



Transitioning to SAS[®] Enterprise Guide[®]

Michael Todd
Nth Analytics

1

September 15, 2008



Introduction

- Enterprise Guide is a very powerful programming environment
- However, it was designed as a tool for power users to run interactively
- This complicates using it to replace traditional SAS coding
- So we have the question: replacement, or adjunct?



Enterprise Guide Overview

- **Enterprise Guide (EG) is a visually oriented SAS programming environment**
- **Enables programmers to create diagrams linking objects that EG turns into SAS code**
 - **PROC SQL instead of data step**
 - **Reporting PROCs are PROC TABULATE and PROC PRINT**

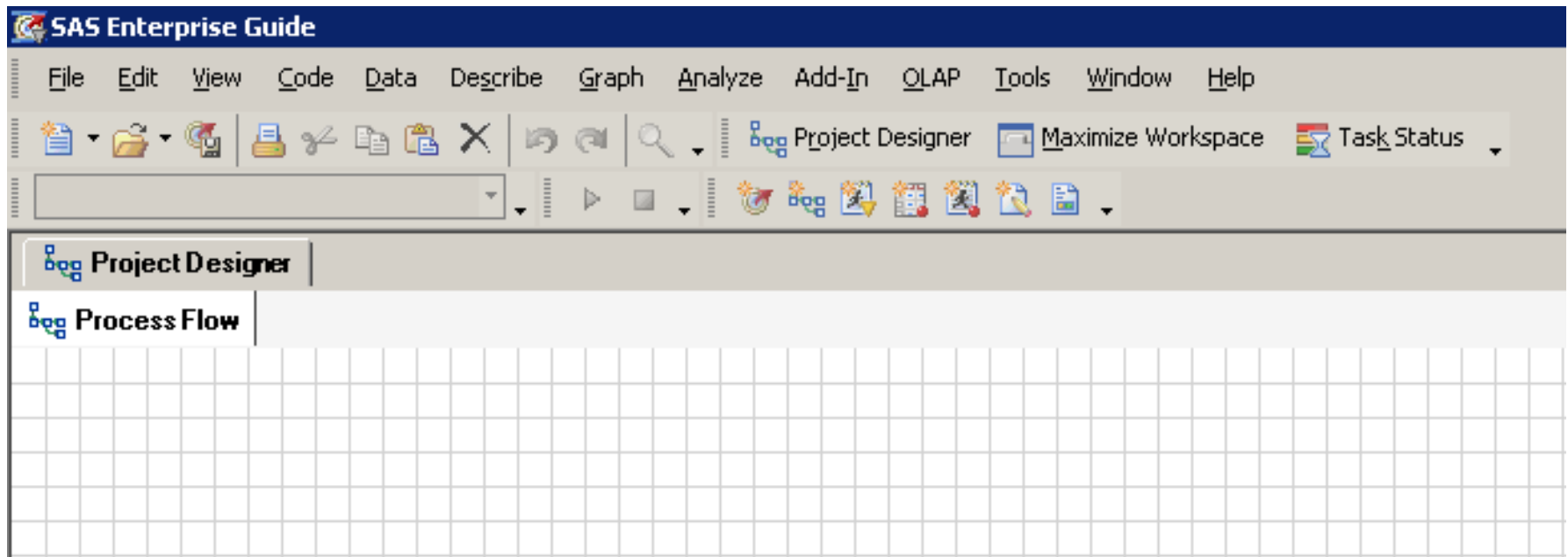


Process Flow

- **The main interface is the process flow.**
- **Initially, it is a blank screen on which to arrange the objects.**
- **The programmer sets properties for objects, and the objects generate SAS code “behind the scenes”.**
- **“Code objects” can also be included, containing user-written custom code**
 - **These can be linked to the built-in objects in a process flow.**

