Sr No	Table Contents	Page No

SVKM's NMIMS Deemed to be UNIVERSITY

SVKM's NMIMS

Mukesh Patel School of Technology Management & Engineering AN

Project Report ON

SAS VA AND SAS STUDIO

Faculty Mentor:

MR SURAJPATIL

(Assistant Professor)

MPSTME, NMIMS

Shirpur Campus

Submitted By:

Charvee Saraiya B220

Batch: BTech CE

Table

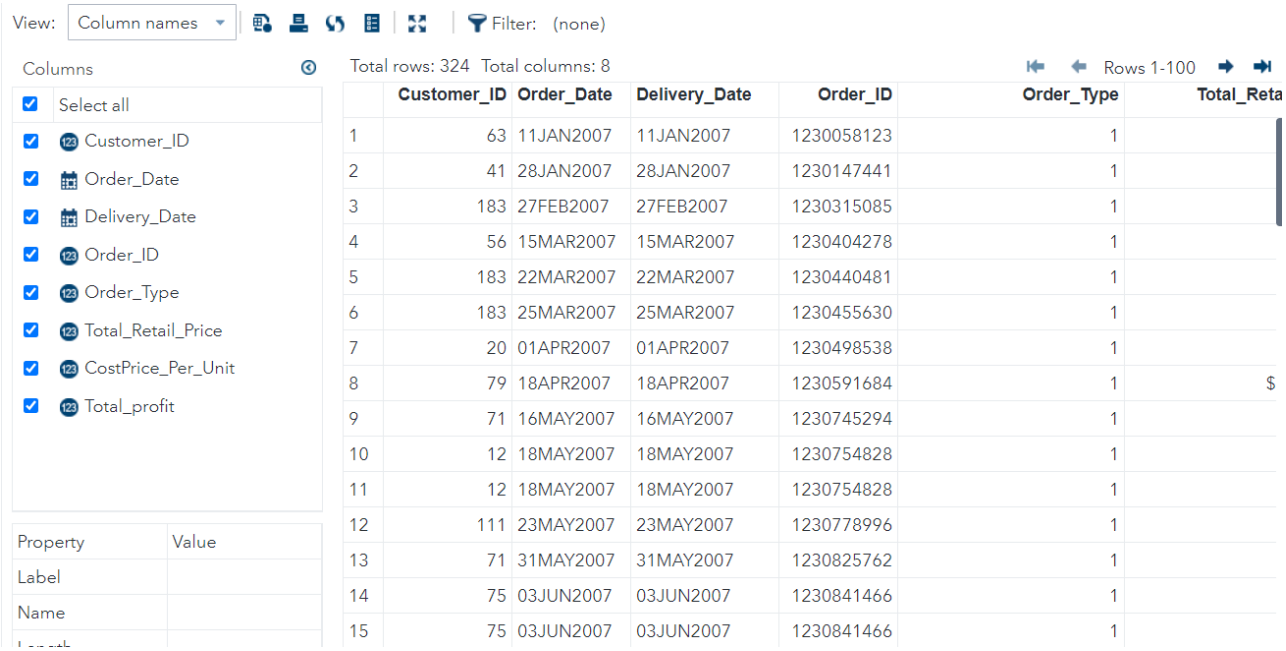
1.INTRODUCTION:

Using Sas Studio to analyse various Scenarios for Various DataSet.It contains various different Variables.

2.DATA SET INFORMATION

2.1 Datasets Used in SAS Studio–

1.Order_Fact



	Customer_ID	Order_Date	Delivery_Date	Order_ID	Order_Type	Total_Reta
1	63	11JAN2007	11JAN2007	1230058123	1	
2	41	28JAN2007	28JAN2007	1230147441	1	
3	183	27FEB2007	27FEB2007	1230315085	1	
4	56	15MAR2007	15MAR2007	1230404278	1	
5	183	22MAR2007	22MAR2007	1230440481	1	
6	183	25MAR2007	25MAR2007	1230455630	1	
7	20	01APR2007	01APR2007	1230498538	1	
8	79	18APR2007	18APR2007	1230591684	1	$
9	71	16MAY2007	16MAY2007	1230745294	1	
10	12	18MAY2007	18MAY2007	1230754828	1	
11	12	18MAY2007	18MAY2007	1230754828	1	
12	111	23MAY2007	23MAY2007	1230778996	1	
13	71	31MAY2007	31MAY2007	1230825762	1	
14	75	03JUN2007	03JUN2007	1230841466	1	
15	75	03JUN2007	03JUN2007	1230841466	1	

2.Sales

View: Column names | Filter: (none)

Columns

- Select all
- Employee_ID
- First_Name
- Last_Name
- Gender
- Salary
- Job_Title
- Country
- Birth_Date
- Hire_Date

Property	Value
Label	
Name	
Length	
Type	
Format	
Informat	

Total rows: 165 Total columns: 9 Rows 1-100

	Employee_ID	First_Name	Last_Name	Gender	Salary	Job_Title
1	120102	Tom	Zhou	M	108255	Sales Manager
2	120103	Wilson	Dawes	M	87975	Sales Manager
3	120121	Irenie	Elvish	F	26600	Sales Rep. II
4	120122	Christina	Ngan	F	27475	Sales Rep. II
5	120123	Kimiko	Hotstone	F	26190	Sales Rep. I
6	120124	Lucian	Daymond	M	26480	Sales Rep. I
7	120125	Fong	Hofmeister	M	32040	Sales Rep. IV
8	120126	Satyakam	Denny	M	26780	Sales Rep. II
9	120127	Sharryn	Clarkson	F	28100	Sales Rep. II
10	120128	Monica	Kletschkus	F	30890	Sales Rep. IV
11	120129	Alvin	Roebuck	M	30070	Sales Rep. III
12	120130	Kevin	Lyon	M	26955	Sales Rep. I
13	120131	Marinus	Surawski	M	26910	Sales Rep. I
14	120132	Fancine	Kaiser	F	28525	Sales Rep. III
15	120133	Petrea	Soltau	F	27440	Sales Rep. II
16	120134	Sian	Shannan	M	28015	Sales Rep. II
17	120135	Alexei	Platts	M	32490	Sales Rep. IV

3.employee_address

View: Column names | Filter: (none)

Columns

- Select all
- Employee_ID
- Employee_Name
- Street_ID
- Street_Number
- Street_Name
- City
- State
- Postal_Code
- Country

Property	Value
Label	
Name	
Length	
Type	
Format	
Informat	

Total rows: 424 Total columns: 9 Rows 1-100

	Employee_ID	Employee_Name	Street_ID
1	121044	Abbott, Ray	9260116912
2	120145	Aisbitt, Sandy	1600101803
3	120761	Akinfolarin, Tameaka	9260121030
4	120656	Amos, Salley	9260123736
5	121107	Anger, Rose	9260120989
6	121038	Anstey, David	9260116991
7	120273	Antonini, Doris	9260116925
8	120759	Apr, Nishan	9260123711
9	120798	Ardskin, Elizabeth	9260116954
10	121030	Areu, Jeryl	9260116937
11	121017	Arizmendi, Gilbert	9260123825
12	121062	Armant, Debra	9260123787
13	121119	Armogida, Bruce	9260121159
14	120812	Arruza, Fauver	9260116937
15	120756	Asta, Wendy	9260121146
16	120754	Atkins, John	9260116816
17	120185	Bahlman, Sharon	1600103028

2.2 Dataset using SAS VISUAL ANALITICS:

1.MegaCrop

2.Telecom

3.1Business Scenarios

Order_Fact:

Business Scenario 1: Manager wants to list out all the data in different formats for Quality.

Business Scenario 2: Manager wants to create a new dataset where we can calculate and analyze the profit from given dataset Order_Fact

Business Scenario 3: Manager wants to calculate the total sales amount for each customer in the Order Fact dataset. here we do data manipulation

Business Scenario 4: Manager wants to know data of all the customers that has Quantity greater than 3

Business Scenario 5: Manager wants to sort the frequency of all the orders

Business Scenario 6: Manager wants to know data of all the customers that has Total_Retail_Price greater than 1000

Business Scenario 7: Categorize data on the basis of order_type and delivery date and Determine the distribution of the Total_Retail_Price variable in the order_fact dataset

Business Scenario 8: Filter for valid orders and calculate sales amount for each order. Then group the data by Customer_ID and calculate total orders and sales amount

SALES DATASET:

Business Scenario 1: Manager wants to know data of all the Sales that has Salary greater than 20000

Business Scenario 2: Manager wants to know data of all the Sales in which job_title contains "Rep";

Business Scenario 3: Manager wants to know the data of sales in country Australia or United States

Business Scenario 4: Manager wants to analyse data graphically using vertical Bar between county and type which is equal to Frequency

Business Scenario 5: Manger wants to analyse data graphically using vertical Bar between Salary and type which is equal to Mean

Business Scenario 6: Manger wants to analyse data graphically using block plot between Country and sumvar which is equal to Salary

Business Scenario 7: Manger wants to Agrregrate data of sales on the basis of salary using tiers Format

EMPLOYEE_ADDRESS DATASET:

Business Scenario 1: Manager wants to sort data of all the Employees on the basis of country.

