

Finally, a Tool for Business Users! A Step-By-Step Practical Approach to Pharma Sales Reporting Using SAS® Enterprise Guide® 4.3

Ramya Purushothaman, Cognizant Technology Solutions
Airaha Chelvakanthan Manickam, Cognizant Technology Solutions

ABSTRACT:

SAS® Enterprise Guide® can be looked at as the Integrated Development Environment (IDE) for SAS which can be used by business users even with minimal SAS programming experience to derive meaningful reports. This paper explains a step-by-step representation of a real-time scenario in the Pharmaceutical industry addressed through SAS Enterprise Guide. The example illustrated here covers importing the raw sales data purchased from an external vendor, perform business transformations, summarization and creating visual representation like charts, reports for Executive dashboard presentation, scheduling and e-mailing the reports to the end users without manual intervention. All these needs are easily accomplished using a series of simple clicks in SAS Enterprise Guide.

INTRODUCTION:

Since the release of SAS 9, SAS Enterprise Guide is the new face of SAS for general purpose use in a variety of situations. Enterprise Guide is now included with SAS Display Manager for Windows, sold with almost every BI Server sale, bundled with many of the SAS Solutions, and loved by many non-programmers and programmers too.

SAS Enterprise Guide provides users with very powerful data management capabilities. Frequently used data management capabilities include a sophisticated query builder, data sampling, ranking, transposing and even creating and editing data.

Since the introduction of SAS Enterprise Guide, the non-SAS programmers also benefit from simple to complicated transformations & reporting using SAS EG. It is clearly evident from this paper.

PHARMA SALES DATA REPORTING USING SAS ENTERPRISE GUIDE:

Of all the papers, the authors have come across from various SAS forums, there is very few that target the real intended use of SAS Enterprise Guide, which is to make the job easier for business users. The authors, having Pharma Sales data processing background, have defined the goal of this paper to provide an easy approach with simple point and click events to process and report Pharma Sales data for the benefit of business users.

This paper explains the following key components of Pharma Sales processing and how they can be easily completed using SAS Enterprise Guide 4.3.

1) Data Extraction & Transformation

- Extraction uses the Import wizard and Transformation uses query builder components of SAS Enterprise Guide to read and apply business transformations on data.

2) Reporting

- Reporting uses various summary and frequency statistics and report wizards in charts.

3) Automatic Scheduling using inbuilt Scheduler

4) Distribution via Email

Image (1) shows the common data flow of Pharma Sales data and the role of Data Warehouse in processing the data.

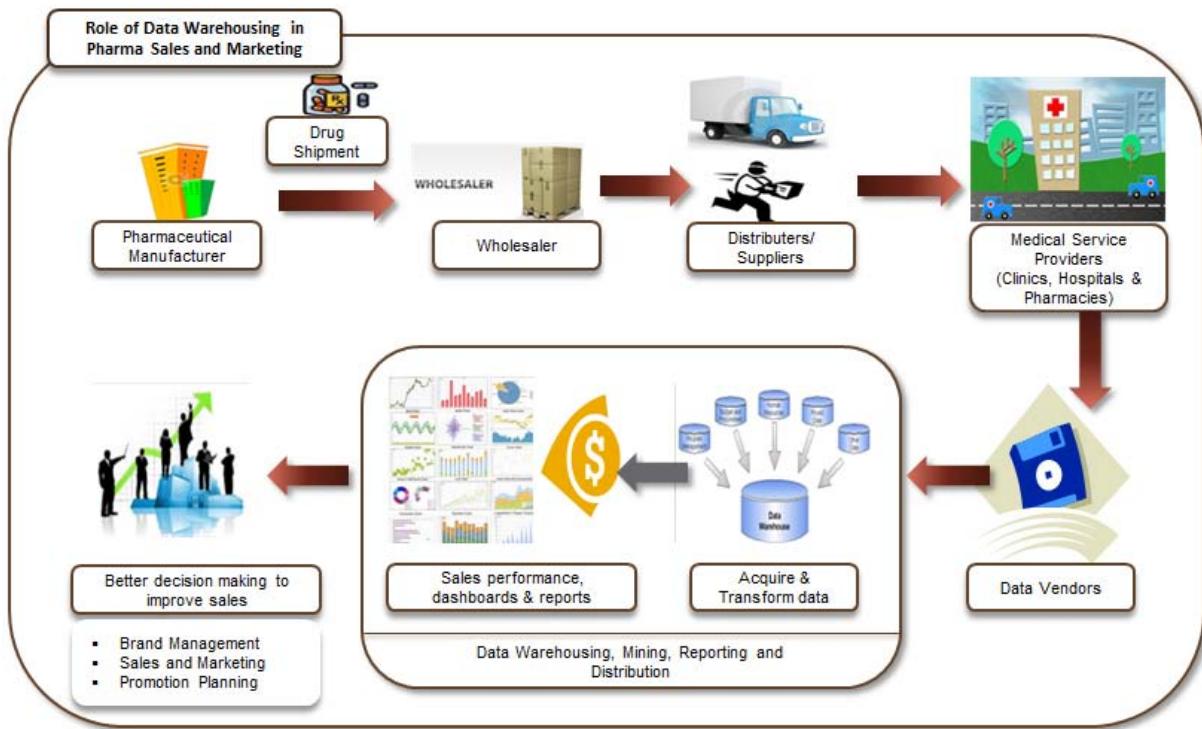


Image 1: Essence of the role of Data Warehousing in Pharma Sales Reporting

Image (2) shows the various functions of the Data Warehouse in processing and reporting Pharma Sales data and how they are covered by SAS Enterprise Guide 4.3.

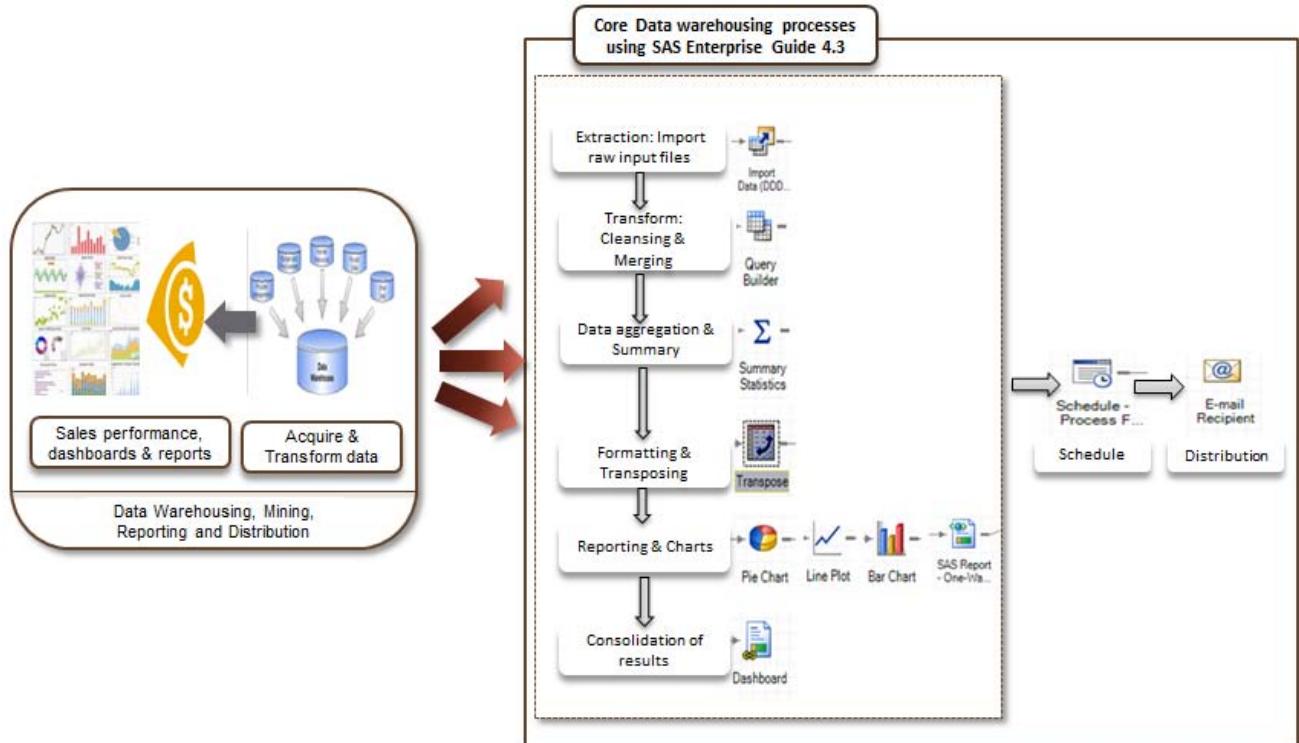


Image 2: Components of Data Warehousing achieved through SAS Enterprise Guide

Performing these core data warehousing functions in SAS Enterprise Guide 4.3 is taken as a case-study in this paper and explained.

