



## EXPERIMENT 1

**SAP ID:** 60004190013**Name:** Aryan Parekh**Division:** A**Batch:** A1**Subject:** Business Analytics**Semester:** VI**Date of Submission:** 29th March, 2022**Aim:** To import data in SAS Studio from Excel or CSV file.**Code and Output:**

```
PROC IMPORT DATAFILE='/home/u61126016/My_Dataset/Salary_Data.csv'  
    DBMS=CSV  
    OUT=SALARY REPLACE;  
    GETNAMES=YES;  
  
RUN;  
PROC PRINT DATA=SALARY;  
    TITLE 'SALARY DATA IMPORT FROM CSV';  
RUN;
```

**SALARY DATA IMPORT FROM CSV**

Obs	YearsExperience	Salary
1	1.1	39343
2	1.3	46205
3	1.5	37731
4	2	43525
5	2.2	39891
6	2.9	56642
7	3	60150
8	3.2	54445
9	3.2	64445
10	3.7	57189
11	3.9	63218
12	4	55794
13	4	56957
14	4.1	57081
15	4.5	61111
16	4.9	67938
17	5.1	66029
18	5.3	83088
19	5.9	81363
20	6	93940
21	6.8	91738
22	7.1	98273
23	7.9	101302
24	8.2	113812
25	8.7	109431
26	9	105582
27	9.5	116969
28	9.6	112635
29	10.3	122391
30	10.5	121872



```
PROC IMPORT DATAFILE='/home/u61126016/My_Dataset/Salary_Data.xlsx'
  DBMS=XLSX
  OUT=SALARY2 REPLACE;
  GETNAMES=YES;
RUN;
PROC PRINT DATA=SALARY2;
  TITLE 'SALARY DATA IMPORT FROM XLSX';
RUN;
```

**SALARY DATA IMPORT FROM XLSX**

Obs	YearsExperience	Salary
1	1.1	39343
2	1.3	46205
3	1.5	37731
4	2	43525
5	2.2	39891
6	2.9	56642
7	3	60150
8	3.2	54445
9	3.2	64445
10	3.7	57189
11	3.9	63218
12	4	55794
13	4	56957
14	4.1	57081
15	4.5	61111
16	4.9	67938
17	5.1	66029
18	5.3	83088
19	5.9	81363
20	6	93940
21	6.8	91738
22	7.1	98273
23	7.9	101302
24	8.2	113812
25	8.7	109431
26	9	105582
27	9.5	116969
28	9.6	112635
29	10.3	122391
30	10.5	121872

**Conclusion:** From this experiment, we learn how to import a CSV file and an XLSX file in SAS Studio.



## EXPERIMENT 2

**SAP ID:** 60004190013**Name:** Aryan Parekh**Division:** A**Batch:** A1**Subject:** Business Analytics**Semester:** VI**Date of Submission:** 29th March, 2022**Aim:** To create summary statistical data in SAS Studio.**Code and Output:**

```
PROC IMPORT DATAFILE='/home/u61126016/My_Dataset/Student_Marks.csv'  
      DBMS=CSV  
      OUT=Student_Marks REPLACE;  
      GETNAMES=YES;  
RUN;  
PROC PRINT DATA=Student_Marks;  
      TITLE 'STUDENT MARKS TABLE';  
RUN;
```

**STUDENT MARKS TABLE**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	60004190013	Aryan Parekh	m	97	93	82
2	60004190060	Kartik Suvarna	m	96	91	85
3	60004190066	Mayav Antani	m	92	78	89
4	60004190059	Karishni Mehta	f	95	82	71
5	60004190061	Khushi Chavan	f	47	96	83
6	60004190048	Ishita Kheria	f	75	98	93
7	60004190034	Divyashree Vaidya	f	89	91	99
8	60004190067	Megh Dedhia	m	10	15	20
9	60004190011	Akshath Mahajan	m	50	60	55
10	60004190003	Adithya Sanyal	m	65	70	73
11	60004190025	Deap Daru	m	75	78	83
12	60004190004	Aditi Kamath	f	45	35	30



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

```
PROC MEANS DATA=Student_Marks N MEAN MEDIAN MODE MIN MAX RANGE  
QRANGE CV SKEWNESS MAXDEC=2;  
    TITLE 'STATISTICAL SUMMARY OF STUDENT MARKS TABLE (MEANS)';  
    VAR Maths_Score Physics_Score Chemistry_Score;  
RUN;
```

#### STATISTICAL SUMMARY OF STUDENT MARKS TABLE (MEANS)

The MEANS Procedure

Variable	N	Mean	Median	Mode	Minimum	Maximum	Range	Quartile Range	Coeff of Variation	Skewness
Maths_Score	12	69.67	75.00	75.00	10.00	97.00	87.00	45.00	38.96	-0.94
Physics_Score	12	73.92	80.00	78.00	15.00	98.00	83.00	27.00	34.87	-1.43
Chemistry_Score	12	71.92	82.50	83.00	20.00	99.00	79.00	24.00	34.39	-1.28

```
PROC UNIVARIATE DATA=Student_Marks;  
    TITLE 'STATISTICAL SUMMARY OF STUDENT MARKS TABLE  
(UNIVARIATE)';  
    VAR Maths_Score Physics_Score Chemistry_Score;  
RUN;
```



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

## STATISTICAL SUMMARY OF STUDENT MARKS TABLE (UNIVARIATE)

The UNIVARIATE Procedure  
Variable: Maths\_Score

Moments			
N	12	Sum Weights	12
Mean	69.6666667	Sum Observations	836
Std Deviation	27.1404875	Variance	736.606061
Skewness	-0.9449848	Kurtosis	0.40991548
Uncorrected SS	66344	Corrected SS	8102.66667
Coeff Variation	38.9576375	Std Error Mean	7.83478388

Basic Statistical Measures			
Location		Variability	
Mean	69.66667	Std Deviation	27.14049
Median	75.00000	Variance	736.60606
Mode	75.00000	Range	87.00000
		Interquartile Range	45.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	8.89197	Pr >  t	<.0001
Sign	M	6	Pr >=  M	0.0005
Signed Rank	S	39	Pr >=  S	0.0005

Quantiles (Definition 5)	
Level	Quantile
100% Max	97.0
99%	97.0
95%	97.0
90%	96.0
75% Q3	93.5
50% Median	75.0
25% Q1	48.5
10%	45.0
5%	10.0
1%	10.0
0% Min	10.0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
10	8	89	7
45	12	92	3
47	5	95	4
50	9	96	2
65	10	97	1



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

## STATISTICAL SUMMARY OF STUDENT MARKS TABLE (UNIVARIATE)

The UNIVARIATE Procedure  
Variable: Physics\_Score

Moments			
N	12	Sum Weights	12
Mean	73.9166667	Sum Observations	887
Std Deviation	25.7768889	Variance	664.44697
Skewness	-1.425991	Kurtosis	1.41414757
Uncorrected SS	72873	Corrected SS	7308.91667
Coeff Variation	34.8728779	Std Error Mean	7.4411411

Basic Statistical Measures			
Location		Variability	
Mean	73.91667	Std Deviation	25.77687
Median	80.00000	Variance	664.44697
Mode	78.00000	Range	83.00000
		Interquartile Range	27.00000

Note: The mode displayed is the smallest of 2 modes with a count of 2.

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	9.933512	Pr >  t	<.0001
Sign	M	6	Pr >=  M	0.0005
Signed Rank	S	39	Pr >=  S	0.0005

Quantiles (Definition 5)	
Level	Quantile
100% Max	98
99%	98
95%	98
90%	96
75% Q3	92
50% Median	80
25% Q1	65
10%	35
5%	15
1%	15
0% Min	15

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
15	8	91	2
35	12	91	7
60	9	93	1
70	10	96	5
78	11	98	6



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

## STATISTICAL SUMMARY OF STUDENT MARKS TABLE (UNIVARIATE)

The UNIVARIATE Procedure  
Variable: Chemistry\_Score

Moments			
N	12	Sum Weights	12
Mean	71.9166667	Sum Observations	863
Std Deviation	24.7329678	Variance	611.719697
Skewness	-1.2849724	Kurtosis	0.73537868
Uncorrected SS	68793	Corrected SS	6728.91667
Coeff Variation	34.3911488	Std Error Mean	7.13979281

Basic Statistical Measures			
Location		Variability	
Mean	71.91667	Std Deviation	24.73297
Median	82.50000	Variance	611.71970
Mode	83.00000	Range	79.00000
		Interquartile Range	24.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	10.07265	Pr >  t	<.0001
Sign	M	6	Pr >=  M	0.0005
Signed Rank	S	39	Pr >=  S	0.0005

Quantiles (Definition 5)	
Level	Quantile
100% Max	99.0
99%	99.0
95%	99.0
90%	93.0
75% Q3	87.0
50% Median	82.5
25% Q1	63.0
10%	30.0
5%	20.0
1%	20.0
0% Min	20.0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
20	8	83	11
30	12	85	2
55	9	89	3
71	4	93	6
73	10	99	7



PROC FREQ DATA=Student\_Marks;

TITLE 'STATISTICAL SUMMARY OF STUDENT MARKS TABLE (FREQ)';

TABLES Name Maths\_Score Physics\_Score Chemistry\_Score;

RUN;

**STATISTICAL SUMMARY OF STUDENT MARKS TABLE (FREQ)**

The FREQ Procedure

Name	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Adithya Sanyal	1	8.33	1	8.33
Aditi Kamath	1	8.33	2	16.67
Akshath Mahajan	1	8.33	3	25.00
Aryan Parekh	1	8.33	4	33.33
Deap Daru	1	8.33	5	41.67
Divyashree Vaidya	1	8.33	6	50.00
Ishita Kheria	1	8.33	7	58.33
Karishni Mehta	1	8.33	8	66.67
Kartik Suvarna	1	8.33	9	75.00
Khushi Chavan	1	8.33	10	83.33
Mayav Antani	1	8.33	11	91.67
Megh Dedhia	1	8.33	12	100.00



Shri Vile Parle Kelavani Mandal's

**Dwarkadas J. Sanghvi College of Engineering**

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

Maths_Score	Frequency	Percent	Cumulative Frequency	Cumulative Percent
10	1	8.33	1	8.33
45	1	8.33	2	16.67
47	1	8.33	3	25.00
50	1	8.33	4	33.33
65	1	8.33	5	41.67
75	2	16.67	7	58.33
89	1	8.33	8	66.67
92	1	8.33	9	75.00
95	1	8.33	10	83.33
96	1	8.33	11	91.67
97	1	8.33	12	100.00

Physics_Score	Frequency	Percent	Cumulative Frequency	Cumulative Percent
15	1	8.33	1	8.33
35	1	8.33	2	16.67
60	1	8.33	3	25.00
70	1	8.33	4	33.33
78	2	16.67	6	50.00
82	1	8.33	7	58.33
91	2	16.67	9	75.00
93	1	8.33	10	83.33
96	1	8.33	11	91.67
98	1	8.33	12	100.00

Chemistry_Score	Frequency	Percent	Cumulative Frequency	Cumulative Percent
20	1	8.33	1	8.33
30	1	8.33	2	16.67
55	1	8.33	3	25.00
71	1	8.33	4	33.33
73	1	8.33	5	41.67
82	1	8.33	6	50.00
83	2	16.67	8	66.67
85	1	8.33	9	75.00
89	1	8.33	10	83.33
93	1	8.33	11	91.67
99	1	8.33	12	100.00



Shri Vile Parle Kelavani Mandal's

## Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

**Conclusion:** From this experiment, we learn how to create summary statistical data in SAS Studio.



## EXPERIMENT 3

**SAP ID:** 60004190013**Name:** Aryan Parekh**Division:** A**Batch:** A1**Subject:** Business Analytics**Semester:** VI**Date of Submission:** 29th March, 2022**Aim:** To export data in SAS Studio to Excel and PDF file.**Code and Output:**

```
PROC IMPORT DATAFILE='/home/u61126016/My_Dataset/Salary_Data.csv'  
    DBMS=CSV  
    OUT=SALARY REPLACE;  
RUN;  
PROC PRINT DATA=SALARY;  
    TITLE 'SALARY DATA TABLE';  
RUN;
```

**SALARY DATA TABLE**

Obs	YearsExperience	Salary
1	1.1	39343
2	1.3	46205
3	1.5	37731
4	2	43525
5	2.2	39891
6	2.9	56642
7	3	60150
8	3.2	54445
9	3.2	64445
10	3.7	57189
11	3.9	63218
12	4	55794
13	4	56957
14	4.1	57081
15	4.5	61111
16	4.9	67938
17	5.1	66029
18	5.3	83088
19	5.9	81363
20	6	93940
21	6.8	91738
22	7.1	98273
23	7.9	101302
24	8.2	113812
25	8.7	109431
26	9	105582
27	9.5	116969
28	9.6	112635
29	10.3	122391
30	10.5	121872



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

```
ODS PDF FILE='/home/u61126016/My_Dataset/Salary_Data_EXPORT.pdf';
PROC PRINT DATA=Salary;
  TITLE 'SALARY DATA TABLE';
RUN;
PROC MEANS DATA=Salary N MEAN MEDIAN MODE MIN MAX RANGE
MAXDEC=2;
  TITLE 'STATISTICAL SUMMARY OF SALARY DATA TABLE';
RUN;
ODS PDF CLOSE;
```

Obs	YearsExp	Salary
1	1.1	39343
2	1.3	46205
3	1.5	37731
4	2	43525
5	2.2	39891
6	2.9	56421
7	3	60150
8	3.2	54445
9	3.2	64445
10	3.7	57389
11	3.9	63218
12	4	50794
13	4	50794
14	4.1	57000
15	4.5	61111
16	4.9	67938
17	5.1	66029
18	5.3	83088
19	5.9	81363
20	6	95840
21	6.8	91396
22	7	91396
23	7.9	101302
24	8.2	113812
25	8.7	109431
26	9	105582
27	9.5	116969
28	9.6	112635
29	10.3	122391
30	10.5	121872

STATISTICAL SUMMARY OF SALARY DATA TABLE					
The MEANS Procedure					
Variable	N	Mean	Median	Minimum	Maximum
YearsExp	30	5.7	62217.30	3.0	37731.00
Salary	30	76033.00	62217.30	22391.00	86612.00

```
ODS CSVALL FILE='/home/u61126016/My_Dataset/Salary_Data_EXPORT.csv';
PROC PRINT DATA=Salary;RUN;
ODS CSVALL CLOSE;
```