Business Scenario 2: Manager wants to Add bonus of 10% to each Employee whose country is Australia and hire date is greater than 1 Jan 2000

Business Scenario 3: Manager wants to compensation to each Employee

3.IMPLEMENTATION USING SAS STUDIO:

3.1Business Scenarios:

Order_Fact:

Business Scenario 1: Manager wants to list out all the data in different formats for Quality.

Business Scenario 2: Manager wants to create a new dataset where we can calculate and analyze the profit from given dataset Order_Fact

Business Scenario 3: Manager wants to calculate the total sales amount for each customer in the Order Fact dataset. here we do data manipulation

Business Scenario 4: Manager wants to know data of all the customers that has Quantity greater than 3

Business Scenario 5: Manager wants to sort the frequency of all the orders

Business Scenario 6: Manager wants to know data of all the customers that has Total_Retail_Price greater than 1000

Business Scenario 7: Categorize data on the basis of order_type and delivery date and Determine the distribution of the Total_Retail_Price variable in the order_fact dataset

Business Scenario 8: Filter for valid orders and calculate sales amount for each order. Then group the data by Customer_ID and calculate total orders and sales amount

SALES DATASET:

Business Scenario 1: Manager wants to know data of all the Sales that has Salary greater than 20000

Business Scenario 2: Manager wants to know data of all the Sales in which job_title contains "Rep";

Business Scenario 3: Manager wants to know the data of sales in country Australia or United States

Business Scenario 4: Manager wants to analyse data graphically using vertical Bar between county and type which is equal to Frequency

Business Scenario 5: Manger wants to analyse data graphically using vertical Bar between Salary and type which is equal to Mean

Business Scenario 6: Manger wants to analyse data graphically using block plot

between Country and sumvar which is equal to Salary

Business Scenario 7: Manger wants to Agrregrate data of sales on the basis of salary using tiers Format

EMPLOYEE_ADDRESS DATASET:

Business Scenario 1: Manager wants to sort data of all the Employees on the basis of country.

Business Scenario 2:Manager wants to Add bonus of 10% to each Employee whose country is Australia and hire date is greater than 1 Jan 2000

Business Scenario 3: Manager wants to compensation to each Employee

3.2 Report Building Using SAS Studio:

ORDER_FACT

Code:-

Business Scenario 1: Manager wants to list out all the data in different formats for Quality.

```
proc print data=orion.order_fact;
run;


proc format;
   value qty 1='One' 2='Two' 3='Three' 4='Four' other='Miscoded';
run;

proc print data=orion.order_fact;
   var Customer_ID Employee_ID Street_ID Order_Date Delivery_Date Order_ID
      Order_Type Product_ID Quantity Total_Retail_Price CostPrice_Per_Unit Discount;
   format Order_Date monyy7. Delivery_Date ddmmyy10.
      Quantity qty. Total_Retail_Price dollar10.3 CostPrice_Per_Unit Discount
      dollar12.3;
run;

data ordertype2 ordertype3 other;
   drop Street_ID Order_Date Delivery_Date Order_ID Product_ID;
   set orion.order_fact;

   if order_type=2 then
      output ordertype2;
   else if order_type=3 then
      output ordertype3;
   else
      output other;
run;

proc print data=ordertype2 noobs;
run;

proc print data=ordertype3 noobs;
run;
```

Obs	Customer_ID	Employee_ID	Street_ID	Order_Date	Delivery_Date	Order_ID	Order_Type	Product_ID	Quantity	Total_Retail_Price	CostPrice_Per_Unit	Discount
1	63	121039	9260125492	JAN2007	11/01/2007	1230058123	1	220101300017	One	$16.500	$7.450	.
2	5	99999999	9260114570	JAN2007	19/01/2007	1230080101	2	230100500026	One	$247.500	$109.550	.
3	45	99999999	9260104847	JAN2007	22/01/2007	1230106883	2	240600100080	One	$28.300	$8.550	.
4	41	120174	1600101527	JAN2007	28/01/2007	1230147441	1	240600100010	Two	$32.000	$6.500	.
5	183	120134	1600100760	FEB2007	27/02/2007	1230315085	1	240200200039	Three	$83.600	$8.800	.
6	79	99999999	9260101874	MAR2007	03/03/2007	1230333319	2	240100400005	One	$234.600	$115.950	.
7	23	99999999	9260126679	MAR2007	06/03/2007	1230338566	2	240800200062	One	$35.400	$16.150	.
8	23	99999999	9260126679	MAR2007	06/03/2007	1230338566	2	240800200063	Two	$73.800	$16.850	.
9	45	99999999	9260104847	MAR2007	11/03/2007	1230371142	2	240500100004	Two	$127.000	$28.000	.
10	45	99999999	9260104847	MAR2007	11/03/2007	1230371142	2	240500200003	One	$23.200	$11.100	.
11	56	121059	9260111871	MAR2007	15/03/2007	1230404278	1	220200300002	Two	$75.000	$17.050	.
12	183	120149	1600100760	MAR2007	22/03/2007	1230440481	1	230100600005	One	$129.800	$63.200	.
13	16	99999999	3940105865	MAR2007	26/03/2007	1230450371	2	230100600018	Two	$128.400	$32.200	.
14	79	99999999	9260101874	MAR2007	25/03/2007	1230453723	2	240500200056	One	$24.100	$11.550	.
15	183	120134	1600100760	MAR2007	25/03/2007	1230455630	1	240200100233	Two	$91.800	$22.450	.
16	2788	99999999	9050100001	MAR2007	30/03/2007	1230478006	2	230100100025	Two	$60.600	$13.150	.
17	20	121066	9260118934	APR2007	01/04/2007	1230498538	1	230100300006	One	$68.500	$34.350	.
18	70046	99999999	2600100017	APR2007	03/04/2007	1230500669	3	240200100131	Two	$148.600	$41.350	.
19	12386	99999999	4750100001	APR2007	03/04/2007	1230503155	2	220101400310	One	$31.800	$14.200	.
20	23	99999999	9260126679	APR2007	23/04/2007	1230591673	2	220200200024	Three	$178.500	$32.250	.
21	36	99999999	9260128237	APR2007	20/04/2007	1230591675	3	240500100039	One	$34.500	$15.400	.
22	79	121045	9260101874	APR2007	18/04/2007	1230591684	1	240200100076	Four	$1,796.000	$246.550	.
23	61	99999999	3940108887	APR2007	29/04/2007	1230619748	2	220200100092	One	$83.000	$41.600	.
24	61	99999999	3940108887	APR2007	29/04/2007	1230619748	2	220200300005	Three	$345.000	$52.350	.
25	13	99999999	3940105189	APR2007	03/05/2007	1230642273	2	230100500082	One	$126.100	$58.600	.
26	171	99999999	1600101555	MAY2007	04/05/2007	1230657844	3	240100100646	One	$109.900	$46.800	.
27	11171	99999999	3600100032	MAY2007	09/05/2007	1230690733	3	240200100043	Two	$282.400	$69.400	.

