY*STATE_CODE/ OUT=MED_SUMMARY NOCUM NOPERCENT;
RUN;

/* PROC REPORT */

PROC REPORT DATA=SSN6.MED_2023;
COLUMNS STATE_CODE COMPANY NO_OF_TRIPS SPENT_AMOUNT;
DEFINE STATE_CODE / WIDTH=15 GROUP;
DEFINE COMPANY / WIDTH=15 GROUP;
DEFINE NO_OF_TRIPS / WIDTH=15 'VISITS';
DEFINE SPENT_AMOUNT / WIDTH=15 'SPENT';
RBREAK AFTER/SUMMARIZE DOL DUL ;
FORMAT NO_OF_TRIPS COMMA10. SPENT_AMOUNT DOLLAR12.;
RUN;

ODS PDF FILE='/home/debendra330/BATCH_202402/SESSION_6/A4.SAS_OUTPUT/MED_SUMMARY_2023_V1.PDF';
PROC REPORT DATA=SSN6.MED_2023;
COLUMNS STATE_CODE COMPANY NO_OF_TRIPS SPENT_AMOUNT;
DEFINE STATE_CODE / WIDTH=15 GROUP;
DEFINE COMPANY / WIDTH=15 GROUP;
DEFINE NO_OF_TRIPS / WIDTH=15 'VISITS';
DEFINE SPENT_AMOUNT / WIDTH=15 'SPENT';
RBREAK AFTER/SUMMARIZE DOL DUL ;
FORMAT NO_OF_TRIPS COMMA10. SPENT_AMOUNT DOLLAR12.;
RUN;
ODS PDF CLOSE;

/* PROC TABULATE */

PROC TABULATE DATA=SSN6.MED_2023;
CLASS STATE_CODE COMPANY;
CLASSLEV STATE_CODE COMPANY;
VAR SPENT_AMOUNT;
TABLE STATE_CODE ALL='TOTAL_SALES', COMPANY*SPENT_AMOUNT*SUM ALL*SPENT_AMOUNT*SUM;
RUN;
```

```
PROC TABULATE DATA=SSN6.MED_2023 FORMAT=DOLLAR12. S=[FOREGROUND=BLACK JUST=C];
CLASS COMPANY STATE_CODE;
CLASSLEV COMPANY STATE_CODE / S=[BACKGROUND=RED FOREGROUND=WHITE];
VAR SPENT_AMOUNT;
TABLE COMPANY ALL='TOTAL_SALES' , STATE_CODE*SPENT_AMOUNT*SUM ALL*SPENT_AMOUNT*SUM={S=[FOREGROUND=WHITE BACKGROUND=RED]};
RUN;

ODS PDF FILE='/home/debendra330/BATCH_202402/SESSION_6/A4.SAS_OUTPUT/MED_TABULATE_V1.PDF';
PROC TABULATE DATA=SSN6.MED_2023 FORMAT=DOLLAR12. S=[FOREGROUND=BLACK JUST=C];
CLASS COMPANY STATE_CODE;
CLASSLEV COMPANY STATE_CODE / S=[BACKGROUND=RED FOREGROUND=WHITE];
VAR SPENT_AMOUNT;
TABLE COMPANY ALL='TOTAL_SALES' , STATE_CODE*SPENT_AMOUNT*SUM ALL*SPENT_AMOUNT*SUM={S=[FOREGROUND=WHITE BACKGROUND=RED]};
RUN;
ODS PDF CLOSE;

/* TO SELECT THE VARIABLES */

/* 1. PROC MEANS:CATEGORICAL VARIABLE WE USE CLASS STATEMENT, NUMERICAL VARIABLE WE USED VAR STATEMENT */
/* 2. PROC SUMMARY: CATEGORICAL VARIABLE WE USE CLASS, NUMERIC VARIABLE WE USE VAR */
/* 3. PROC UNIVARIATE : CATEGORICAL VARIABLE WE USE CLASS, NUMERIC VARIABLE WE USE VAR */
/* 4. PROC FREQ : CATEGORICAL VARIABLE WE USE TABLES */
/* 5. PROC REPORT : FOR DECLARING THE COLUMNS WE USE COLUMN STATEMENT */
/* 6. PROC TABULATE - CATEGORICAL VARIABLE USE CLASS AND CLASSLEV, NUMERIC VARIABLE WE USE VAR */

/* CHARTS IN SAS */
/* ======================== */

PROC PRINT DATA=SSN6.LAPTOP_SALES (OBS=10);
RUN;

/* HORIZONTAL BAR CHART */
TITLE1 "SALES BY LAPTOP";
PROC GCHART DATA=SSN6.LAPTOP_SALES;
HBAR LAPTOP / TYPE=SUM SUMVAR=SALES;
RUN;

