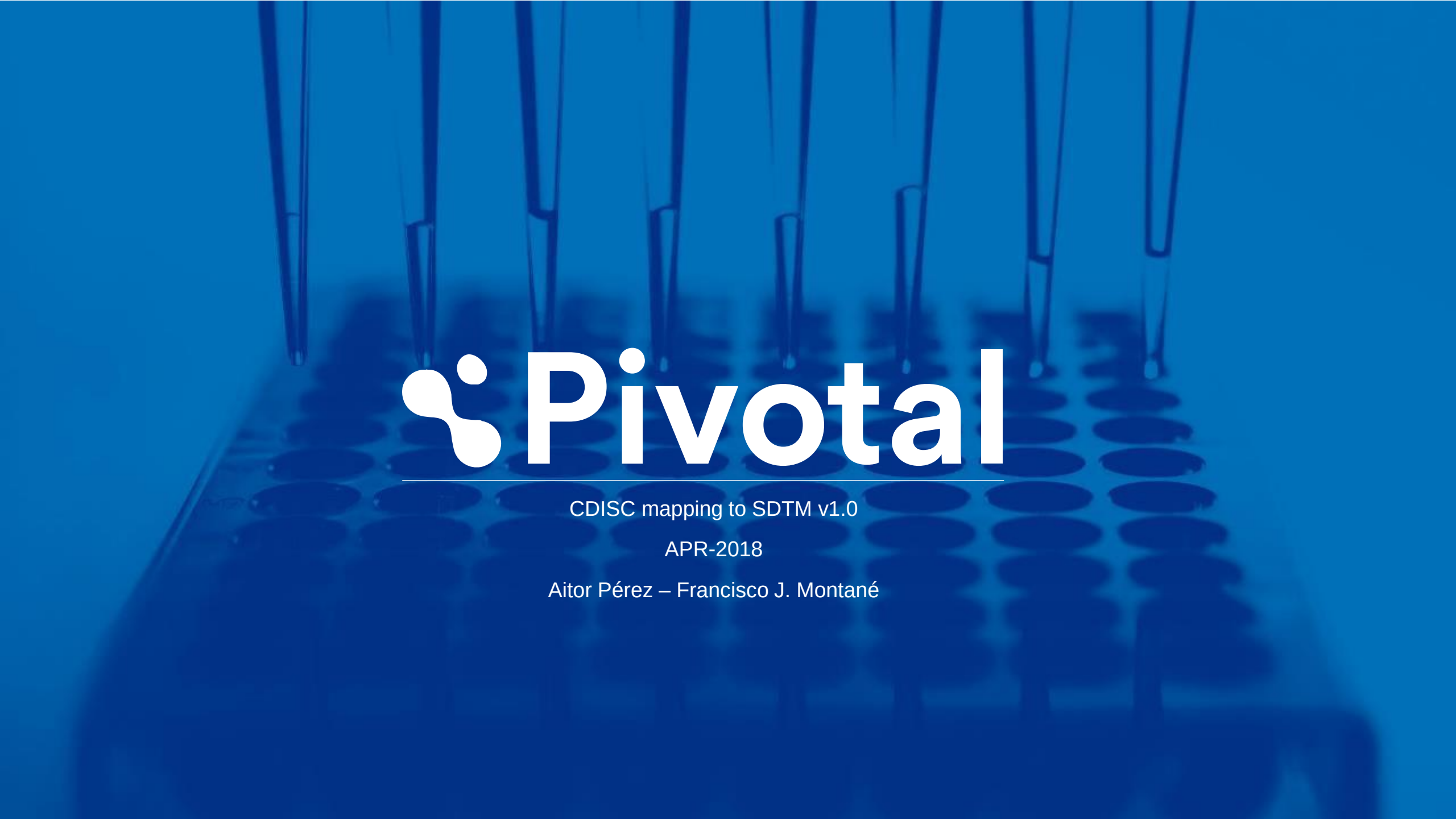


CDISC mapping to SDTM v1.0

APR-2018

Aitor Pérez – Francisco J. Montané

Disclaimer

*This document is the **property of** **PIVOTAL S.L.** **and its affiliates** and/or the corresponding client/sponsor (the "Company") and is strictly **confidential.** It contains information intended **only for the person to whom it is transmitted.** With receipt of this information, the recipient acknowledges and agrees that:*

- *This document is **not intended to be distributed**, and if **distributed inadvertently**, will be **returned to the Company** as soon as possible;*
- *The recipient **will not copy, fax, reproduce, divulge, or distribute** this confidential information, in whole or in part, **without the prior express written consent** of the Company; and,*
- *All of the information herein will be treated **as confidential material** with no less care than that afforded to its own confidential material.*

The opinions expressed are in good faith and while every care has been taken in preparing this document, the Company makes no representations and gives no warranties of whatever nature in respect of this document, including but not limited to the accuracy or completeness of any information, facts and opinions contained therein.

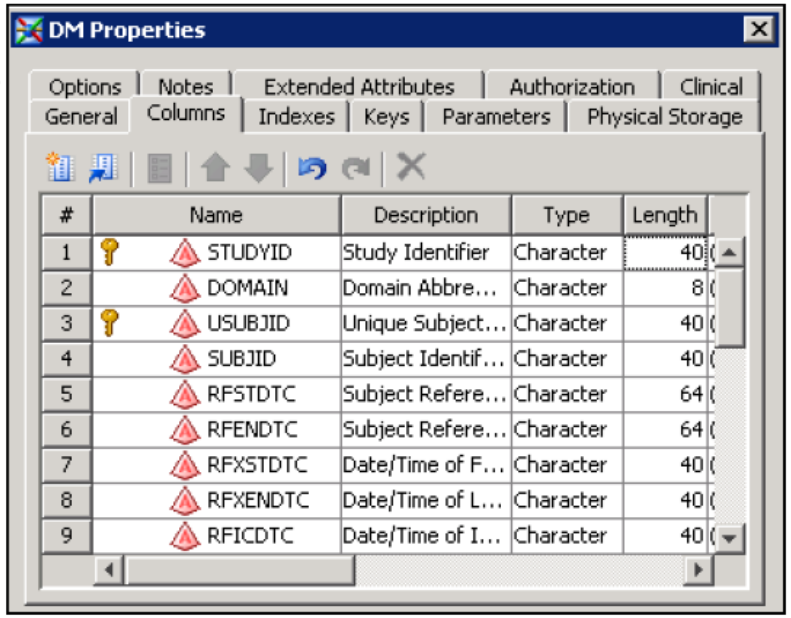
Agenda

1. CDISC
2. CDASH
3. SDTM/SEND
4. ADaM
5. SAS CDI
6. Pinnacle21
7. Domains
8. Implementation

CDISC Standards

SDTM	Study Data Tabulation Model
ADaM	Analysis Data Model
SEND	Standard for the Exchange of Non-Clinical Data

DM domain metadata definition supplied by SAS Clinical Data Integration



Domain metadata
- table-level attributes
- column level attributes

- Global, open, multidisciplinary, non-profit organization.
- Mission: develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare

- CDISC standards are vendor-neutral, platform-independent and freely available
- Standards to support clinical research data:
  - Acquisition
  - Exchange
  - Submission
  - Archive

CDASH

Clinical Data Acquisition Standards Harmonization

Adquisition

- **CDASH** establishes a standard way to **collect data in a similar way across studies and sponsors** so that data collection formats and structures provide clear traceability of submission data

https://www.cdisc.org/standards/foundational/cdash

SDTM

Study Data Tabulation Model

Exchange

- **SDTM** provides a standard for **organizing and formatting data** to streamline processes in collection, management, analysis and reporting.

- One of the required standards for data submission to FDA (U.S.) and PMDA (Japan).

https://www.cdisc.org/standards/foundational/sdtm

SEND

Standard for Exchange of Nonclinical Data

Exchange

- **SEND** is an implementation of the SDTM standard for nonclinical studies. SEND specifies a way to collect and present nonclinical data in a consistent format.

https://www.cdisc.org/standards/foundational/send

ADaM

Analysis Data Model

Submission

- **ADaM** defines dataset and metadata standards that support:
  - **efficient generation, replication, and review** of clinical trial statistical analyses
  - **traceability** between analysis results, analysis data, and SDTM data

https://www.cdisc.org/standards/foundational/adam

ADaM will begins to be required from 15MAR2019

SAS CDI

SAS Clinical Data Integration

- SAS Clinical Data Integration (CDI)
  - helps organize, standardize and manage clinical research data
  - Provides repeatability and automation
  - facilitates the process of transforming data to meet industry-mandated data standards such as CDISC

Pinnacle21

- Pinnacle 21 Community:
  - Free software
  - Includes Validator, Define.xml Generator, Data Converter, and ClinicalTrials.gov Miner
  - Includes the last 5 releases of CDISC Terminology.

https://www.pinnacle21.com/downloads

Special-Purpose Domains

SDTM Domains

Special Purpose Domains is an SDTM category in its own right as they provide specific standardized structures to represent additional important information that does not fit any of the General Observation Classes.

- **CO (Comments):** collected on topical case report form (CRF) pages or on a separate page specifically dedicated to comments.
- **DM (Demographics):** standard variables that describe each subject in a clinical study
- **SE (Subject Elements):** actual order of elements followed by the subject
- **SV (Subject Visits)**: actual start and end data/time for each visit of each individual subject

General Observation Classes

SDTM Domains

Most subject-level observations collected during the study should be represented according to one of the three SDTM general observation classes.

- **Interventions:**
  - **CM (Concomitant and Prior Medications):** Concomitant and Prior Medications/Therapies used by the subject
  - **EX (Exposure):** details of a subject's exposure to protocol-specified study treatment
  - **EC (Exposure as Collected):** protocol-specified study treatment administrations, as collected
  - **PR (Procedures)**: collected details describing a subject's therapeutic and diagnostic procedures
  - **SU (Substance Use):** substance use information that may be used to assess the efficacy and/or safety of therapies that look to mitigate the effects of chronic substance use.

General Observation Classes

SDTM Domains

- **Events:**
  - **AE (Adverse Events):** Adverse events recorded in CRF either as free text or a pre-specified list of terms
  - **CE (Clinical Evens):** clinical events of interest that would not be classified as adverse events.
  - **DS (Disposition):** information representing data, vocabulary or records related to disposition.
  - **DV (Protocol Deviations)**: protocol violations and deviations during the course of the study (not a response to each violation or deviation)
  - **HO (Healthcare Encounters):** inpatient and outpatient healthcare events (e.g., hospitalizations, ambulatory surgery, …).
  - **MH (Medical Hystory):** subject's prior history at the start of the trial.

General Observation Classes

SDTM Domains

- Findings:
  - DA (Drug Accountability)
  - DD (Death Details)
  - EG (Electrocardiogram Results)
  - IE (Inclusion/Exclusion Criteria Not Met)
  - IS (Immunogenicity Specimen Assessments)
  - LB (Laboratory Test Findings)
  - MB (Microbiology specimen findings)
  - MS (microbiology susceptibility test results)
  - MI (Microscopic Findings)
  - MO (Morphology)
  - PC (Concentrations of drugs/metabolites in fluids or tissues)
  - PP (Pharmacokinetic parameters derived from PC data)
  - PE (Physical Exam)
  - QS (Questionaries)
  - RP (Reproductive System Findings)
  - SC (Subject Characteristics)
  - SS (Subject Status)
  - TU (Tumor Identification)
  - TR (Tumor Response)
  - SR (Disease Response)
  - VS (Vital Signs)
  - FA (Findings About)

Trial Design Datasets

SDTM Domains

- Experimental Design:
  - TA (Trial Arms): describes each planned arm in the trial
  - TE (Trial Elements): definitions of the elements that appear in the Trial A
- Scheduling of Assessments:
  - TV (Trial Visits): describes the planned Visits in a trial
  - TD (Trial Disease Assessments): information on the protocol-specified disease assessment schedule
- Trial Summary and Eligibility:
  - TI (Trial Inclusion/Exclusion Criteria): It contains all the inclusion and exclusion criteria for the trial (not subject oriented)
  - TS (Trial Summary): summary of the trial in a structured format

Implementation

Extract from OC

Data can came from several sources; Oracle, Medidata, SQL but also from excel or SAS Datasets

Transform in SAS CDI

SAS CDI allows transform the original data from the different sources to standard format, either SDTM, ADaM, SEND,.. through a set of predefined process.