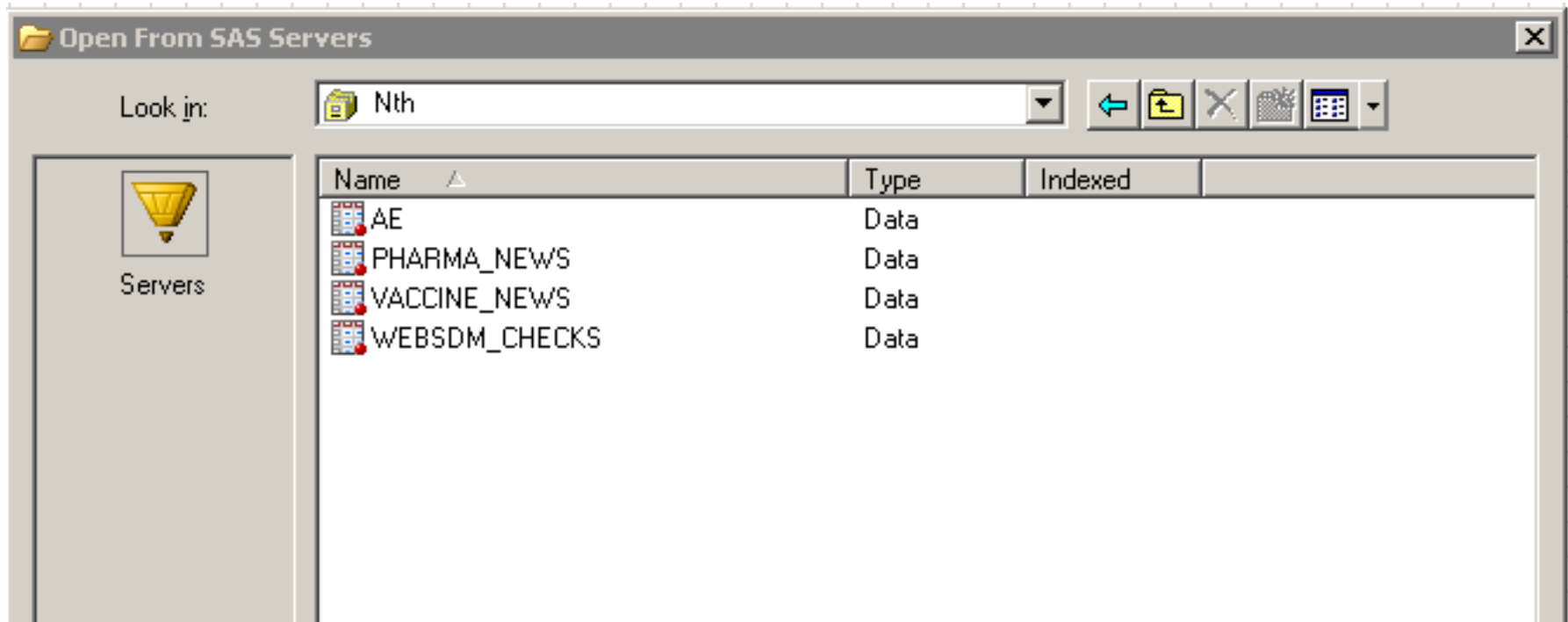
Enterprise Guide Demo

- **Start with a blank process flow**



Enterprise Guide Demo

- Load the data from one of the libraries



Enterprise Guide Demo

- **Build the query**

Query name: Output name:

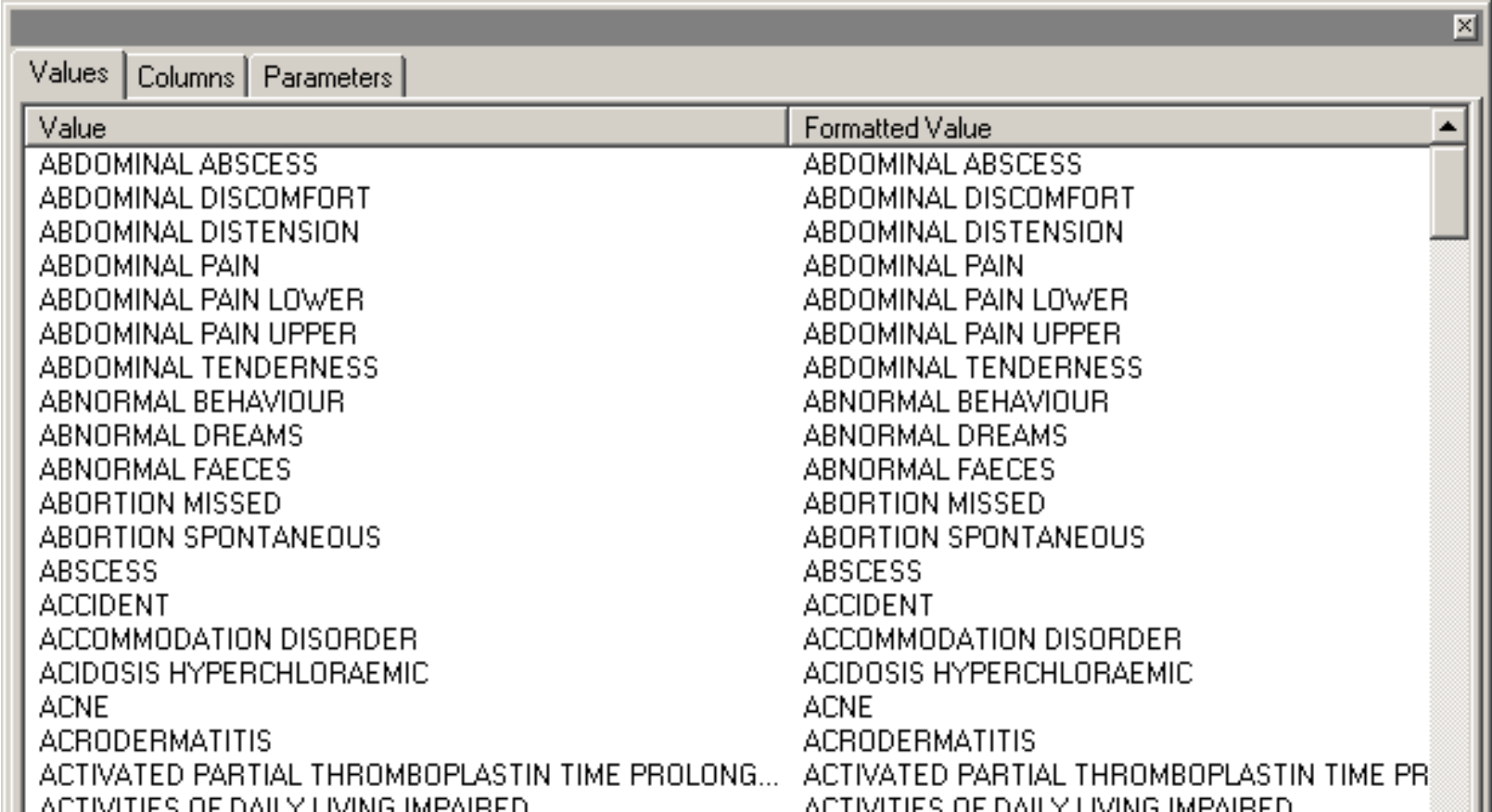
Select Data | Filter Data | Sort Data

Column Name	Input	Summary
STUDYID (Study Identifier)	AE.STUDYID	
USUBJID (Unique Subject Identifier)	AE.USUBJID	
SEX (Sex)	AE.SEX	
AGE (Age)	AE.AGE	
SAFETY (Safety)	AE.SAFETY	
ARM (Description of Planned Arm)	AE.ARM	
AEBODSYS (Body System or Organ System)	AE.AEBODSYS	
AEDECOD (Dictionary-Derived Term)	AE.AEDECOD	
AESEV (Severity/Intensity)	AE.AESEV	
AESER (Serious Event)	AE.AESER	
AEACN (Action Taken)	AE.AEACN	
AEOU (Outcome of Adverse Event)	AE.AEOU	
PHASE (Phase)	AE.PHASE	