Obs	YearsExp	Salary
1	1.1	39343
2	1.3	46205
3	1.5	37731
4	2	43525
5	2.2	39891
6	2.9	56421
7	3	60150
8	3.2	54445
9	3.2	64445
10	3.7	57389
11	3.9	63218
12	4	55794
13	4	56957
14	4.1	57081
15	4.5	61111
16	4.9	67938
17	5.1	66029
18	5.3	83088
19	5.9	81363
20	6	95840
21	6.8	91396
22	7	91396
23	7.9	101302
24	8.2	113812
25	8.7	109431
26	9	105582
27	9.5	116969
28	9.6	112635
29	10.3	122391
30	10.5	121872



Shri Vile Parle Kelavani Mandal's

## Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

**Conclusion:** From this experiment, we learn how to export data in SAS Studio to Excel and PDF file.



## EXPERIMENT 4

**SAP ID:** 60004190013**Name:** Aryan Parekh**Division:** A**Batch:** A1**Subject:** Business Analytics**Semester:** VI**Date of Submission:** 29th March, 2022**Aim:** To manipulate data in SAS Studio using functions.**Code and Output:**

```
DATA MATH_FUNCTIONS;
  V1=21; V2=42; V3=13; V4=15; V5=29;
  MAXIMUM = MAX(v1,v2,v3,v4,v5);
  MINIMUM = MIN (v1,v2,v3,v4,v5);
  MEDIAN = MEDIAN (v1,v2,v3,v4,v5);
  RANDOM_VALUE = RANUNI(0);
  SUM = SUM(v1,v2,v3,v4,v5);
  SQUARE_ROOT_OF_SUM = SQRT(SUM(v1,v2,v3,v4,v5));
RUN;
PROC PRINT DATA = MATH_FUNCTIONS;
  TITLE 'MATH FUNCTIONS';
RUN;
```

**MATH FUNCTIONS**

Obs	V1	V2	V3	V4	V5	MAXIMUM	MINIMUM	MEDIAN	RANDOM_VALUE	SUM	SQUARE_ROOT_OF_SUM
1	21	42	13	15	29	42	13	21	0.42693	120	10.9545

```
DATA DATE_FUNCTIONS;
  INPUT @1 date1 date9. @11 date2 date9. ;
  FORMAT date1 date9. date2 date9. ;
  YEAR_DIFFERENCE = INTCK('YEAR',date1,date2);
  MONTH_DIFFERENCE = INTCK('MONTH',date1,date2);
  WEEKDAY = WEEKDAY(date1);
  TODAY = TODAY();
  TIME = time();
  DATALINES;
01MAR2009 11JUL2012
21OCT2000 16AUG1998
;
PROC PRINT DATA = DATE_FUNCTIONS;
  TITLE 'DATE FUNCTIONS';
RUN;
```

**DATE FUNCTIONS**

Obs	date1	date2	YEAR_DIFFERENCE	MONTH_DIFFERENCE	WEEKDAY	TODAY	TIME
1	01MAR2009	11JUL2012	3	40	1	22733	68148.92
2	21OCT2000	16AUG1998	-2	-26	7	22733	68148.92

**DATA CHARACTER\_FUNCTIONS;**

ORIGINAL = 'hElLo wORld';

LOWER CASE = LOWERCASE('hElLo wORld');

UPPER CASE = UPCASE('hElLo wORld');

REVERSE = REVERSE('hElLo wORld');

AFTER\_SPLIT\_1ST\_WORD = SCAN('hElLo wORld',1);

AFTER\_SPLIT\_2ND\_WORD = SCAN('hElLo wORld',2);

RUN;

PROC PRINT DATA = CHARACTER\_FUNCTIONS;

TITLE 'CHARACTER FUNCTIONS';

RUN;

**CHARACTER FUNCTIONS**

Obs	ORIGINAL	LOWER CASE	UPPER CASE	REVERSE	AFTER_SPLIT_1ST_WORD	AFTER_SPLIT_2ND_WORD
1	hElLo wORld	hello world	HELLO WORLD	dIROW oLIEh	hElLo	wORld

**DATA NUMBER\_TRUNCATE\_FUNCTIONS;**

ORIGINAL = 12.59;

CEIL = CEIL(12.59);

FLOOR = FLOOR(12.59);

INT = INT(12.59);

ROUND = ROUND(12.59);

RUN;

PROC PRINT DATA = NUMBER\_TRUNCATE\_FUNCTIONS;

TITLE 'NUMBER TRUNCATE FUNCTIONS';

RUN;

**NUMBER TRUNCATE FUNCTIONS**

Obs	ORIGINAL	CEIL	FLOOR	INT	ROUND
1	12.59	13	12	12	13

**Conclusion:** From this experiment, we learn how to manipulate data in SAS Studio using functions.



## EXPERIMENT 5

---

**SAP ID:** 60004190013**Name:** Aryan Parekh**Division:** A**Batch:** A1**Subject:** Business Analytics**Semester:** VI**Date of Submission:** 7th May, 2022

---

**Aim:** To use data with formats like charts and graphs in SAS Visual Analytics.**Theory:**

**Pie Chart:** A pie chart (or a circle chart) is a circular statistical graphic, which is divided into slices to illustrate numerical proportion. In a pie chart, the arc length of each slice (and consequently its central angle and area) is proportional to the quantity it represents. While it is named for its resemblance to a pie which has been sliced, there are variations on the way it can be presented.

**Word Cloud:** Word Clouds (also known as wordle, word collage or tag cloud) are visual representations of words that give greater prominence to words that appear more frequently. For Word Clouds, the words that are added most frequently by audience members using their smartphones. This type of visualization can help presenters to quickly collect data from their audience, highlight the most common answers and present the data in a way that everyone can understand.

**Tree Map:** A tree map chart provides a hierarchical view of your data and makes it easy to spot patterns, such as which items are a store's best sellers. The tree branches are represented by rectangles and each sub-branch is shown as a smaller rectangle. The tree map chart displays categories by color and proximity and can easily show lots of data which would be difficult with other chart types.

**Line chart:**

- Line chart is a type of chart that displays information as a series of data points connected by straight line segments.
- A line chart is a way of visually representing an asset's price history using a single, continuous line.
- A line chart is easy to understand and simple in form, typically only depicting only changes in an asset's closing price over time.
- Because line charts usually only show closing prices, they reduce noise from less critical times in the trading day, such as the open, high, and low prices.
- Because of its simplicity, however, traders looking to identify patterns or trends may opt for chart types with more information, such as a candlestick.



Shri Vile Parle Kelavani Mandal's

## Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

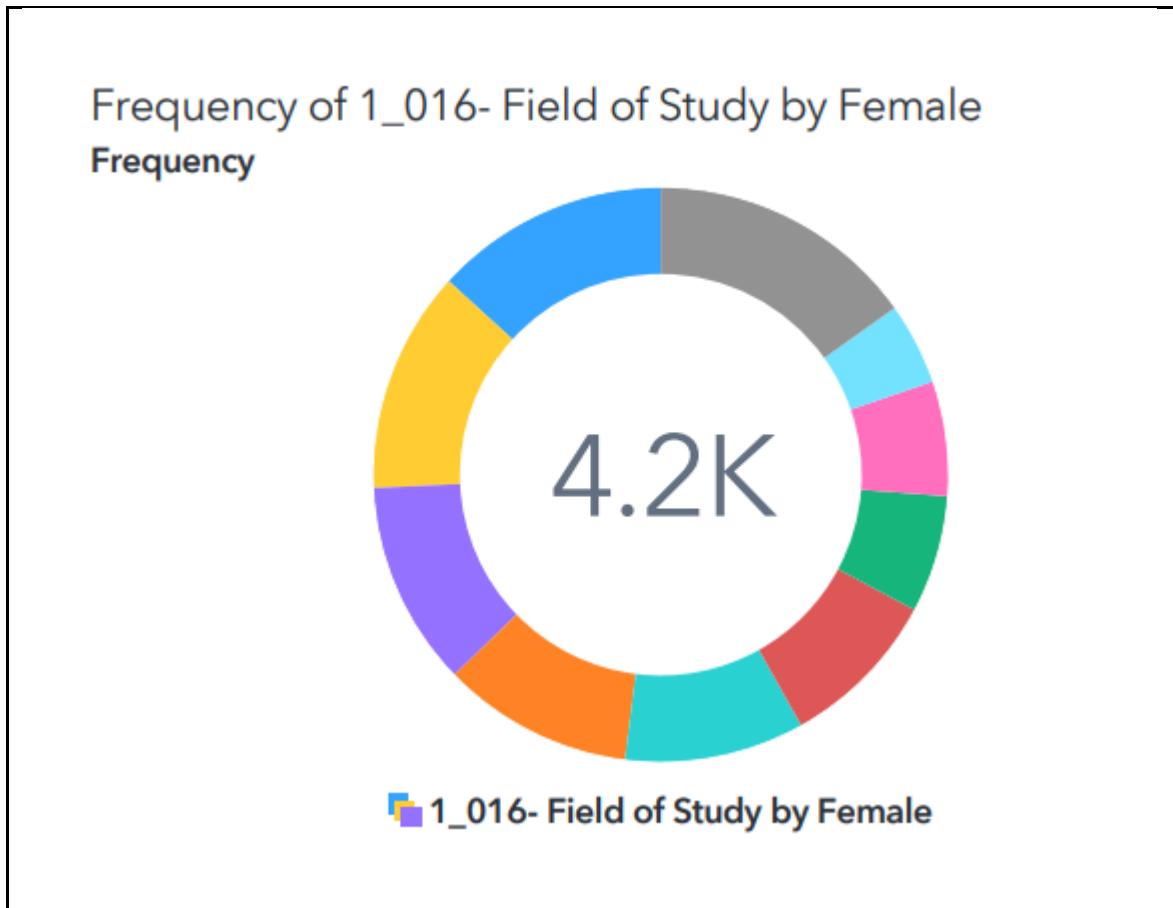
### Geo Chart:

A geo chart is a map of a country, a continent, or a region with areas identified in one of three ways:

- The region mode colors whole regions, such as countries, provinces, or states.
- The markers mode uses circles to designate regions that are scaled according to a value that you specify.
- The text mode labels the regions with identifiers (e.g., "Russia" or "Asia").

A geo chart is rendered within the browser using SVG or VML. Note that the geo chart is not scrollable or draggable, and it's a line drawing rather than a terrain map; if you want any of that, consider a map visualization instead.

### Code and Output:





Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

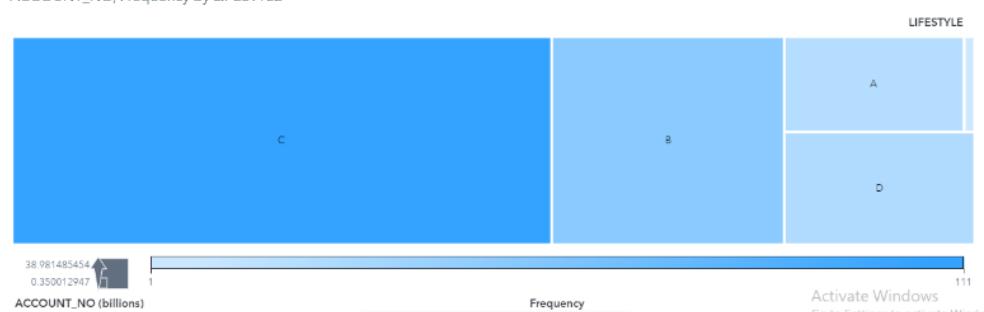
NAAC Accredited with "A" Grade (CGPA : 3.18)

Frequency of 1\_017- Field of Study by Male



[CHK] [MM] [SAV]

ACCOUNT\_NO, Frequency by LIFESTYLE

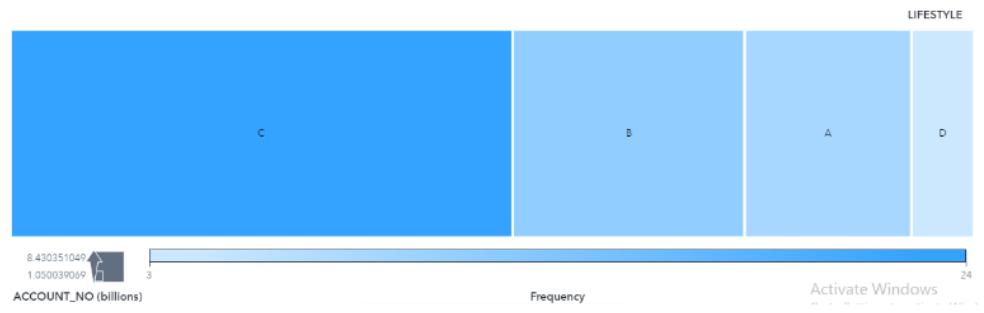


LIFESTYLE

Activate Windows

[CHK] [MM] [SAV]