Load SDTM Domains

Once we have our data standardized are load in the templates of standardized domains.

Validate SDTM Domains

The information in the built domains must been validated to ensure that meet the standardization criteria.

Implementation

- Requirements
  - Define a study
  - Define a library
  - Original datasets
  - Load the standard

- Steps
  - Registry the original datasets
  - Create a new Job
  - Identify the variables and datasets to include in the domain
  - Do the necessary transformation to conform the final domain to upload into the standard
  - Domain validation

Implementation

- For carry out the mapping we use the SDTM Implementation guideline V3.2 (V3.3 will be release at 18th May 2018)
- In the implementation guideline we can find all information about the domain like:
  - The description: The Demographics domain includes a set of essential standard variables that describe each subject in a clinical study. It is the parent domain for all other observations for human clinical subjects.
  - Structure definition:

DM – Specification for the Demographics Domain Model

dm.xpt, Demographics — Version 3.2. One record per subject, Tabulation

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	DM	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product. This must be a unique number, and could be a compound identifier formed by concatenating STUDYID-SITEID-SUBJID.	Req
SUBJID	Subject Identifier for the Study	Char		Topic	Subject identifier, which must be unique within the study. Often the ID of the subject as recorded on a CRF.	Req
RFSTDTC	Subject Reference Start Date/Time	Char	ISO 8601	Record Qualifier	Reference Start Date/time for the subject in ISO 8601 character format. Usually equivalent to date/time when subject was first exposed to study treatment. Required for all randomized subjects; will be null for all subjects who did not meet the milestone the date requires, such as screen failures or unassigned subjects.	Exp
RFENDTC	Subject Reference End Date/Time	Char	ISO 8601	Record Qualifier	Reference End Date/time for the subject in ISO 8601 character format. Usually equivalent to the date/time when subject was determined to have ended the trial, and often equivalent to date/time of last exposure to study treatment. Required for all randomized subjects; null for screen failures or unassigned subjects.	Exp
RFXSTDTC	Date/Time of First Study Treatment	Char	ISO 8601	Record Qualifier	First date of exposure to any protocol-specified treatment or therapy, equal to the earliest value of EXSTDTC.	Exp
RFXENDTC	Date/Time of Last Study Treatment	Char	ISO 8601	Record Qualifier	Last date of exposure to any protocol-specified treatment or therapy, equal to the latest value of EXENDTC (or the latest value of EXSTDTC if EXENDTC was not collected or is missing).	Exp

Implementation

CDISC Core Variables

- A **Required** variable is any variable that is basic to the identification of a data record (i.e., essential key variables and a topic variable) or is necessary to make the record meaningful. Required variables must always be included in the dataset and cannot be null for any record.
- An **Expected** variable is any variable necessary to make a record useful in the context of a specific domain. Expected variables may contain some null values, but in most cases will not contain null values for every record. When no data has been collected for an expected variable, however, a null column must still be included in the dataset, and a comment must be included in the define.xml to state that data was not collected.
- A **Permissible** variable should be used in a domain as appropriate when collected or derived. Except where restricted by specific domain assumptions, any SDTM Timing and Identifier variables, and any Qualifier variables from the same general observation class are permissible for use in a domain based on that general observation class. The Sponsor can decide whether a Permissible variable should be included as a column when all values for that variable are null. The sponsor does not have the discretion to exclude permissible variables when they contain data.

Implementation

- Assumptions:

DM – Assumptions for the Demographics Domain Model

1. Investigator and site identification: Companies use different methods to distinguish sites and investigators. CDISC assumes that SITEID will always be present, with INVID and INVNAM used as necessary. This should be done consistently and the meaning of the variable made clear in the define.xml.
2. Every subject in a study must have a subject identifier (SUBJID). In some cases a subject may participate in more than one study. To identify a subject uniquely across all studies for all applications or submissions involving the product, a unique identifier (USUBJID) must be included in all datasets. Subjects occasionally change sites during the course of a clinical trial. The sponsor must decide how to populate variables such as USUBJID, SUBJID and SITEID based on their operational and analysis needs, but only one DM record should be submitted for the subject. The Supplemental Qualifiers dataset may be used if appropriate to provide additional information.
3. Concerns for subject privacy suggest caution regarding the collection of variables like BRTHDTC. This variable is included in the Demographics model in the event that a sponsor intends to submit it; however, sponsors should follow regulatory guidelines and guidance as appropriate.
4. The values of ARM and ARMCD in DM must match entries in the Trial Arms (TA) dataset, except for subjects who were not fully assigned to an Arm. Subjects who did not receive the treatments to which they were assigned will still have the values of ARM and ARMCD to which they were assigned. SE/DM Examples 1 and 2 in *Section 5 - SE Domain: SE - Examples for the SUBJECT ELEMENTS Domain Model* show examples of subjects whose actual treatment did not match their planned treatment.

   Some subjects may leave the trial before they can be assigned to an Arm, or, in the case of trials where Arm is assigned by two or more successive allocation processes, may leave before the last of these processes. Such subjects will not be assigned to one of the planned Arms described in the Trial Arms dataset, and must have special values of ARM and ARMCD assigned.

   - Data for screen failure subjects, if submitted, should be included in the Demographics dataset, with ARMCD = "SCRNFAIL" and ARM = "Screen Failure". Sponsors may include a record in the Disposition dataset indicating when the screen failure event occurred. DM/SE Example 6 shows an example of data submitted for a screen failure subject.
     - Some trial designs include Elements after screening but before Arm assignments are made, and so may have subjects who are not screen failures, but are not assigned to an Arm. Subjects withdrawn from a trial before assignment to an Arm, if they are not screen failures, should have ARMCD = "NOTASSGN" and ARM = "Not Assigned". Example Trial 1in *Section 7.2 - Experimental Design: Example Trial 1, A Parallel Trial, TA – Examples For Trial Arms Dataset*, which includes a screening Epoch and a run-in Epoch before randomization, is an example of such a trial; data for a subject who passed screening but was not randomized in this trial are shown in DM/SE Example 6.
     - In trials where Arm assignment is done by means of two or more allocation processes at separate points in time, subjects who drop out after the first allocation process but before the last allocation process, should be assigned values of ARMCD that reflect only the allocation processes they underwent. Example Trial 3, *Section 7.2 - Experimental Design: Example Trial 3, A Trial With Multiple Branch Points, TA - Examples for Trial Arms Dataset*, is such a trial. DM/SE Example 7 shows sample data for subjects in this trial.
5. When study population flags are included in SDTM, they are treated as Supplemental Qualifiers *[see Section 8: 8.4, Relating Non-Standard Variables Values To A Parent Domain]* to DM and placed in the SUPPDM dataset. Controlled terms for these subject-level population flags, (e.g., COMPLT, SAFETY, ITT and PPROT) are listed in *Appendix C2 - Supplemental Qualifier Name Codes*. See ICH E9 for more information and definitions. Note that the ADaM subject-level analysis dataset (ADSL) includes population flags; consult the ADaM Implementation Guide for more information about these variables.

Implementation

- Different examples:

DM – Examples for the Demographics Domain Model

Examples of using the DM domain for typical scenarios are provided below. Example 1 displays the all Required and Expected variables; in examples 2 - 6, certain Required or Expected variables have been omitted in consideration of space and clarity. Example 1 is a general Demographics example showing typical data recorded for a clinical trial. Examples 2 through 5 display various scenarios for representing race and ethnicity information. Example 6 shows the handling of ARMCD for Subjects Withdrawn before Assignment to an Arm, and Example 7 shows the handling ARMCD for Subjects Withdrawn when assignment to an Arm is Incomplete.

DM Example 1 – General Demographics

dm.xpt

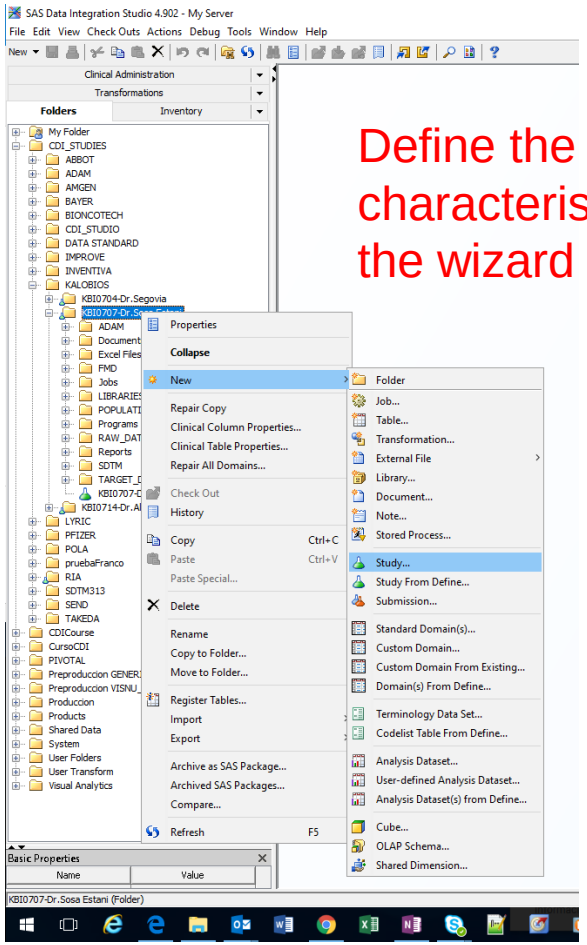
Row	STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC	RFXSTDTC	RFXENDTC	RFICTTC	RFPENDTC
1	ABC123	DM	ABC12301001	001	2006-01-12	2006-03-10	2006-01-12	2006-03-10	2006-01-03	2006-04-01
2	ABC123	DM	ABC12301002	002	2006-01-15	2006-02-28	2006-01-15	2006-02-28	2006-01-04	2006-03-26
3	ABC123	DM	ABC12301003	003	2006-01-16	2006-03-19	2006-01-16	2006-03-19	2006-01-02	2006-03-19
4	ABC123	DM	ABC12301004	004					2006-01-07	2006-01-08
5	ABC123	DM	ABC12302001	001	2006-02-02	2006-03-31	2006-02-02	2006-03-31	2006-01-15	2006-04-12
6	ABC123	DM	ABC12302002	002	2006-02-03	2006-04-05	2006-02-03	2006-04-05	2006-01-10	2006-04-25

Row	SITEID	INVNAM	BRTHDTC	AGE	AGEU	SEX	RACE	ETHNIC
1 (cont)	01	JOHNSON, M	1948-12-13	57	YEARS	M	WHITE	HISPANIC OR LATINO
2 (cont)	01	JOHNSON, M	1955-03-22	50	YEARS	M	WHITE	NOT HISPANIC OR LATINO
3 (cont)	01	JOHNSON, M	1938-01-19	68	YEARS	F	BLACK OR AFRICAN AMERICAN	NOT HISPANIC OR LATINO
4 (cont)	01	JOHNSON, M	1941-07-02			M	ASIAN	NOT HISPANIC OR LATINO
5 (cont)	02	GONZALEZ, E	1950-06-23	55	YEARS	F	AMERICAN INDIAN OR ALASKA NATIVE	NOT HISPANIC OR LATINO
6 (cont)	02	GONZALEZ, E	1956-05-05	49	YEARS	F	NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDERS	NOT HISPANIC OR LATINO