Summary groups

Enterprise Guide Demo

- **Build the SELECT statement by picking values from a list**



The screenshot shows a dialog box titled 'Values' with three tabs: 'Values', 'Columns', and 'Parameters'. The 'Values' tab is active, displaying a table with two columns: 'Value' and 'Formatted Value'. The table contains a list of medical terms, each appearing in both columns. The terms are: ABDOMINAL ABSCESS, ABDOMINAL DISCOMFORT, ABDOMINAL DISTENSION, ABDOMINAL PAIN, ABDOMINAL PAIN LOWER, ABDOMINAL PAIN UPPER, ABDOMINAL TENDERNESS, ABNORMAL BEHAVIOUR, ABNORMAL DREAMS, ABNORMAL FAECES, ABORTION MISSED, ABORTION SPONTANEOUS, ABSCESS, ACCIDENT, ACCOMMODATION DISORDER, ACIDOSIS HYPERCHLORAEMIC, ACNE, ACRODERMATITIS, ACTIVATED PARTIAL THROMBOPLASTIN TIME PROLONG..., and ACTIVITIES OF DAILY LIVING IMPAIRED.

Value	Formatted Value
ABDOMINAL ABSCESS	ABDOMINAL ABSCESS
ABDOMINAL DISCOMFORT	ABDOMINAL DISCOMFORT
ABDOMINAL DISTENSION	ABDOMINAL DISTENSION
ABDOMINAL PAIN	ABDOMINAL PAIN
ABDOMINAL PAIN LOWER	ABDOMINAL PAIN LOWER
ABDOMINAL PAIN UPPER	ABDOMINAL PAIN UPPER
ABDOMINAL TENDERNESS	ABDOMINAL TENDERNESS
ABNORMAL BEHAVIOUR	ABNORMAL BEHAVIOUR
ABNORMAL DREAMS	ABNORMAL DREAMS
ABNORMAL FAECES	ABNORMAL FAECES
ABORTION MISSED	ABORTION MISSED
ABORTION SPONTANEOUS	ABORTION SPONTANEOUS
ABSCESS	ABSCESS
ACCIDENT	ACCIDENT
ACCOMMODATION DISORDER	ACCOMMODATION DISORDER
ACIDOSIS HYPERCHLORAEMIC	ACIDOSIS HYPERCHLORAEMIC
ACNE	ACNE
ACRODERMATITIS	ACRODERMATITIS
ACTIVATED PARTIAL THROMBOPLASTIN TIME PROLONG...	ACTIVATED PARTIAL THROMBOPLASTIN TIME PR
ACTIVITIES OF DAILY LIVING IMPAIRED	ACTIVITIES OF DAILY LIVING IMPAIRED

Enterprise Guide Demo

- **EG generates the SELECT statement for you**

Edit Filter

Column: AE.AEDECOD

Operator: In a list of values

Values:

- ABDOMINAL PAIN UPPER
- ABDOMINAL TENDERNESS
- ABNORMAL BEHAVIOUR
- ABNORMAL DREAMS
- ABNORMAL FAECES

Add Remove

Filter definition

```
AE.AEDECOD IN ('ABDOMINAL PAIN UPPER','ABDOMINAL TENDERNESS','ABNORMAL BEHAVIOUR','ABNORMAL DREAMS','ABNORMAL FAECES','ABORTION MISSED','ABORTION SPONTANEOUS','ABSCESS','ACCIDENT','ACCOMMODATION DISORDER','ACIDOSIS HYPERCHLORAEMIC','ACNE','ACRODERMATITIS','ACTIVATED PARTIAL THROMBOPLASTIN TIME PROLONGED','ACTIVITIES OF DAILY LIVING IMPAIRED','ACUTE SINUSITIS','ACUTE TONSILLITIS','ADENOIDAL HYPERTROPHY','ADENOIDITIS','ADVERSE DRUG REACTION','AEROPHAGIA','AFFECT LABILITY','AFFECTIVE DISORDER','AGEUSIA','AGGRESSION','AGITATION','AGNOSIA','AKINESIA','ALANINE AMINOTRANSFERASE','ALANINE AMINOTRANSFERASE INCREASED','ALCOHOL INTOLERANCE','ALCOHOL POISONING','ALCOHOL WITHDRAWAL SYNDROME','ALLERGIC BRONCHITIS','ALLERGIC SINUSITIS','ALLERGY TO ARTHROPOD STING','ALLERGY TO VACCINE','ALOPECIA','ALOPECIA AREATA','ALTERED VISUAL DEPTH PERCEPTION','AMBLYOPIA','AMENORRHOEA','AMNESIA','ANAEMIA','ANAL INFLAMMATION','ANAL ULCER','ANGER','ANGINA PECTORIS','ANHEDONIA','ANHIDROSIS','ANIMAL BITE','ANIMAL SCRATCH','ANKLE FRACTURE','ANOREXIA','ANDORGASMIA','ANOSMIA','ANTICONVULSANT DRUG LEVEL INCREASED','ANTICONVULSANT TOXICITY','ANTINUCLEAR ANTIBODY POSITIVE','ANXIETY','APATHY','APHASIA','APHONIA','APHTHOUS STOMATITIS','APPENDICITIS','APPETITE DISORDER','AREFLEXIA','ARRHYTHMIA','ARRHYTHMIA'
```

Enterprise Guide Demo

The screenshot displays the SAS Enterprise Guide Project Designer interface. The main workspace shows a process flow diagram with four nodes: 'AE', 'Query for AE', 'Query_for...', and 'Log'. Arrows indicate the flow from 'AE' to 'Query for AE', and from 'Query for AE' to both 'Query_for...' and 'Log'. A context menu is open over the 'Query for AE' node, listing various actions. The 'Run Process Flow' option is highlighted in blue. A callout box with a red dot and the text 'Run the process flow' points to this menu item. The status bar at the bottom shows 'Ready' and the user 'sas1\miket as Mike, connected to sas2.nthanalytics.com:8561/Foundation'.