/* VERTICAL BAR CHART */
TITLE1 "SALES BY COUNTRY";
PROC GCHART DATA=SSN6.LAPTOP_SALES;
VBAR COUNTRY / TYPE=SUM SUMVAR=SALES;
RUN;

/* VERICAL BAR CHART WITH DATA LABEL */
TITLE1 "SALES BY COUNTRY";
PROC GCHART DATA=SSN6.LAPTOP_SALES;
VBAR COUNTRY / TYPE=SUM SUMVAR=SALES OUTSIDE=SUM;
RUN;


/* VERICAL BAR CHART WITH DATA LABEL IN DESCENDING ORDER*/
TITLE1 "SALES BY COUNTRY";
PROC GCHART DATA=SSN6.LAPTOP_SALES;
VBAR COUNTRY / TYPE=SUM SUMVAR=SALES DESCENDING OUTSIDE=SUM;
RUN;

/* VBAR WITH AXIS NAME */
TITLE1 "SALES BY COUNTRY";
GOPTIONS HTEXT=13PT HTITLE=15PT;
AXIS1 LABEL=NONE;
AXIS2 LABEL=('SALES IN DOLLAR');
PROC GCHART DATA=SSN6.LAPTOP_SALES;
VBAR COUNTRY / TYPE=SUM SUMVAR=SALES DESCENDING
MAXIS=AXIS1 RAXIS=AXIS2;
RUN;
QUIT;

/* VBAE WITH COLOUR GRIDIING */
PATTERN1 V=SOLID COLOR=RED;
PROC GCHART DATA=SSN6.WW_LAPTOP_SALES;
VBAR COUNTRY / WIDTH= 25 TYPE=MEAN SUMVAR=SALES DESCENDING
MAXIS=AXIS1 RAXIS=AXIS2 OUTSIDE=MEAN COUTLINE=GRAY;
RUN;
QUIT;

/* VBAR WITH MULTIP COLOURS */
TITLE1 "AVERAGE LAPTOP SALES";
PATTERN1 VALUE=SOLID COLOR=PINK;
PATTERN2 VALUE=SOLID COLOR=CX42C0FB;
PROC GCHART DATA=SSN6.LAPTOP_SALES;
VBAR LAPTOP / WIDTH= 25 TYPE=MEAN SUMVAR=SALES DESCENDING
MAXIS=AXIS1 RAXIS=AXIS2 OUTSIDE=MEAN COUTLINE=GRAY SUBGROUP=LAPTOP;
RUN;
QUIT;


/* VBAR WITH FREQ IN PERCENTAGE */
TITLE1 "% OF LAPTOP SALES";
GOPTIONS HTEXT=13PT HTITLE=15PT;
AXIS1 LABEL=NONE  VALUE=(F="ARIAL/BOLD" "LAPTOP"); ;
AXIS2 LABEL=(A=90 F="ARIAL/BOLD" 'PERCENTAGE OF LAPTOP SALES') MINOR=NONE  OFFSET=(0,0);
PROC GCHART DATA=SSN6.LAPTOP_SALES;
VBAR LAPTOP / WIDTH= 25 TYPE=PCT DESCENDING
MAXIS=AXIS1 RAXIS=AXIS2 INSIDE= FREQ OUTSIDE=PCT;
RUN;
```

```
QUIT;


/* VBAR WITH MULTI-BARS */
TITLE;
AXIS1 LABEL=('MSRP') MINOR=NONE OFFSET=(0,0);
AXIS2 LABEL=NONE;
AXIS3 LABEL=NONE OFFSET=(7,7);
PROC GCHART DATA=SSN6.LAPTOP_SALES;
VBAR LAPTOP  / DISCRETE TYPE=MEAN SUMVAR=SALES
GROUP=COUNTRY COUTLINE=GRAY RAXIS=AXIS1 MAXIS=AXIS2 GAXIS=AXIS3 NOFRAME;
RUN;
QUIT;

/* STACKED BAT CHARTS */
TITLE;
AXIS1 LABEL=('MSRP') MINOR=NONE OFFSET=(0,0);
AXIS2 LABEL=NONE OFFSET=(7,7);
PROC GCHART DATA=SSN6.WW_LAPTOP_SALES;
VBAR COUNTRY / DISCRETE TYPE=MEAN SUMVAR=SALES
SUBGROUP=LAPTOP COUTLINE=GRAY WIDTH=10
RAXIS=AXIS1 MAXIS=AXIS2 NOFRAME;
RUN;
QUIT;