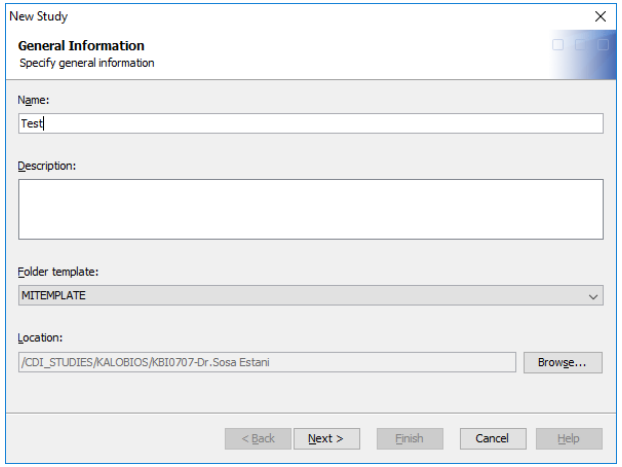
Row	ARMCD	ARM	ACTARMCD	ACTARM	COUNTRY
1 (cont)	A	Drug A	A	Drug A	USA
2 (cont)	P	Placebo	P	Placebo	USA
3 (cont)	P	Placebo	P	Placebo	USA
4 (cont)	SCRNFAIL	Screen Failure	SCRNFAIL	Screen Failure	USA
5 (cont)	P	Placebo	P	Placebo	USA
6 (cont)	A	Drug A	A	Drug A	USA

Implementation

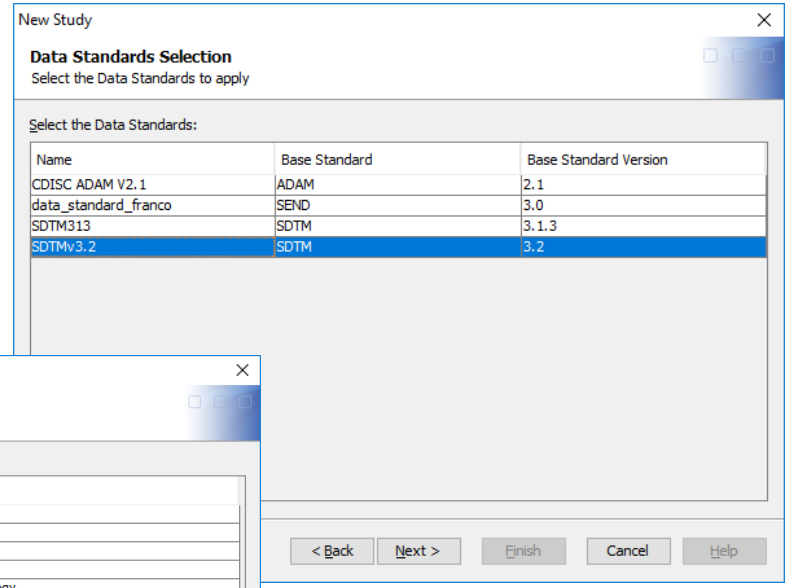
Working with SAS Data Integration Studio: Creating a new Study

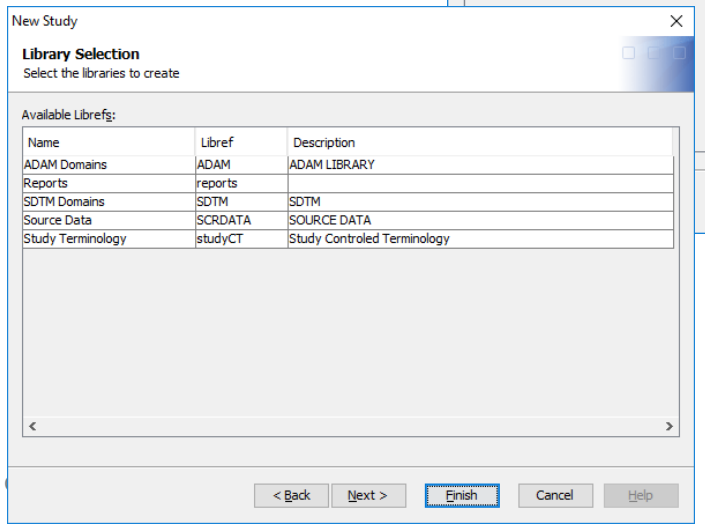


Define the study characteristics following the wizard



Choose the standard to be used in the study

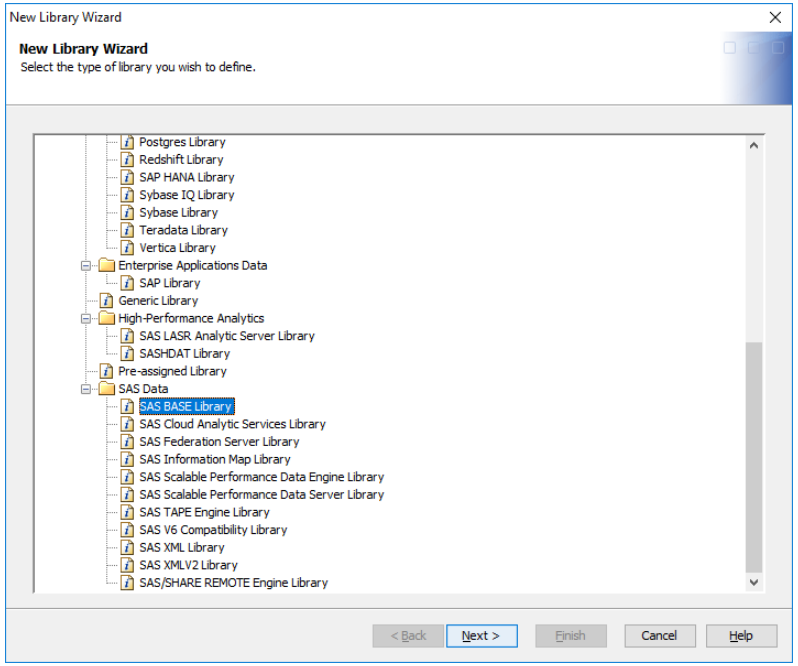


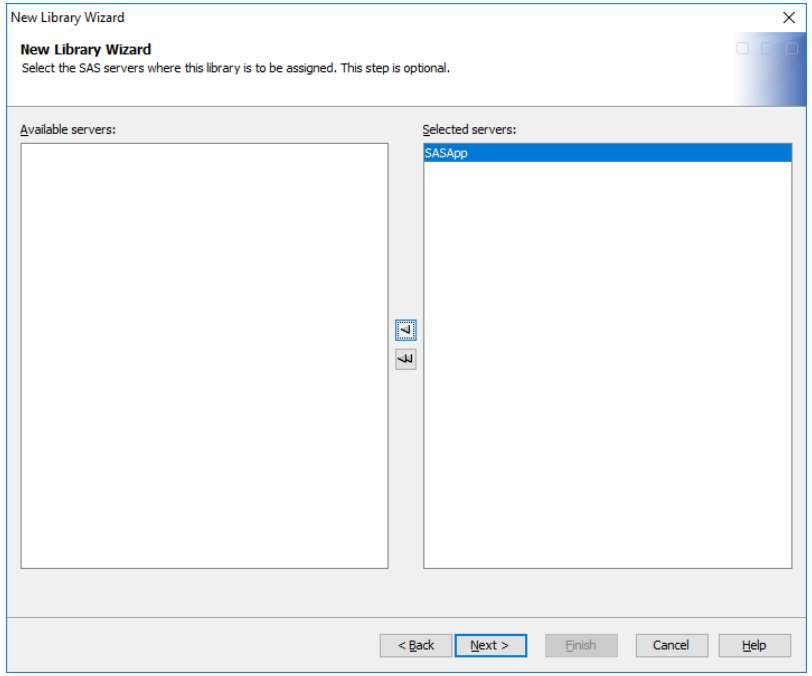


Implementation

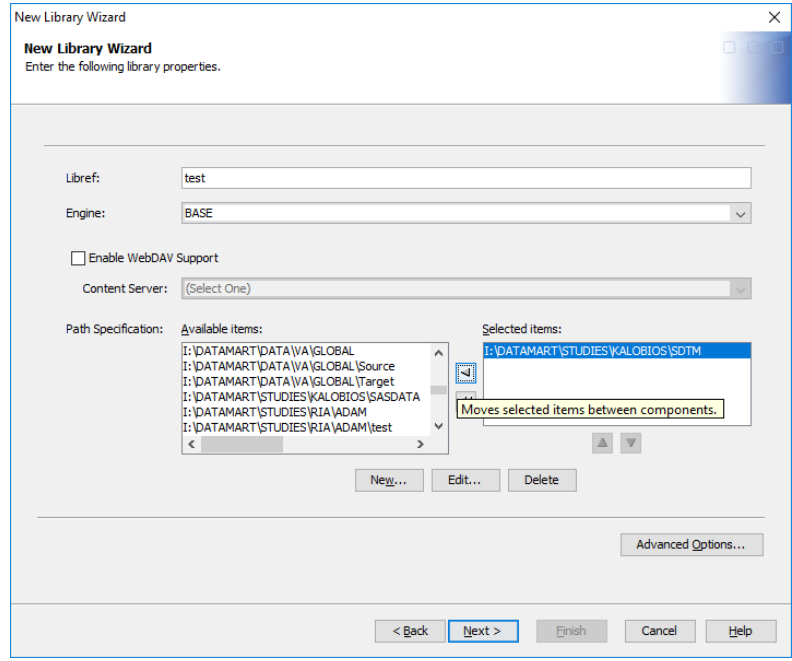
Working with SAS Data Integration Studio: Creating a new Study

Create a SAS BASE library





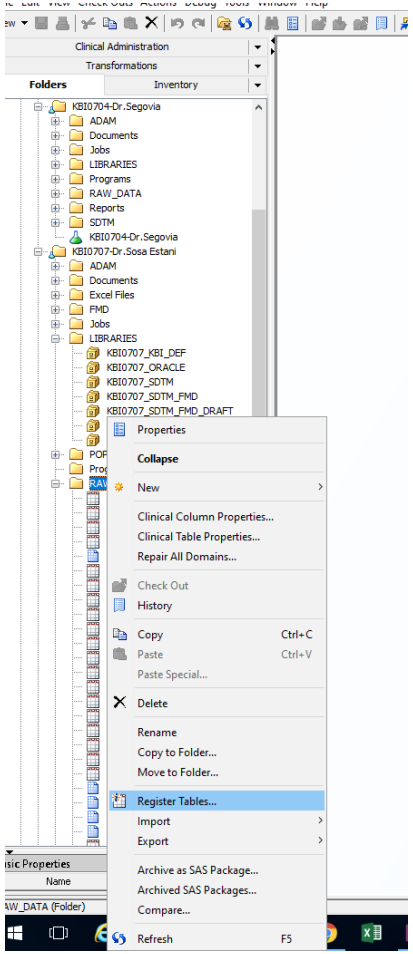
Choose a library reference and root for physical files