EXTRACTION AND TRANSFORMATION OF SALES DATA:

In this section, raw Pharma Sales data is acquired, imported into SAS Enterprise Guide and transformed using Query Builder.

Image (3) shows the flow in SAS Enterprise Guide 4.3 with the extraction and transformation components in it. Details of each component are given below as separate steps.

For the purpose of the case-study, DDD (Drug Data Distribution) data is taken as an example.

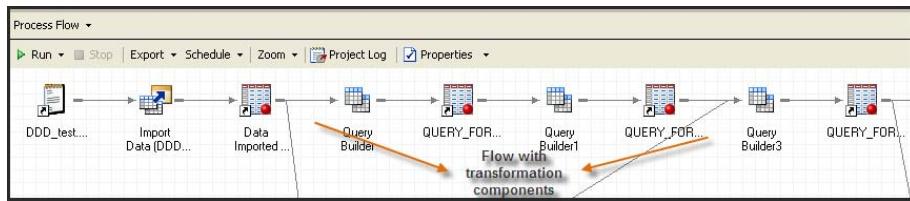


Image 3: Extraction & transformation components in SAS Enterprise Guide

Step: 1 – Importing Dollar/Unit demand data

This step imports raw purchased DDD data from a third party vendor using the import feature available in SAS Enterprise Guide. As seen in the picture below, the first record of the file is the header record.

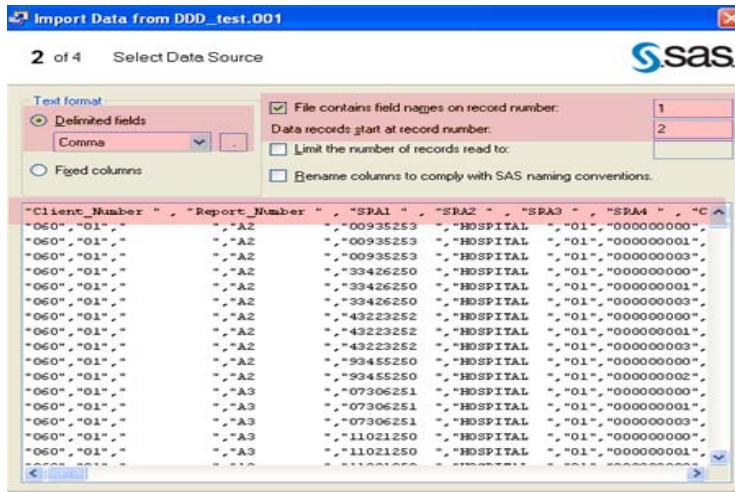


Image 4: Importing data into SAS Enterprise Guide

The import feature in SAS Enterprise Guide automatically detects the best informat/format for every data element in the file and it also allows the user to change the informat/format if he wants to. In this example, the informat/format of Zip code is changed to "string" to show all five digits – as shown below.

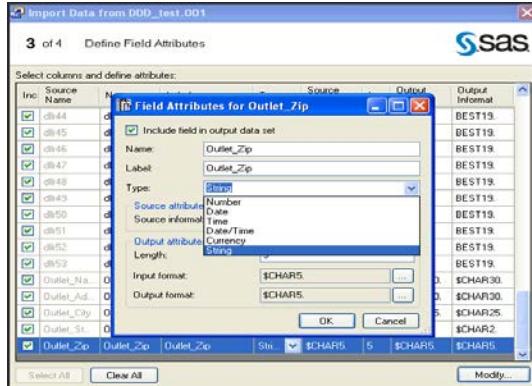


Image 5: Importing data – changing field type

If needed, the user can also change the name of the variable in the SAS dataset to be created. For example in the screen below the input variable “SRA4” form the raw data is changed to “Distribution_channel” to make it more descriptive for reporting purposes.



Image 6: Importing data – changing field name

Step: 2 – Applying transformations on data

Query builder is a powerful component of SAS Enterprise Guide 4.3 and it allows applying business transformations and joins on the data very easily. In step: 2, we are extracting only the last 3 bytes from the 9 byte product group field to enable merge with other sources down the line.

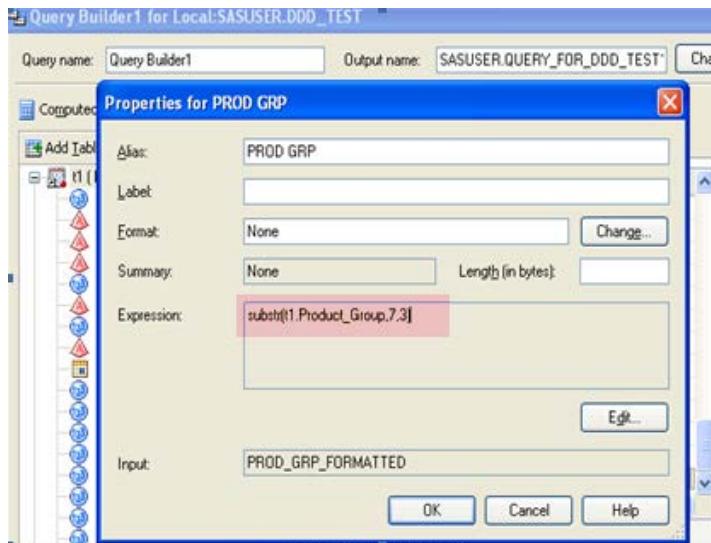


Image 7: Query Builder – adding substring function

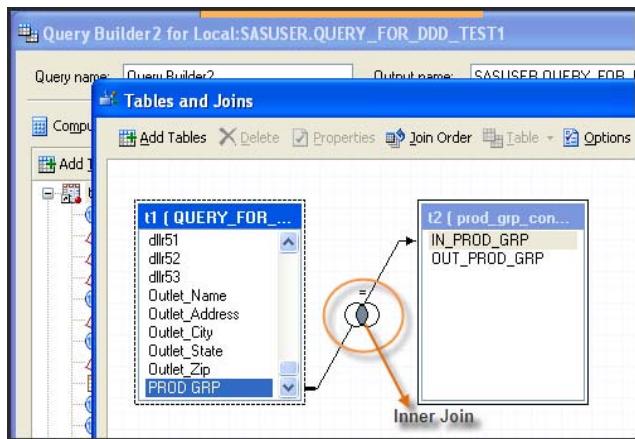
The resultant data is shown below with the new field created for “prod_grp”.

Outlet_Address	Outlet_City	Outlet_State	Outlet_Zip	prod_grp
RENOS CEN...	SAN JUAN	PR	00935	000
RENOS CEN...	SAN JUAN	PR	00935	001
RENOS CEN...	SAN JUAN	PR	00935	003
S CONGRES...	BOYNTON BEA...	FL	33426	000
S CONGRES...	BOYNTON BEA...	FL	33426	001
S CONGRES...	BOYNTON BEA...	FL	33426	003
W BROAD S...	COLUMBUS	OH	43223	000
W BROAD S...	COLUMBUS	OH	43223	001

Image 8: Query Builder – result of substring function

Step: 3 – Join to get OUT PROD GRP

The inbound data contains product group number which needs to be converted to a different product group number that is unique and meaningful within the organization and recognized by the in-house systems. So this step joins the input data with a look-up table called “prod_grp_conversion” on PROD_GRP field to get OUT_PROD_GRP. The additional field OUT_PRD_GRP is appended with the input data to the right.