ACCOUNT\_NO, Frequency by LIFESTYLE



LIFESTYLE

Activate Windows



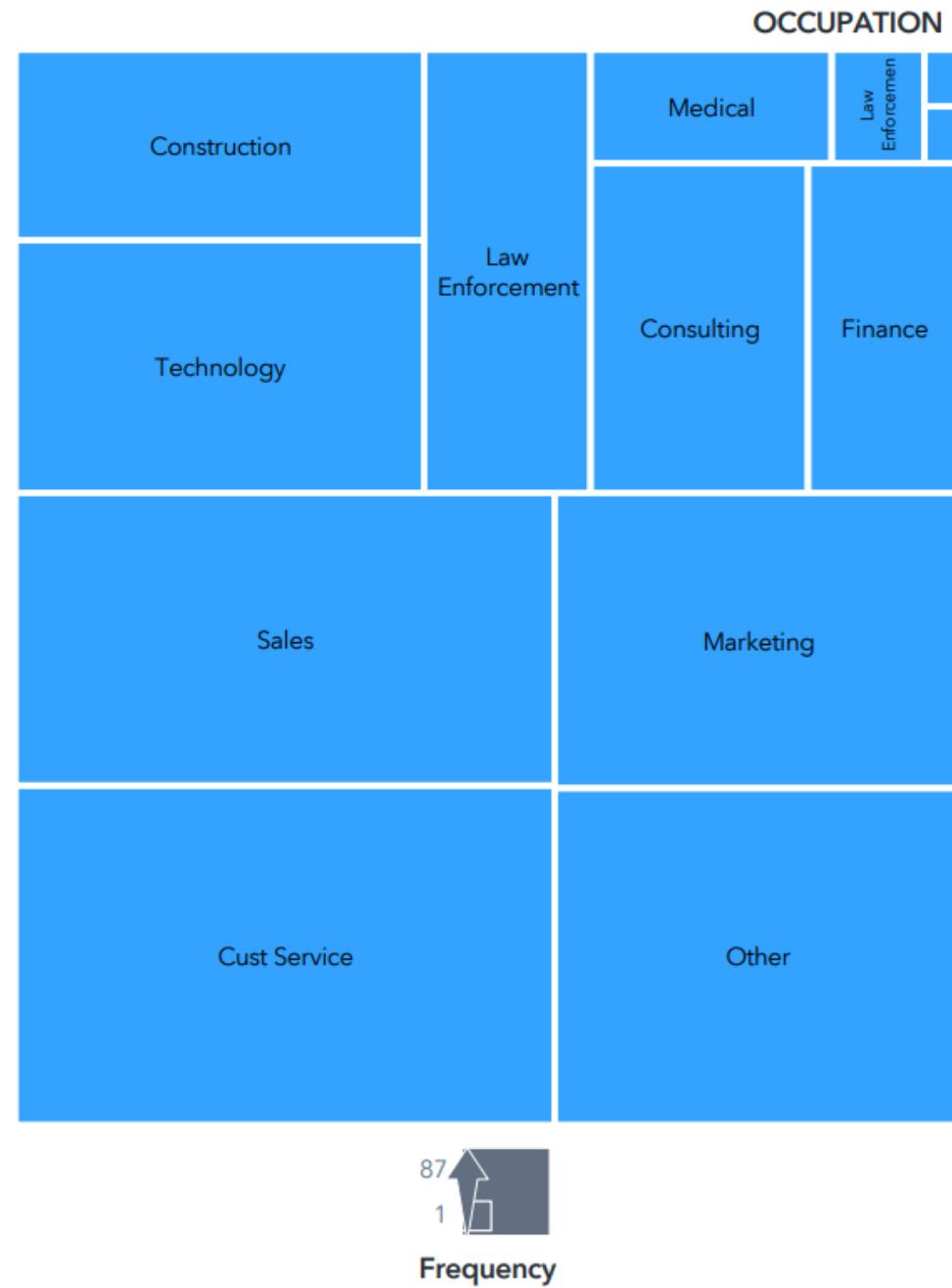
Shri Vile Parle Kelavani Mandal's

## Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

### Frequency of OCCUPATION



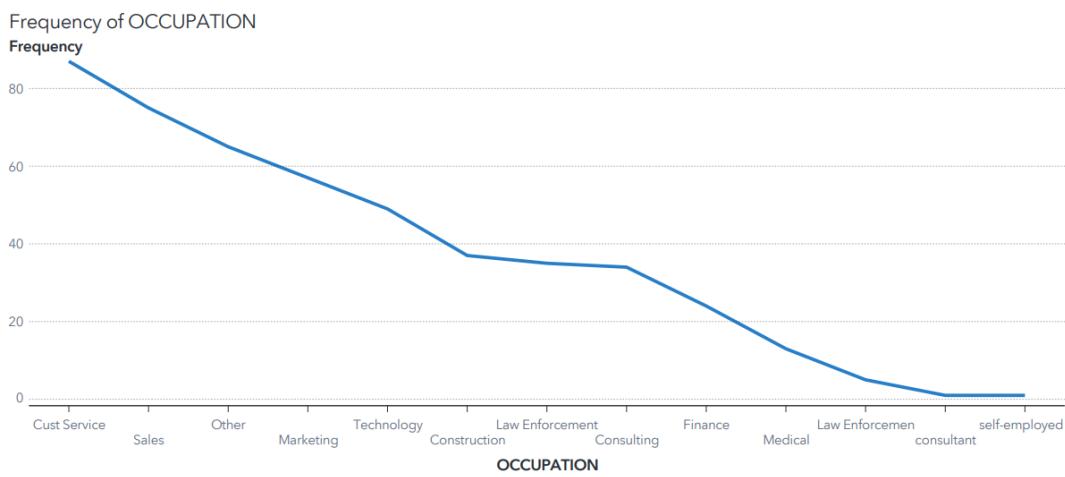
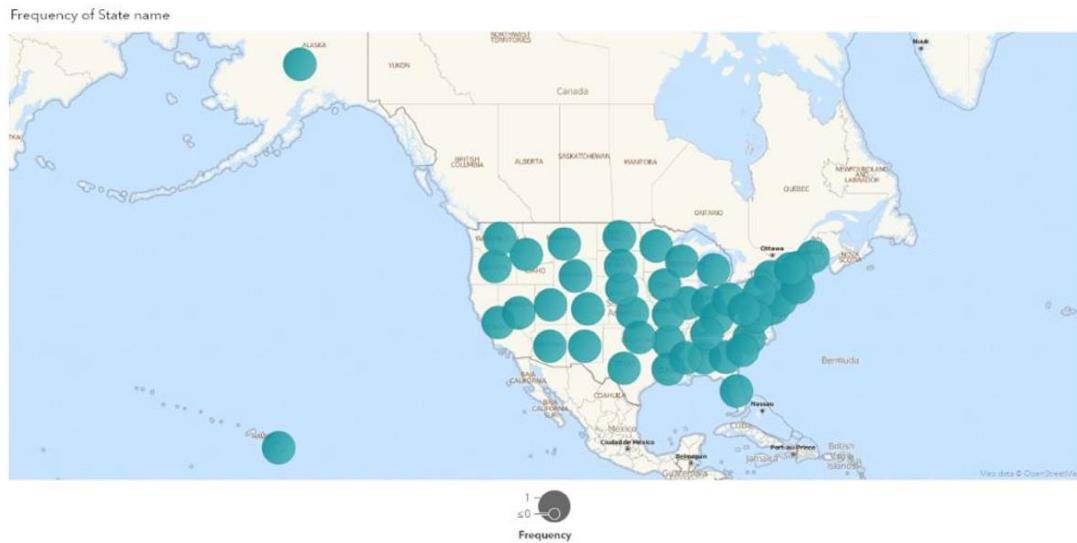


Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)



**Conclusion:** In this experiment, we learnt and explored SAS visual analytics, we also visualized the bank customer's dataset and made various graphs and charts like bar chart, pie chart, geographical chart, tree map and word cloud.



## EXPERIMENT 6

---

**SAP ID:** 60004190013

**Name:** Aryan Parekh

**Division:** A

**Batch:** A1

**Subject:** Business Analytics

**Semester:** VI

**Date of Submission:** 6th April, 2022

---

**Aim:** To create data by applying filters in SAS Studio and performing data analysis on it.

**Theory:**

The WHERE= option:

Selects observations from SAS data sets that meet a particular condition. The WHERE= option allows one to select only those observations from a SAS data set that meet a certain condition. Just as is true for the KEEP= and DROP= options, the WHERE= option can be attached to the SET statement or the DATA statement. If the WHERE= option is attached to the SET statement, SAS selects the observations that meet the condition as it reads in the data. If the WHERE= option is attached to the DATA statement, SAS selects the observations as it writes the data from the program data vector to the output data set.

The DROP= and KEEP= options:

The DROP= option tells SAS which variables you want to drop from a data set. If you place the DROP= option on the SET statement, SAS drops the specified variables when it reads the input data set. On the other hand, if you place the DROP= option on the DATA statement, SAS drops the specified variables when it writes to the output data set.

The KEEP= option tells SAS which variables you want to keep in a data set. If you place the KEEP= option on the SET statement, SAS keeps the specified variables when it reads the input data set. On the other hand, if you place the KEEP= option on the DATA statement, SAS keeps the specified variables when it writes to the output data set.

The IN= option:

The IN= option tells SAS to create an "indicator variable" that takes on either the value 0 or 1 depending on whether or not the current observation comes from the input data set. If the observation does not come from the input data set, then the indicator variable takes on the value 0. If the observation does come from the input data set, then the indicator takes on the value 1. The IN= option is especially useful when merging and concatenating data sets which we'll study in the next two lessons.

**Macro Variables:**

A macro variable in SAS is a string variable that allows you to dynamically modify the text in a SAS program through symbolic substitution.

**PROC SORT:**

The SORT procedure orders SAS data set observations by the values of one or more character or numeric variables. The SORT procedure either replaces the original data set or creates a new data set. PROC SORT produces only an output data set.

**Code and Output:**

```
PROC IMPORT DATAFILE='/home/u61126016/My_Dataset/Student_Marks.csv'
  DBMS=CSV
  OUT=Student_Marks REPLACE;
  GETNAMES=YES;
RUN;
PROC PRINT DATA=Student_Marks;
  TITLE 'STUDENT MARKS TABLE';
RUN;
```

**STUDENT MARKS TABLE**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	60004190013	Aryan Parekh	m	97	93	82
2	60004190060	Kartik Suvarna	m	96	91	85
3	60004190066	Mayav Antani	m	92	78	89
4	60004190059	Karishni Mehta	f	95	82	71
5	60004190061	Khushi Chavan	f	47	96	83
6	60004190048	Ishita Kheria	f	75	98	93
7	60004190034	Divyashree Vaidya	f	89	91	99
8	60004190067	Megh Dedhia	m	10	15	20
9	60004190011	Akshath Mahajan	m	50	60	55
10	60004190003	Adithya Sanyal	m	65	70	73
11	60004190025	Deap Daru	m	75	78	83
12	60004190004	Aditi Kamath	f	45	35	30

```
/* Using WHERE statement*/
```

```
PROC PRINT DATA=Student_Marks;
  TITLE 'STUDENT MARKS TABLE (MATHS SCORE MISSING)';
  WHERE Maths_Score IS MISSING;
RUN;
PROC PRINT DATA=Student_Marks;
  TITLE 'STUDENT MARKS TABLE (MATHS SCORE NOT MISSING)';
  WHERE Maths_Score IS NOT MISSING;
RUN;
PROC PRINT DATA=Student_Marks;
  TITLE 'STUDENT MARKS TABLE (USING IN STATEMENT)';
  WHERE Name IN ('Aryan Parekh','Divyashree Vaidya');
```



Shri Vile Parle Kelavani Mandal's

**Dwarkadas J. Sanghvi College of Engineering**

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

RUN;

**STUDENT MARKS TABLE (MATHS SCORE MISSING)**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
5	60004190061	Khushi Chavan	f	.	96	83
8	60004190067	Megh Dedhia	m	.	15	20

**STUDENT MARKS TABLE (MATHS SCORE NOT MISSING)**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	60004190013	Aryan Parekh	m	97	93	82
2	60004190060	Kartik Suvarna	m	96	91	85
3	60004190066	Mayav Antani	m	92	78	89
4	60004190059	Karishni Mehta	f	95	82	71
6	60004190048	Ishita Kheria	f	75	98	93
7	60004190034	Divyashree Vaidya	f	89	91	.
9	60004190011	Akshath Mahajan	m	50	60	55
10	60004190003	Adithya Sanyal	m	65	70	73
11	60004190025	Deap Daru	m	75	78	83
12	60004190004	Aditi Kamath	f	45	35	30

**STUDENT MARKS TABLE (USING IN STATEMENT)**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	60004190013	Aryan Parekh	m	97	93	82
7	60004190034	Divyashree Vaidya	f	89	91	.

/\* Using WHERE AND DROP IN SASHELP.IRIS \*/

DATA IRIS;

```
SET SASHELP.IRIS;
WHERE SPECIES IN ('Setosa','Versicolor');
DROP SEPALLENGTH SEPALWIDTH;
Creator='Aryan';
PetalDifference = PETALLENGTH-PETALWIDTH;
FORMAT PetalDifference DOLLAR10.;
```

RUN;

PROC PRINT DATA=IRIS;

TITLE 'IRIS DATASET (WHERE,DROP,NEW COLUMN,FORMAT)';

RUN;

**IRIS DATASET (WHERE, DROP, NEW COLUMN, FORMAT)**

Obs	Species	PetalLength	PetalWidth	Creator	PetalDifference
1	Setosa	14	2	Aryan	\$12
2	Setosa	14	3	Aryan	\$11
3	Setosa	10	2	Aryan	\$8
4	Setosa	17	5	Aryan	\$12
5	Setosa	13	2	Aryan	\$11
6	Setosa	16	2	Aryan	\$14
7	Setosa	14	2	Aryan	\$12
8	Setosa	14	1	Aryan	\$13
9	Setosa	13	2	Aryan	\$11
10	Setosa	16	6	Aryan	\$10
11	Setosa	13	2	Aryan	\$11
12	Setosa	16	2	Aryan	\$14
13	Setosa	14	3	Aryan	\$11
14	Setosa	16	2	Aryan	\$14
15	Setosa	19	2	Aryan	\$17
16	Setosa	16	2	Aryan	\$14
17	Setosa	12	2	Aryan	\$10
18	Setosa	11	1	Aryan	\$10
19	Setosa	12	2	Aryan	\$10
20	Setosa	19	4	Aryan	\$15
21	Setosa	14	2	Aryan	\$12
22	Setosa	14	2	Aryan	\$12
23	Setosa	16	4	Aryan	\$12
24	Setosa	14	2	Aryan	\$12
25	Setosa	15	4	Aryan	\$11
26	Setosa	14	2	Aryan	\$12
27	Setosa	15	4	Aryan	\$11
28	Setosa	15	1	Aryan	\$14
29	Setosa	14	2	Aryan	\$12
30	Setosa	15	2	Aryan	\$13
31	Setosa	17	4	Aryan	\$13
32	Setosa	15	2	Aryan	\$13
33	Setosa	14	2	Aryan	\$12
34	Setosa	13	2	Aryan	\$11
35	Setosa	15	2	Aryan	\$12

Note: This table continues up until 100 observations.

```
/* Using macro variable */
%LET student_gender=m;
PROC PRINT DATA=Student_Marks;
  TITLE 'STUDENT MARKS TABLE FOR GENDER=M (USING MACRO
VARIABLE)';
  WHERE Gender = "&student_gender";
  VAR Name Gender Maths_Score Physics_Score Chemistry_Score;
```



```
RUN;  
PROC MEANS DATA=Student_Marks;  
    TITLE 'STATISTICAL SUMMARY OF STUDENT MARKS TABLE (FREQ) FOR  
GENDER=M (USING MACRO VARIABLE)';  
    WHERE Gender = "&student_gender";  
    VAR Maths_Score Physics_Score Chemistry_Score;  
RUN;
```