Implementation

Working with SAS Data Integration Studio: Registry of tables

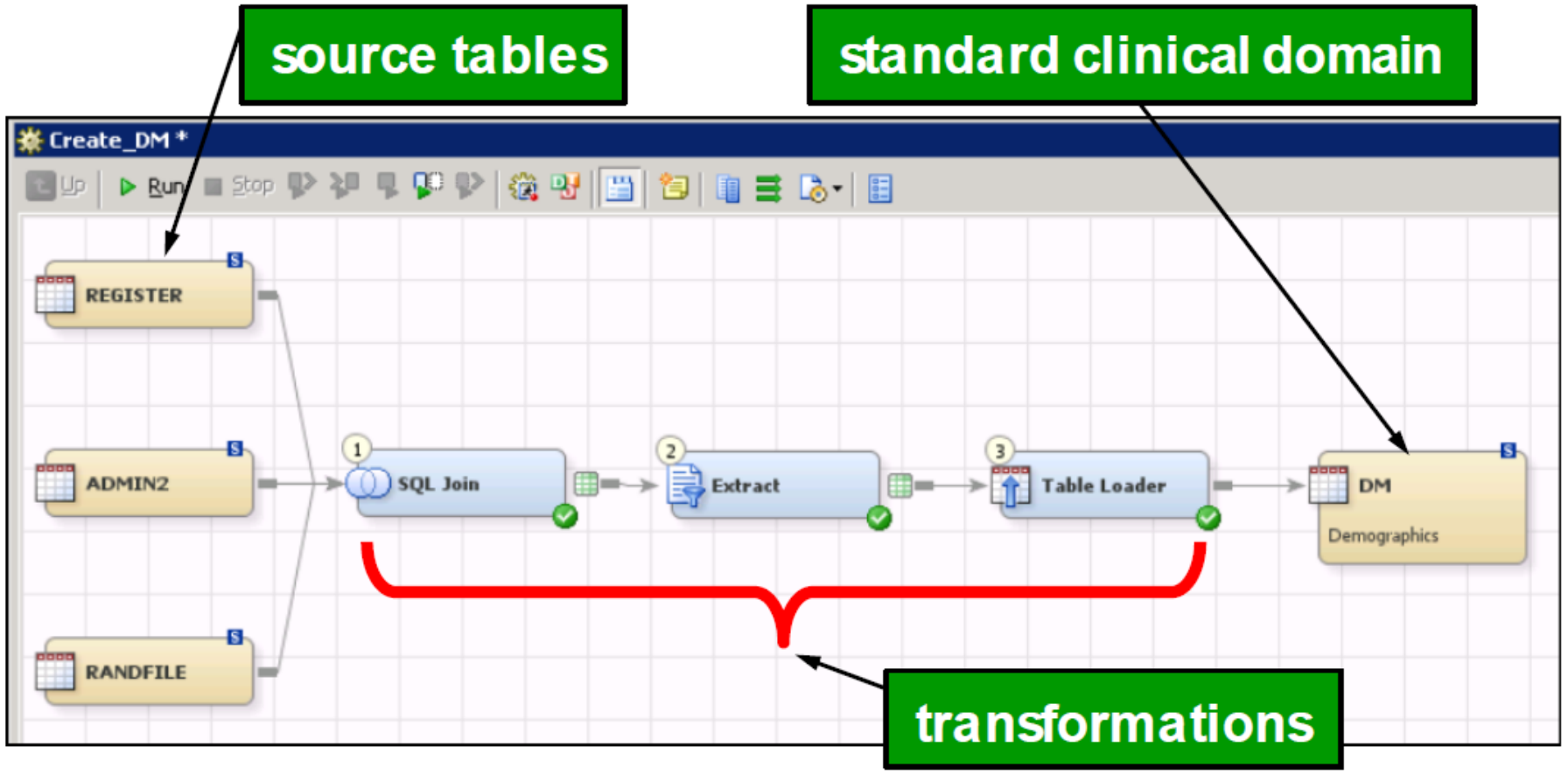
Registry SAS datasets tables



Now the datasets are created in the metadata server

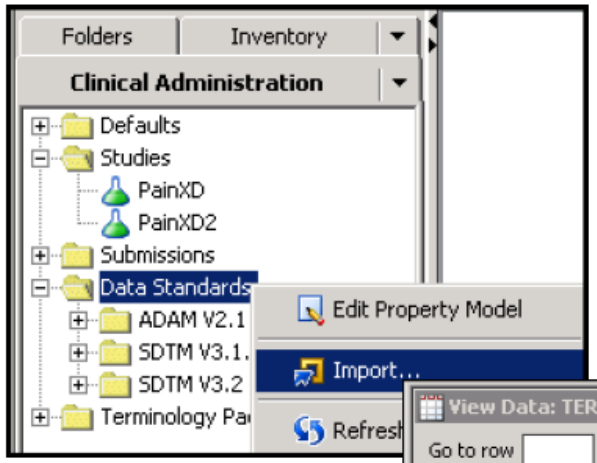
Implementation

- Visually design jobs to transform clinical data into standard domains



Implementation

- Import CDISC-controlled terminology tables



View Data: TERMS (8,145 rows)

Go to row Go

#	sourcesystem	sourcesystemversion	description	codelist	codelist_code	codelist_name	codeli
1	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACN	C66767	Action Taken with S...	No
2	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACN	C66767	Action Taken with S...	No
3	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACN	C66767	Action Taken with S...	No
4	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACN	C66767	Action Taken with S...	No
5	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACN	C66767	Action Taken with S...	No
6	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACN	C66767	Action Taken with S...	No
7	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACN	C66767	Action Taken with S...	No
8	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACSPCAT	C101865	Acute Coronary Sy...	No
9	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACSPCAT	C101865	Acute Coronary Sy...	No
10	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACSPCAT	C101865	Acute Coronary Sy...	No
11	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACSPCAT	C101865	Acute Coronary Sy...	No
12	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	ACSPCAT	C101865	Acute Coronary Sy...	No
13	NCI Thesaurus	2014-06-27	CDISC SDTM Co...	AESEV	C66769	Severity/Intensity ...	No

Implementation

- Import and manage compliance checks

Manage Compliance Checks: SDTM V3.2

Available checks:

Check ID	Description
SDTM0004	Source metadata includes domain data set not found in reference metadata
SDTM0005	Custom domain data set does not adhere to specification naming guidelines
SDTM0006	Source data library contains domain data not found in study metadata
SDTM0011	Identifies a column that was described in the domain description but not included in the SAS dataset for that domain
SDTM0014	Identifies a column listed in the domain description as Permissible ('Perm') but not included in the SAS dataset for that doma
SDTM0018	Identifies a domain that appears to be Associated Persons data but does not contain the Associated Persons Identifier (AP
SDTM0022	Column length < length defined in standard
SDTM0023	Column length > length defined in standard
SDTM0031	Column format found but column not subject to controlled terminology
SDTM0032	Column format found but format name mismatch with standard controlled terminology name
SDTM0202	Identifies a null (empty) value found in a column where (Standard) Core attribute is 'Exp'
SDTM0203	Character column value is not correctly upcased per spec
SDTM0204	Character column value contains the numeric missing '.' or any special missing value like '.N'
SDTM0205	Column value is not left-justified
SDTM0271	Value for column defined as a data set key is null

SAS CDI Functionality

- Execute compliance checks for domains column structure and data values

SAS Clinical Standards Toolkit 1.7
CDISC-SDTM 3.2 VALIDATION

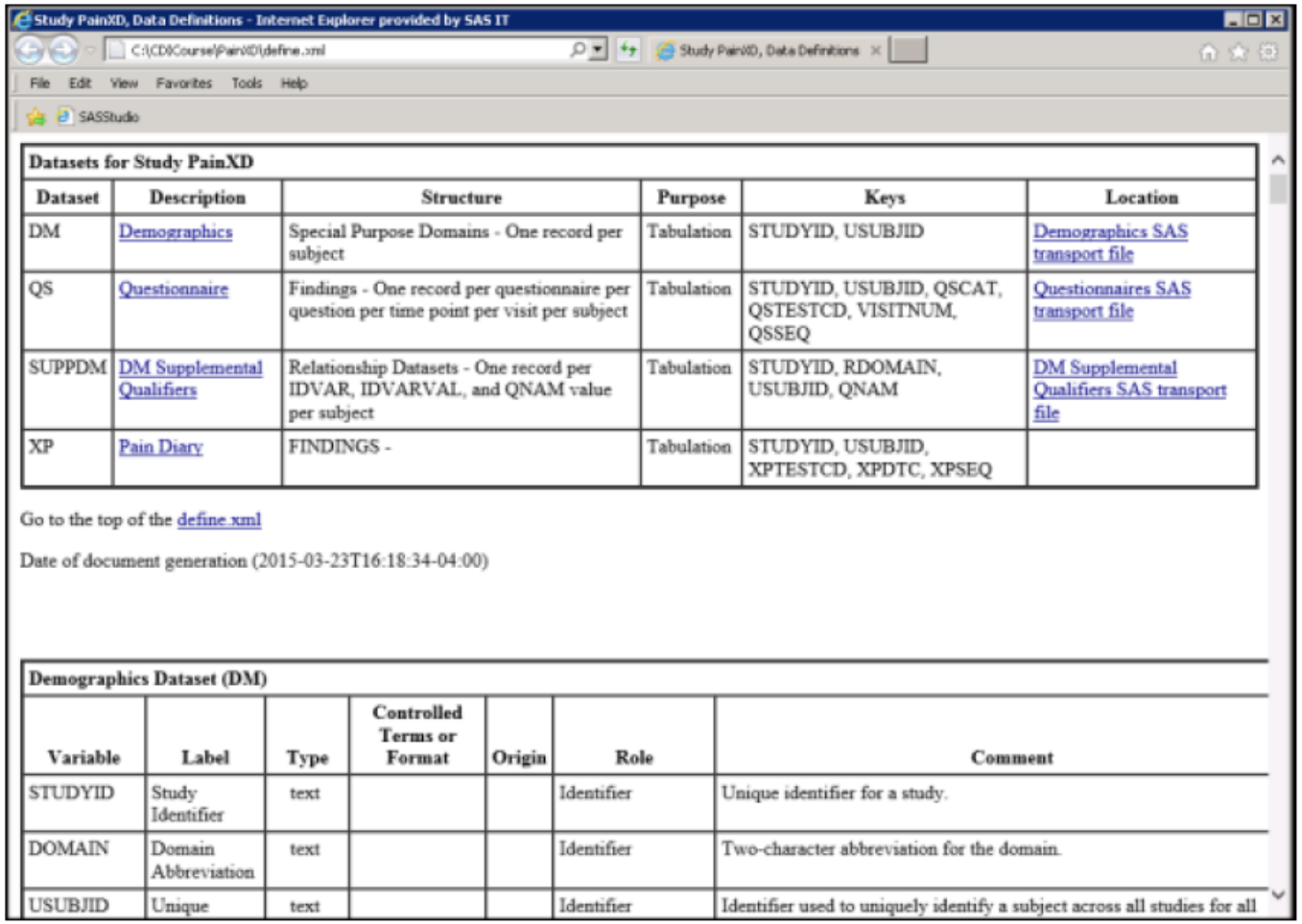
Process Results, CheckID: SDTM0204

Description: Character column value contains the numeric missing '.' or any special missing value like '.N'
Check scope: (Tables) _ALL_, (Columns) _ALL_
Source: SAS (SAS0011)
Validation check macro: cstcheck_column, using source metadata

Check Invocation	Seq #	Source Data	Result Identifier	Message	Severity	Problem Detected?	Actual Value	Keys
1	3	sdtm.QS	SDTM0204	Column value contains numeric missing value	Note	Yes	QSORRES=.	STUDYID=PXD,USUBJID=PXD-130-4006,QSCAT=BRIEF PAIN INVENTORY: MODIFIED SHORT FORM,QSTESTCD=BPIQ03,VISITNUM=3,QSSEQ=25
1	3	sdtm.QS	SDTM0204	Column value contains numeric missing value	Note	Yes	QSORRES=.	STUDYID=PXD,USUBJID=PXD-130-4006,QSCAT=BRIEF PAIN INVENTORY: MODIFIED SHORT FORM,QSTESTCD=BPIQ03,VISITNUM=3,QSSEQ=25
1	3	sdtm.QS	SDTM0204	Column value contains numeric missing value	Note	Yes	QSORRES=.	STUDYID=PXD,USUBJID=PXD-130-4006,QSCAT=BRIEF PAIN INVENTORY: MODIFIED SHORT

SAS CDI Functionality

- Generate CDISC standard define.xml describing domain and study metadata for clinical submissions



Datasets for Study PainXD

Dataset	Description	Structure	Purpose	Keys	Location
DM	Demographics	Special Purpose Domains - One record per subject	Tabulation	STUDYID, USUBJID	Demographics SAS transport file
QS	Questionnaire	Findings - One record per questionnaire per question per time point per visit per subject	Tabulation	STUDYID, USUBJID, QSCAT, QSTESTCD, VISITNUM, QSSEQ	Questionnaires SAS transport file
SUPPDM	DM Supplemental Qualifiers	Relationship Datasets - One record per IDVAR, IDVARVAL, and QNAM value per subject	Tabulation	STUDYID, RDOMAIN, USUBJID, QNAM	DM Supplemental Qualifiers SAS transport file
XP	Pain Diary	FINDINGS -	Tabulation	STUDYID, USUBJID, XPTESTCD, XPDTC, XPSEQ	