Business Scenario 2: Manager wants to create a new dataset where we can calculate and analyze the profit from given dataset Order_Fact

```
data orion.orderfact2;
  set orion.order_fact;
  where order_type=1 ;
  Total_profit =Total_Retail_Price-CostPrice_Per_Unit ;
  drop Employee_ID Street_ID Product_ID Quantity Discount;
run;

proc print data = orion.orderfact2;
run;
```

Obs	Customer_ID	Order_Date	Delivery_Date	Order_ID	Order_Type	Total_Retail_Price	CostPrice_Per_Unit	Total_profit
1	63	11JAN2007	11JAN2007	1230058123	1	$16.50	$7.45	9.05
2	41	28JAN2007	28JAN2007	1230147441	1	$32.00	$6.50	25.50
3	183	27FEB2007	27FEB2007	1230315085	1	$63.60	$8.80	54.80
4	56	15MAR2007	15MAR2007	1230404278	1	$75.00	$17.05	57.95
5	183	22MAR2007	22MAR2007	1230440481	1	$129.80	$63.20	66.60
6	183	25MAR2007	25MAR2007	1230455630	1	$91.80	$22.45	69.35
7	20	01APR2007	01APR2007	1230498538	1	$68.50	$34.35	34.15
8	79	18APR2007	18APR2007	1230591684	1	$1,796.00	$246.55	1549.45
9	71	16MAY2007	16MAY2007	1230745294	1	$134.50	$67.35	67.15
10	12	18MAY2007	18MAY2007	1230754828	1	$68.40	$11.50	56.90
11	12	18MAY2007	18MAY2007	1230754828	1	$268.00	$29.80	238.20
12	111	23MAY2007	23MAY2007	1230778996	1	$35.50	$17.25	18.25
13	71	31MAY2007	31MAY2007	1230825762	1	$265.60	$55.85	209.75
14	75	03JUN2007	03JUN2007	1230841466	1	$109.20	$23.85	85.35
15	75	03JUN2007	03JUN2007	1230841466	1	$56.00	$11.70	44.30
16	56	11JUN2007	11JUN2007	1230885738	1	$33.00	$7.45	25.55
17	56	11JUN2007	11JUN2007	1230885738	1	$97.20	$14.85	82.35
18	18	16JUN2007	16JUN2007	1230912536	1	$50.80	$12.10	38.70
19	215	07JUL2007	07JUL2007	1231023774	1	$22.90	$10.85	12.05
20	215	07JUL2007	07JUL2007	1231023774	1	$78.20	$39.20	39.00
21	12	16JUL2007	16JUL2007	1231077006	1	$193.80	$40.00	153.80
22	79	27JUL2007	27JUL2007	1231135703	1	$92.00	$16.05	75.95
23	79	27JUL2007	27JUL2007	1231135703	1	$39.00	$8.95	30.05
24	31	02AUG2007	02AUG2007	1231169108	1	$220.20	$36.65	183.55
25	215	04AUG2007	04AUG2007	1231176288	1	$45.40	$10.85	34.55
26	71	08AUG2007	08AUG2007	1231204878	1	$111.40	$27.95	83.45

Business Scenario 3: Manager wants to calculate the total sales amount for each customer in the Order Fact dataset. here we do data manipulation

```
data sales_by_customer;
  format customer_id $10. total_sales dollar12.2;
  keep customer_id total_sales;
  drop product_id date_key quantity;
  set order_fact;
  by customer_id;
  retain total_sales 0;
  if first.customer_id then total_sales=0;
  total_sales+sales;
  if last.customer_id then output;
run;
```

Obs	Customer_ID	Employee_ID	Street_ID	Order_Date	Delivery_Date	Order_ID	Order_Type	Product_ID	Quantity	Total_Retail_Price	CostPrice_Per_Unit	Discount
22	79	121045	9260101874	18APR2007	18APR2007	1230591684	1	240200100076	4	$1,796.00	$246.55	.
38	12	121039	9260103713	18MAY2007	18MAY2007	1230754828	1	220101400269	4	$268.00	$29.80	.
54	2806	99999999	8010100089	04JUL2007	08JUL2007	1231014780	2	240100400043	4	$1,064.00	$131.50	.
61	61	99999999	3940108887	20JUL2007	25JUL2007	1231094514	2	230100500074	4	$544.00	$51.30	.
109	183	120127	1600100760	12DEC2007	12DEC2007	1231898348	1	220101400289	4	$240.80	$27.40	.
112	111	120127	1600102072	18DEC2007	18DEC2007	1231930216	1	220100300008	4	$342.80	$34.35	.
119	19	99999999	3940106547	26DEC2007	30DEC2007	1231976710	3	240300100020	4	$56.40	$6.05	.
129	53	120143	1600103258	13JAN2008	13JAN2008	1232087464	1	230100600022	4	$336.40	$42.15	.
156	4	99999999	9260106519	08APR2008	09APR2008	1232654929	3	240500100017	4	$214.00	$23.75	.
207	195	120148	1600101663	03OCT2008	03OCT2008	1233837302	1	230100500026	4	$990.00	$109.55	.
227	92	121109	9260117676	11DEC2008	11DEC2008	1234301319	1	240700200010	4	$105.60	$10.10	.
236	17	121109	9260123306	29DEC2008	29DEC2008	1234419240	1	240700100013	5	$119.95	$9.75	.
238	36	99999999	9260128237	11JAN2009	14JAN2009	1234534069	3	240800100026	4	$525.20	$58.55	.
283	92	121051	9260117676	08JUL2009	08JUL2009	1236349115	1	240200100221	4	$396.00	$49.60	.
287	70108	99999999	2600100046	22JUL2009	02AUG2009	12361835?6	2	240200200071	4	$74.80	$10.00	.
318	5	121041	9260114570	08DEC2009	08DEC2009	1237890730	1	220101400387	4	$50.40	$5.85	.
405	16	99999999	3940105865	19JUL2010	20JUL2010	1239713046	2	230100500058	4	$66.80	$7.25	.
413	65	99999999	3940100176	28JUL2010	03AUG2010	1239785436	3	230100600031	4	$354.00	$42.45	.
419	16	99999999	3940105865	16AUG2010	17AUG2010	1239932984	2	220200300129	5	$406.00	$37.05	.
467	52	121044	9260116235	14DEC2010	14DEC2010	1240903120	1	220100100629	4	$213.20	$26.75	.
468	16	99999999	3940105865	22DEC2010	23DEC2010	1240961599	2	230100200025	4	$1,103.60	$125.25	.
471	89	121135	9260116551	03JAN2011	04JAN2011	1241063739	1	240800200035	6	$160.80	$12.15	.
488	171	120127	1600101555	28FEB2011	28FEB2011	1241561055	1	220101400088	5	$192.00	$17.50	.
506	45	121040	9260104847	18APR2011	18APR2011	1242012269	1	220100300019	4	$142.40	$16.70	.
515	90	121032	9260111614	03MAY2011	03MAY2011	1242149082	1	210200900033	4	$56.80	$6.45	.
529	34	99999999	9260111379	07JUN2011	13JUN2011	1242467585	3	240100100737	5	$239.50	$20.65	.
535	17023	99999999	2600100021	13JUN2011	18JUN2011	1242515373	3	240200100057	4	$168.00	$20.25	.

Business Scenario 4: Manager wants to know data of all the customers that has Quantity greater than 3

```
proc print data=orion.order_fact;
where Quantity>3;
run;
```

Obs	Customer_ID	Employee_ID	Street_ID	Order_Date	Delivery_Date	Order_ID	Order_Type	Product_ID	Quantity	Total_Retail_Price	CostPrice_Per_Unit	Discount
22	79	121045	9260101874	18APR2007	18APR2007	1230591684	1	240200100076	4	$1,796.00	$246.55	.
38	12	121039	9260103713	18MAY2007	18MAY2007	1230754828	1	220101400269	4	$268.00	$29.80	.
54	2806	99999999	8010100089	04JUL2007	08JUL2007	1231014780	2	240100400043	4	$1,064.00	$131.50	.
61	61	99999999	3940108887	20JUL2007	25JUL2007	1231094514	2	230100500074	4	$544.00	$51.30	.
109	183	120127	1600100760	12DEC2007	12DEC2007	1231898348	1	220101400289	4	$240.80	$27.40	.
112	111	120127	1600102072	18DEC2007	18DEC2007	1231930216	1	220100300008	4	$342.80	$34.35	.
119	19	99999999	3940106547	26DEC2007	30DEC2007	1231976710	3	240300100020	4	$56.40	$6.05	.
129	53	120143	1600103258	13JAN2008	13JAN2008	1232087464	1	230100600022	4	$336.40	$42.15	.
156	4	99999999	9260106519	08APR2008	09APR2008	1232654929	3	240500100017	4	$214.00	$23.75	.
207	195	120148	1600101663	03OCT2008	03OCT2008	1233837302	1	230100500026	4	$990.00	$109.55	.
227	92	121109	9260117676	11DEC2008	11DEC2008	1234301319	1	240700200010	4	$105.60	$10.10	.
236	17	121109	9260123306	29DEC2008	29DEC2008	1234419240	1	240700100013	5	$119.95	$9.75	.
238	36	99999999	9260128237	11JAN2009	14JAN2009	1234534069	3	240800100026	4	$525.20	$58.55	.
283	92	121051	9260117676	08JUL2009	08JUL2009	1236349115	1	240200100221	4	$396.00	$49.60	.
287	70108	99999999	2600100046	22JUL2009	02AUG2009	12361835?6	2	240200200071	4	$74.80	$10.00	.
318	5	121041	9260114570	08DEC2009	08DEC2009	1237890730	1	220101400387	4	$50.40	$5.85	.
405	16	99999999	3940105865	19JUL2010	20JUL2010	1239713046	2	230100500058	4	$66.80	$7.25	.
413	65	99999999	3940100176	28JUL2010	03AUG2010	1239785436	3	230100600031	4	$354.00	$42.45	.
419	16	99999999	3940105865	16AUG2010	17AUG2010	1239932984	2	220200300129	5	$406.00	$37.05	.
467	52	121044	9260116235	14DEC2010	14DEC2010	1240903120	1	220100100629	4	$213.20	$26.75	.
468	16	99999999	3940105865	22DEC2010	23DEC2010	1240961599	2	230100200025	4	$1,103.60	$125.25	.
471	89	121135	9260116551	03JAN2011	04JAN2011	1241063739	1	240800200035	6	$160.80	$12.15	.
488	171	120127	1600101555	28FEB2011	28FEB2011	1241561055	1	220101400088	5	$192.00	$17.50	.
506	45	121040	9260104847	18APR2011	18APR2011	1242012269	1	220100300019	4	$142.40	$16.70	.
515	90	121032	9260111614	03MAY2011	03MAY2011	1242149082	1	210200900033	4	$56.80	$6.45	.
529	34	99999999	9260111379	07JUN2011	13JUN2011	1242467585	3	240100100737	5	$239.50	$20.65	.
535	17023	99999999	2600100021	13JUN2011	18JUN2011	1242515373	3	240200100057	4	$168.00	$20.25	.