**STUDENT MARKS TABLE FOR GENDER=M (USING MACRO VARIABLE)**

Obs	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	Aryan Parekh	m	97	93	82
2	Kartik Suvarna	m	96	91	85
3	Mayav Antani	m	92	78	89
8	Megh Dedhia	m	.	15	20
9	Akshath Mahajan	m	50	60	55
10	Adithya Sanyal	m	65	70	73
11	Deap Daru	m	75	78	83

**STATISTICAL SUMMARY OF STUDENT MARKS TABLE (FREQ) FOR GENDER=M (USING MACRO VARIABLE)**

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
Maths_Score	6	79.1666667	19.1563740	50.0000000	97.0000000
Physics_Score	7	69.2857143	26.5186368	15.0000000	93.0000000
Chemistry_Score	7	69.5714286	24.6025550	20.0000000	89.0000000

```
/* Using PROC SORT */  
PROC SORT DATA=Student_Marks;  
    BY SAP_ID;  
RUN;  
PROC PRINT DATA=Student_Marks;  
    TITLE 'STUDENT MARKS TABLE (ASCENDING SAP ID)';  
RUN;  
PROC SORT DATA=Student_Marks;  
    BY DESCENDING SAP_ID;  
RUN;  
PROC PRINT DATA=Student_Marks;  
    TITLE 'STUDENT MARKS TABLE (DESCENDING SAP ID)';  
RUN;  
PROC SORT DATA=Student_Marks;  
    BY DESCENDING NAME;  
RUN;  
PROC PRINT DATA=Student_Marks;  
    TITLE 'STUDENT MARKS TABLE (DESCENDING NAME)';  
RUN;
```



Shri Vile Parle Kelavani Mandal's

**Dwarkadas J. Sanghvi College of Engineering**

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

**STUDENT MARKS TABLE (ASCENDING SAP ID)**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	60004190003	Adithya Sanyal	m	65	70	73
2	60004190004	Aditi Kamath	f	45	35	30
3	60004190011	Akshath Mahajan	m	50	60	55
4	60004190013	Aryan Parekh	m	97	93	82
5	60004190025	Deap Daru	m	75	78	83
6	60004190034	Divyashree Vaidya	f	89	91	.
7	60004190048	Ishita Kheria	f	75	98	93
8	60004190059	Karishni Mehta	f	95	82	71
9	60004190080	Kartik Suvarna	m	96	91	85
10	60004190081	Khushi Chavan	f	.	96	83
11	60004190086	Mayav Antani	m	92	78	89
12	60004190087	Megh Dedhia	m	.	15	20

**STUDENT MARKS TABLE (DESCENDING SAP ID)**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	60004190087	Megh Dedhia	m	.	15	20
2	60004190086	Mayav Antani	m	92	78	89
3	60004190081	Khushi Chavan	f	.	96	83
4	60004190080	Kartik Suvarna	m	96	91	85
5	60004190059	Karishni Mehta	f	95	82	71
6	60004190048	Ishita Kheria	f	75	98	93
7	60004190034	Divyashree Vaidya	f	89	91	.
8	60004190025	Deap Daru	m	75	78	83
9	60004190013	Aryan Parekh	m	97	93	82
10	60004190011	Akshath Mahajan	m	50	60	55
11	60004190004	Aditi Kamath	f	45	35	30
12	60004190003	Adithya Sanyal	m	65	70	73

**STUDENT MARKS TABLE (DESCENDING NAME)**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	60004190087	Megh Dedhia	m	.	15	20
2	60004190086	Mayav Antani	m	92	78	89
3	60004190081	Khushi Chavan	f	.	96	83
4	60004190080	Kartik Suvarna	m	96	91	85
5	60004190059	Karishni Mehta	f	95	82	71
6	60004190048	Ishita Kheria	f	75	98	93
7	60004190034	Divyashree Vaidya	f	89	91	.
8	60004190025	Deap Daru	m	75	78	83
9	60004190013	Aryan Parekh	m	97	93	82
10	60004190011	Akshath Mahajan	m	50	60	55
11	60004190004	Aditi Kamath	f	45	35	30
12	60004190003	Adithya Sanyal	m	65	70	73



```
/* Using SORT and saving in another SAS table */
PROC SORT DATA=Student_Marks out=Student_Marks_2;
    BY Maths_Score DESCENDING Sap_Id;
RUN;
PROC PRINT DATA=Student_Marks_2;
    TITLE 'STUDENT MARKS TABLE (ASCENDING MATHS SCORE +
DESCENDING SAP ID)';
RUN;
```

**STUDENT MARKS TABLE (ASCENDING MATHS SCORE + DESCENDING SAP ID)**

Obs	SAP_ID	Name	Gender	Maths_Score	Physics_Score	Chemistry_Score
1	60004190067	Megh Dedhia	m	.	15	20
2	60004190061	Khushi Chavan	f	.	96	83
3	60004190004	Aditi Kamath	f	45	35	30
4	60004190011	Akshath Mahajan	m	50	60	55
5	60004190003	Adithya Sanyal	m	65	70	73
6	60004190048	Ishita Kheria	f	75	98	93
7	60004190025	Deap Daru	m	75	78	83
8	60004190034	Divyashree Vaidya	f	89	91	.
9	60004190066	Mayav Antani	m	92	78	89
10	60004190059	Karishni Mehta	f	95	82	71
11	60004190060	Kartik Suvarna	m	96	91	85
12	60004190013	Aryan Parekh	m	97	93	82

**Conclusion:** From this experiment, we learn how to create data by applying filters in SAS Studio and performing data analysis on it.



## EXPERIMENT 7

**SAP ID:** 60004190013**Name:** Aryan Parekh**Division:** A**Batch:** A1**Subject:** Business Analytics**Semester:** VI**Date of Submission:** 10<sup>th</sup> May, 2022**Aim:** To work with graph level display rule in SAS Studio/Visual Analytics.**Code and Output:**

The screenshot shows the SAS Visual Analytics interface. On the left, there are navigation panels for Data, Objects, Suggest, Outline, and Review. The main area displays a table titled 'Report 2' with columns: ACCOUNT\_NO, CUSTOMER\_NAME, GENDER, JOB\_STATUS, and HH\_INCOME. The data consists of 25 rows of customer information. On the right, a 'Display Rules' panel is open, showing a dropdown menu 'List table - ACCOUNT\_NO 2'. Below it, there are sections for 'Table Rows' and 'Report'. Under 'Report', there is a section for 'GENDER' with three color-coded boxes: blue for 'M', pink for 'F', and cyan for 'Other'. A note says 'Drop a color mapped value rule here to make it a report level rule'. The bottom of the interface shows a toolbar with various icons.

ACCOUNT_NO	CUSTOMER_NAME	GENDER	JOB_STATUS	HH_INCOME
350001131	Stephen seiters		parttime	15000
350003321	Doug mArtisclan	M	fulltime	26000
350003541	Johnnie Rheaves	M	fulltime	13000
350004321	Daniel Zill	M	fulltime	99000
350004421	Patty Bernasconi	F	fulltime	85000
350005461	Donald Sherretts	M	fulltime	30000
350005521	Ed Gompf	M	fulltime	92000
350007931	Danny Dales	M	parttime	44000
350008843	Garth Skiffington		fulltime	32668
350008855	Alejandra Prenatt	F	fulltime	159000
350008867	Brendan Qin	M	parttime	31000
350008870	Maryanne Whymar	F	fulltime	21000
350008879	Chase Furler	M	fulltime	12000
350008891	Cristopher Wiget	M	fulltime	84000
350008903	Marc Wanger	M	fulltime	27368
350008915	Quintin Isaacson	M	fulltime	23000
350008927	Leslie Cackowski	F	fulltime	30000
350008939	Marietta Bjornberg	F	fulltime	94000
350008951	Lacy woodfin	F	fulltime	21000
350008963	Lucius Winchester	M	fulltime	10000
350008975	Damion Matkin	M	fulltime	28913



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

SAS® Visual Analytics - Explore and Visualize

## Report 2

Editing

Table Level Display Rule (Categorical) Table Level Display Rule (Numerical) Graph Level Display Rule (Numerical) Graph Level Display Rule 2 (Numerical) +

Data Objects Suggest Outline Review

Display Rules

List table - ACCOUNT\_NO

+ New rule Table Rows

HH\_INCOME

HH\_INCOME > 60000

Rules Actions Filters Ranks

ACCOUNT_NO	CUSTOMER_NAME	OCCUPATION	HH_INCOME
350011961	Eileen Bauchter	Sales	98000
350011975	Vivian Brzostowski	Technology	14000
350011987	Milton Kuhlman	Consulting	20000
350011999	Mary Maurizio	Construction	28000
350012011	Elmo Gabour	Law Enforcement	17000
350012023	Rachel Larison	Technology	18000
350012035	Dwayne Maddalena	Medical	57000
350012047	Coleman Cuneo	Cust Service	24521
350012059	Caitlin Reigner	Other	32668
350012071	Logan Muhl	Technology	12000
350012083	Callie Leboeuf	Cust Service	50922
350012095	Mitzi Ihenyen	Technology	17000
350012107	Alyssa Biasotti	Technology	96000
350012119	Janet Schaffter	Technology	23000
350012131	Randal Gansen	Marketing	55000
350012143	Gil Scarpa	Construction	62000
350012155	Adela Cervantsz	Law Enforcement	47309
350012167	Armando Kolm	Marketing	203000
350012179	Vanessa Lewallen	Sales	17000
350012191	Rosalia Kennemur	Marketing	203000
350012203	Tammie Schwartzen		15000

15 Days Left

SAS® Visual Analytics - Explore and Visualize

## Report 2

Editing

Table Level Display Rule (Categorical) Table Level Display Rule (Numerical) Graph Level Display Rule (Numerical) Graph Level Display Rule 2 (Numerical) +

Data Objects Suggest Outline Review

Display Rules

Bar - STATE 1

+ New rule Object

HH\_INCOME

HH\_INCOME <= 75000

HH\_INCOME

HH\_INCOME > 75000

Rules Actions Filters Ranks

HH\_INCOME by STATE

STATE

STATE	HH_INCOME (millions)
WY	~0.2
WA	~0.5
TN	~0.8
SC	~0.3
OK	~0.6
north carolina	~1.0
New Jersey	~0.2
NC	~0.3
MN	~0.2
LA	~0.5
IL	~1.3
GA	~0.7
DE	~0.4
California	~2.8
AL	~0.7

HH\_INCOME (millions)

15 Days Left



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)



SAS® Visual Analytics - Explore and Visualize

### Report 2

Editing

Table Level Display Rule (Numerical) Graph Level Display Rule (Numerical) Graph Level Display Rule 2 (Numerical) Table Level Display Rule (Numeric)(Guage) +

ACCOUNT\_NO ▲ CUSTOMER\_NAME HH\_INCOME

ACCOUNT_NO	CUSTOMER_NAME	HH_INCOME
350010667	George Tukis	19000
350010679	Jimmie Hardgrove	33000
350010691	Tisha Gorder	28913
350010703	Clarice Knower	81000
350010715	Renae Eldridge	79000
350010727	Floyd Veazey	127000
350010739	Sybil Marmerchant	91000
350010751	Reyna Bangle	47000
350010763	Rodrigo Wildrick	25000
350010775	Taylor Fogerty	62000
350010787	Reva Lecates	25000
350010799	Eli Bettner	14000
350010811	Owen Sparacino	16000
350010823	Forest Orea	56000
350010835	Martin Carley	26000
350010847	Derek Kreutzbender	72000
350010859	Deanna Gerbi	115000
350010871	Alfonso Griglen	48000
350010883	Sadie Rowlett	54000
350010895	Ella Pahnke	44000
350010907	Miranda Hammitt	35000

Display Rules

List table - ACCOUNT\_NO 3

+ New rule Column - HH\_INCOME

HH\_INCOME Left of text, with HH\_INCOME as x

0 <= x < 50,000

50,000 <= x < 100,000

100,000 <= x <= 200,000



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

FINAL REPORT:

## Display Rules

Creation Date: Tuesday, 10 May 2022 11:22:10  
Author: aryanparekh2412@gmail.com

6 Pages



Table Level Display Rule (Categorical)

ACCOUNT_NO ▲	CUSTOMER_NAME	GENDER	JOB_STATUS	HH_INCOME
350001131	Stephen seiters		parttime	15000
350003321	Doug mAtrisciano	M	fulltime	26000
350003541	Johnnie Rheaves	M	fulltime	13000
350004321	Daniel Zill	M	fulltime	99000
350004421	Patty Bernasconi	F	fulltime	85000
350005461	Donald Sherretts	M	fulltime	30000
350005521	Ed Gompf	M	fulltime	92000
350007931	Danny Dales	M	parttime	44000
350008843	Garth Skiffington		fulltime	32668
350008855	Alejandra Prenatt	F	fulltime	159000
350008867	Brendan Qin	M	parttime	31000
350008870	Maryanne Whyman	F	fulltime	21000
350008879	Chase Furler	M	fulltime	12000
350008891	Cristophe Wiget	M	fulltime	84000
350008903	Marc Wanger	M	fulltime	27368
350008915	Quintin Isacson	M	fulltime	23000
350008927	Leslie Cackowski	F	fulltime	30000
350008939	Marietta Bjornberg	F	fulltime	94000
350008951	Lacy woodfin	F	fulltime	21000
350008963	Lucius Winchester	M	fulltime	10000
350008975	Damion Matkin	M	fulltime	28913
350008987	Petra Mcnichol	F	fulltime	25000
350008999	Robin Grotz	U	fulltime	26000
350009011	Reina Wolchesky		fulltime	40000
350009023	Katina Ramano	F	fulltime	78000

A1.1



Table Level Display Rule (Numerical)

ACCOUNT_NO ▲	CUSTOMER_NAME	OCCUPATION	HH_INCOME
350011951	Efren Baucher	Other	13000
350011963	Naomi Greenly	Sales	98000
350011975	Vivian Brzostowski	Technology	14000
350011987	Milton Kuhlman	Consulting	20000
350011999	Mary Maurizio	Construction	28000
350012011	Elmo Gabouer	Law Enforcement	17000
350012023	Rachel Garrison	Technology	18000
350012035	Dwayne Maddalena	Medical	57000
350012047	Coleman Cuneo	Cust Service	24521
350012059	Caitlin Reiniger	Other	32668
350012071	Logan Muhl	Technology	12000
350012083	Callie Leboeuf	Cust Service	50922
350012095	Mitzi Ihenyen	Technology	17000
350012107	Alyssa Biasotti	Technology	96000
350012119	Janet Schaffter	Technology	23000
350012131	Randal Gansen	Marketing	55000
350012143	Gil Scarpa	Construction	62000
350012155	Adela Cervantsz	Law Enforcement	47309
350012167	Armando Kolm	Marketing	203000
350012179	Vanessa Lewallen	Sales	17000
350012191	Rosalia Kennemur	Marketing	203000
350012203	Tammie Schwartzwal	Other	15000
350012215	Darrin Neiss	Finance	37000
350012227	Rosalie Kriger	Cust Service	129000
350012239	Burton Brining	Marketing	28000