Go to the top of the define.xml

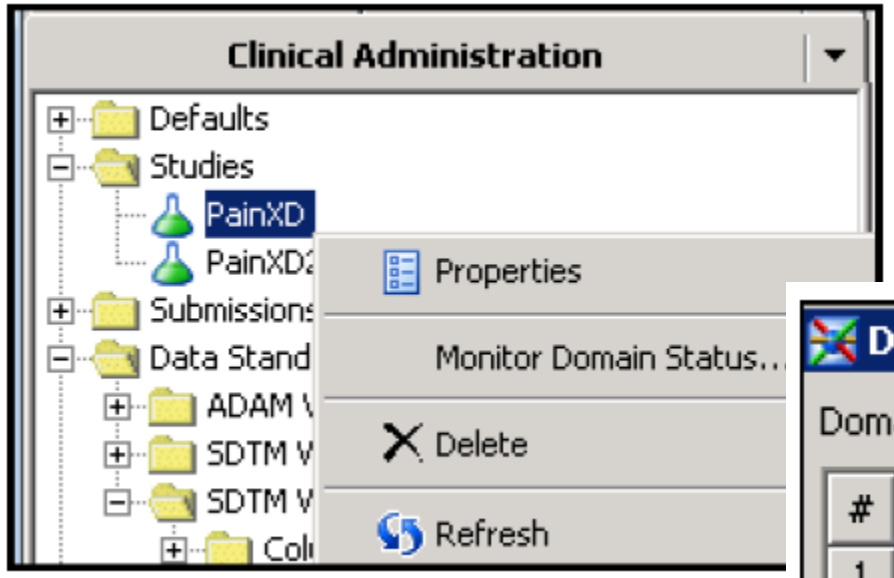
Date of document generation (2015-03-23T16:18:34-04:00)

Demographics Dataset (DM)

Variable	Label	Type	Controlled Terms or Format	Origin	Role	Comment
STUDYID	Study Identifier	text			Identifier	Unique identifier for a study.
DOMAIN	Domain Abbreviation	text			Identifier	Two-character abbreviation for the domain.
USUBJID	Unique	text			Identifier	Identifier used to uniquely identify a subject across all studies for all

SAS CDI Functionality

- Monitor the progress of clinical data integration projects



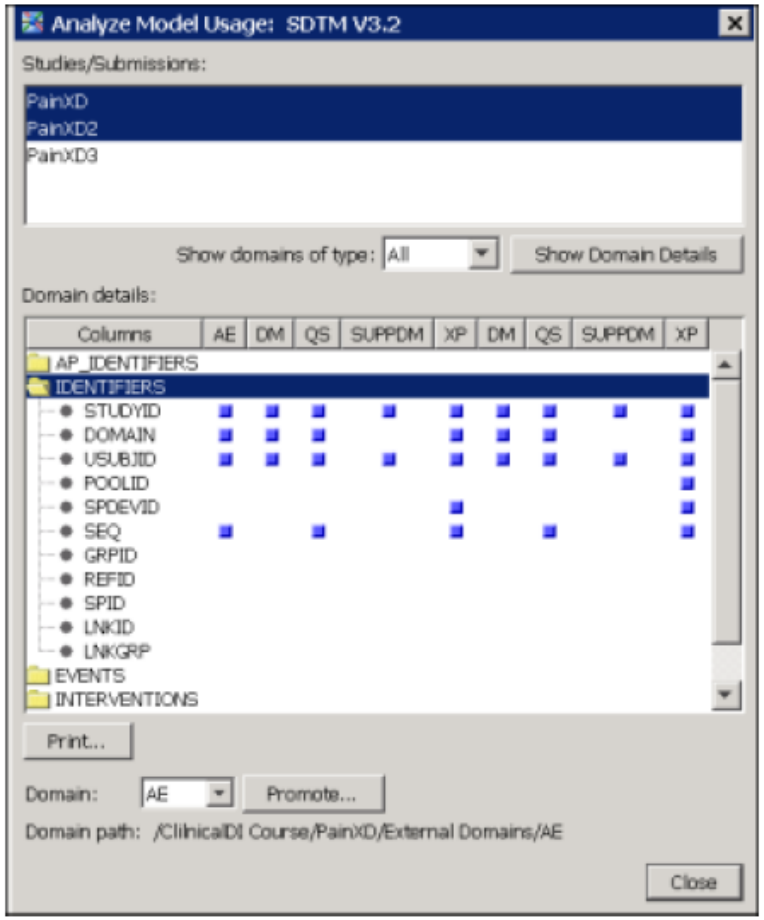
Domain Status: PainXD

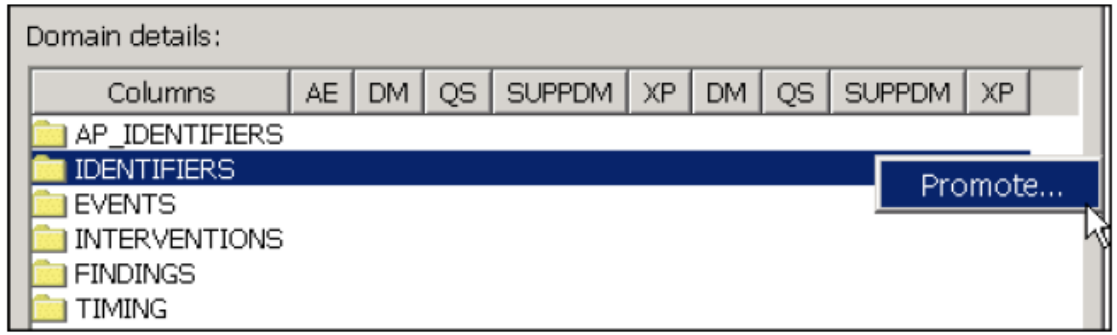
Domains:

#	Name	ID	Description	Is Mapp...	Is Validated?	Locked By
1	QS	QS	Questionnaire	Yes	Yes	
2	XP	XP	Pain Diary	Yes	No	
3	DM	DM	Demographics	Yes	Yes	
4	AE	AE	Adverse Events	No	Yes	
5	SUPPDM	SUPP...	DM Supplemental Qualifiers	Yes	No	

SAS CDI Functionality

- Compare domain usage across clinical studies and promote custom domains for general use





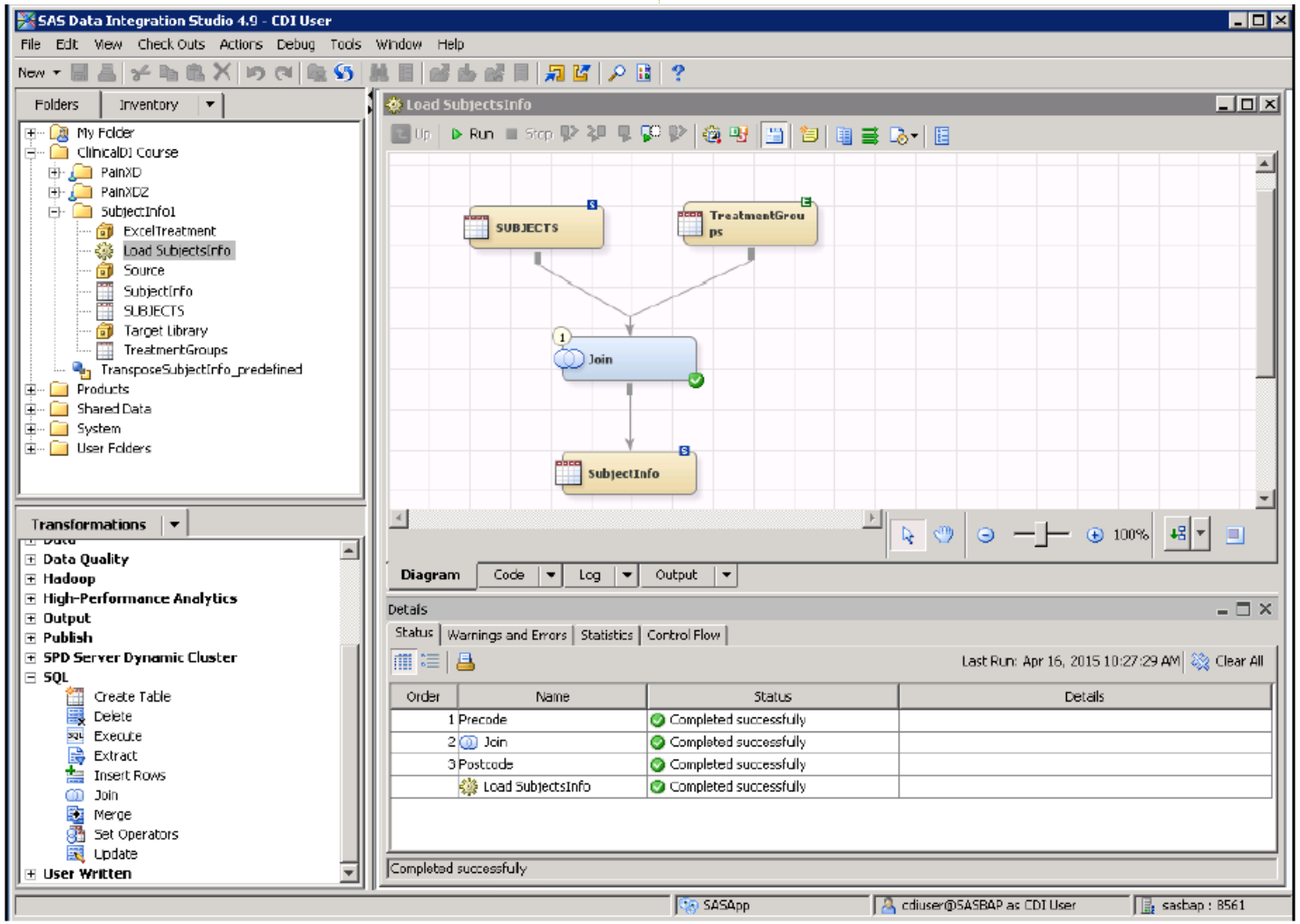
Pivotal

SAS CDI Functionality

- Integration with Electronic Data Capture (EDC) systems
  - Data capture from RAVE
  - Data capture from Oracle