/* LABEL CHARTS */
TITLE;
GOPTIONS HTEXT=10PT HTITLE=12PT;
PROC GCHART DATA=SSN6.WW_LAPTOP_SALES;
VBAR LAPTOP / SPACE=1 WIDTH=10 OUTSIDE=FREQ LEVELS=4 RANGE;
RUN; QUIT;

/* SCATTER CHART */
SYMBOL1 VALUE=CIRCLE HEIGHT=3 INTERPOL=NONE COLOR=BLUE;
SYMBOL2 VALUE=CIRCLE HEIGHT=3 INTERPOL=NONE COLOR=RED;
PROC GPLOT DATA=SSN6.WW_LAPTOP_SALES;
PLOT UNITS*SALES=COUNTRY;
RUN;

/* DISTRIBUTION CHART-1 */
TITLE "SALES DISTRIBUTION BY COUNTRY";
SYMBOL1 INTERPOL=BOXT BWIDTH=4 COLOR=BLUE;
AXIS1 MINOR=NONE OFFSET=(0,0);
AXIS2 OFFSET=(20,20);
PROC GPLOT DATA=SSN6.WW_LAPTOP_SALES;
PLOT PRICE*SALES=1 /
VAXIS=AXIS1 HAXIS=AXIS2 NOFRAME;
RUN;

/* DISTRIBUTION CHART-2 */
TITLE "HEIGHT DISTRIBUTION BY SEX";
SYMBOL1 INTERPOL=BOXT BWIDTH=4 COLOR=BLUE;
AXIS1 MINOR=NONE OFFSET=(0,0);
AXIS2 OFFSET=(20,20);
PROC GPLOT DATA=SSN6.WW_LAPTOP_SALES;
PLOT UNITS*PRICE=1 /
VAXIS=AXIS1 HAXIS=AXIS2 NOFRAME;
RUN;

/* LINE CHART */
TITLE "LINE CHART";
SYMBOL1 VALUE=NONE INTERPOL=SM COLOR=BLUE;
PROC GPLOT DATA=SASHELP.FAILURE;
WHERE CAUSE ="Contamination" AND PROCESS = "Process A";
PLOT COUNT*DAY;
RUN;
QUIT;


/* PIE CHART-1 */
PROC SQL;
create table CARS1 as
SELECT make, model, type, invoice, horsepower, length, weight
 FROM
SASHELP.CARS
WHERE make in ('Audi','BMW')
;
RUN;

PROC TEMPLATE;
   DEFINE STATGRAPH pie;
      BEGINGRAPH;
         LAYOUT REGION;
            PIECHART CATEGORY = type /
            DATALABELLOCATION = OUTSIDE
            CATEGORYDIRECTION = CLOCKWISE
            START = 180 NAME = 'pie';
            DISCRETELEGEND 'pie' /
            TITLE = 'Car Types';
         ENDLAYOUT;
      ENDGRAPH;
   END;
RUN;
PROC SGRENDER DATA = cars1
            TEMPLATE = pie;
```

```
RUN;


/* PIE CHART-2 */
PROC TEMPLATE;
DEFINE STATGRAPH simplepie;
BEGINGRAPH;
LAYOUT REGION;
PIECHART CATEGORY = type / Group = origin
DATALABELLOCATION = INSIDE
DATALABELCONTENT=ALL
CATEGORYDIRECTION = CLOCKWISE
DATASKIN= SHEEN
START = 180 NAME = 'pie';
DISCRETELEGEND 'pie' /
TITLE = 'Car models by origin';
ENDLAYOUT;
ENDGRAPH;
END;
RUN;
PROC SGRENDER DATA = cars1
TEMPLATE = pie;
RUN;

/* PIE CHART-3 */
PROC TEMPLATE;
DEFINE STATGRAPH pie;
BEGINGRAPH;
LAYOUT REGION;
PIECHART CATEGORY = type /
DATALABELLOCATION = INSIDE
DATALABELCONTENT = ALL
CATEGORYDIRECTION = CLOCKWISE
DATASKIN = SHEEN
START = 180 NAME = 'pie';
DISCRETELEGEND 'pie' /
TITLE = 'Car Types';
ENDLAYOUT;
ENDGRAPH;
END;
RUN;
PROC SGRENDER DATA = cars1
TEMPLATE = pie;
RUN;

/* PIE CHART-4 */
PROC TEMPLATE;
DEFINE STATGRAPH pie;
BEGINGRAPH;
LAYOUT REGION;
PIECHART CATEGORY = type / Group = make
DATALABELLOCATION = INSIDE
DATALABELCONTENT = ALL
CATEGORYDIRECTION = CLOCKWISE
```