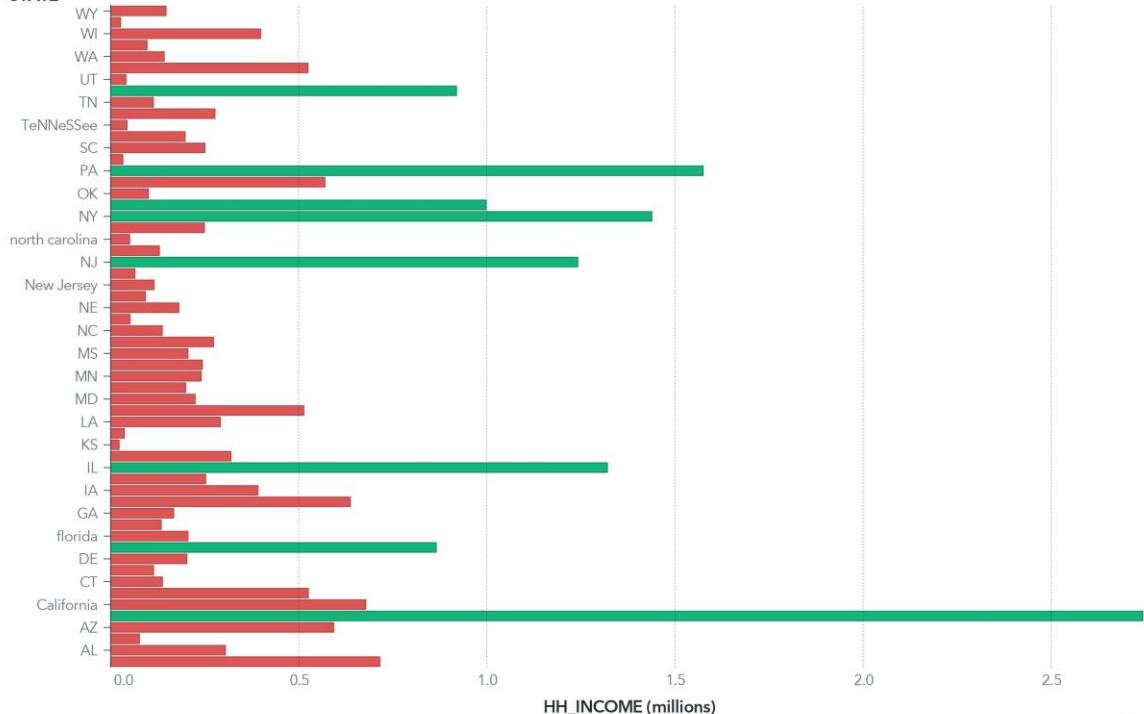
A2.1



## Graph Level Display Rule (Numerical)

## HH\_INCOME by STATE

## STATE

[A3.1](#)



Shri Vile Parle Kelavani Mandal's

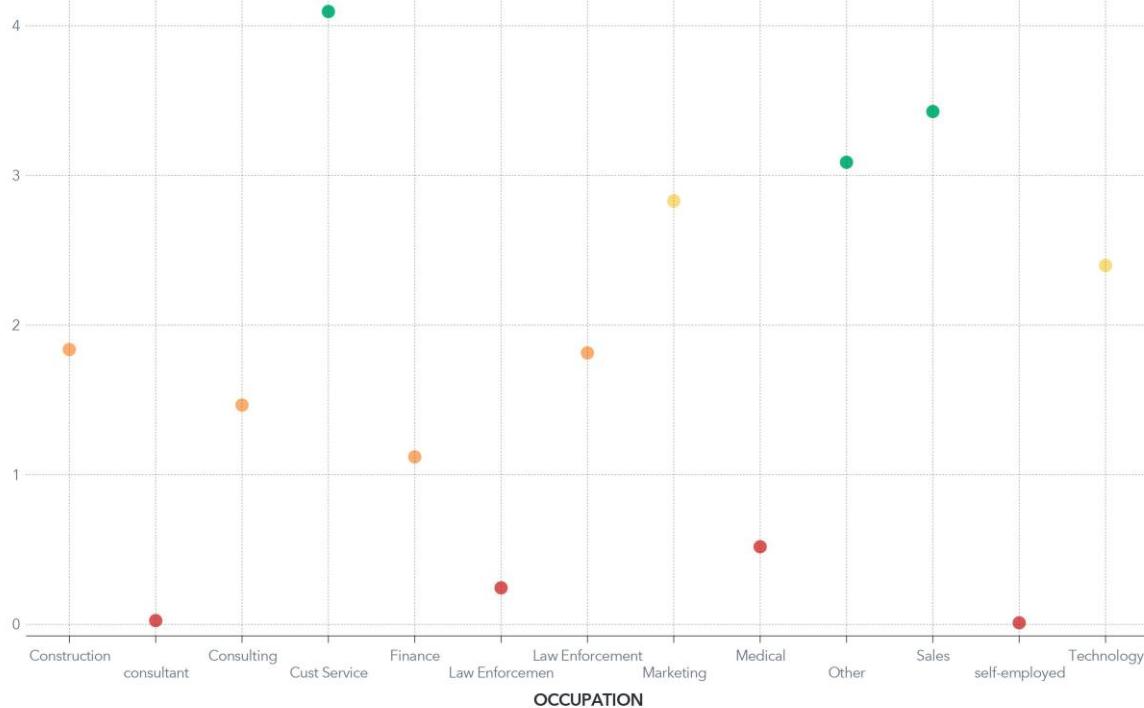
# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

Graph Level Display Rule 2 (Numerical)

HH\_INCOME by OCCUPATION  
HH\_INCOME (millions)



A4.1



Table Level Display Rule (Numeric)(Guage)

ACCOUNT_NO ▲	CUSTOMER_NAME	HH_INCOME
350010655	Emilio Lampkin	43360
350010667	George Tukis	19000
350010679	Jimmie Hardgrove	33000
350010691	Tisha Gorder	28913
350010703	Clarice Knower	81000
350010715	Renae Eldridge	79000
350010727	Floyd Veazey	127000
350010739	Sybil Marmerchant	91000
350010751	Reyna Bangle	47000
350010763	Rodrigo Wildrick	25000
350010775	Taylor Fogerty	62000
350010787	Reva Lecates	25000
350010799	Eli Bettner	14000
350010811	Owen Sparacino	16000
350010823	Forest Orea	56000
350010835	Martin Carley	26000
350010847	Derek Kreutzbender	72000
350010859	Deanna Gerbi	115000
350010871	Alfonso Griglen	48000
350010883	Sadie Rowlett	54000
350010895	Ella Pahnke	44000
350010907	Miranda Hammitt	35000
350010919	Vernon Engelman	13000
350010931	Veronica Achorn	28000
350010943	Kory Wooldridge	62000

[A5.1](#)



Shri Vile Parle Kelavani Mandal's

# Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

## Appendix

### A1.1 List table - ACCOUNT\_NO 2

Display Rules: GENDER

█ M

█ F

█ Other

### A2.1 List table - ACCOUNT\_NO 1

Display Rules: HH\_INCOME

abc HH\_INCOME > 60000

### A3.1 HH\_INCOME by STATE

Display Rules: HH\_INCOME

█ HH\_INCOME ≤ 750000

█ HH\_INCOME > 750000

### A4.1 HH\_INCOME by OCCUPATION

Display Rules: HH\_INCOME

█ 3000000 ≤ HH\_INCOME ≤ 4095104

█ 2000000 ≤ HH\_INCOME ≤ 3000000

█ 1000000 ≤ HH\_INCOME ≤ 2000000

█ 0 ≤ HH\_INCOME ≤ 1000000

### A5.1 List table - ACCOUNT\_NO 3

Display Rules: HH\_INCOME

Left of text, with HH\_INCOME as x

█ 0 ≤ x < 50000

█ 50000 ≤ x < 100000

█ 100000 ≤ x ≤ 200000

**Conclusion:** From this experiment, we learn how to work with graph level display rule in SAS Studio/Visual Analytics.

**MINI PROJECT****SAP ID:** 60004190013**Name:** Aryan Parekh**SAP ID:** 60004190015**Name:** Aryan Trivedi**Division:** A**Batch:** A1**Subject:** Business Analytics**Semester:** VI**Date of Submission:** 2<sup>nd</sup> June, 2022

**Aim:** To analyze the data obtained from an E-commerce website and get meaningful insights out of it.

**Idea:** We have 1425 rows of E-commerce data with attributes like: InvoiceNo, Description, Quantity, InvoiceDateTime, UnitPrice, CustomerID, and Country.

This data will first be passed through SAS Studio for processing, and then the data obtained from there will be passed on to SAS Visual Analytics for further analysis in forms of different charts and tables.

Originally the data looks like:

InvoiceNo	Description	Quantity	InvoiceDateTime	UnitPrice	CustomerID	Country
536365	WHITE HANGING HEART T-LIGHT HOLDER	6	12-01-2010 08:26	2.55	17850	United Kingdom
536365	WHITE METAL LANTERN	6	12-01-2010 08:26	3.39	17850	United Kingdom
536365	CREAM CUPID HEARTS COAT HANGER	8	12-01-2010 08:26	2.75	17850	United Kingdom
536365	KNITTED UNION FLAG HOT WATER BOTTLE	6	12-01-2010 08:26	3.39	17850	United Kingdom
536365	RED WOOLLY HOTTIE WHITE HEART.	6	12-01-2010 08:26	3.39	17850	United Kingdom
536365	SET 7 BABUSHKA NESTING BOXES	2	12-01-2010 08:26	7.65	17850	United Kingdom
536365	GLASS STAR FROSTED T-LIGHT HOLDER	6	12-01-2010 08:26	4.25	17850	United Kingdom
536366	HAND WARMER UNION JACK	6	12-01-2010 08:28	1.85	17850	United Kingdom
536366	HAND WARMER RED POLKA DOT	6	12-01-2010 08:28	1.85	17850	United Kingdom
536367	ASSORTED COLOUR BIRD ORNAMENT	32	12-01-2010 08:34	1.69	13047	United Kingdom
536367	POPPY'S PLAYHOUSE BEDROOM	6	12-01-2010 08:34	2.1	13047	United Kingdom
536367	POPPY'S PLAYHOUSE KITCHEN	6	12-01-2010 08:34	2.1	13047	United Kingdom
536367	FELTCRAFT PRINCESS CHARLOTTE DOLL	8	12-01-2010 08:34	3.75	13047	United Kingdom
536367	IVORY KNITTED MUG COSY	6	12-01-2010 08:34	1.65	13047	United Kingdom
536367	BOX OF 6 ASSORTED COLOUR TEASPOONS	6	12-01-2010 08:34	4.25	13047	United Kingdom
536367	BOX OF VINTAGE JIGSAW BLOCKS	3	12-01-2010 08:34	4.95	13047	United Kingdom
536367	BOX OF VINTAGE ALPHABET BLOCKS	2	12-01-2010 08:34	9.95	13047	United Kingdom
536367	HOME BUILDING BLOCK WORD	3	12-01-2010 08:34	5.95	13047	United Kingdom
536367	LOVE BUILDING BLOCK WORD	3	12-01-2010 08:34	5.95	13047	United Kingdom
536367	RECIPE BOX WITH METAL HEART	4	12-01-2010 08:34	7.95	13047	United Kingdom
536367	DOORMAT NEW ENGLAND	4	12-01-2010 08:34	7.95	13047	United Kingdom
536368	JAM MAKING SET WITH JARS	6	12-01-2010 08:34	4.25	13047	United Kingdom
536368	RED COAT RACK PARIS FASHION	3	12-01-2010 08:34	4.95	13047	United Kingdom
536368	YELLOW COAT RACK PARIS FASHION	3	12-01-2010 08:34	4.95	13047	United Kingdom
536368	BLUE COAT RACK PARIS FASHION	3	12-01-2010 08:34	4.95	13047	United Kingdom
536369	BATH BUILDING BLOCK WORD	3	12-01-2010 08:35	5.95	13047	United Kingdom
536370	ALARM CLOCK BAKELIKE PINK	24	12-01-2010 08:45	3.75	12583	France
536370	ALARM CLOCK BAKELIKE RED	24	12-01-2010 08:45	3.75	12583	France
536370	ALARM CLOCK BAKELIKE GREEN	12	12-01-2010 08:45	3.75	12583	France

**Code and Output (PART 1):**

```
/* Importing the E-commerce Dataset */
PROC IMPORT DATAFILE='/home/u61126016/My_Dataset/Ecommerce_data.xlsx'
    DBMS=XLSX
    OUT=ECOMMERCE REPLACE;
    GETNAMES=YES;
RUN;

/* Printing the 1st 10 observations */
PROC PRINT DATA=ECOMMERCE (OBS=10);
    TITLE 'E-COMMERCE DATA';
RUN;
```

**E-COMMERCE DATA**

Obs	InvoiceNo	Description	Quantity	InvoiceDateTime	UnitPrice	CustomerID	Country
1	536365	WHITE HANGING HEART T-LIGHT HOLDER	6	12JAN10:08:26:00	2.55	17850	United Kingdom
2	536365	WHITE METAL LANTERN	6	12JAN10:08:26:00	3.39	17850	United Kingdom
3	536365	CREAM CUPID HEARTS COAT HANGER	8	12JAN10:08:26:00	2.75	17850	United Kingdom
4	536365	KNITTED UNION FLAG HOT WATER BOTTLE	6	12JAN10:08:26:00	3.39	17850	United Kingdom
5	536365	RED WOOLLY HOTTIE WHITE HEART.	6	12JAN10:08:26:00	3.39	17850	United Kingdom
6	536365	SET 7 BABUSHKA NESTING BOXES	2	12JAN10:08:26:00	7.65	17850	United Kingdom
7	536365	GLASS STAR FROSTED T-LIGHT HOLDER	6	12JAN10:08:26:00	4.25	17850	United Kingdom
8	536366	HAND WARMER UNION JACK	6	12JAN10:08:28:00	1.85	17850	United Kingdom
9	536366	HAND WARMER RED POLKA DOT	6	12JAN10:08:28:00	1.85	17850	United Kingdom
10	536367	ASSORTED COLOUR BIRD ORNAMENT	32	12JAN10:08:34:00	1.69	13047	United Kingdom

```
/* Removing rows with No CustomerID */
DATA ECOMMERCE;
    SET WORK.ECOMMERCE;
    WHERE CustomerID IS NOT MISSING;

/* Adding Dollar Format to UnitPrice */
DATA ECOMMERCE;
    SET WORK.ECOMMERCE;
    FORMAT UnitPrice DOLLAR10.2;
RUN;

/* Printing the 1st 10 observations */
PROC PRINT DATA=ECOMMERCE (OBS=10);
    TITLE 'E-COMMERCE DATA (DOLLAR FORMAT)';
RUN;
```