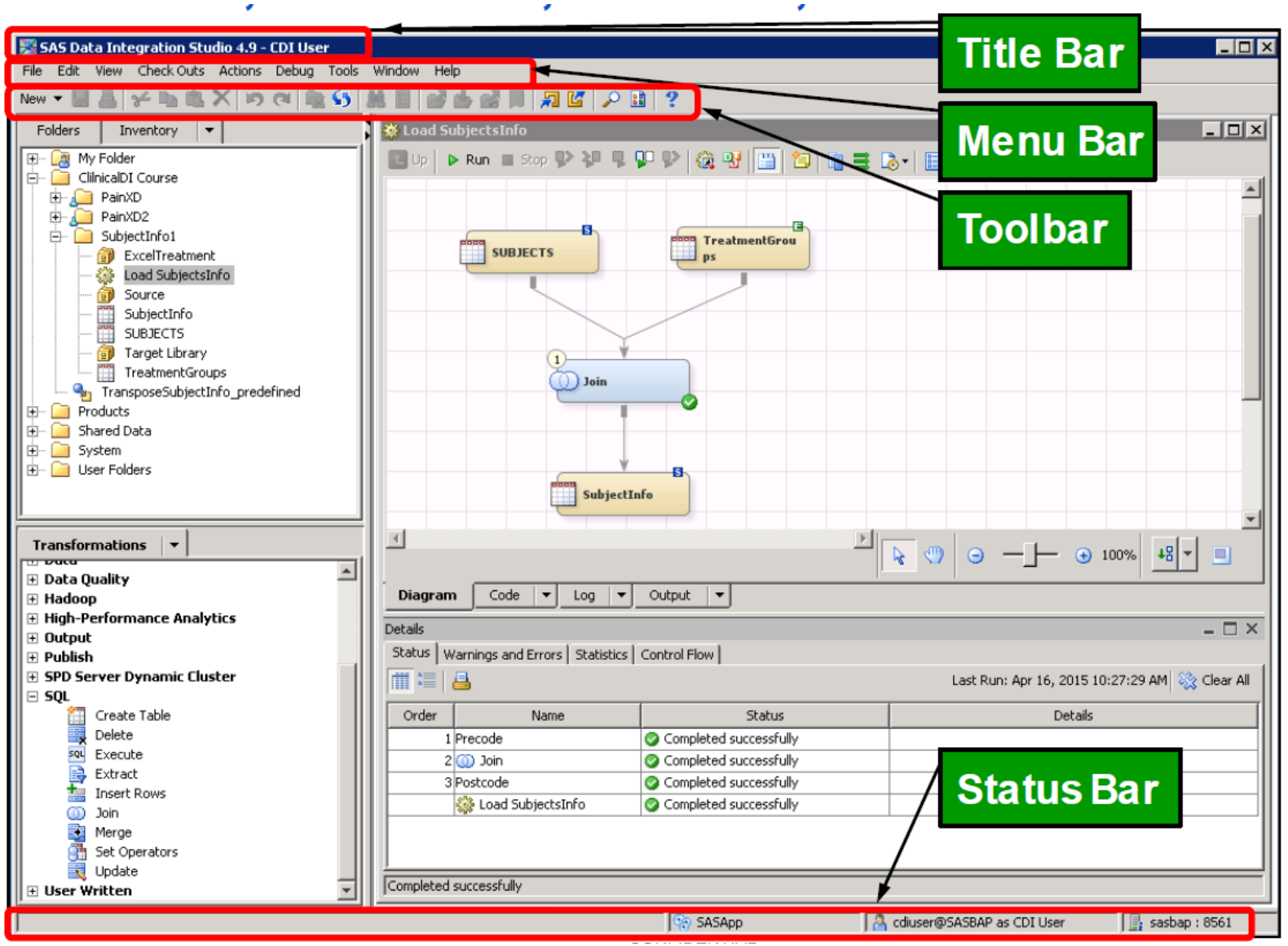
SAS Data Integration Studio Interface

- SAS Data Integration Studio has many components available in the desktop interface