Run the process flow

Enterprise Guide Demo

- **Generated SAS code**

```
PROC SQL;
  CREATE TABLE WORK.Query_for_AE AS SELECT AE.STUDYID FORMAT=$15.,
    AE.USUBJID,
    AE.SEX,
    AE.AGE FORMAT=BEST22.,
    AE.SAFETY,
    AE.ARM,
    AE.AEBODSYS,
    AE.AEDECOD,
    AE.AESEV,
    AE.AESER,
    AE.AEACN,
    AE.AEOUT,
    AE.PHASE
  FROM NTH.AE AS AE
  WHERE AE.SAFETY = "YES" AND AE.AEDECOD IN ("AXILLARY PAIN", "BACK INJURY", "BACK PAIN"
" BENIGN PROSTATIC HYPERPLASIA", "BEREAVEMENT REACTION", "BLADDER DISORDER", "BLADDER SP
" BLOOD BILIRUBIN INCREASED", "BLOOD CALCIUM INCREASED", "BLOOD CHLORIDE DECREASED", "BL
" BLOOD LACTATE DEHYDROGENASE INCREASED", "BLOOD OSMOLARITY DECREASED", "BLOOD PHOSPHORU
" BLOOD TRIGLYCERIDES ABNORMAL", "BLOOD TRIGLYCERIDES INCREASED", "BLOOD UREA INCREASED"
" BOWEL SOUNDS ABNORMAL", "BRADYCARDIA", "BRADYKINESIA", "BRADYPHRENIA", "BRAIN NEOPLASM
" BREATH ODOUR", "BREATH SOUNDS ABNORMAL", "BRONCHIOLITIS", "BRONCHITIS", "BRONCHITIS AC
" BURSA DISORDER", "BURSITIS", "BUTTOCK PAIN", "CALCINOSIS", "CALCULUS URETERIC", "CALCU
" CARTILAGE INJURY", "CAT SCRATCH DISEASE", "CATARACT", "CATARACT OPERATION", "CATHETER
" CERVICOBRACHIAL SYNDROME", "CHALAZION", "CHANGE OF BOWEL HABIT", "CHEST DISCOMFORT");
QUIT;
```

EG as a Programming Environment



- **Advantages:**

- **Enterprise Guide generates better and more reliable code than programmers produce manually**
- **Enterprise Guide enforces a standard programming style**
- **SAS programmers can be developers while end users run the code.**
- **Using stored processes, the source code can be secured and hidden from the end users.**
- **Enterprise Guide can be deployed worldwide, allowing users to connect to a central SAS repository and run SAS programs without having SAS installed on their systems.**

EG as a Programming Environment



- **However, there are many pitfalls to go along with the advantages.**
 - **Introduction of a new programming environment requires substantial IT support.**
 - **Installation, Validation, FDA 21 CFR Part 11 compliance.**
 - **Move away from stable programming environment**
 - **Can't assume easy upgrade path**
 - **Requires archive copies be maintained**



Considerations

- **This is a completely new environment**
- **Not like installing a new version of SAS**
- **Will require full IQ/PQ/OQ**
- **What do to with old programs**
 - Is it too big a transition?
 - Start fresh with a new project?
 - Incremental?
 - Same problem as with installing any new reporting system
- **Steep learning curve**



More considerations

- **EG is designed as a tool for power users to run interactively.**
- **Therefore, the easiest transition is to provide a new environment for statisticians to run their programs interactively.**
- **EG can replace the traditional production batch environment, but this takes some doing**
 - **You are asking EG to do something it was not designed to do**
 - **But there are many advantages to doing so**



Advantages: PROCs

PROC Tabulate and PROC SQL



PROC TABULATE

- **One of the built-in tasks generates PROC TABULATE code**
 - Summary Tables
- **One could say Enterprise Guide rescued PROC TABULATE from oblivion**
- **Can anyone actually write PROC TABULATE code?**
 - Bizarre syntax makes it a daunting task
- **EG makes it possible to use it easily and effectively**



PROC SQL

- **PROC SQL has been around since 1990, but many programmers still cannot use it**
- **Lack of widespread knowledge makes it difficult to deploy this tool**
- **By generating the code with an easy-to-use front end, EG makes it feasible to deploy PROC SQL with a minimum of pain**



Uses of EG

- **Validation**
 - Statisticians
 - Throw-away code
- **Ad-hoc analyses**
 - Statisticians
 - One-shot
 - May be preserved if part of submission
- **Production tables and listings**
 - Programmers
 - Full production/validation/documentation requirements



Statisticians: Ad-hoc Analyses



Ad-hoc Analyses

- **Usually, programmers do standard tables for a study, and statisticians may be expected to handle non-standard analyses**
- **Non-standard analyses can evolve into big projects**
- **Need some heavy-duty tools**