Business Scenario 5: Manager wants to sort the frequency of all the orders

Secenerio: Company wants to know the frequency of all the orders

```
proc freq data=orion.order_fact;
run;
```

The FREQ Procedure

Customer ID				
Customer_ID	Frequency	Percent	Cumulative Frequency	Cumulative Percent
4	9	1.46	9	1.46
5	19	3.08	28	4.54
9	6	0.97	34	5.51
10	32	5.19	66	10.70
11	2	0.32	68	11.02
12	16	2.59	84	13.61
13	6	0.97	90	14.59
16	25	4.05	115	18.64
17	6	0.97	121	19.61
18	3	0.49	124	20.10
19	7	1.13	131	21.23
20	6	0.97	137	22.20
23	11	1.78	148	23.99
24	7	1.13	155	25.12
27	8	1.30	163	26.42
29	10	1.62	173	28.04
31	13	2.11	186	30.15
34	11	1.78	197	31.93
36	9	1.46	206	33.39
39	2	0.32	208	33.71
41	20	3.24	228	36.95
45	22	3.57	250	40.52
49	14	2.27	264	42.79
50	3	0.49	267	43.27

Business Scenario 6: Manger wants to Agrregrate data of sales on the basis of salary using tiers Format

```
proc print data=orion.order_fact;
where Total_Retail_Price>1000;
run;
```

Obs	Customer_ID	Employee_ID	Street_ID	Order_Date	Delivery_Date	Order_ID	Order_Type	Product_ID	Quantity	Total_Retail_Price	CostPrice_Per_Unit	Discount
22	79	121045	9260101874	18APR2007	18APR2007	1230691684	1	240200100076	4	$1,796.00	$246.55	.
41	88	99999999	9260100179	25MAY2007	28MAY2007	1230793366	2	240400200097	3	$1,250.40	$124.20	.
54	2806	99999999	8010100089	04JUL2007	08JUL2007	1231014780	2	240100400043	4	$1,064.00	$131.50	.
72	70187	99999999	2600100035	13AUG2007	18AUG2007	1231227910	2	240200200013	3	$1,266.00	$42.20	.
120	19	99999999	3940106547	26DEC2007	30DEC2007	1231976710	3	240300100032	2	$1,200.20	$300.00	.
206	49	121053	9260104510	10SEP2008	10SEP2008	1233689304	1	240300300070	3	$1,514.40	$229.65	.
262	16	99999999	3940105865	10MAY2009	12MAY2009	1235744141	2	230100700009	3	$1,687.50	$287.10	.
288	9	99999999	3940106659	10AUG2009	15AUG2009	1236673732	3	230100700008	3	$1,542.60	$250.70	.
293	183	120136	1600100760	21AUG2009	21AUG2009	1236783056	1	240300300090	3	$1,561.80	$237.05	.
408	71	121073	9260124130	21JUL2010	21JUL2010	1239735632	1	230100700009	2	$1,136.20	$289.95	.
468	16	99999999	3940105865	22DEC2010	23DEC2010	1240961599	2	230100200025	4	$1,103.60	$125.25	.
541	70100	99999999	2600100015	20JUN2011	24JUN2011	1242578860	2	240200100173	4	$1,937.20	$247.70	.
591	10	121053	9260129395	01NOV2011	01NOV2011	1243797399	1	240300100028	2	$1,066.40	$251.35	.

SALES DATASET:

Business Scenario 1: Manager wants to know data of all the Sales that has Salary greater than 20000

```
proc print data=orion.sales;
where Salary>20000;
run;
```

Obs	Employee_ID	First_Name	Last_Name	Gender	Salary	Job_Title	Country	Birth_Date	Hire_Date
1	120102	Tom	Zhou	M	108255	Sales Manager	AU	4971	12205
2	120103	Wilson	Dawes	M	87975	Sales Manager	AU	-2535	6575
3	120121	Irenie	Elvish	F	26600	Sales Rep. II	AU	-4169	6575
4	120122	Christina	Ngan	F	27475	Sales Rep. II	AU	-523	8217
5	120123	Kimiko	Hotstone	F	26190	Sales Rep. I	AU	3193	10866
6	120124	Lucian	Daymond	M	26480	Sales Rep. I	AU	1228	8460
7	120125	Fong	Hofmeister	M	32040	Sales Rep. IV	AU	-391	8460
8	120126	Satyakam	Denny	M	26780	Sales Rep. II	AU	11951	18475
9	120127	Sharryn	Clarkson	F	28100	Sales Rep. II	AU	8404	15645
10	120128	Monica	Kletschkus	F	30890	Sales Rep. IV	AU	11152	18567
11	120129	Alvin	Roebuck	M	30070	Sales Rep. III	AU	3248	10866
12	120130	Kevin	Lyon	M	26955	Sales Rep. I	AU	10575	18383
13	120131	Marinus	Surawski	M	26910	Sales Rep. I	AU	8668	17167
14	120132	Fancine	Kaiser	F	28525	Sales Rep. III	AU	-2462	8309
15	120133	Petrea	Soltau	F	27440	Sales Rep. II	AU	11069	18536
16	120134	Sian	Shannan	M	28015	Sales Rep. II	AU	-2400	6575
17	120135	Alexei	Platts	M	32490	Sales Rep. IV	AU	4774	15249
18	120136	Atul	Leyden	M	26605	Sales Rep. I	AU	8659	17198
19	120137	Marina	Iyengar	F	29715	Sales Rep. III	AU	8471	18322
20	120138	Shani	Duckett	F	25795	Sales Rep. I	AU	8592	18444
21	120139	Fang	Wilson	F	26810	Sales Rep. II	AU	11187	18506
22	120140	Michael	Minas	M	26970	Sales Rep. I	AU	11903	18536

Business Scenario 2: Manager wants to know data of all the Sales in which job_title contains "Rep";

```
proc print data=orion.sales;
where job_title contains"Rep";
run;
```