The screenshot shows the 'Tables and Joins' interface in the Query Builder. A blue dashed box encloses the left table, 't1 (QUERY_FOR_DDD_TEST1)', which contains columns: dl151, dl152, dl153, Outlet_Name, Outlet_Address, Outlet_City, Outlet_State, Outlet_Zip, and PROD_GRP. An orange dashed box encloses the right table, 't2 (prod_grp_con...)', which contains columns: IN_PROD_GRP and OUT_PROD_GRP. A line with a circular arrow icon connects the two tables, labeled 'Inner Join'. Below the interface is a result table with the following data:

Outlet_City	Outlet_State	Outlet_Zip	PROD_GRP	OUT_PROD_GRP
SAN JUAN	PR	00935	001	010
SAN JUAN	PR	00935	003	030
BOYNTON BEA...	FL	33426	001	010
BOYNTON BEA...	FL	33426	003	030
COLUMBUS	OH	43223	001	010
COLUMBUS	OH	43223	003	030
SANTA MARIA	CA	93455	002	020
JERSEY CITY	NJ	07306	001	010
JERSEY CITY	NJ	07306	003	030
GREAT NECK	NY	11021	001	010
GREAT NECK	NY	11021	003	030
MANHASSET	NY	11030	001	010
MANHASSET	NY	11030	003	030
BRANDON	FL	22511	001	010

An orange arrow points from the 'OUT_PROD_GRP' column in the table to the text 'New field from prod_grp_conversion table'.

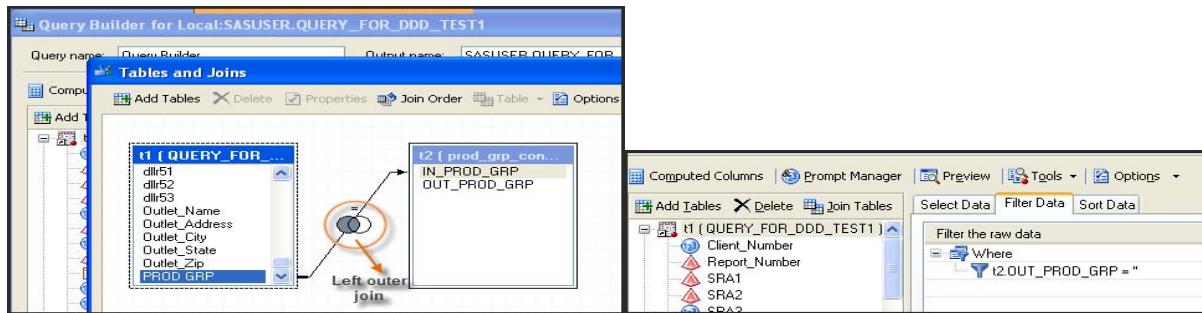
Image 9: Query Builder – join on product group table & the result with OUT_PROD_GRP

Step: 4 – Capturing exception records from the JOIN

The join at step: 3 is an “inner join”, which means any unmatched input records (which did not get a hit on joining with “prod_grp_conversion” table) are dropped. It is imperative that the dropped records need to be captured for any further analysis and further corrective action. This is explained in this step.

This step joins same tables as in step: 3, but with a “left outer join”. This way all the records of the left table, which is the input data, are written at the output of join. Then a filter condition is added to include only records with OUT_PROD_GRP = blank. The resultant data are only unmatched records from the “inner join”.

The output is routed to an exception file, which can then be routed to the appropriate department by exporting to an MS excel.



The screenshot shows the 'Tables and Joins' interface in the Query Builder. A blue dashed box encloses the left table, 't1 (QUERY_FOR_DDD_TEST1)', which contains columns: dl151, dl152, dl153, Outlet_Name, Outlet_Address, Outlet_City, Outlet_State, Outlet_Zip, and PROD_GRP. An orange dashed box encloses the right table, 't2 (prod_grp_con...)', which contains columns: IN_PROD_GRP and OUT_PROD_GRP. A line with a circular arrow icon connects the two tables, labeled 'Left outer join'. To the right is a 'Select Data' panel with a 'Filter Data' tab selected, containing the condition 'Where t2.OUT_PROD_GRP = '''. Below the interface is a result table with the following data:

Client_Number	Report_Number	SRA1	SRA2
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30
31	31	31	31
32	32	32	32
33	33	33	33
34	34	34	34
35	35	35	35
36	36	36	36
37	37	37	37
38	38	38	38
39	39	39	39
40	40	40	40
41	41	41	41
42	42	42	42
43	43	43	43
44	44	44	44
45	45	45	45
46	46	46	46
47	47	47	47
48	48	48	48
49	49	49	49
50	50	50	50
51	51	51	51
52	52	52	52
53	53	53	53
54	54	54	54
55	55	55	55
56	56	56	56
57	57	57	57
58	58	58	58
59	59	59	59
60	60	60	60
61	61	61	61
62	62	62	62
63	63	63	63
64	64	64	64
65	65	65	65
66	66	66	66
67	67	67	67
68	68	68	68
69	69	69	69
70	70	70	70
71	71	71	71
72	72	72	72
73	73	73	73
74	74	74	74
75	75	75	75
76	76	76	76
77	77	77	77
78	78	78	78
79	79	79	79
80	80	80	80
81	81	81	81
82	82	82	82
83	83	83	83
84	84	84	84
85	85	85	85
86	86	86	86
87	87	87	87
88	88	88	88
89	89	89	89
90	90	90	90
91	91	91	91
92	92	92	92
93	93	93	93
94	94	94	94
95	95	95	95
96	96	96	96
97	97	97	97
98	98	98	98
99	99	99	99
100	100	100	100

Image 10: Query Builder – Join to get exception records

Step: 5 – JOIN to obtain TC key and Product Name

This step joins data from step: 3 with a look up table called “tc_mkt_prod_mapping” which is the product hierarchy maintained in-house, on two columns – Report_Number and OUT_PROD_GRP to obtain Therapeutic Class, Market basket and product names. The join performed at this step is an inner join. Image (11) explains the join and resultant data.

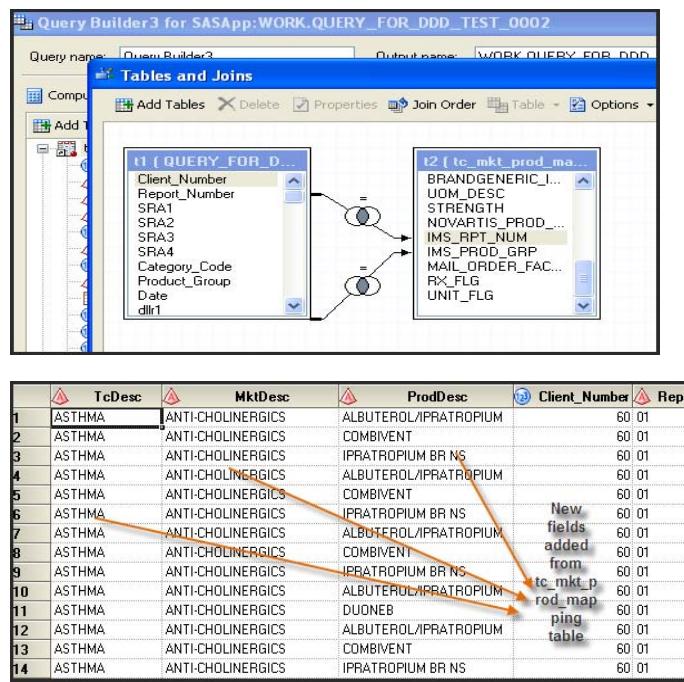


Image 11: Query Builder – Join to get Therapeutic Codes

At the end of Step: 5, the raw data has been imported, transformed to the needs of the business user and ready for reporting.

REPORTING OF SALES DATA:

In this section, the transformed Pharma Sales data from the previous section is used to produce meaningful charts for knowing the insights of the data and for management presentations.

Following picture shows the flow in SAS Enterprise Guide 4.3 with the Reporting components in it. Details of each component are given below as separate steps.

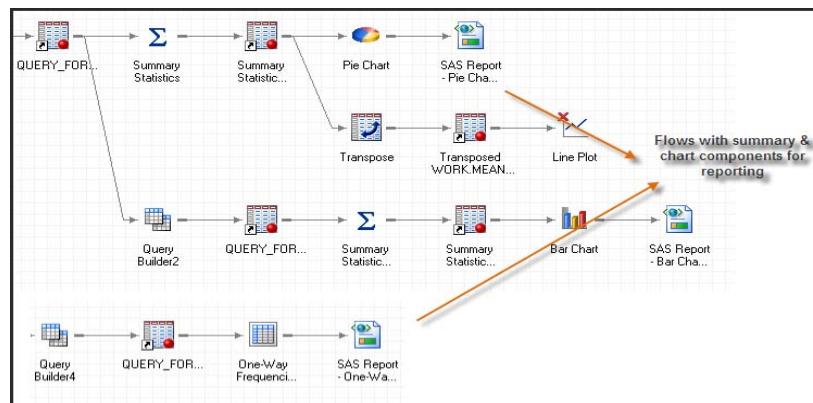


Image 12: Reporting components in SAS Enterprise Guide

Step: 6 – One-way frequency reporting on distribution channels

This is the first report produced from this exercise. This step reads output of Step: 5 to include only certain unique fields on the data using a query builder component and performs a one-way frequency component on the distribution channel to report the frequency on different channels across country that sell in that Therapeutic class taken.