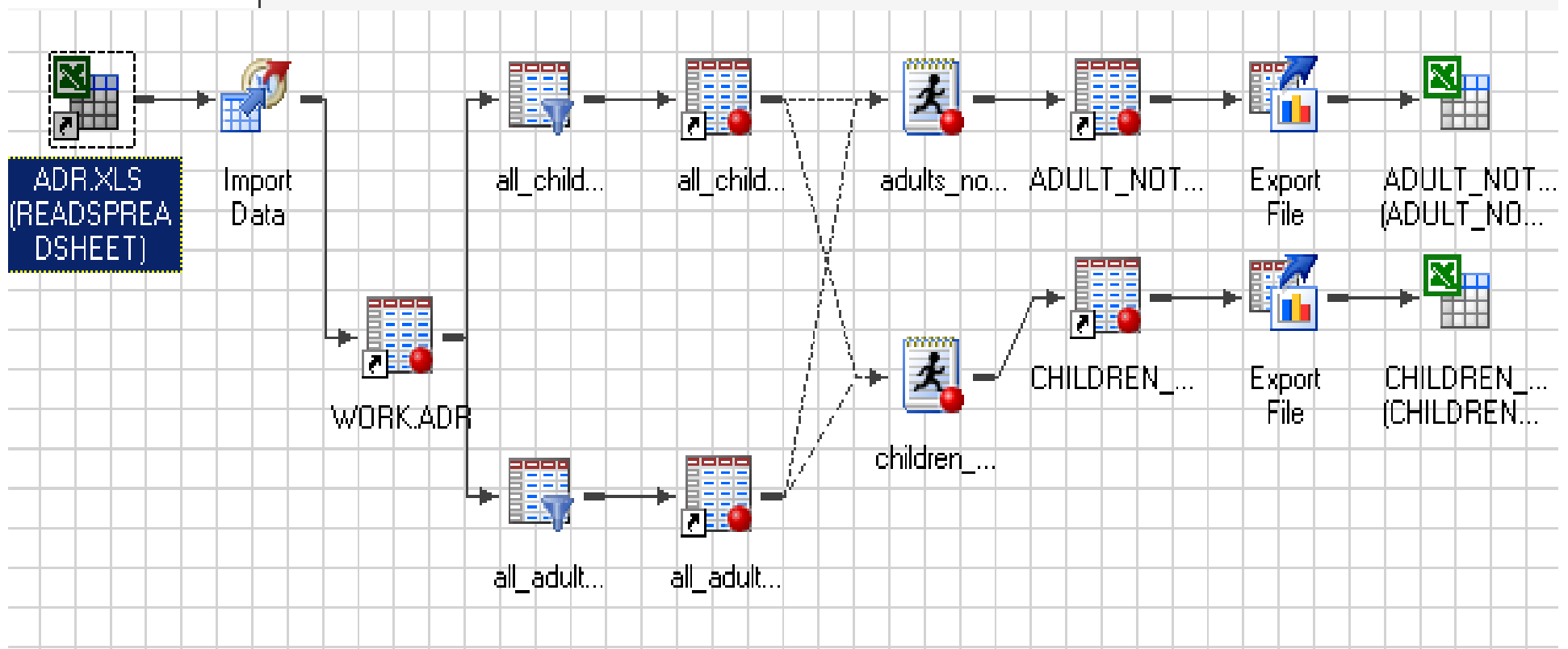


Ad-hoc Analyses

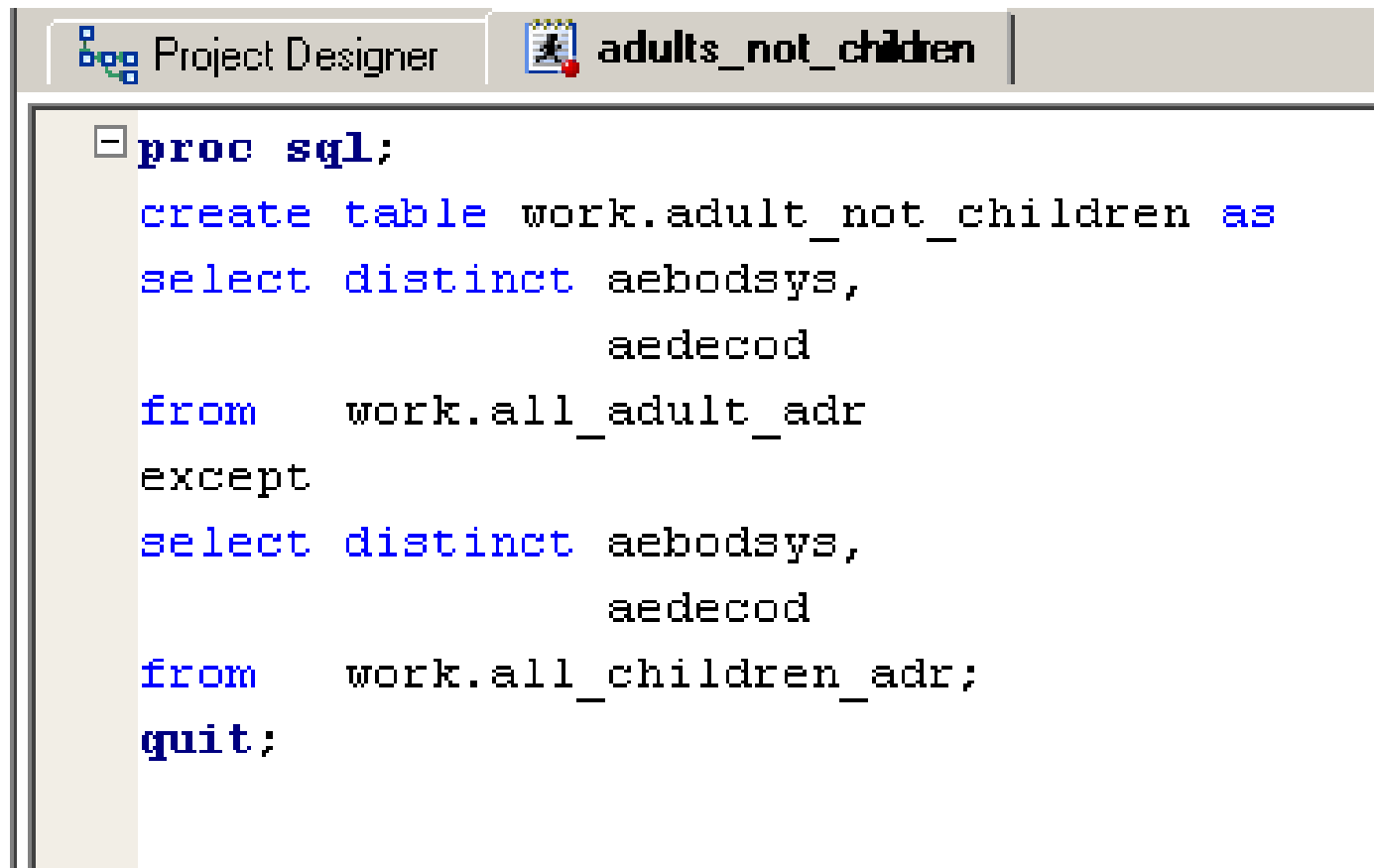
- **While statisticians understand the analyses, they may lack the technical skills to produce maintainable code**
- **Advantages of EG come into play:**
 - Easier to get other statisticians involved
 - Transfer to programming group
 - Use EG to produce tables and graphs