Obs	Employee_ID	First_Name	Last_Name	Gender	Salary	Job_Title	Country	Birth_Date	Hire_Date
3	120121	Irenie	Elvish	F	26600	Sales Rep. II	AU	-4169	6575
4	120122	Christina	Ngan	F	27475	Sales Rep. II	AU	-523	8217
5	120123	Kimiko	Hotstone	F	26190	Sales Rep. I	AU	3193	10866
6	120124	Lucian	Daymond	M	26480	Sales Rep. I	AU	1228	8460
7	120125	Fong	Hofmeister	M	32040	Sales Rep. IV	AU	-391	8460
8	120126	Satyakam	Denny	M	26780	Sales Rep. II	AU	11951	18475
9	120127	Sharryn	Clarkson	F	28100	Sales Rep. II	AU	8404	15645
10	120128	Monica	Kletschkus	F	30890	Sales Rep. IV	AU	11152	18567
11	120129	Alvin	Roebuck	M	30070	Sales Rep. III	AU	3248	10866
12	120130	Kevin	Lyon	M	26955	Sales Rep. I	AU	10575	18383
13	120131	Marinus	Surawski	M	26910	Sales Rep. I	AU	8668	17167
14	120132	Fancine	Kaiser	F	28525	Sales Rep. III	AU	-2462	8309
15	120133	Petrea	Soltau	F	27440	Sales Rep. II	AU	11069	18536
16	120134	Sian	Shannan	M	28015	Sales Rep. II	AU	-2400	6575
17	120135	Alexei	Platts	M	32490	Sales Rep. IV	AU	4774	15249
18	120136	Atul	Leyden	M	26605	Sales Rep. I	AU	8659	17198
19	120137	Marina	Iyengar	F	29715	Sales Rep. III	AU	8471	18322
20	120138	Shani	Duckett	F	25795	Sales Rep. I	AU	8592	18444
21	120139	Fang	Wilson	F	26810	Sales Rep. II	AU	11187	18506
22	120140	Michael	Minas	M	26970	Sales Rep. I	AU	11903	18536
23	120141	Amanda	Liebman	F	27465	Sales Rep. II	AU	11759	18383
24	120142	Vincent	Eastley	M	29695	Sales Rep. III	AU	11122	18444

Business Scenario 3: Manager wants to know the data of sales in country Australia or United States

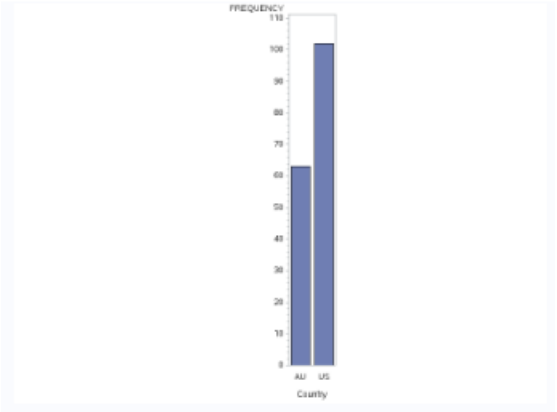
```
proc print data=orion.sales;
where country in ("AU","US");
run;
```

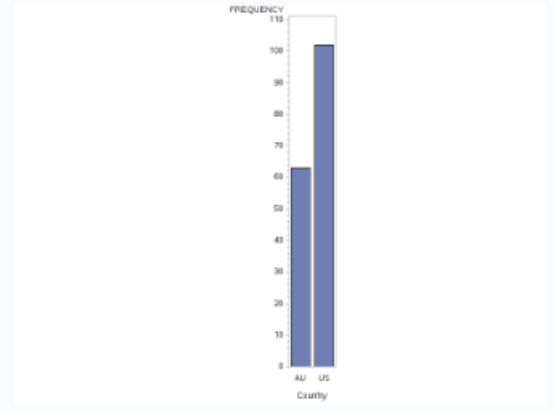
47	120165	Tadashi	Pretorius	M	27050	Sales Rep. I	AU	10305	18322
48	120166	Fadi	Nowd	M	30660	Sales Rep. IV	AU	-4218	6575
49	120167	Kimiko	Tilley	F	25185	Sales Rep. I	AU	-607	6606
50	120168	Selina	Barcoe	F	25275	Sales Rep. I	AU	10310	18567
51	120169	Cos	Tannous	M	28135	Sales Rep. III	AU	10228	18383
52	120170	Alban	Kingston	M	28830	Sales Rep. III	AU	6630	13423
53	120171	Alena	Moody	F	26205	Sales Rep. II	AU	10427	18506
54	120172	Edwin	Comber	M	28345	Sales Rep. III	AU	-4292	6575
55	120173	Hernani	Osborn	M	26715	Sales Rep. I	AU	-677	7822
56	120174	Doungkamol	Simms	F	26850	Sales Rep. I	AU	-4374	6575
57	120175	Andrew	Conolly	M	25745	Sales Rep. I	AU	11918	18536
58	120176	Koavea	Pa	M	26095	Sales Rep. I	AU	11270	18567
59	120177	Franca	Kierce	F	28745	Sales Rep. III	AU	8495	15372
60	120178	Billy	Plested	M	26165	Sales Rep. II	AU	-404	6665
61	120179	Matsuoka	Wills	M	28510	Sales Rep. III	AU	6648	17532
62	120180	Vino	George	M	26970	Sales Rep. II	AU	-553	8370
63	120198	Meera	Body	F	28025	Sales Rep. III	AU	11708	18597
64	120261	Harry	Highpoint	M	243190	Chief Sales Officer	US	4800	11535
65	121018	Julienne	Magolan	F	27560	Sales Rep. II	US	-4381	6575
66	121019	Scott	Desanctis	M	31320	Sales Rep. IV	US	11133	17684
67	121020	Cherda	Ridley	F	31750	Sales Rep. IV	US	10280	16922
68	121021	Priscilla	Farren	F	32985	Sales Rep. IV	US	6918	13939
69	121022	Robert	Stevens	M	32210	Sales Rep. IV	US	8701	16833

Business Scenario 4:Manager wants to analyse data graphically using vertical Bar between county and type which is equal to Frequency

```
proc gchart data=orion.sales;
vbar country/type = freq;
```

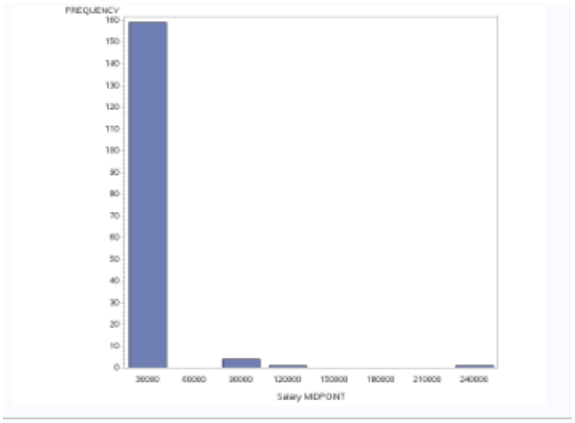
run;

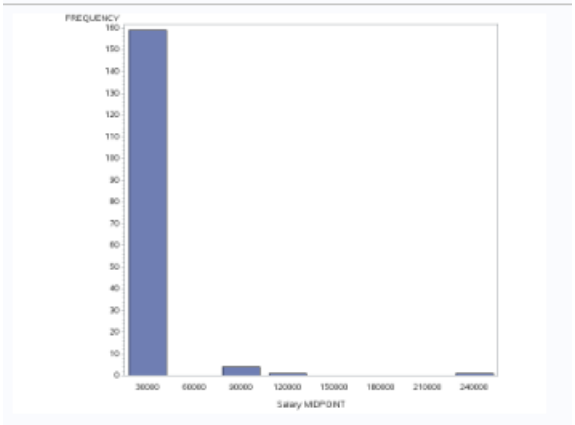




Business Scenario 5: Manger wants to analyse data graphically using vertical Bar between Salary and type which is equal to Mean

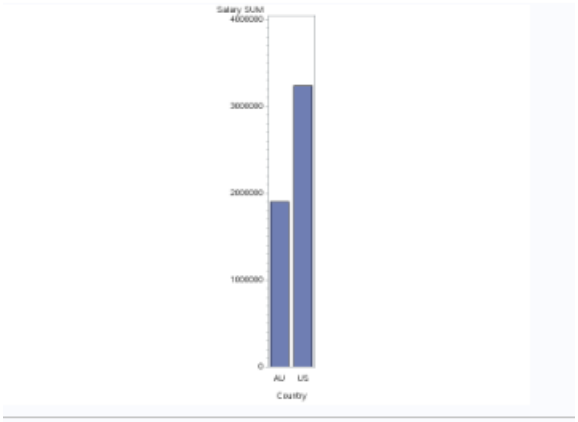
```
proc gchart data=orion.sales;
vbar Salary/type=mean;
run;
```