**E-COMMERCE DATA (DOLLAR FORMAT)**

Obs	InvoiceNo	Description	Quantity	InvoiceDateTime	UnitPrice	CustomerID	Country
1	536365	WHITE HANGING HEART T-LIGHT HOLDER	6	12JAN10:08:26:00	\$2.55	17850	United Kingdom
2	536365	WHITE METAL LANTERN	6	12JAN10:08:26:00	\$3.39	17850	United Kingdom
3	536365	CREAM CUPID HEARTS COAT HANGER	8	12JAN10:08:26:00	\$2.75	17850	United Kingdom
4	536365	KNITTED UNION FLAG HOT WATER BOTTLE	6	12JAN10:08:26:00	\$3.39	17850	United Kingdom
5	536365	RED WOOLLY HOTTIE WHITE HEART.	6	12JAN10:08:26:00	\$3.39	17850	United Kingdom
6	536365	SET 7 BABUSHKA NESTING BOXES	2	12JAN10:08:26:00	\$7.65	17850	United Kingdom
7	536365	GLASS STAR FROSTED T-LIGHT HOLDER	6	12JAN10:08:26:00	\$4.25	17850	United Kingdom
8	536366	HAND WARMER UNION JACK	6	12JAN10:08:28:00	\$1.85	17850	United Kingdom
9	536366	HAND WARMER RED POLKA DOT	6	12JAN10:08:28:00	\$1.85	17850	United Kingdom
10	536367	ASSORTED COLOUR BIRD ORNAMENT	32	12JAN10:08:34:00	\$1.69	13047	United Kingdom

/\* Calculating TotalPrice using UnitPrice and Quantity \*/

DATA ECOMMERCE;

SET WORK. ECOMMERCE;

TotalPrice = Quantity\*UnitPrice;

FORMAT TotalPrice DOLLAR10.2;

RUN;

/\* Printing the 1st 10 observations \*/

PROC PRINT DATA=ECOMMERCE (OBS=10);

TITLE 'E-COMMERCE DATA (NEW FIELD ADDED)';

RUN;

**E-COMMERCE DATA (NEW FIELD ADDED)**

Obs	InvoiceNo	Description	Quantity	InvoiceDateTime	UnitPrice	CustomerID	Country	TotalPrice
1	536365	WHITE HANGING HEART T-LIGHT HOLDER	6	12JAN10:08:26:00	\$2.55	17850	United Kingdom	\$15.30
2	536365	WHITE METAL LANTERN	6	12JAN10:08:26:00	\$3.39	17850	United Kingdom	\$20.34
3	536365	CREAM CUPID HEARTS COAT HANGER	8	12JAN10:08:26:00	\$2.75	17850	United Kingdom	\$22.00
4	536365	KNITTED UNION FLAG HOT WATER BOTTLE	6	12JAN10:08:26:00	\$3.39	17850	United Kingdom	\$20.34
5	536365	RED WOOLLY HOTTIE WHITE HEART.	6	12JAN10:08:26:00	\$3.39	17850	United Kingdom	\$20.34
6	536365	SET 7 BABUSHKA NESTING BOXES	2	12JAN10:08:26:00	\$7.65	17850	United Kingdom	\$15.30
7	536365	GLASS STAR FROSTED T-LIGHT HOLDER	6	12JAN10:08:26:00	\$4.25	17850	United Kingdom	\$25.50
8	536366	HAND WARMER UNION JACK	6	12JAN10:08:28:00	\$1.85	17850	United Kingdom	\$11.10
9	536366	HAND WARMER RED POLKA DOT	6	12JAN10:08:28:00	\$1.85	17850	United Kingdom	\$11.10
10	536367	ASSORTED COLOUR BIRD ORNAMENT	32	12JAN10:08:34:00	\$1.69	13047	United Kingdom	\$54.08

/\* Splitting InvoiceDateTime to InvoiceDate and InvoiceTime \*/

DATA ECOMMERCE;

SET WORK. ECOMMERCE;

InvoiceDate=datepart(InvoiceDateTime);

InvoiceTime=timepart(InvoiceDateTime);

DROP InvoiceDateTime;

FORMAT InvoiceDate DDMMYYY10.;

FORMAT InvoiceTime TIMEAMPM.;

RUN;



/\* Printing the 1st 10 observations \*/

PROC PRINT DATA=ECOMMERCE (OBS=10);

    TITLE 'E-COMMERCE DATA (DATETIME FIELD SPLIT INTO DATE AND TIME)';

RUN;

**E-COMMERCE DATA (DATETIME FIELD SPLIT INTO DATE AND TIME)**

Obs	InvoiceNo	Description	Quantity	UnitPrice	CustomerID	Country	TotalPrice	InvoiceDate	InvoiceTime
1	536365	WHITE HANGING HEART T-LIGHT HOLDER	6	\$2.55	17850	United Kingdom	\$15.30	12/01/2010	8:26:00 AM
2	536365	WHITE METAL LANTERN	6	\$3.39	17850	United Kingdom	\$20.34	12/01/2010	8:26:00 AM
3	536365	CREAM CUPID HEARTS COAT HANGER	8	\$2.75	17850	United Kingdom	\$22.00	12/01/2010	8:26:00 AM
4	536365	KNITTED UNION FLAG HOT WATER BOTTLE	6	\$3.39	17850	United Kingdom	\$20.34	12/01/2010	8:26:00 AM
5	536365	RED WOOLLY HOTTIE WHITE HEART.	6	\$3.39	17850	United Kingdom	\$20.34	12/01/2010	8:26:00 AM
6	536365	SET 7 BABUSHKA NESTING BOXES	2	\$7.65	17850	United Kingdom	\$15.30	12/01/2010	8:26:00 AM
7	536365	GLASS STAR FROSTED T-LIGHT HOLDER	6	\$4.25	17850	United Kingdom	\$25.50	12/01/2010	8:26:00 AM
8	536366	HAND WARMER UNION JACK	6	\$1.85	17850	United Kingdom	\$11.10	12/01/2010	8:28:00 AM
9	536366	HAND WARMER RED POLKA DOT	6	\$1.85	17850	United Kingdom	\$11.10	12/01/2010	8:28:00 AM
10	536367	ASSORTED COLOUR BIRD ORNAMENT	32	\$1.69	13047	United Kingdom	\$54.08	12/01/2010	8:34:00 AM

/\* Creating a new table with transactions of country mentioned in MACRO variable \*/

%LET chosen\_country=Australia;

DATA ECOMMERCE PARTICULAR\_COUNTRY;

    SET WORK.ECOMMERCE;

    WHERE Country = "&chosen\_country";

RUN;

/\* Printing observations \*/

PROC PRINT DATA=ECOMMERCE PARTICULAR\_COUNTRY;

    TITLE "E-COMMERCE DATA FOR &chosen\_country";

RUN;

**E-COMMERCE DATA FOR Australia**

Obs	InvoiceNo	Description	Quantity	UnitPrice	CustomerID	Country	TotalPrice	InvoiceDate	InvoiceTime
1	536389	CHRISTMAS LIGHTS 10 REINDEER	6	\$8.50	12431	Australia	\$51.00	12/01/2010	10:03:00 AM
2	536389	VINTAGE UNION JACK CUSHION COVER	8	\$4.95	12431	Australia	\$39.60	12/01/2010	10:03:00 AM
3	536389	VINTAGE HEADS AND TAILS CARD GAME	12	\$1.25	12431	Australia	\$15.00	12/01/2010	10:03:00 AM
4	536389	SET OF 3 COLOURED FLYING DUCKS	6	\$5.45	12431	Australia	\$32.70	12/01/2010	10:03:00 AM
5	536389	SET OF 3 GOLD FLYING DUCKS	4	\$6.35	12431	Australia	\$25.40	12/01/2010	10:03:00 AM
6	536389	RED RETROSPOT UMBRELLA	6	\$5.95	12431	Australia	\$35.70	12/01/2010	10:03:00 AM
7	536389	BLACK/BLUE POLKADOT UMBRELLA	3	\$5.95	12431	Australia	\$17.85	12/01/2010	10:03:00 AM
8	536389	RED DINER WALL CLOCK	2	\$8.50	12431	Australia	\$17.00	12/01/2010	10:03:00 AM
9	536389	ALARM CLOCK BAKELIKE GREEN	4	\$3.75	12431	Australia	\$15.00	12/01/2010	10:03:00 AM
10	536389	ALARM CLOCK BAKELIKE RED	4	\$3.75	12431	Australia	\$15.00	12/01/2010	10:03:00 AM
11	536389	BLUE DINER WALL CLOCK	2	\$8.50	12431	Australia	\$17.00	12/01/2010	10:03:00 AM
12	536389	IVORY DINER WALL CLOCK	2	\$8.50	12431	Australia	\$17.00	12/01/2010	10:03:00 AM
13	536389	LARGE HEART MEASURING SPOONS	24	\$1.65	12431	Australia	\$39.60	12/01/2010	10:03:00 AM
14	536389	SMALL HEART MEASURING SPOONS	24	\$0.85	12431	Australia	\$20.40	12/01/2010	10:03:00 AM



```
/* Calculating Statistical Summary (MEAN) of TotalPrice */
```

```
PROC MEANS DATA=ECOMMERCE N MEAN MEDIAN MODE RANGE QRANGE CV SKEWNESS  
MAXDEC=2;
```

```
    TITLE 'STATISTICAL SUMMARY OF E-COMMERCE TABLE FOR TotalPrice (MEANS)';  
    VAR TotalPrice;
```

```
RUN;
```

### STATISTICAL SUMMARY OF E-COMMERCE TABLE FOR TotalPrice (MEANS)

#### The MEANS Procedure

Analysis Variable : TotalPrice							
N	Mean	Median	Mode	Range	Quartile Range	Coeff of Variation	Skewness
1424	22.17	11.10	15.00	1668.60	16.20	288.39	14.85

```
/* Sorting in order of CustomerID first, then InvoiceNo */
```

```
PROC SORT DATA=ECOMMERCE;
```

```
    BY CustomerID InvoiceNo;
```

```
RUN;
```

```
/* Printing the 1st 10 observations */
```

```
PROC PRINT DATA=ECOMMERCE (OBS=10);
```

```
    TITLE 'SORTED E-COMMERCE DATA';
```

```
RUN;
```

#### SORTED E-COMMERCE DATA

Obs	InvoiceNo	Description	Quantity	UnitPrice	CustomerID	Country	TotalPrice	InvoiceDate	InvoiceTime
1	536389	CHRISTMAS LIGHTS 10 REINDEER	6	\$8.50	12431	Australia	\$51.00	12/01/2010	10:03:00 AM
2	536389	VINTAGE UNION JACK CUSHION COVER	8	\$4.95	12431	Australia	\$39.60	12/01/2010	10:03:00 AM
3	536389	VINTAGE HEADS AND TAILS CARD GAME	12	\$1.25	12431	Australia	\$15.00	12/01/2010	10:03:00 AM
4	536389	SET OF 3 COLOURED FLYING DUCKS	6	\$5.45	12431	Australia	\$32.70	12/01/2010	10:03:00 AM
5	536389	SET OF 3 GOLD FLYING DUCKS	4	\$6.35	12431	Australia	\$25.40	12/01/2010	10:03:00 AM
6	536389	RED RETROSPOT UMBRELLA	6	\$5.95	12431	Australia	\$35.70	12/01/2010	10:03:00 AM
7	536389	BLACK/BLUE POLKA DOT UMBRELLA	3	\$5.95	12431	Australia	\$17.85	12/01/2010	10:03:00 AM
8	536389	RED DINER WALL CLOCK	2	\$8.50	12431	Australia	\$17.00	12/01/2010	10:03:00 AM
9	536389	ALARM CLOCK BAKELIKE GREEN	4	\$3.75	12431	Australia	\$15.00	12/01/2010	10:03:00 AM
10	536389	ALARM CLOCK BAKELIKE RED	4	\$3.75	12431	Australia	\$15.00	12/01/2010	10:03:00 AM

```
/* Exporting the modified table to csv */
```

```
ODS CSVALL FILE='/home/u61126016/My_Dataset/Ecommerce_Data_EXPORT.csv';
```

```
    TITLE 'FINAL E-COMMERCE DATA';
```

```
PROC PRINT DATA=ECOMMERCE;
```

```
    RUN;
```

```
ODS CSVALL CLOSE;
```

## Code and Output (PART 2):

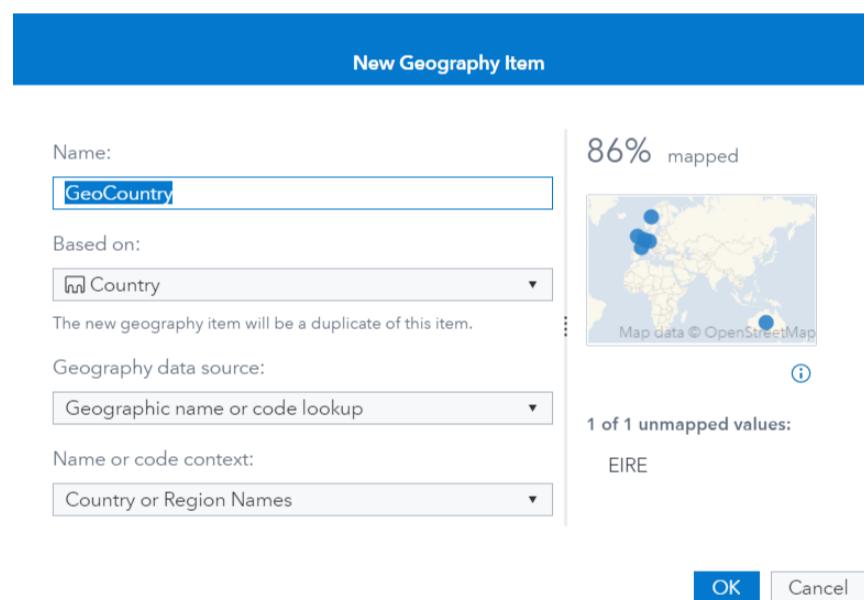
Now we have imported the above dataset in SAS Visual Analytics.