Ad-hoc Analyses

Process Flow





SAS Code Object



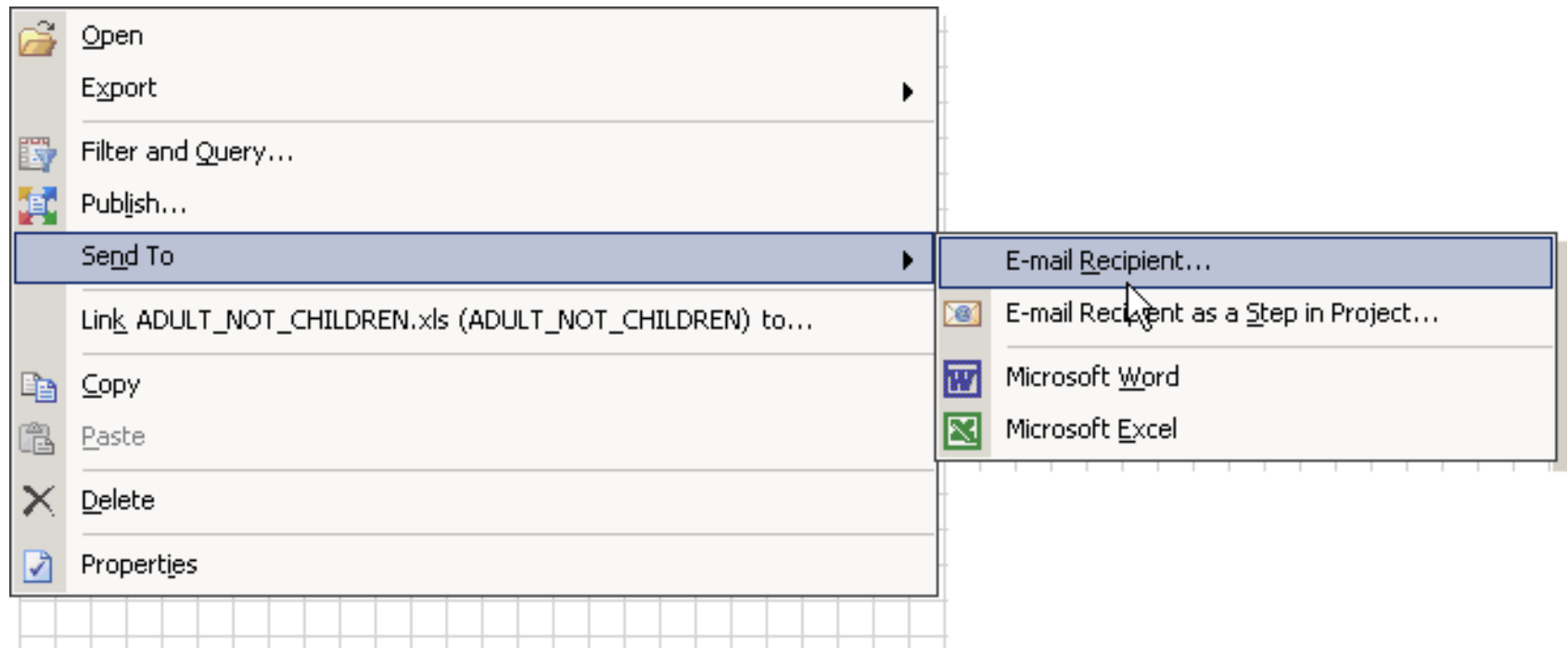
The screenshot shows the SAS Project Designer interface. The title bar includes 'Project Designer' and the name of the code object, 'adults_not_children'. The code object is expanded to show the following SQL code:

```
proc sql;  
  create table work.adult_not_children as  
  select distinct aebodsys,  
                 aedecod  
  from    work.all_adult_adr  
  except  
  select distinct aebodsys,  
                 aedecod  
  from    work.all_children_adr;  
quit;
```