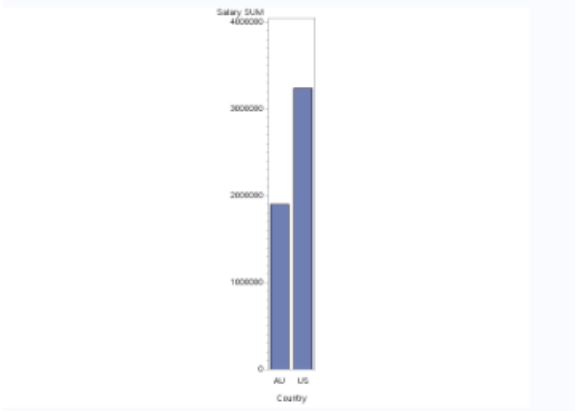




Business Scenario 6: Manger wants to analyse data graphically using block plot between Country and sumvar which is equal to Salary

```
proc gchart data=orion.sales;
vbar country/Sumvar=salary;
run;
```





Business Scenario 7: Manger wants to Agrregrate data of sales on the basis of salary using tiers Format

```
proc print data=orion.sales;
var employee_id salary country birth_Date
hire_date;
format salary tiers. country $ctry. birth_date date9. ;
run;

proc format;
value tiers low-<50000 = Tier1
          50000-<100000 = Tier2
          100000-high= Tier3;
run;
```

Obs	Employee_ID	Salary	Country	Birth_Date	Hire_Date
1	120102	Tier3	AU	11AUG1973	12205
2	120103	Tier2	AU	22JAN1953	6575
3	120121	Tier1	AU	02AUG1948	6575
4	120122	Tier1	AU	27JUL1958	8217
5	120123	Tier1	AU	28SEP1968	10866
6	120124	Tier1	AU	13MAY1963	8460
7	120125	Tier1	AU	06DEC1958	8460
8	120126	Tier1	AU	20SEP1992	18475
9	120127	Tier1	AU	04JAN1983	15645
10	120128	Tier1	AU	14JUL1990	18567
11	120129	Tier1	AU	22NOV1968	10866
12	120130	Tier1	AU	14DEC1988	18383
13	120131	Tier1	AU	25SEP1983	17167
14	120132	Tier1	AU	05APR1953	8309
15	120133	Tier1	AU	22APR1990	18536
16	120134	Tier1	AU	06JUN1953	6575
17	120135	Tier1	AU	26JAN1973	15249
18	120136	Tier1	AU	16SEP1983	17198
19	120137	Tier1	AU	12MAR1983	18322
20	120138	Tier1	AU	11JUL1983	18444
21	120139	Tier1	AU	18AUG1990	18506

EMPLOYEE_ADDRESS DATASET:

Business Scenario 1: Manager wants to sort data of all the Employees on the basis of country.

```
data us au other;
set orion.employee_addresses;
if country="AU" then output au;
  else if
  country="US" then output us;
  else
     output other;
     run;
     proc print data= work.AU;
     run;
     proc print data=work.US;
     run;
     proc print data=work.other;
    run;
```

Obs	Employee_ID	Employee_Name	Street_ID	Street_Number	Street_Name	City	State	Postal_Code	Country
1	120145	Aisbitt, Sandy	1600101803	30	Bingera Street	Melbourne		2001	AU
2	120185	Bahlman, Sharon	1600103028	24	LaTrobe Street	Sydney		2165	AU
3	120109	Baker, Gabriele	1600103074	166	Toorak Road	Sydney		2119	AU
4	120188	Baran, Shanmuganathan	1600102985	207	Canterbury Road	Sydney		1225	AU
5	120144	Barbis, Viney	1600102974	3	Alice Street	Sydney		2114	AU
6	120168	Barcoe, Selina	1600101850	435	Sherwood Rd	Melbourne		8003	AU
7	120182	Barreto, Geok-Seng	1600103062	241	Royal Perade	Sydney		1115	AU
8	120104	Billington, Kareen	1600103066	40	Smith Street	Sydney		1670	AU
9	120183	Blanton, Brig	1600101844	6	Palmiston Cresent	Melbourne		3150	AU
10	120198	Body, Meera	1600103038	51	Martin Place	Sydney		1131	AU
11	120114	Buddery, Jeannette	1600103023	12	Hunua Road Papakura	Sydney		1004	AU
12	120181	Cantatore, Lorian	1600102985	207	Canterbury Road	Sydney		1225	AU
13	120187	Catenacci, Reyne	1600101829	44	Gow Street	Melbourne		2066	AU
14	120146	Cederlund, Wendall	1600103054	1	Pacific Rise Mount Wellington	Sydney		20000	AU
15	120149	Chantharasy, Judy	1600101820	1	Embarcadero Center Site 1500	Melbourne		3000	AU
16	120127	Clarkson, Sharryn	1600102972	21	Albert St	Sydney		1170	AU
17	120172	Comber, Edwin	1600102978	121	Blackburn Road	Sydney		2580	AU

Business Scenario 2:Manager wants to Add bonus of 10% to each Employee whose

country is Australia and hire date is greater than 1 Jan 2000

```
data Subset1;
set orion.sales;
Bonus=salary*0.10;
where country in ('AU') and hire_date>'1JAN2000'd;
drop last_name birth_date employee_id gender;
keep first_name Salary country hire_date Bonus;
run;
```

Table: WORK.SUBSET1 | View: Column names | Filter: (none)

Total rows: 32 Total columns: 5 | Rows 1-32

	First_Name	Salary	Country	Hire_Date	Bonus
1	Satyakam	26780	AU	18475	2678
2	Sharryn	28100	AU	15645	2810
3	Monica	30890	AU	18567	3089
4	Kevin	26955	AU	18383	2695.5
5	Marinus	26910	AU	17167	2691
6	Petrea	27440	AU	18536	2744
7	Alexei	32490	AU	15249	3249
8	Atul	26605	AU	17198	2660.5
9	Marina	29715	AU	18322	2971.5
10	Shani	25795	AU	18444	2579.5
11	Fang	26810	AU	18506	2681
12	Michael	26970	AU	18536	2697
13	Amanda	27465	AU	18383	2746.5
14	Vincent	29695	AU	18444	2969.5
15	Viney	30265	AU	18536	3026.5
16	Skev	26580	AU	18536	2658

Business Scenario 3: Manager wants to compensation to each Employee

```
data Company;
set Orion.sales;
bonus=500;
compensation=salary+bonus;
bonus_month=month(hire_date);
keep employee_id salary bonus compensation bonus_month;
run;
proc print
data=company;
run;
```

Obs	Employee_ID	Salary	bonus	compensation	bonus_month
1	120102	108255	500	108755	6
2	120103	87975	500	88475	1
3	120121	26600	500	27100	1
4	120122	27475	500	27975	7
5	120123	26190	500	26690	10
6	120124	26480	500	26980	3
7	120125	32040	500	32540	3
8	120126	26780	500	27280	8
9	120127	28100	500	28600	11
10	120128	30890	500	31390	11
11	120129	30070	500	30570	10
12	120130	26955	500	27455	5
13	120131	26910	500	27410	1
14	120132	28525	500	29025	10
15	120133	27440	500	27940	10
16	120134	28015	500	28515	1
17	120135	32490	500	32990	10
18	120136	26605	500	27105	2
19	120137	29715	500	30215	3
20	120138	25795	500	26295	7
21	120139	26810	500	27310	9
22	120140	26970	500	27470	10

4. DASHBOARD DESIGN USING SAS VISUAL ANALYTICS

4.1 Bussiness Scenario:

Business Scenario 1: The company wants to analyze its facility regions to improve its future profitability along with the data of different product brand and product lines and products.

Business Scenario 2: The company needs a dashboard to analyze its various facility regions based on profitability vs expensesso that it can study where expenses are more than profit and improve precisely on the areas of losses based on the visualization.

Business Scenario 3: The company needed a dashboard to analyze its product lines and its products in terms of profitability,product price and target so that it can avoid losses, learn from past data and secure more profit on its products in future.

Business Scenario 4: A telecom company needs a dashboard to analyze and know more about the call duration, number of calls by subscriber region and call start date, hour so that the company knows where to launch its nextproduct based on highest call duration subscriber region for most profitability.

Business Scenario 5: A telecom company needs a dashboard to analyze different switch makes with respect to number of dropcalls and call type so that it can get a better insight of which switch to improve based on the call type andmaximum no. of calls dropped.