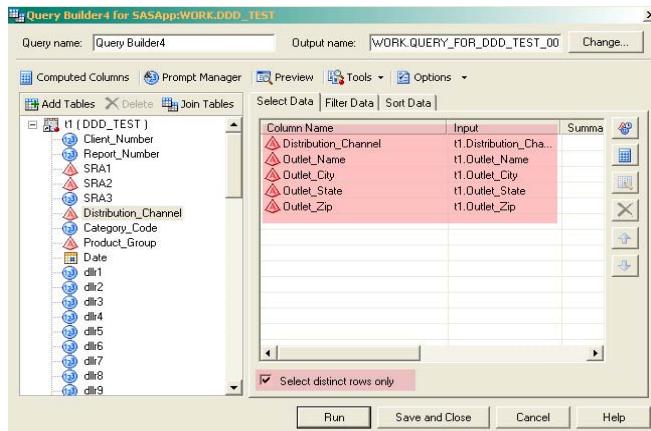


Image 13: Query Builder to select fields for one-way frequency reporting

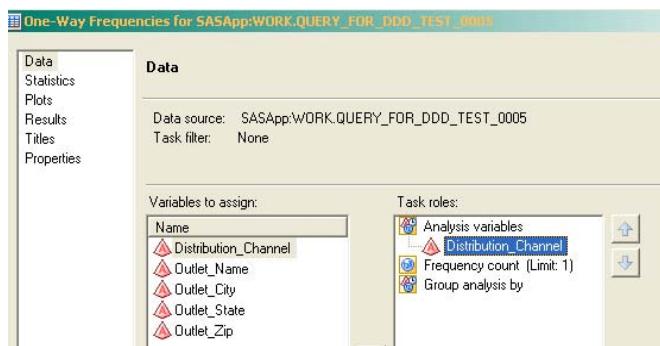


Image 14: One-way frequency reporting

Image (15) shows the result of one-way-frequency distribution as counts and as a vertical bar Chart.

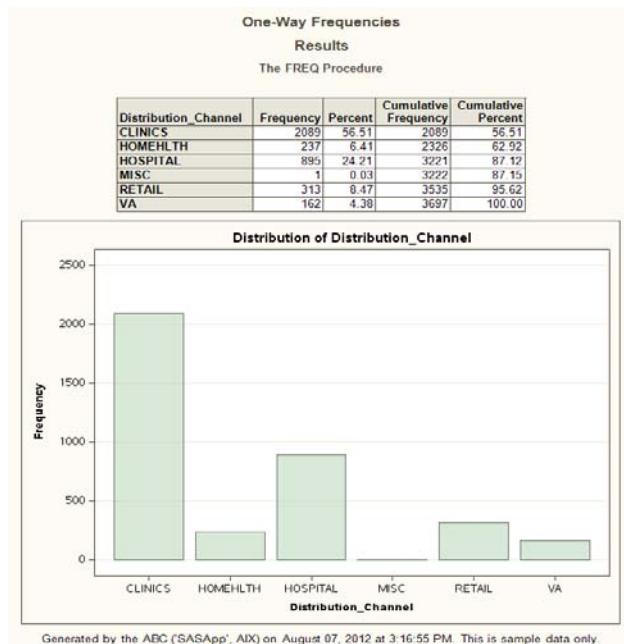


Image 15: Resultant graph from one-way frequency reporting

Step: 7 – Adding Summary Statistics and Pie-Charting on sales dollars by Product

At this step, a summary statistics component is added to perform summary of all 53 weeks of sales dollars by products and their therapeutic class at a nation level.

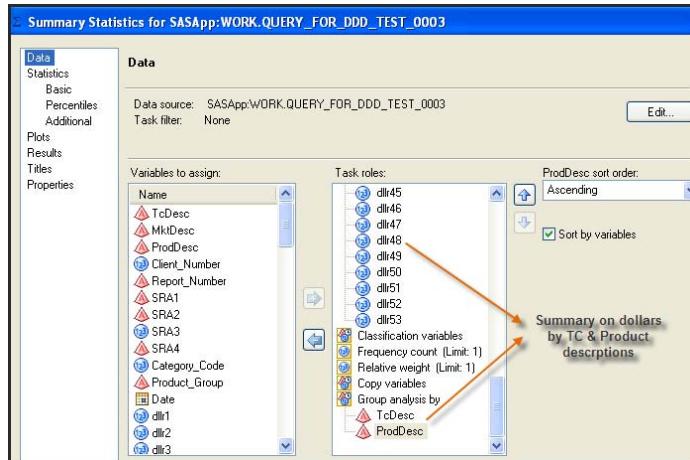


Image 16: Summary statistics wizard – selecting fields to report

Next, a pie-chart component is added to represent the distribution of first-week's dollars by products for each Therapeutic Category (TC).

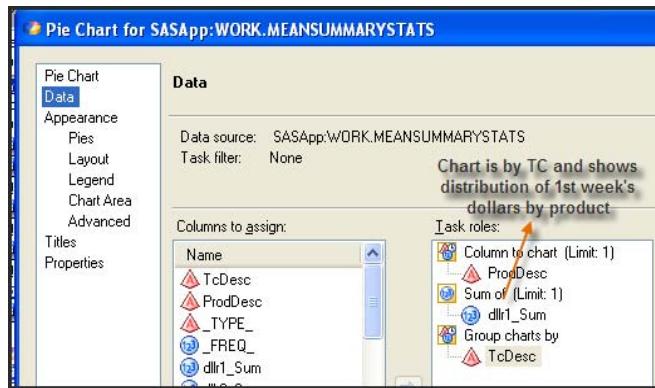


Image 17: Pie chart wizard

As there is only one Therapeutic Category (TC) – ASTHMA – in the sales data, one pie-chart is produced.

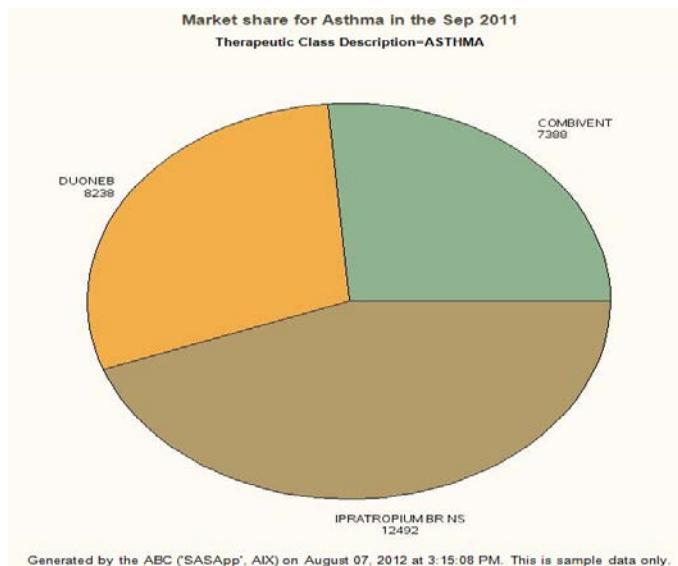


Image 18: Resultant Pie chart by Therapeutic Category

Step: 8 – Transpose and Line-Chart on sales dollars by week for all products

Line-charting is the suitable option to show multiple weeks/months of sales data for multiple products to analyze the trending of sales across a wide time period. The pre-requisite for performing such line-charting is to have the sales dollars in one column (normalized data). This can be achieved by using a transpose wizard in the Enterprise guide.

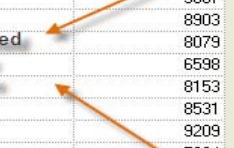
The Transpose component takes output of the summary statistics from step: 7 and transposes de-normalized sales dollars to normalized sales dollars.

Data before and after transpose component & the selected options of transpose component are shown in the Images (19) & (20) below.