Results

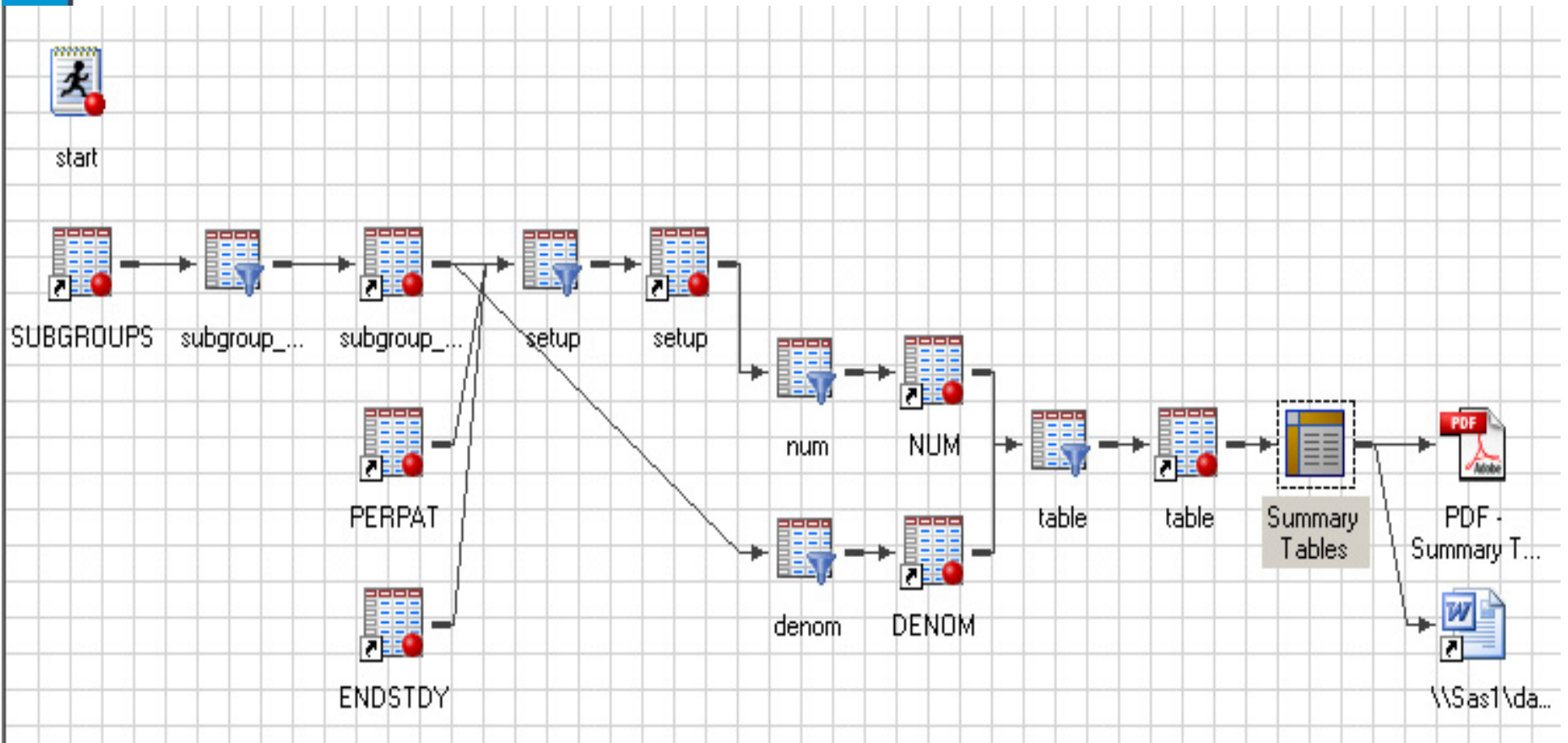
Project Designer		CHILDREN_NOT_ADULT.xls (CHILDREN_NOT_ADULT) (read-only)	
	 AEBODSYS	 AEDECOD	
1	BLOOD AND LYMPHATIC SYSTEM DISORDER	EOSINOPHILIA	
2	EAR AND LABYRINTH DISORDERS	VERTIGO	
3	GASTROINTESTINAL DISORDERS	VOMITING	
4	GENERAL DISORDERS AND ADMINISTRATION	HYPERTHERMIA	
5	GENERAL DISORDERS AND ADMINISTRATION	PYREXIA	
6	NERVOUS SYSTEM DISORDERS	PSYCHOMOTOR HYPERACTIVITY	
7	SOCIAL CIRCUMSTANCES	LEARNING DISABILITY	

Email Results





Project Flow



Denominator

denom - Query Builder

Query name: denom Output name: WORK.DENOM Change...

Computed Columns Parameters Validate Preview Options

Add Tables... Delete Join..

SUBGROUP_SELECT

- PROTOCOL (Protocol (BER342))
- RANDID (Randomization #)
- POP (Population)
- POPX (Decoded Population Flag)
- SEX (Gender)
- SEXX (Decoded Gender)
- RACEH (Ethnic Group)
- RACEHX (Decoded Ethnic Group)
- ETHCAT (Ethnic Group Category)
- AGE (Age in years)
- AGECAT (Age Category)
- RXGRP (Treatment Group)
- DURROSA (Previous Duration of F
- PREV_ROSA
- SKINDRY (Pre-TMT Facial Signs/S
- SUBGROUP (Disease Quartile)
- RXGRPX

Computed Columns

- DENOM
- COLHEAD

Select Data | Filter Data | Sort Data

Column Name	Input	Summary
RXGRP (Treatment Gr...	SUBGROUP_SELECT.R...	
RXGRPX (Decoded Tr...	SUBGROUP_SELECT.R...	
SUBGROUP	SUBGROUP_SELECT.S...	
DENOM	SUBGROUP_SELECT.R...	COUNT DISTINCT
COLHEAD	COLHEAD	

Summary groups

Automatically select groups Edit Groups...

SUBGROUP_SELECT.RXGRP, SUBGROUP_SELECT.RXGRPX,
SUBGROUP_SELECT.SUBGROUP

Select distinct rows only

Run Save and Close Cancel Help

Custom Code

Advanced Expression Editor

Expression text:

```
TRIM(SUBGROUP_SELECT.RXGRPX) || '^R'\line "" || compress( 'N=' || put(CALCULATED denom , 5.) || '')
```

+ - * / ** AND OR NOT = <> < <= > >= =* (X) 'X' [] { } ,

Clear Expression Add to Expression

Data Functions

Available variables:

- SUBGROUP_SELECT
 - PROTOCOL
 - RANDID
 - POP
 - POPX
 - SEX
 - SEX
 - RACEH
 - RACEHX
 - ETHCAT
 - AGE
 - AGECAT
 - RXGRP
 - DURROSA
 - PREV_ROSA
 - SKINDRY
 - SUBGROUP
 - RXGRP
- Current Query
 - DENOM
 - COLHEAD

Variable values:

Variable Names	Variable Labels
PROTOCOL	Protocol (BER342)
RANDID	Randomization #
POP	Population
POPX	Decoded Population Flag
SEX	Gender
SEX	Decoded Gender
RACEH	Ethnic Group
RACEHX	Decoded Ethnic Group
ETHCAT	Ethnic Group Category
AGE	Age in years
AGECAT	Age Category
RXGRP	Treatment Group
DURROSA	Previous Duration of Rosacea (months)
PREV_RO...	Pre-TMT Facial Signs/Symptoms: Skindry
SKINDRY	Pre-TMT Facial Signs/Symptoms: Skindry
SUBGROUP	Disease Quartile
RXGRP	

Get More Values

Custom Code: Up Close

Expression text:

```
TRIM(SUBGROUP_SELECT.RXGRPX ) || '^R'\line "' || compress( 'N=' || put(CALCULATED denom , 5.) || ' )'
```

Proc Tabulate

Summary Tables

Available variables:

- ALL (universal ...)
- COLHEAD
- ITEM11_recode
- ITEM12_recode
- RXGRP
- THERA_recode


Preview:

COLHEAD			
N	ColPctN	N	ColPctN

Available statistics:

Name	Descript...	
ColP...	Column ...	
ColP...	Column ...	
ColP...	Column ...	