First, we will create one Calculated Item TotalPriceNum and one Geography Item GeoCountry as follows.

TotalPriceNum is a Calculated Measure which is created using TotalPrice which is a Categorical Data Item.



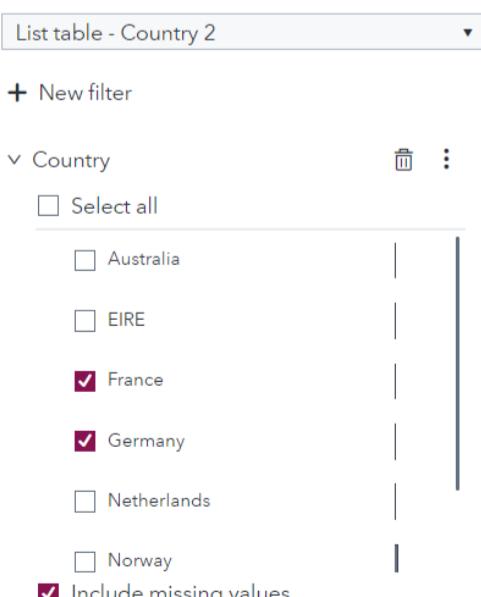
GeoCountry is a Geography Data Item which is created based on Country which is a Categorical Data Item.



Now we will start making the charts and graphs.

First object is a List Table. We use the below filter on the List Table.

### Filters



That is, we are including data of only two countries France and Germany. Now we will apply Display Rules on the List Table as shown below.

Edit Display Rule

**Quantity**

Add an interval  + ⋮

1	<input type="text" value="0"/>	$\leq x <$	<input type="text" value="10"/>	<span style="color: red;">█</span>	<span style="font-size: small;">Delete</span>
2	<input type="text" value="10"/>	$\leq x <$	<input type="text" value="20"/>	█	<span style="font-size: small;">Delete</span>
3	<input type="text" value="20"/>	$\leq x \leq$	<input type="text" value="60"/>	█	<span style="font-size: small;">Delete</span>

Gauge type: Icon ▼

Specify column and placement of gauge:

Country Right of text

OK Cancel

### Display Rules

List table - Country 2 ▼

+ New rule

Column - Country

Quantity

Right of text, with Quantity as x

<span style="color: red;">█</span>	0 $\leq x <$ 10
<span style="color: yellow;">█</span>	10 $\leq x <$ 20
<span style="color: green;">█</span>	20 $\leq x \leq$ 60

The Display Rule is a Gauge which will display a coloured icon to the right of Country based on the mentioned conditions.

The resultant **List Table** is shown below:

Country	CustomerID ▲ Description	TotalPrice(Num)	Quantity
France	12583 LUNCH BOX I LOVE LONDON	46.80	24
France	12583 PANDA AND BUNNIES STICKER SHEET	10.20	12
France	12583 SPACEBOY LUNCH BOX	46.80	24
France	12583 SET 2 TEA TOWELS I LOVE LONDON	70.80	24
France	12583 RED TOADSTOOL LED NIGHT LIGHT	39.60	24
France	12583 MINI JIGSAW CIRCUS PARADE	10.08	24
France	12583 ALARM CLOCK BAKELIKE PINK	90.00	24
France	12583 MINI JIGSAW SPACEBOY	10.08	24
France	12583 CIRCUS PARADE LUNCH BOX	46.80	24
France	12583 SET/2 RED RETROSPOT TEA TOWELS	53.10	18
France	12583 ALARM CLOCK BAKELIKE RED	90.00	24
France	12583 ROUND SNACK BOXES SET OF4 WOODLAND	70.80	24
France	12583 INFLATABLE POLITICAL GLOBE	40.80	48
France	12583 POSTAGE	54.00	3
France	12583 CHARLOTTE BAG DOLLY GIRL DESIGN	17.00	20
France	12583 ALARM CLOCK BAKELIKE GREEN	45.00	12
France	12583 STARS GIFT TAPE	15.60	24
France	12583 VINTAGE HEADS AND TAILS CARD GAME	30.00	24
France	12583 VINTAGE SEASIDE JIGSAW PUZZLES	45.00	12
France	12583 MINI PAINT SET VINTAGE	23.40	36
Germany	12662 MULTI COLOUR SILVER T-LIGHT HOLDER	10.20	12
Germany	12662 JUMBO BAG WOODLAND ANIMALS	19.50	10

Second Object is a Cross Tab. Here we have placed High Cardinality Data Item Product Description in Rows and Low Cardinality Data Item in Columns. Following Display Rules are applied on the Cross Tab.

## Edit Display Rule

TotalPrice(Num)

Operator:

< ▾

Value:

40 ▾

[Specify Intersections](#)

Format:

AvenirNext ▾ **B** *I* U

Background color:

Placement:

TotalPrice(Num) ▾

Allow alerts for this rule

## Edit Display Rule

TotalPrice(Num)

Operator:

> ▾

Value:

40 ▾

[Specify Intersections](#)

Format:

AvenirNext ▾ **B** *I* U

Background color:

Placement:

TotalPrice(Num) ▾

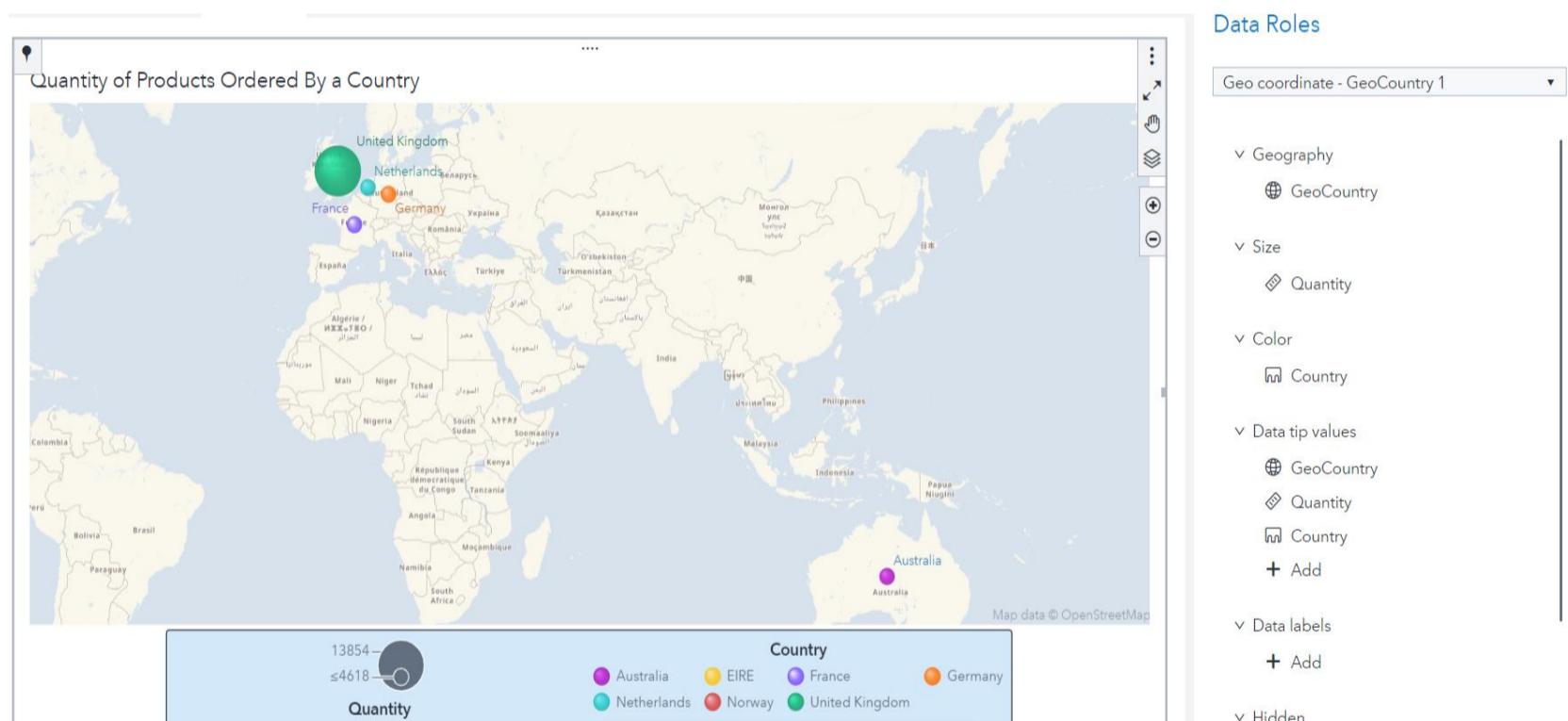
Allow alerts for this rule

The resultant **Cross Tab** is shown below which can be used to find out the expenditure of each Country on each Product:

Product Sale By Country

	Country ▲	Australia	EIRE	France	Germany	Norway	United Kingdom
Description	▲	TotalPrice(Num)	TotalPrice(Num)	TotalPrice(Num)	TotalPrice(Num)	TotalPrice(Num)	TotalPrice(Num)
Total		83.55	41.10	75.00	78.30	97.50	1,138.67
10 COLOUR SPACEBOY PEN		—	—	—	—	—	40.80
200 RED + WHITE BENDY STRAWS		—	—	—	—	15.00	—
3 HOOK HANGER MAGIC GARDEN		—	—	—	23.40	—	—
3 STRIPEY MICE FELTCRAFT		—	23.40	—	—	—	13.65
5 HOOK HANGER MAGIC TOADSTOOL		—	—	—	19.80	—	28.05
5 HOOK HANGER RED MAGIC TOADSTOOL		—	—	—	19.80	—	16.50
60 TEATIME FAIRY CAKE CASES		—	—	—	—	13.20	28.05
72 SWEETHEART FAIRY CAKE CASES		—	—	—	—	13.20	26.40
AGED GLASS SILVER T-LIGHT HOLDER		—	—	—	—	—	94.80
ALARM CLOCK BAKELIKE GREEN		15.00	—	45.00	—	—	78.75
ASSORTED COLOUR BIRD ORNAMENT		—	—	—	—	—	365.04
BISCUIT TIN VINTAGE RED		—	—	—	—	40.50	—
BLACK/BLUE POLKADOT UMBRELLA		17.85	—	—	—	—	23.80
BLUE CHARLIE+LOLA PERSONAL DOORSIGN		—	17.70	—	—	—	—
RED RETROSPOT UMBRELLA		35.70	—	—	—	—	17.85
RETROSPOT CHILDRENS APRON		—	—	—	—	15.60	23.40
ROTATING SILVER ANGELS T-LIGHT HLDR		—	—	—	15.30	—	58.65

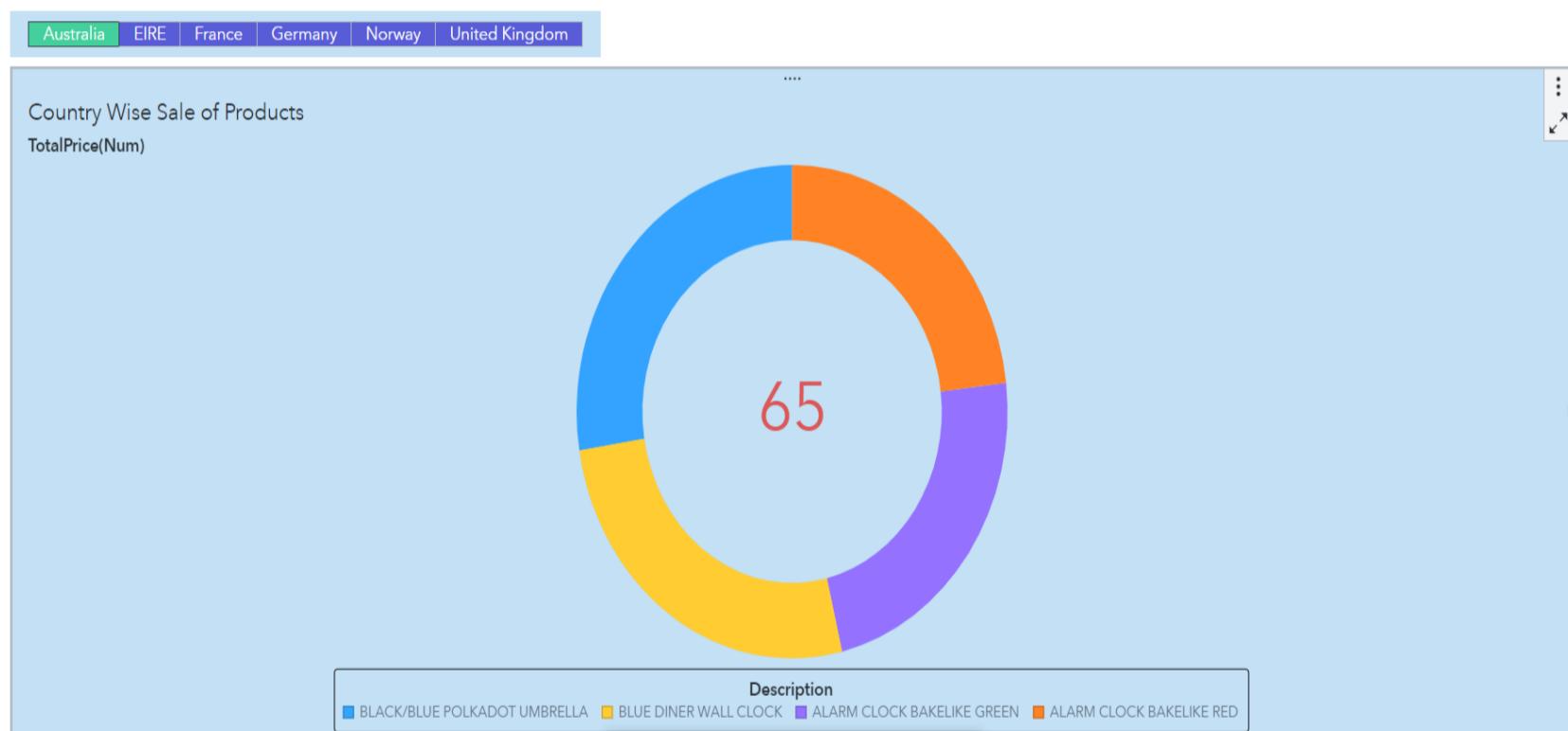
Next object is a **Bubble GeoMap** which can be used to check the quantity of Products purchased by each Country. Also, various changes are done to the Font and Colour to make the GeoMap look aesthetically good.