TcDesc	ProdDesc	_TYPE_	_FREQ_	dllr1_Sum	dllr2_Sum	dllr3_Sum	dllr4_Sum
ASTHMA	ALBUTEROL/IP...	0		1889	7734	6456	6792
ASTHMA	COMBIVENT	0		1731	6617	5687	6090
ASTHMA	DUONEB	0	43	4	0	0	1
ASTHMA	IPRATROPIUM...	0		1337	1113	769	702
De-normalized Sales dollars							

Image 19: Data before transposing



ProdDesc	TRANSPOSED_DOLLARS	TC_ASTHMA
ALBUTEROL/IP...	dllr1_Sum	7734
ALBUTEROL/IP...	dllr2_Sum	6456
ALBUTEROL/IP...	dllr3_Sum	6792
ALBUTEROL/IP...	dllr4_Sum	9067
ALBUTEROL/IP...	dllr5_Sum	8903
ALBUTEROL/IP...	dllr6_Sum	8079
ALBUTEROL/IP...	dllr7_Sum	6598
ALBUTEROL/IP...	dllr8_Sum	8153
ALBUTEROL/IP...	dllr9_Sum	8531
ALBUTEROL/IP...	dllr10_Sum	9209
ALBUTEROL/IP...	dllr11_Sum	7084
ALBUTEROL/IP...	dllr12_Sum	8106
ALBUTEROL/IP...	dllr13_Sum	7568
Normalized Sales dollars		

Image 20: Data after transposing

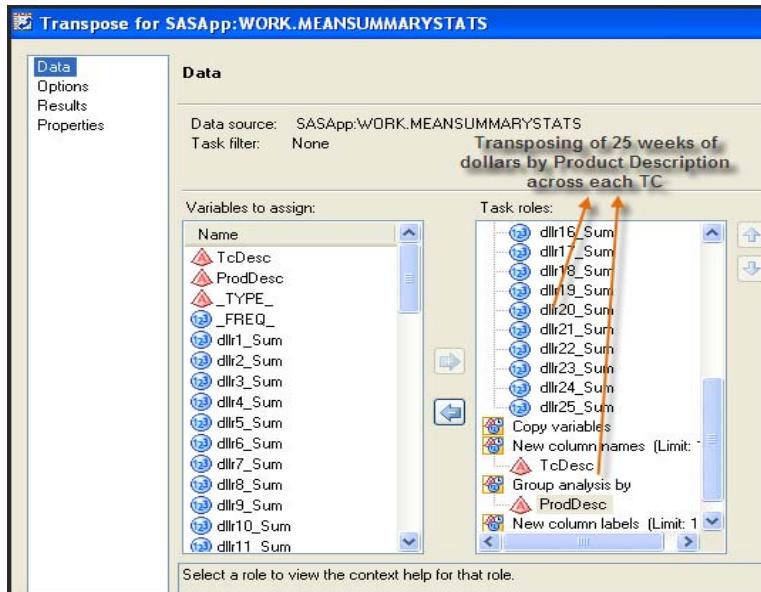


Image 21: Selected transpose options

Transposed & Normalized sales dollars of 25 weeks are passed to the line-chart component. Line-charting is performed for various product groups with dollars in Y axis and months in X axis.

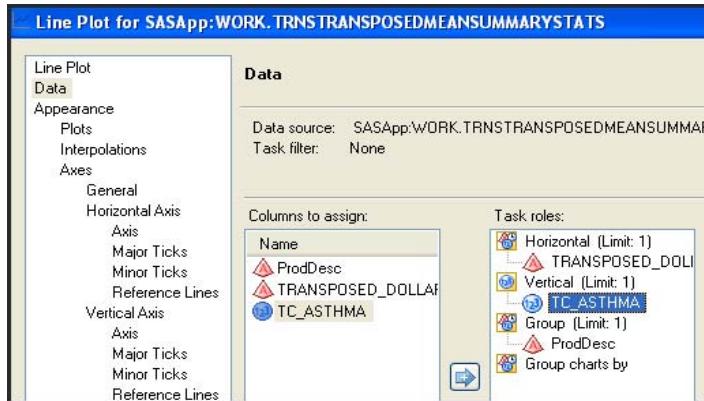


Image 22: Line chart wizard

Resultant line-chart is shown in Image (23) below.

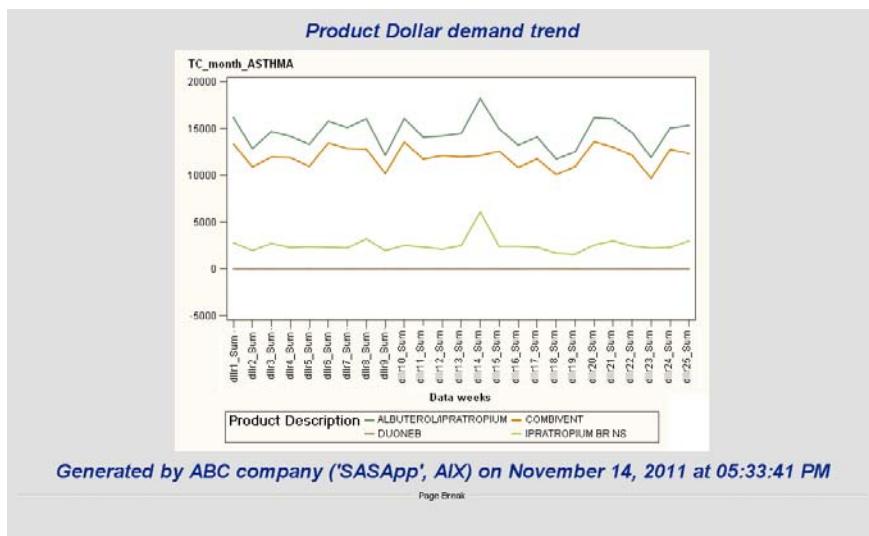


Image 23: Resultant line-chart for Therapeutic Code by product dollar

Step: 9 – Calculation of YTD totals & bar-charting of YTD totals by distribution channel

At this step, YTD (Year To Date) totals are calculated as the sum of all 53 weeks of dollars for every product across distribution channel. The “computed variables” option in query builder component is used to calculate YTD totals.

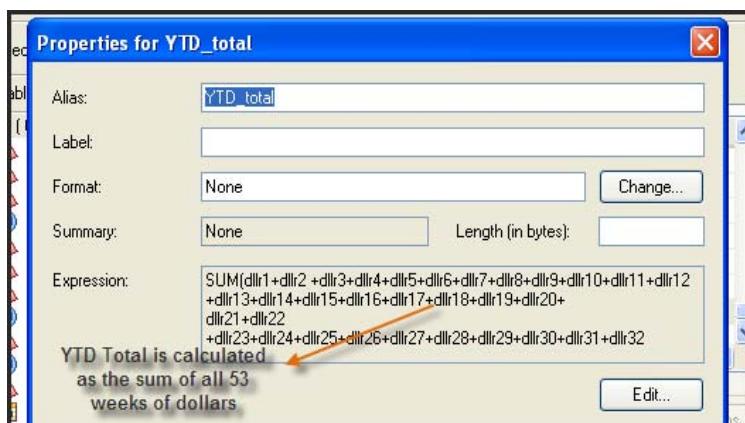


Image 24: Calculation of YTD values

A summary statistics component is added to produce the summary of YTD Sales across distribution channels (on SRA4).

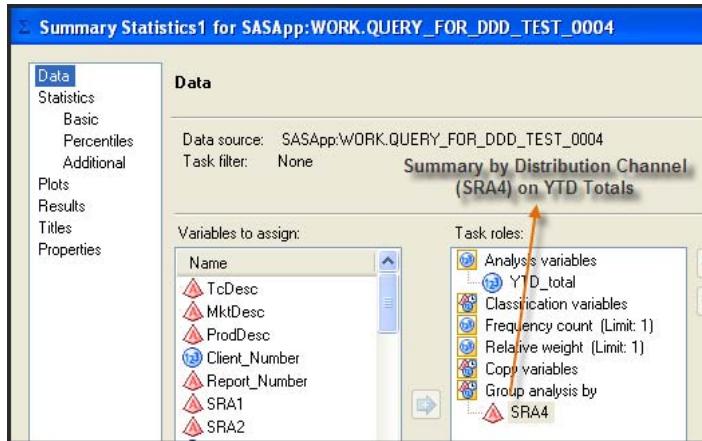


Image 25: Summary wizard for YTD sales

Bar-charting is performed on distribution channel to show trending of YTD totals on various channels. Resultant chart is shown below. Please note that the Title and Foot notes of the chart can be customized as well like in the example below