4.2 Dashboard Design:

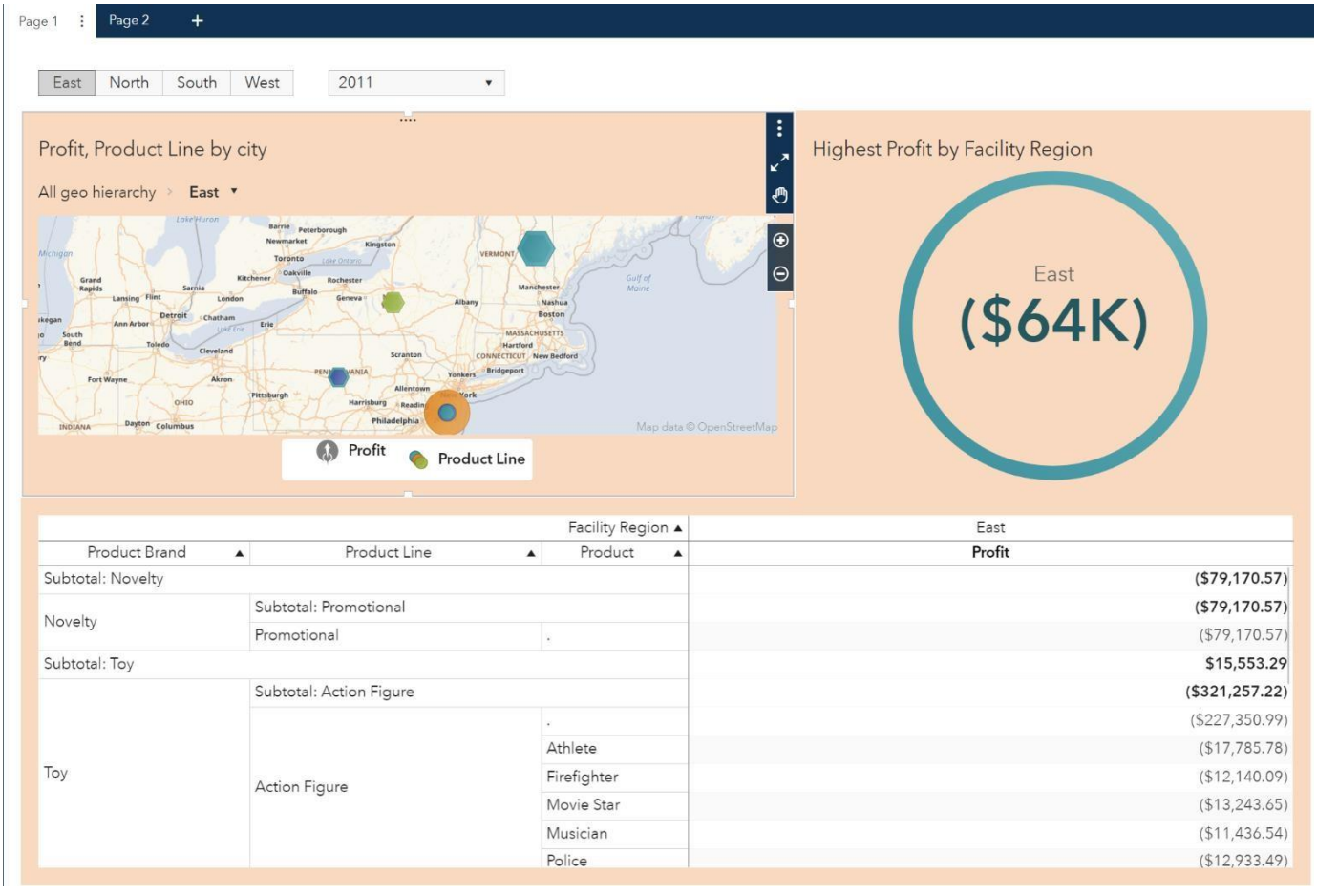
Dataset Mega Corp:

BUSINESS SCENARIO #1

The company wants to analyze its facility regions to improve its future profitability along with the data ofdifferent product brand and product lines and products.

Solution:

We have made a dashboard which contains a geo map with geo data items hierarchy sized by profit and grouped by product line, a key value object which reflects highest profit by facility region and a cross tablevisualization which gives an all over insight of profitability by facility region and product brand, product lineand product hierarchy. All of these charts are controlled by button bar and drop down list to filter by regionand date by year.



			Facility Region ▲	East
Product Brand ▲	Product Line ▲	Product ▲		Profit
Subtotal: Novelty				($79,170.57)
Novelty	Subtotal: Promotional			($79,170.57)
	Promotional	.		($79,170.57)
Subtotal: Toy				$15,553.29
Toy	Subtotal: Action Figure			($321,257.22)
	Action Figure	.		($227,350.99)
		Athlete		($17,785.78)
		Firefighter		($12,140.09)
		Movie Star		($13,243.65)
		Musician		($11,436.54)
		Police		($12,933.49)

In the given page report, one can select years from drop down list and can analyze region-wise and the key value will display Highest profit according to different regions. In the report we have made a tree mapshowing values Product Brand, Product Line, Product and profit. Using Total and subtotal option in the Option on the right side added subtotal to rows and column.

In the recent year, 2011 Highest profit by facility region was observed in East region. Then we wouldsuggest the company to increase sales in East region so as to maximize the profit.

BUSINESS SCENARIO #2

The company needs a dashboard to analyze its various facility regions based on profitability vs expensesso that it can study where expenses are more than profit and improve precisely on the areas of losses based on the visualization.

Solution:

We have made a dashboard where there is a bar chart visualization to show expense vs profit by facilityregion. An animated bubble plot which shows the relation of expenses vs profit over years. A line chart visualization which shows profit and expenses by facility region. a list control based on year which is linked to bar chart and line chart

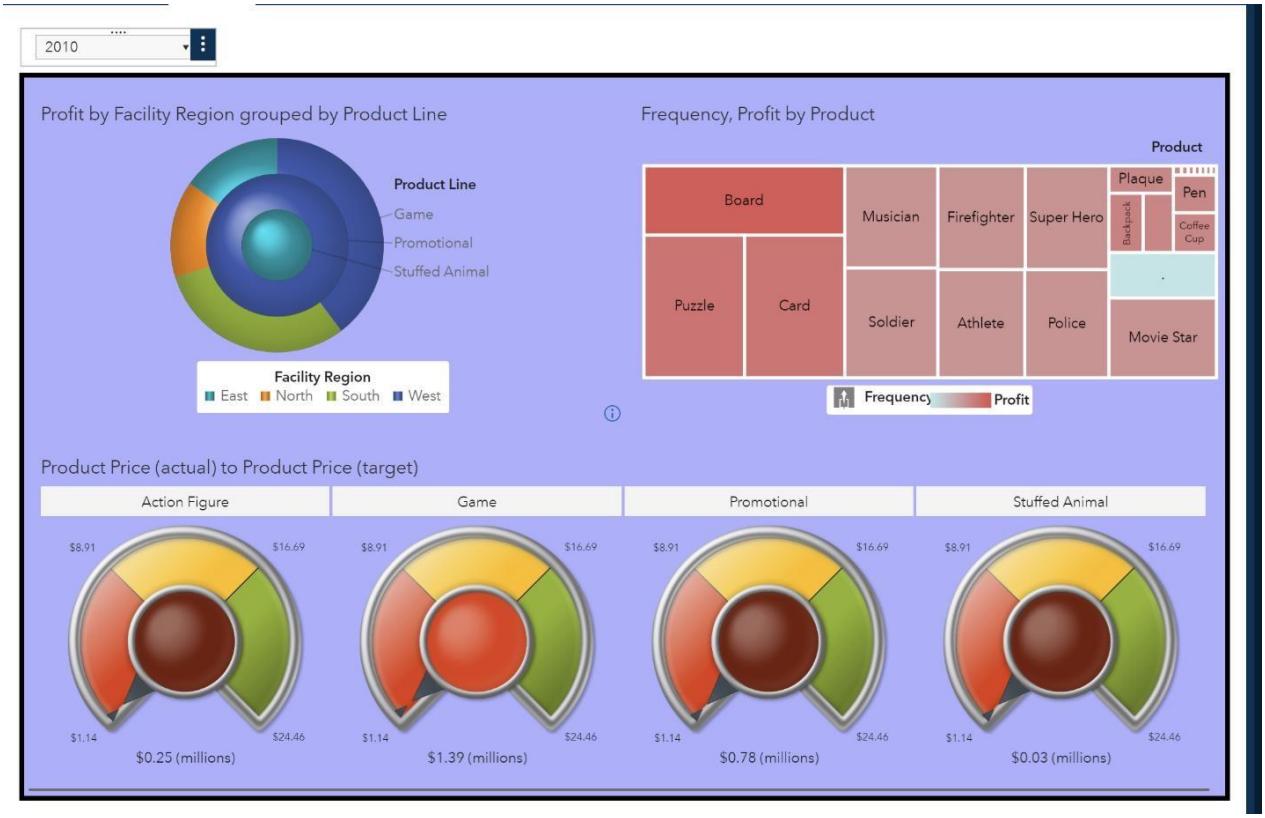
In the given page report, one can select years from drop down list and the bar graph will show the expenses and profit of each product line in given year/years. In the bar graph, we have assigned role such that the expenses is colored as blue and profit as Orange, The one below the reference line(i.e. 0$)shows negative profit. According to the report int the year 2005, the company has maximum profit int the west region with comparatively less expense so it must increase sales to maximize the profit and sale. And we would suggest the company to decrease production in north and east region.