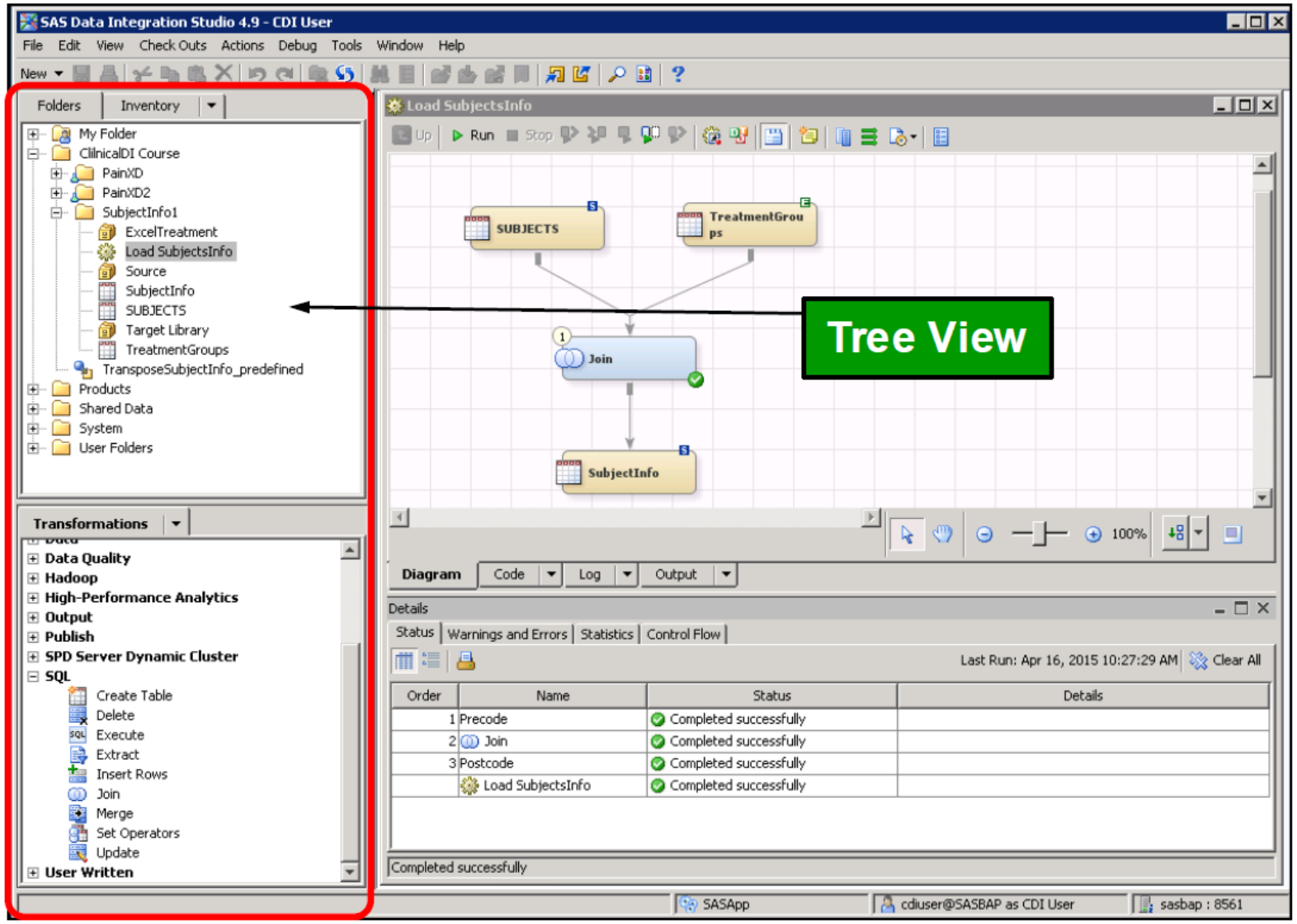


Pivotal

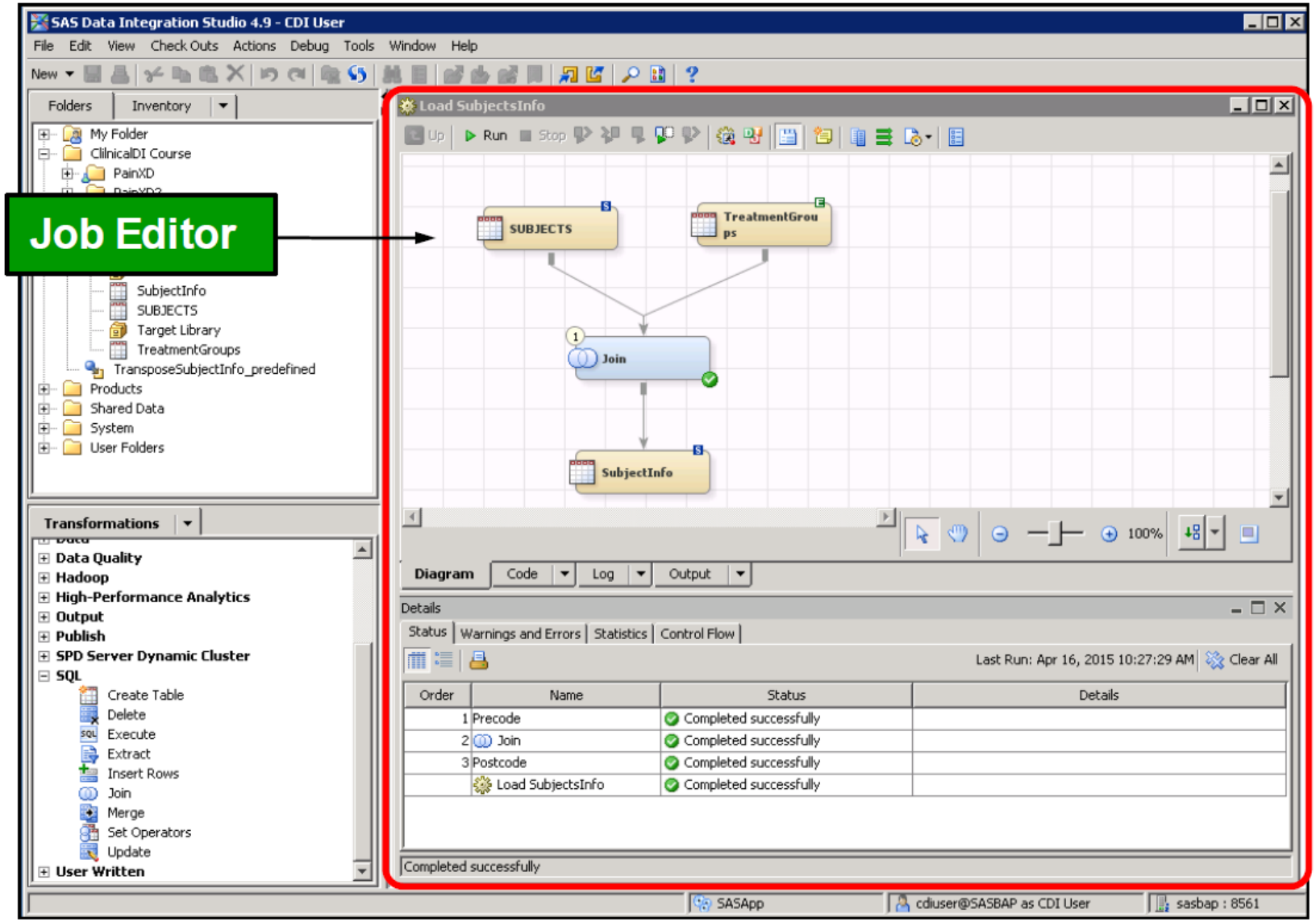
SAS Data Integration Studio Interface



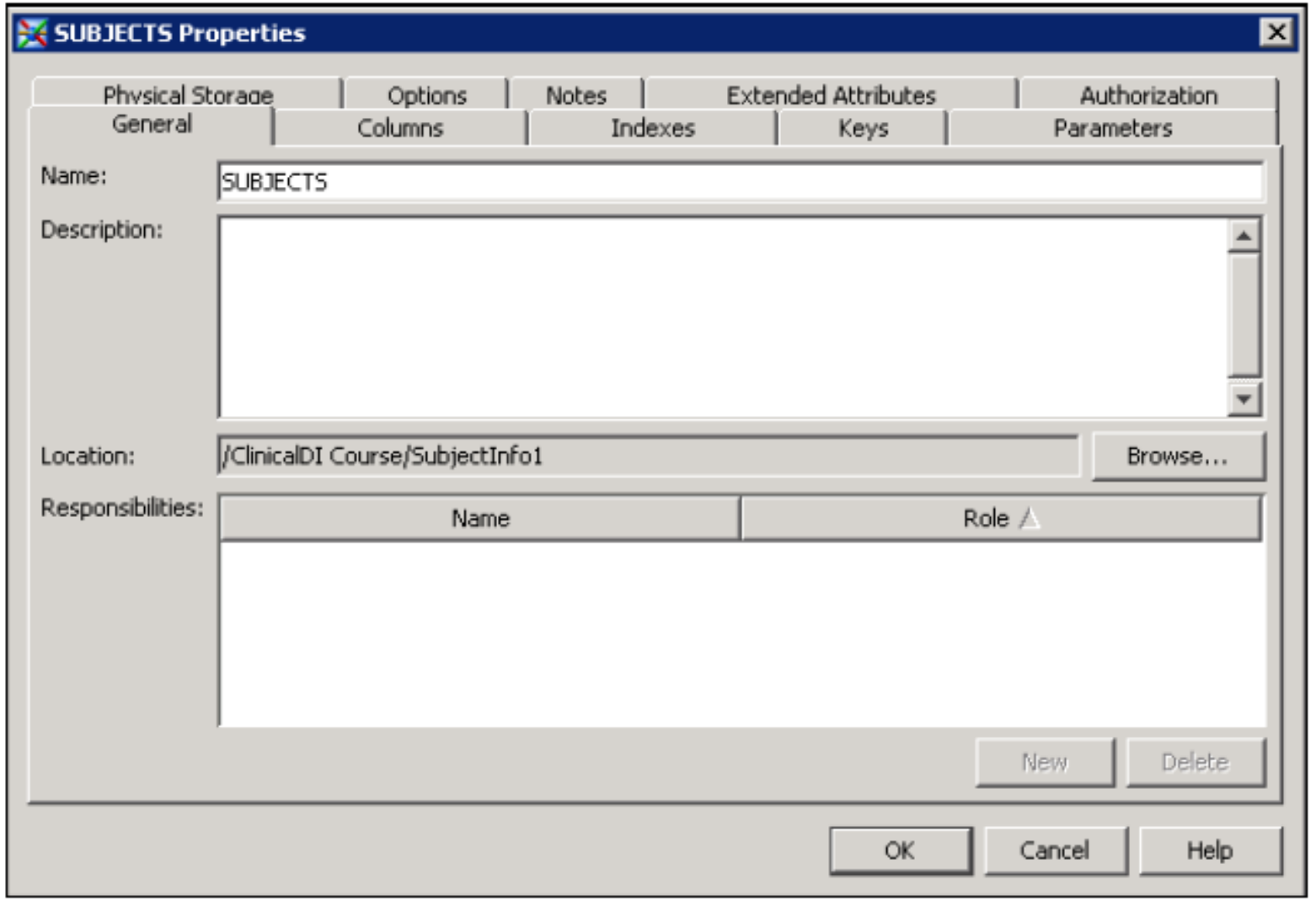
Tree View, basic properties pane



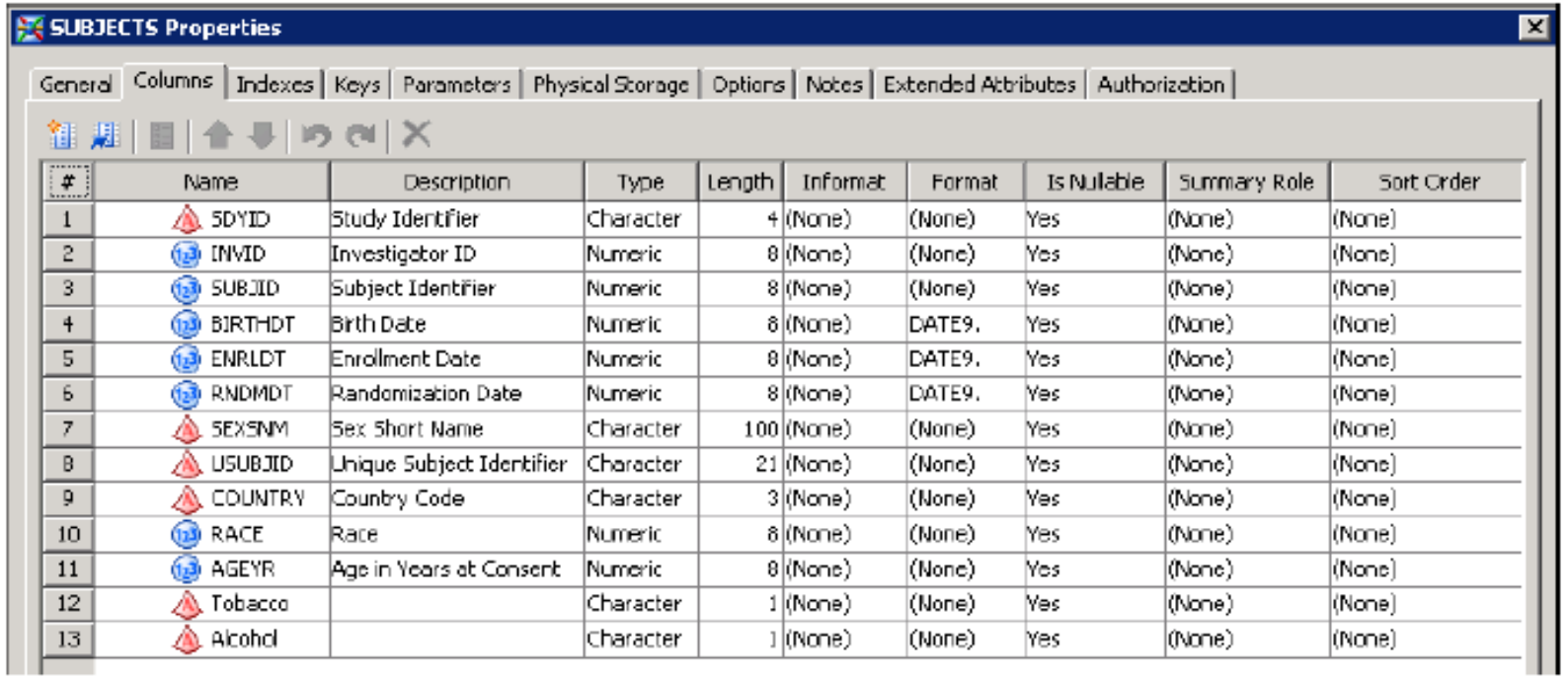
Job Editor


Job Editor
SAS Data Integration Studio 4.9 - CDI User
File Edit View Check Outs Actions Debug Tools Window Help
New
Folders
Inventory
My Folder
ClilnicalDI Course
PainXD
SubjectInfo
SUBJECTS
Target Library
TreatmentGroups
TransposeSubjectInfo_predefined
Products
Shared Data
System
User Folders
Transformations
Data Quality
Hadoop
High-Performance Analytics
Output
Publish
SPD Server Dynamic Cluster
SQL
Create Table
Delete
Execute
Extract
Insert Rows
Join
Merge
Set Operators
Update
User Written
Load SubjectsInfo
Up
Run
Stop
SUBJECTS
TreatmentGroups
Join
SubjectInfo
100%
Diagram
Code
Log
Output
Details
Status
Warnings and Errors
Statistics
Control Flow
Last Run: Apr 16, 2015 10:27:29 AM
Clear All
Order
Name
Status
Details
1 Precode
Completed successfully
2 Join
Completed successfully
3 Postcode
Completed successfully
Load SubjectsInfo
Completed successfully
Completed successfully
SASApp
cdiuser@SASBAP as CDI User
sasbap : 8561

Object properties

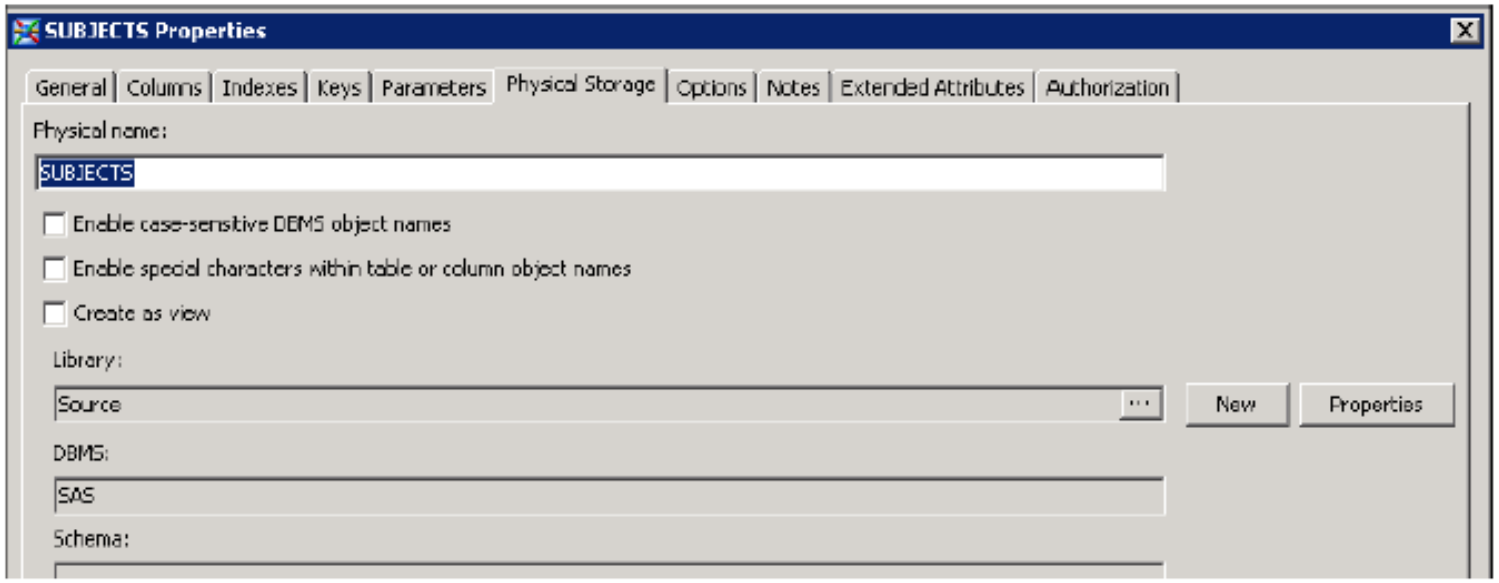


Object properties

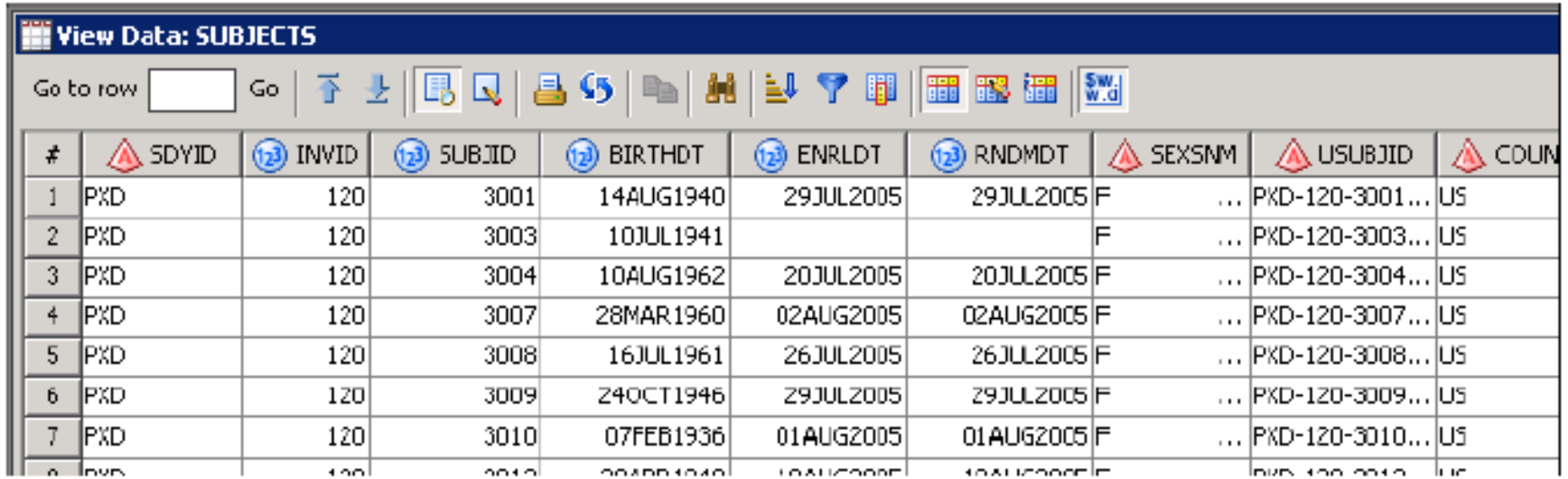


#	Name	Description	Type	Length	Informat	Format	Is Nullable	Summary Role	Sort Order
1	SDYID	Study Identifier	Character	4	(None)	(None)	Yes	(None)	(None)
2	INVID	Investigator ID	Numeric	8	(None)	(None)	Yes	(None)	(None)
3	SUBJID	Subject Identifier	Numeric	8	(None)	(None)	Yes	(None)	(None)
4	BIRTHDT	Birth Date	Numeric	8	(None)	DATE9.	Yes	(None)	(None)
5	ENRLDT	Enrollment Date	Numeric	8	(None)	DATE9.	Yes	(None)	(None)
6	RNDMDT	Randomization Date	Numeric	8	(None)	DATE9.	Yes	(None)	(None)
7	SEXSNM	Sex Short Name	Character	100	(None)	(None)	Yes	(None)	(None)
8	USUBJID	Unique Subject Identifier	Character	21	(None)	(None)	Yes	(None)	(None)
9	COUNTRY	Country Code	Character	3	(None)	(None)	Yes	(None)	(None)
10	RACE	Race	Numeric	8	(None)	(None)	Yes	(None)	(None)
11	AGEYR	Age in Years at Consent	Numeric	8	(None)	(None)	Yes	(None)	(None)
12	Tobacco		Character	1	(None)	(None)	Yes	(None)	(None)
13	Alcohol		Character	1	(None)	(None)	Yes	(None)	(None)