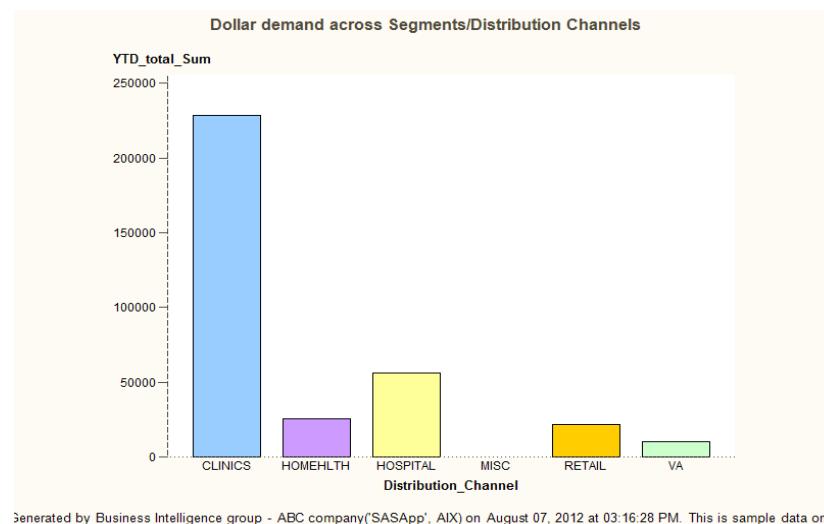
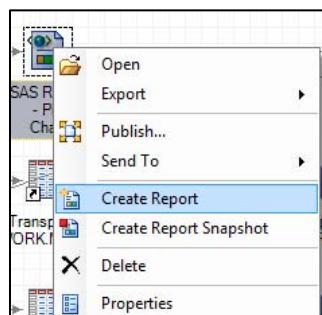


Image 26: Bar charting on YTD values

Step: 10 – Creating a Dashboard

All the reports created above can be combined into one dashboard that gives a comprehensive overview of the product performance.

Right click on one of the reports and say “Create Report”. In the popup window, the reports from left panel can be dragged and dropped into the layout in the right in desired spots to create a Custom dashboard view. Click “Show Preview” to see how it looks in the window below.



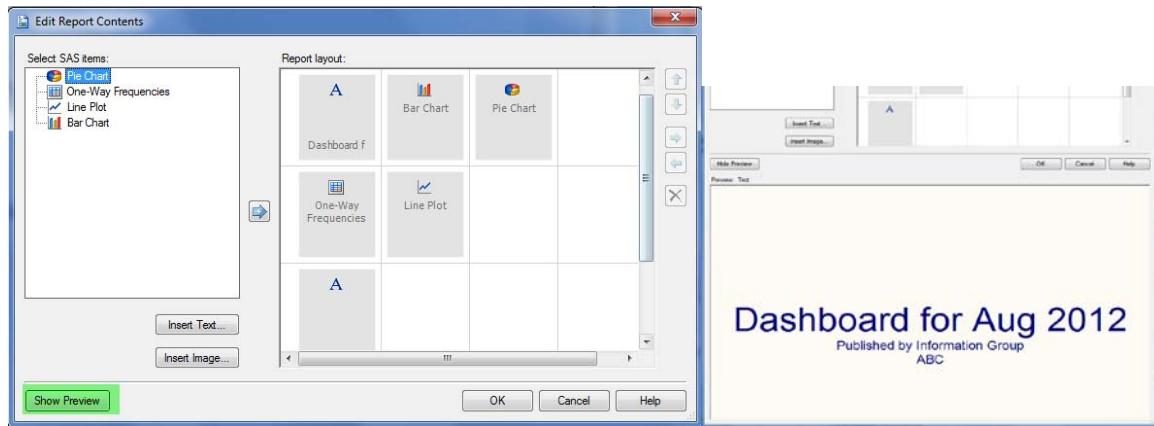


Image 27: Creating Dashboard

The final dashboard will look like this based on the grid layout specified above.

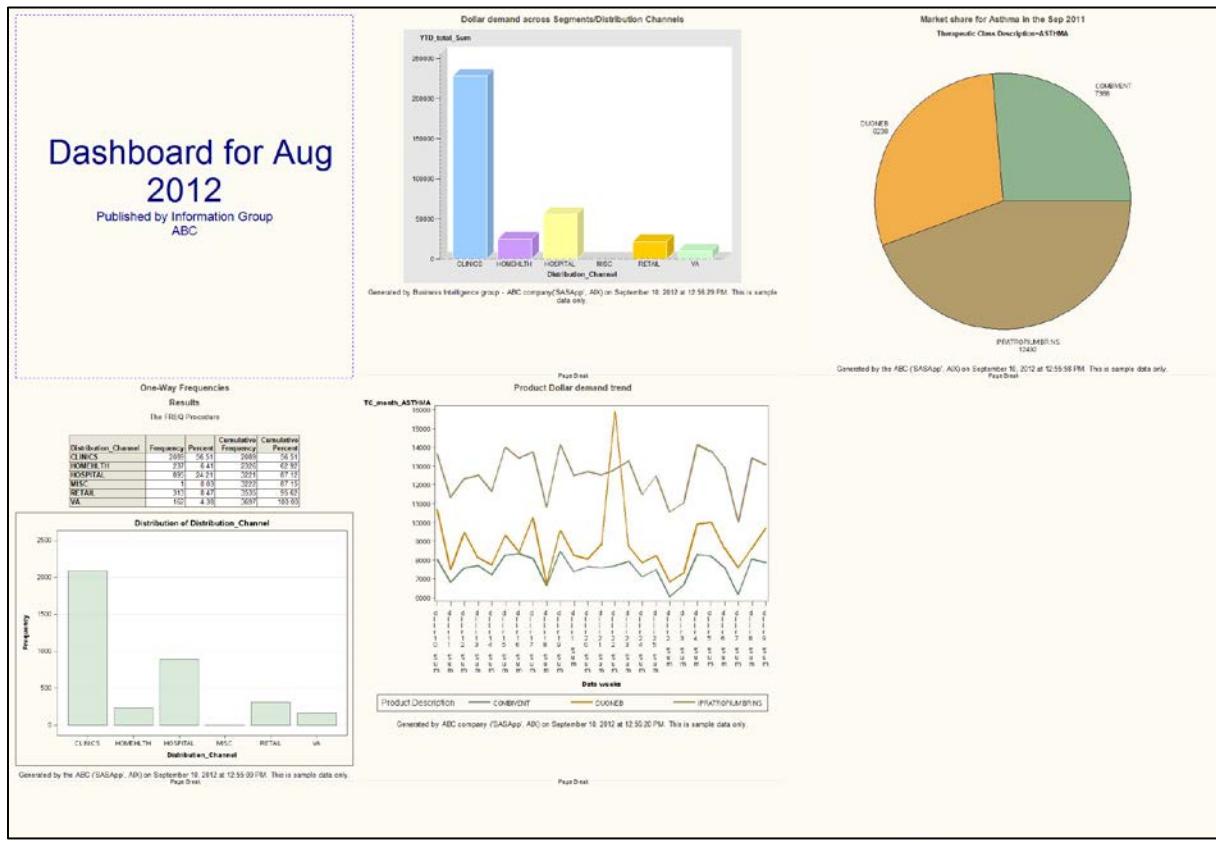


Image 28: Snapshot of the final Dashboard

SCHEDULING AND DISTRIBUTING REPORTS:

Scheduling a project comes handy when the user wants to create certain report(s) regularly without much manual intervention. SAS Enterprise guide offers a menu to schedule the process at a preferred frequency, date and time which creates a Script file with the schedule details and it is triggered when it is due.

Step: 11 – Scheduling in SAS Enterprise Guide 4.3

From the File Menu, Select “Schedule Project” as shown in Image (29).

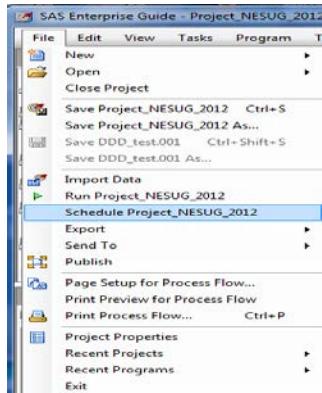


Image 29: File Menu with Schedule option highlighted

In the popup menu, click password and mention “Windows” authentication password. In the schedule tab, select ‘New’ to create a new schedule and specify when the user wants the process to run automatically, either once or recursively. More specifics on the end date till when the recursive schedule should hold good and frequency of run can be set using the *Advanced Schedule Options*.