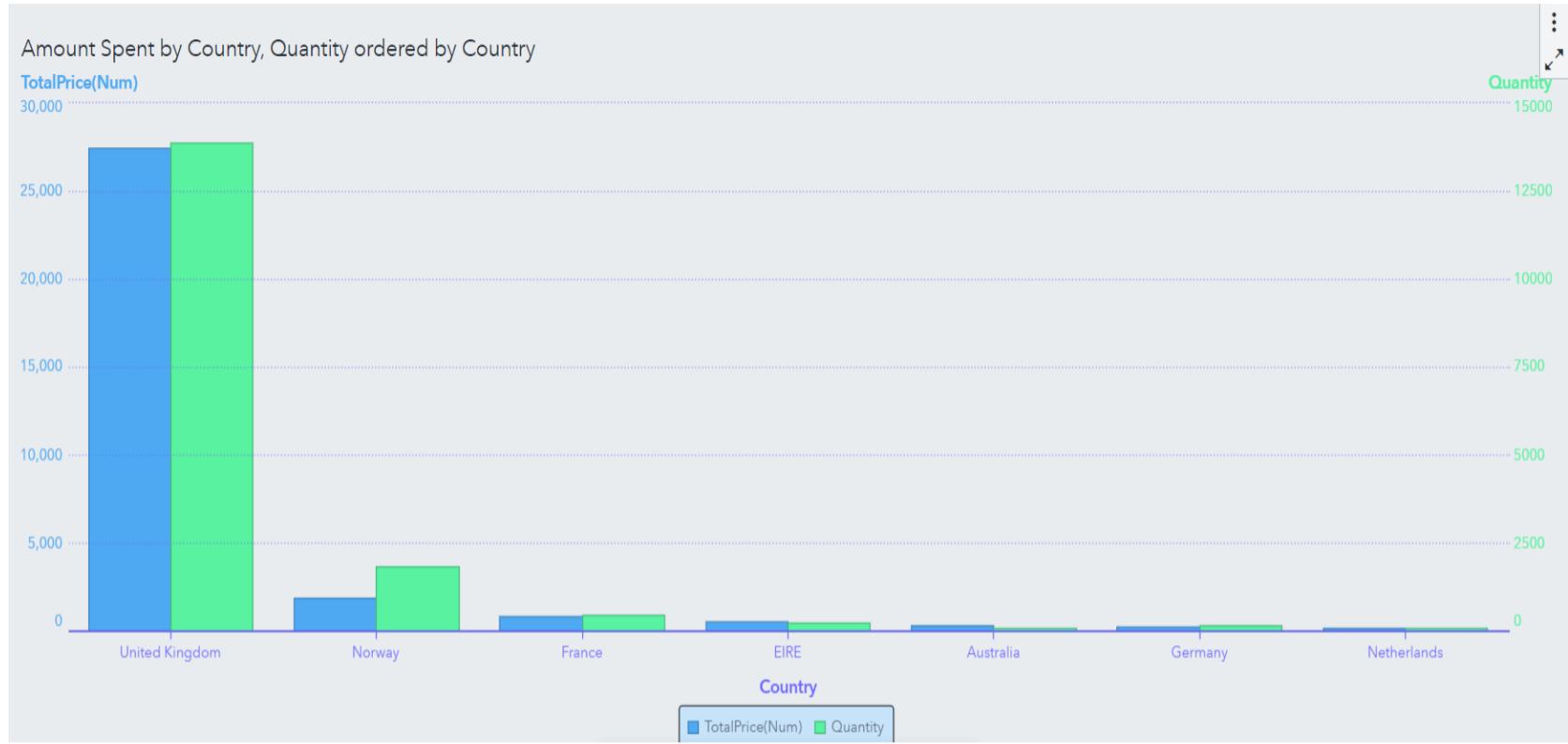
Next object is a **Word Cloud** which displays top 50 Products that have been ordered by the Customers. We have also changed Background, Font, and Colour:



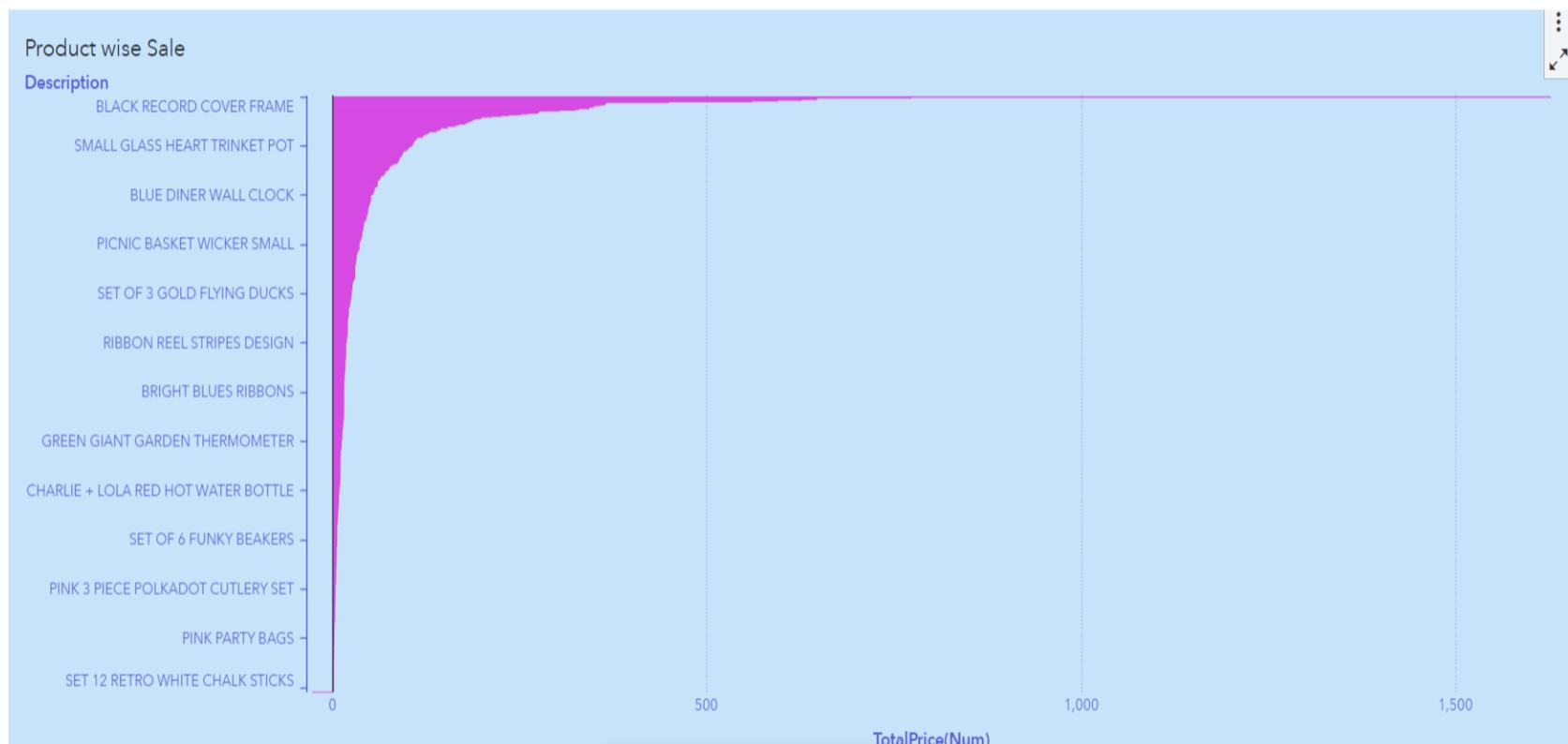
Next object is a Pie Chart which displays most purchased Products by a Country. It displays data of one Country at a time. It is an **Interactive Pie Chart** as we have added Button Bar to it. Using Button Bar, we can see data related to the Country of our choice.



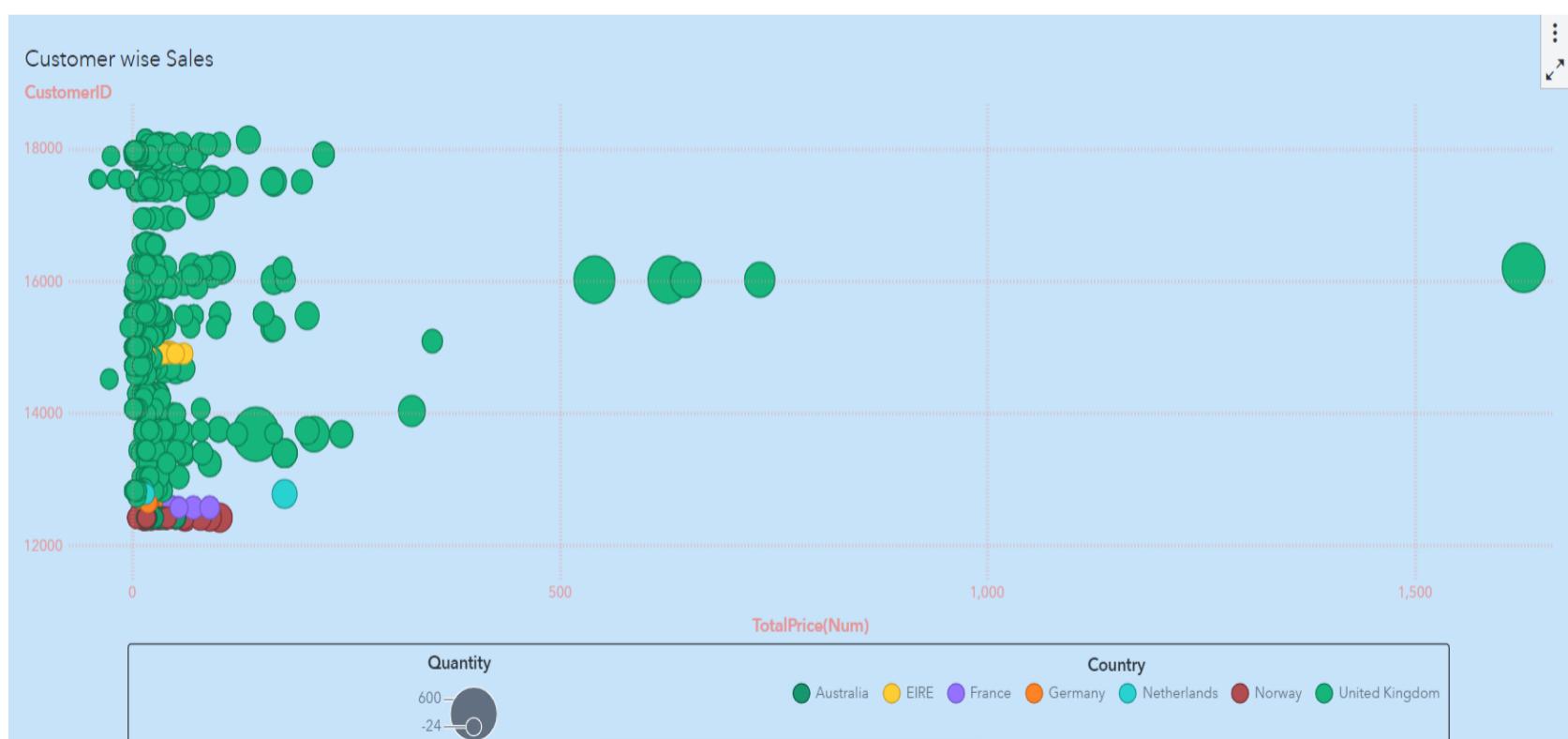
Next object is a **Dual Axis Bar Chart** which represents Country wise Expenditure and Quantity of Products ordered. We have also applied good looking Font and Colour to the graph.



Next object is a **Bar Chart** which depicts the Cost of each Product.



Our last object is a **Bubble Plot** which represents Customer wise Expenditure. It also shows the Quantity of Products purchased by the Customer and even displays the Country to which the Customer belongs.





Shri Vile Parle Kelavani Mandal's

## Dwarkadas J. Sanghvi College of Engineering

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)

### Important Inferences Drawn:

1. From the **List Table** with the Gauge filter on Quantity, we can see that practically all entries in the List table are marked green. This means that the number of each item purchased ranges from 20 to 60. As a result, we can deduce that practically all the purchases are made in bulk. Therefore, users primarily purchase in bulk from this E-commerce business. We might attract more customers if we introduce a bulk buy discount on the website.
2. From the **Cross Tab**, we can figure out that almost all countries make less than \$40 for each product except for United Kingdom. Also, many products are having \$0 sales in a lot of countries. So the export of these products to the respective countries can be reduced in order to increase the savings.
3. From the **Bubble GeoMap** we can readily deduce that the United Kingdom is the largest market for this E-commerce business because it has the largest bubble. As a result, increasing business in the United Kingdom could be extremely advantageous for this e-commerce site.
4. From the **Word Cloud**, we can see the top 50 products purchased by customers are displayed, with the most expensive product being the largest in size. We can see that 'NAMASTE SWAGAT INCENSE' is the most popular product, thus increasing sales of this item could be advantageous.
5. From the **Interactive Pie Chart**, we can examine and get a sense of the most popular products in each country based on total sales for each product. This informs stakeholders about the most popular product choices in each country, allowing sales of these products to be maximized in those countries.
6. From the **Dual Axis Bar Chart**, on the same graph, we may infer country-by-country sales and country-by-country quantity purchases. We can see that the quantity graph in Norway is higher than the price graph. As a result, per-item costs in Norway could be increased to boost profits.
7. From the **Bar Chart** of Product vs TotalPrice, we can see products making lowest and highest revenue and make decisions on its general production.
8. From the **Bubble Plot** of Customer wise Sales, we can see that the majority of bubbles are in the \$0 to \$200 range. This indicates that our target market spends less than \$200. This is the most essential inference, since it provides us with more information about the consumer base with whom we do business, as well as a better understanding of their budget. As a result, by increasing the availability of items in this price range, we may significantly improve sales.

**Conclusion:** From this experiment, we learn about using SAS Studio and SAS Visual Analytics in great detail. We also learn about drawing meaningful inferences from the charts and tables we have in SAS Visual Analytics.