Customize Code

 *<double-click to insert code>*

```
PROC SQL;  
%_SASTASK_DROPDS(WORK.SORTTempTableSorted);  
QUIT;
```

```
/* .....  
Run the tabulate procedure  
..... */
```

 **options nodate nonumber orientation=portrait nobyline;
ods rtf style=NthStyles.Nth_RTF_Table notoc_data BODYTITLE**

file="&dir\tables\%trim(&file).rtf";

TITLE1 J=L "XXX Laboratories, Inc.";

TITLE2 J=L "Clinical Study Report No. &protocol";

TITLE3 J=C ^R"\par" Table 1.1.2 Measures of Patient Satisfaction at End of Study (Per-Protocol Population);

TITLE4 J=C "#BYVAL1";

PROC TABULATE

DATA=WORK.TABLE

 **; by subgroup;**

Proc Tabulate Code

Code Preview for Task

Insert Code...

```
/* End of custom user code. */  
PROC TABULATE  
  DATA=WORK.TABLE  
  
/* Start of custom user code. */  
; by subgroup;  
/* End of custom user code. */  
;  
  CLASS RXGRP / ORDER=UNFORMATTED DESCENDING MISSING;  
  CLASS COLHEAD / ORDER=UNFORMATTED MISSING;  
  CLASS ITEM11_recode / ORDER=UNFORMATTED MISSING;  
  CLASS ITEM12_recode / ORDER=UNFORMATTED MISSING;  
  CLASS THERA_recode / ORDER=UNFORMATTED MISSING;  
  TABLE /* Row Dimension */  
    ITEM11_recode  
    ITEM12_recode  
    THERA_recode,  
  /* Column Dimension */  
  COLHEAD={ LABEL=' ' STYLE={NOBREAKSPACE=ON} } * (  
    N*F=3.0  
    ColPctN={ LABEL='% ' } *F=PCTFMT.)  
    ;  
  ;
```

Table

XXX Laboratories, Inc.
Clinical Study Report No. A1

**Table 1.1.2 Measures of Patient Satisfaction at End of Study (Per-Protocol Population)
Duration of Disease Quartile 1 (<=15 months)**

	Active (N=27)	
	N	%
Cosmetic acceptance of the topical preparation		
(1) Very good	11	40.7
(2) Good	11	40.7
(3) Satisfactory	4	14.8
(4) Poor	1	3.7
Patient tolerability in general		
(1) Good	16	59.3
(2) Acceptable despite minor irritation	9	33.3
(3) Less acceptable due to continuous irritation	1	3.7
(5) No opinion	1	3.7
Patient overall rating of improvement		
(1) Excellent improvement	8	29.6
(2) Good improvement	9	33.3
(3) Moderate improvement	7	25.9
(4) No improvement	2	7.4
(5) Aggravation	1	3.7



Issues

- **As SAS users, we are used to a stable programming environment.**
- **While the rest of the computing world has undergone several revolutions, SAS code for tables, listings, and graphs circa 1990, or earlier, can still be used.**
 - **Change of other hardware/software from 15 years ago still working is remote.**



Backwards Compatibility

- **SAS has made a tremendous commitment to backwards compatibility for SAS programs across versions.**
- **For example, there has been no need to archive the operating system, database, and SAS version along with the code.**
- **SAS programs from 15 years ago still run, with at most minimal modification, and the legacy datasets can easily be updated to the current version.**
- **In short, even with the tremendous changes in the computing environment since the 1990s, the migration path for SAS programs and datasets has been relatively smooth.**



Text-Based Environment

- **One of the main reasons is that environment was text-based, consisting of ASCII text files that can be opened with any basic text editor.**
- **With Enterprise Guide, we are clearly moving away from this paradigm.**
- **Instead, we have an environment that, while it offers much more power, will change from version to version without backwards compatibility.**



EG Migration

- **It is not possible to convert an EG 4 project to EG 3, and it will not be possible to back-convert an Enterprise Guide 7 project to Enterprise Guide 4 in the future.**
 - **SAS products will become like Microsoft products in that sense.**
 - **We will migrate forwards, and the forward migration will not always be smooth.**
 - **Because they are binary files, you cannot automate the conversion of Enterprise Guide projects by running the projects through a batch processor and applying global changes, like you can with text-based SAS programs.**



Archive Copy as a Safety Net

- **Here is where the concept of an archive copy becomes critical.**
- **An archive copy in the Enterprise Guide environment is a SAS program that is generated from the code generated by the Enterprise Guide tasks.**
- **Regardless of what problems arise from version compatibility issues in the programming environment, the archive copy can always be run as traditional SAS programs.**
- **It is critical that the archive copies be generated from an automated, validated process.**



Conclusion

- Despite some limitations, Enterprise Guide offers the evolution in SAS programming, and is an exciting replacement for the traditional Display Manager programming environment.
- However, the programming environment itself, not just the SAS code, will change from version to version.
- This is a major change in the SAS programming world, and has significant implications for regulatory compliance in the pharmaceutical industry.
- Using Enterprise Guide to automatically generate archive copies (i.e., traditional SAS code) of process flows mitigates these risks.



SAS Trademarks

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.



My Contact Info

- **Michael Todd**
- **info@nthanalytics.com**
- **www.nthanalytics.com**