Once scheduled, it is all set and can rest assured that the files and reports as part of the scheduled SAS Enterprise Guide project will be generated at the background while the user can stay focused on something else.

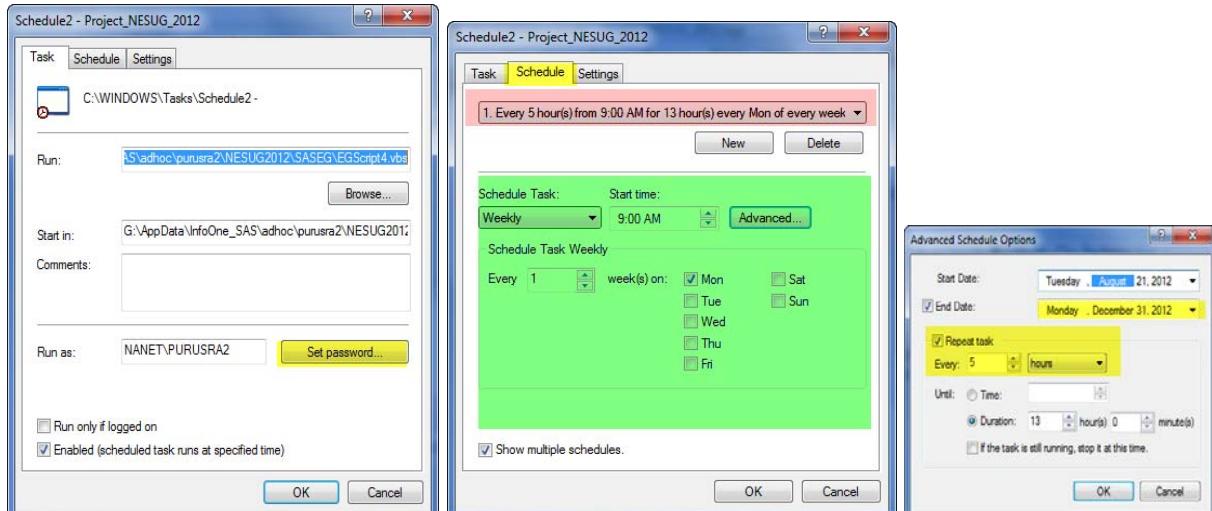


Image30: Scheduling wizard options

The schedule can be reviewed at the Schedules window and also the VB script created by the scheduler can be seen too – as shown in Image (31)

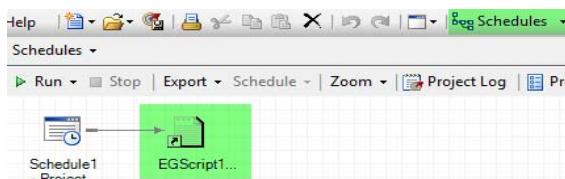


Image 31: Schedule option review in SAS Enterprise Guide

The schedule can be modified anytime by double-clicking on the schedule project icon for that schedule where the user will be taken to the same above set of windows – as shown in Image (30) – to make changes.

Step: 12 – Automated distribution of reports via e-mail:

Once the process is complete and reports are generated, the dashboard, in this case, can be emailed as a step in the project to one or more recipients, automatically.

Before proceeding to the email step, make sure the SMTP details specific to your organization are mentioned under Tools→ Options→ Administration for the emailing to work as step in project as shown in Image (32)

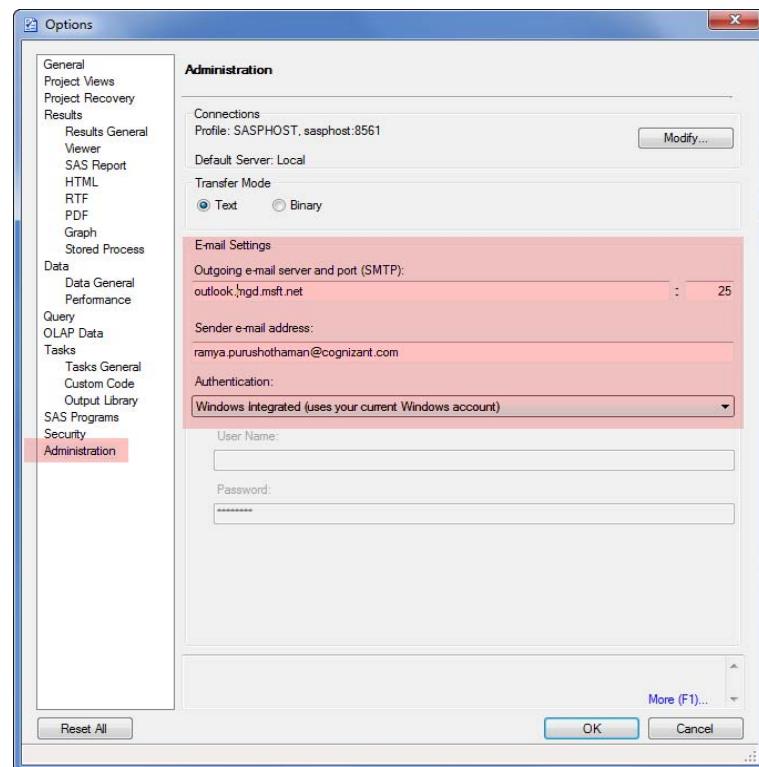


Image 32: SMTP set up for email

Right click on the Dashboard that needs to be emailed and go to Send To → Email as a Step in the Project

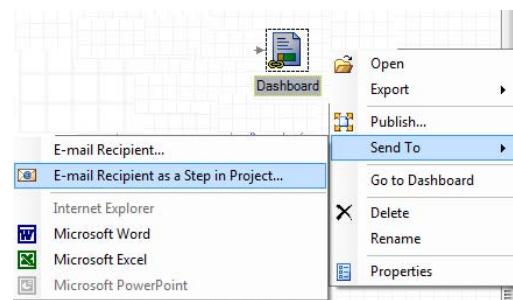


Image 33: Email as Step in Project

In the next window, the file to be sent can be compressed and renamed as needed.

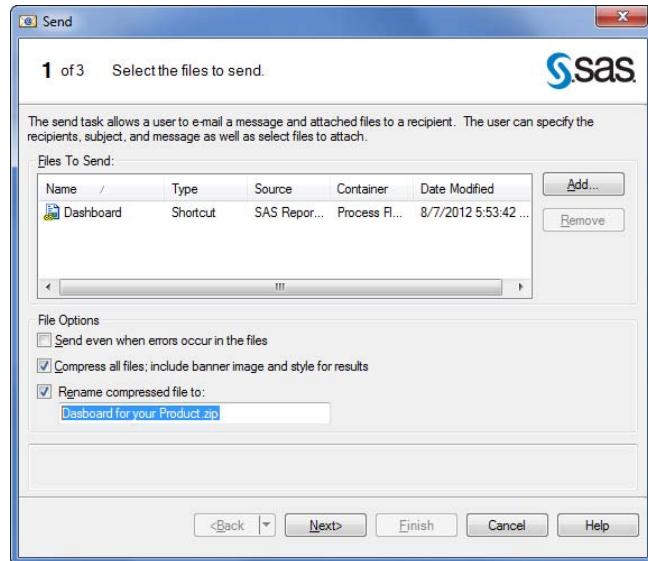


Image 34: Attachment specifications

Then, specify the To, CC, Subject and body of the message fields.