BUSINESS SCENARIO #3

The company needed a dashboard to analyze its product lines and its products in terms of profitability,product price and target so that it can avoid losses, learn from past data and secure more profit on its products in future.

Solution:

We have prepared a dashboard which includes pie chart visualization based on profit by facility region grouped by product line, tree map visualization which reflects profit by product. Gauge dial visualizationwhich analyzes product price actual vs target.



In the given page report, one can select years from drop down list and the pie chart visualization based on profit by facility region grouped by product line, tree map visualization which reflects profit by product.Gauge dial visualization which analyzes product price actual vs target.

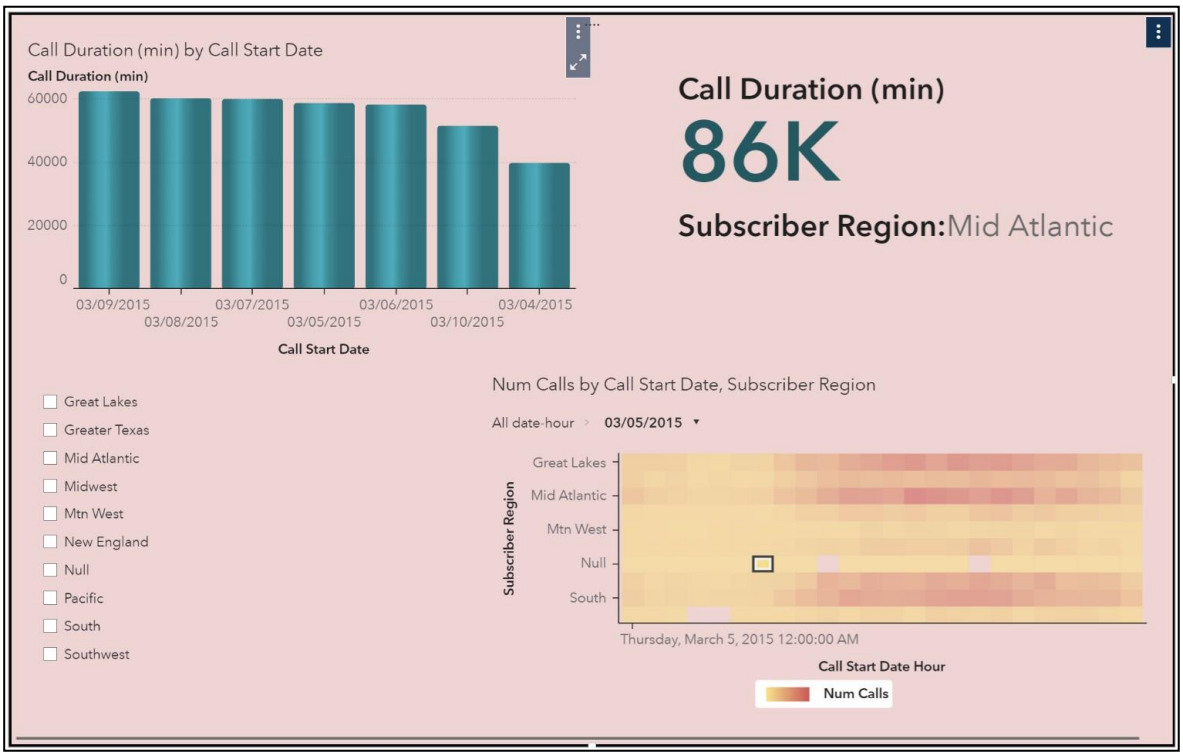
Dataset: Telecom

BUSINESS SCENARIO #4

A telecom company needs a dashboard to analyze and know more about the call duration, number of calls by subscriber region and call start date, hour so that the company knows where to launch its nextproduct based on highest call duration subscriber region for most profitability.

Solution:

We have prepared a dashboard with bar chart which shows call duration with call start date and is filteredby control object. A key value object which shows highest call duration and its subscriber region which is also linked by the control object. A heat map which analyzes number of calls by subscriber region, call start date-hour hierarchy.

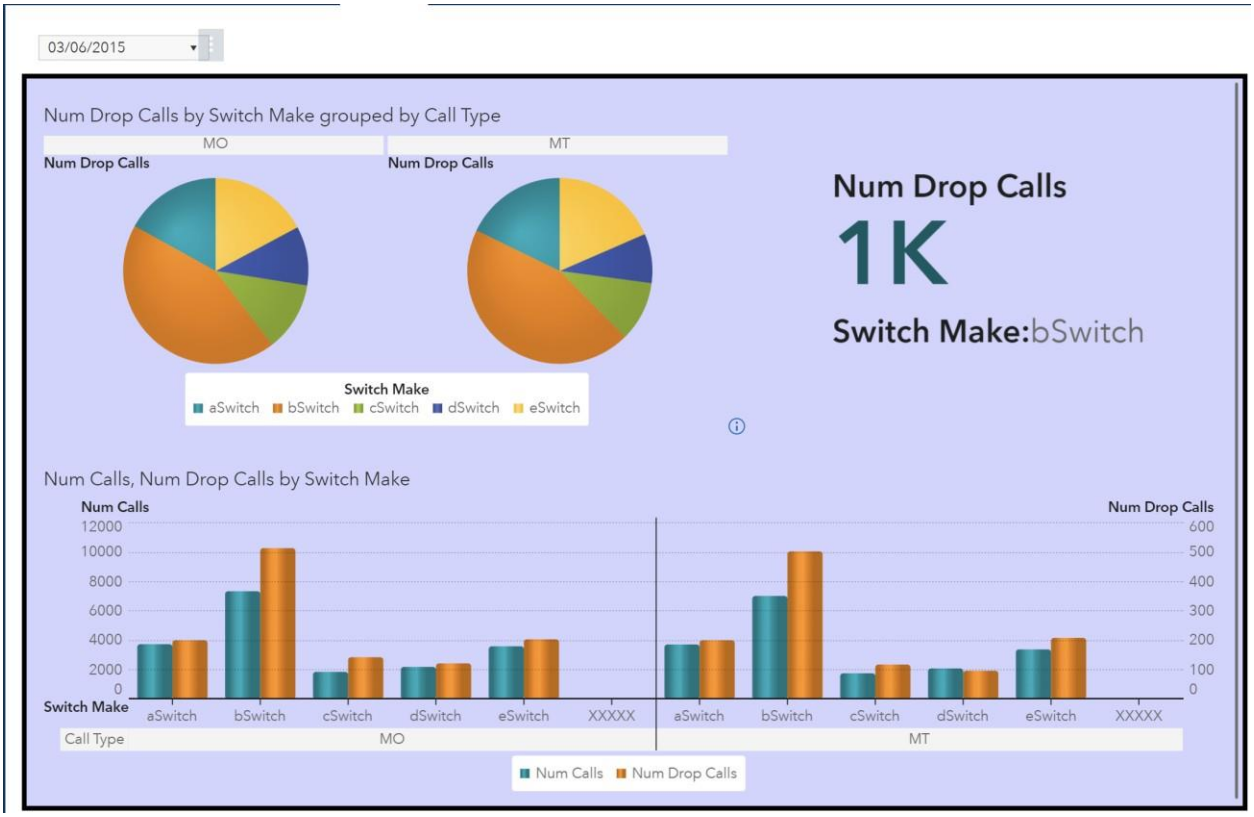


BUSINESS SCENARIO #5

A telecom company needs a dashboard to analyze different switch makes with respect to number of drop calls and call type so that it can get a better insight of which switch to improve based on the call type and maximum no. of calls dropped.

Solution:

We have prepared a dashboard which includes a pie chart visualization of number of calls dropped by switch make and grouped by call type. A key value object which shows highest number of calls droppedand its switch make. Dual axis bar chart which compares number of calls vs number of drop calls by switch make and call type.



Conclusion

The Result of the business Scenarios will help in aggregating Sales ,Employees , order Data offered in dataset .With the help of these strategies will able to analyse the datasets.