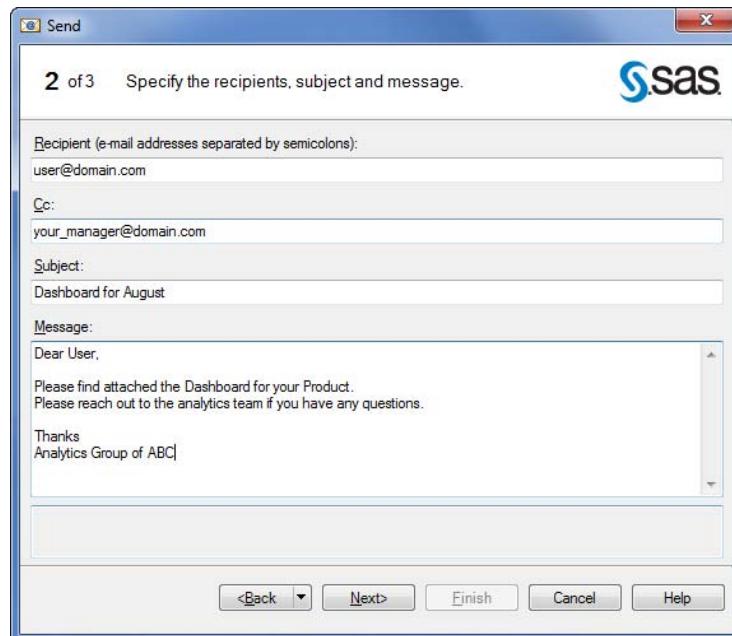


Image 35: Email specifications

Preview the message and click 'Finish'

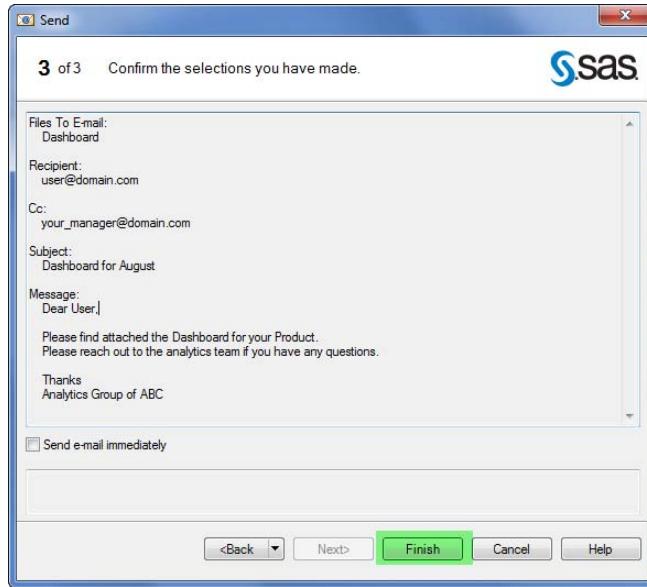


Image 36: Preview email

The email step will show up on the 'Distribute' window like below. If your project is scheduled to run at a preset time, the email with the dashboard will be sent automatically once the project is complete thus not requiring any intervention at all.

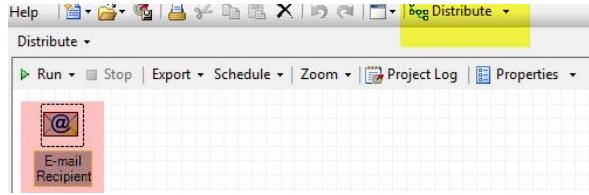


Image 37: Distribute window

If the user wants to have control over emailing, instead of emailing as step in project, he can always export the dashboard as step in the project to automatically write the report out to a specified location so that it can be reviewed before emailed manually.

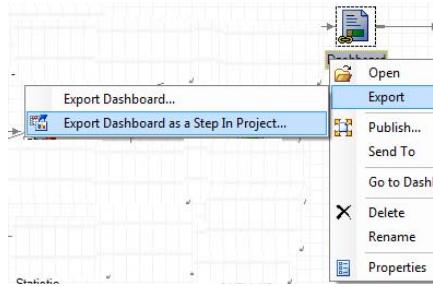


Image 38: Export as Step in Project

COMPLETE OVERVIEW OF ENTERPRISE GUIDE PROJECT:

Following is the complete Project with end – end process flow grouping the series of steps into relevant action performed.

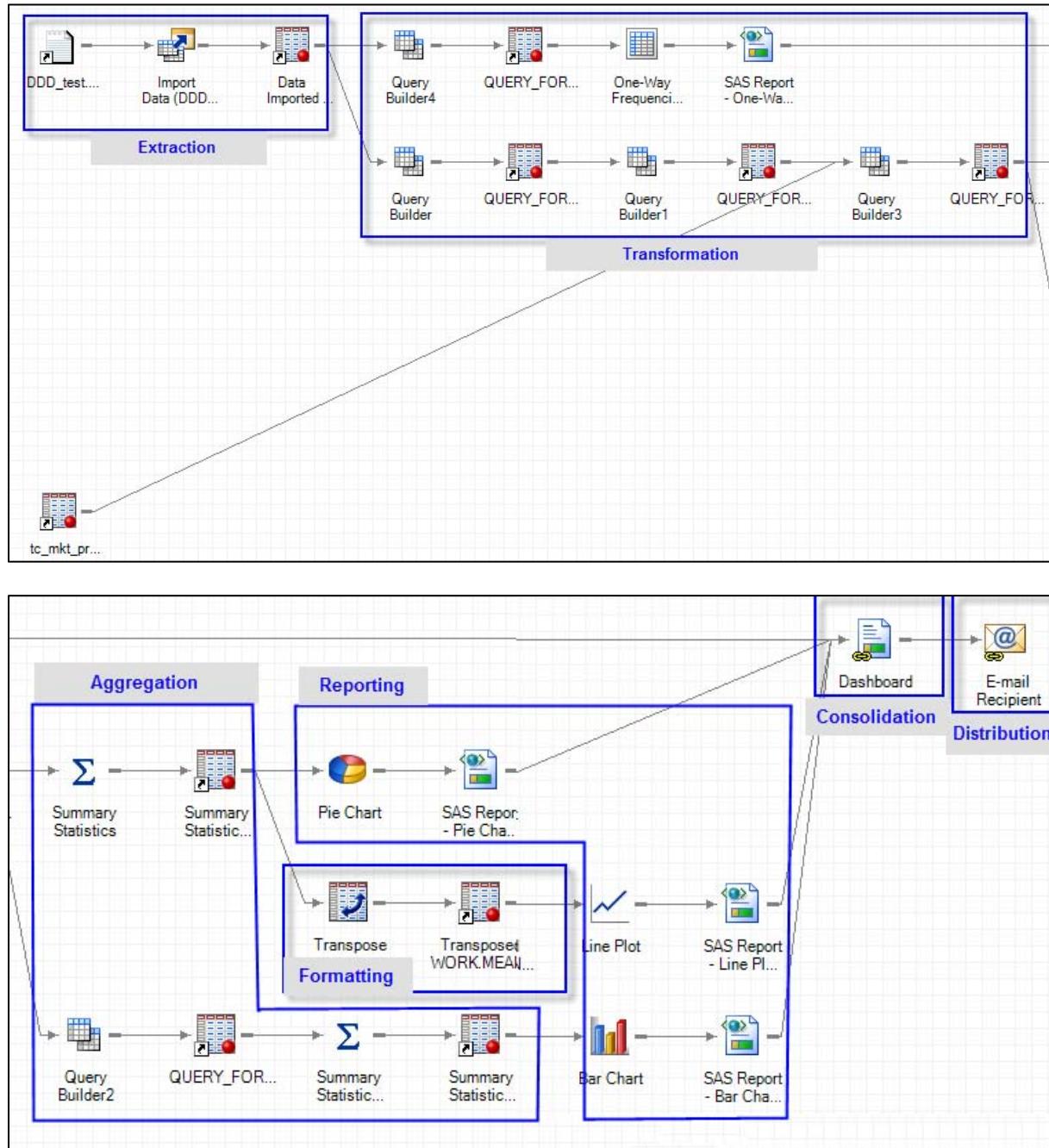


Image 39: Complete Process flow

CONCLUSION

While there may be many user manuals, tips and tricks, how-to's on SAS enterprise guide, we attempted to show you how we put it to work. As the paper explains, the common processing of Pharma Sales data, such as Extraction, Transformation and Reporting with automated scheduling & distribution can be done very efficiently by any business user without writing any SAS code. SAS Enterprise Guide 4.3 is a very powerful tool and can write the tedious SAS code for the business users.

ACKNOWLEDGEMENTS

The authors would like to thank SGF, NESUG and Cognizant Technology Solutions for the opportunity to present the paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name	:	Ramya Purushothaman
Enterprise	:	Cognizant Technology Solutions US Corporation
Address	:	500 Frank W.Burr Blvd
City, State, Zip	:	Teaneck, NJ, 07666
Phone	:	201 699 9438
Mail Id	:	Ramya.Purushothaman@cognizant.com

Name	:	Airaha Chelvakkantan Manickam
Enterprise	:	Cognizant Technology Solutions US Corporation
Address	:	500 Frank W.Burr Blvd
City, State, Zip	:	Teaneck, NJ, 07666
Phone	:	610 316 5780
Mail Id	:	Airahachelvakkantan.Manickam@cognizant.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.