Object properties



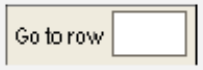
View data window



#	SDYID	INVID	SUBJID	BIRTHDT	ENRLDT	RNDMDT	SEXSNM	USUBJID	COUN
1	PXD	120	3001	14AUG1940	29JUL2005	29JUL2005	F ...	PXD-120-3001...	US
2	PXD	120	3003	10JUL1941			F ...	PXD-120-3003...	US
3	PXD	120	3004	10AUG1962	20JUL2005	20JUL2005	F ...	PXD-120-3004...	US
4	PXD	120	3007	28MAR1960	02AUG2005	02AUG2005	F ...	PXD-120-3007...	US
5	PXD	120	3008	16JUL1961	26JUL2005	26JUL2005	F ...	PXD-120-3008...	US
6	PXD	120	3009	24OCT1946	29JUL2005	29JUL2005	F ...	PXD-120-3009...	US
7	PXD	120	3010	07FEB1936	01AUG2005	01AUG2005	F ...	PXD-120-3010...	US

The functions of the View Data window are controlled by the View Data toolbar

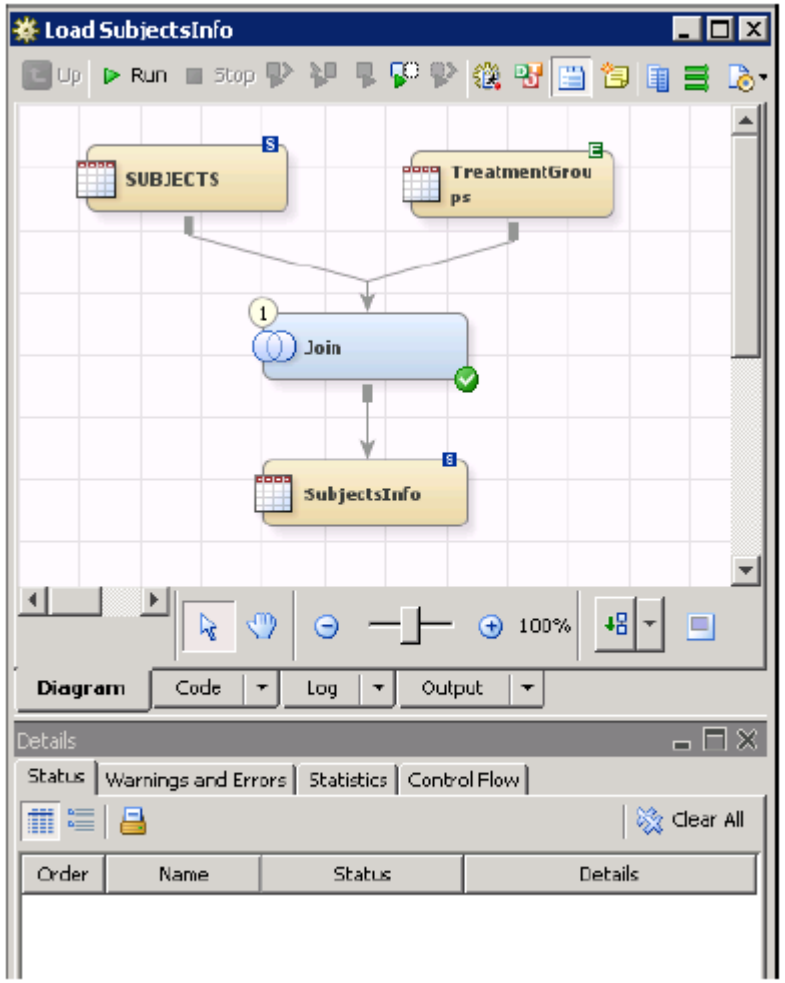
View data window

TOOL	EXPLANATION
	Specifies the number of the first row that is displayed in the table.
	Positions the data with the go-to row as the first data line displayed.
	Navigates to the first record of data in the View Data window.
	Navigates to the last page of data in the View Data window.
	Switches to Browse mode.
	Switches to Edit mode.
	Enables printing.
	Refreshes the view of the data.
	Copies selected data values to the clipboard.
	Displays the Search area.

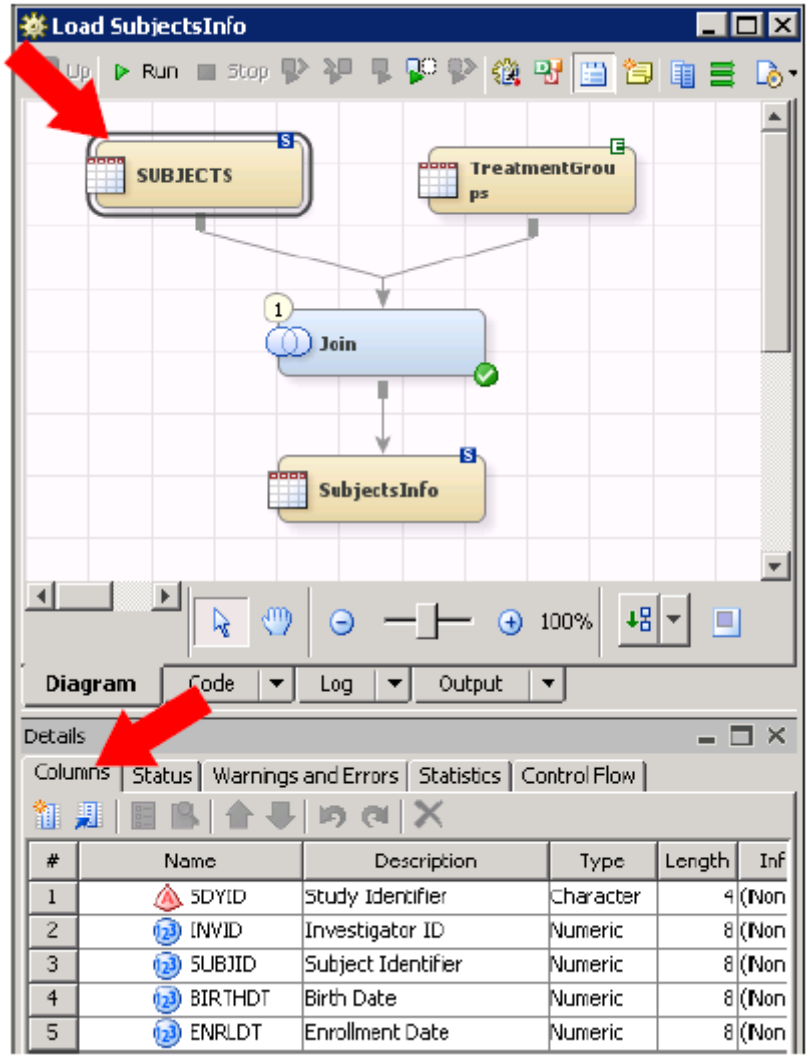
View data window

	Displays the Sort By Column tab in the View Data Options window.
	Displays the Filter tab in the View Data Options window.
	Displays the Columns tab in the View Data Options window.
	Displays physical column names in the column headings. You can display any combination of column metadata, physical column names, and descriptions in the column headings.
	Displays optional descriptions in the column headings.
	Displays optional column metadata in the column headings. This metadata can be entered in some SAS Intelligence Platform applications, such as SAS Information Map Studio.
	Toggles between showing formatted and unformatted data in the View Data window.

Jobs


Load SubjectsInfo
Up
Run
Stop
SUBJECTS
TreatmentGrou
ps
1
Join
SubjectsInfo
100%
Diagram
Code
Log
Output
Details
Status
Warnings and Errors
Statistics
Control Flow
Clear All
Order
Name
Status
Details

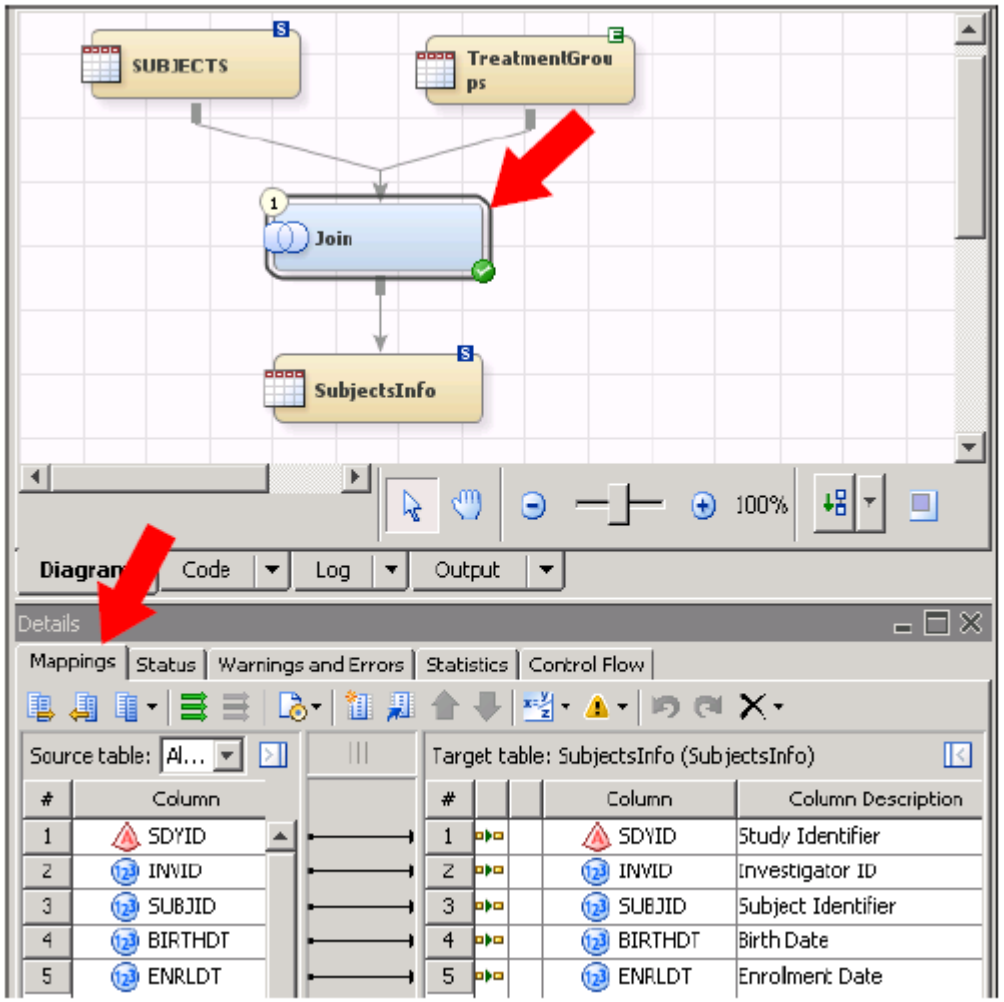
Jobs



The column tab in the Details area displays column attributes for the selected table object.

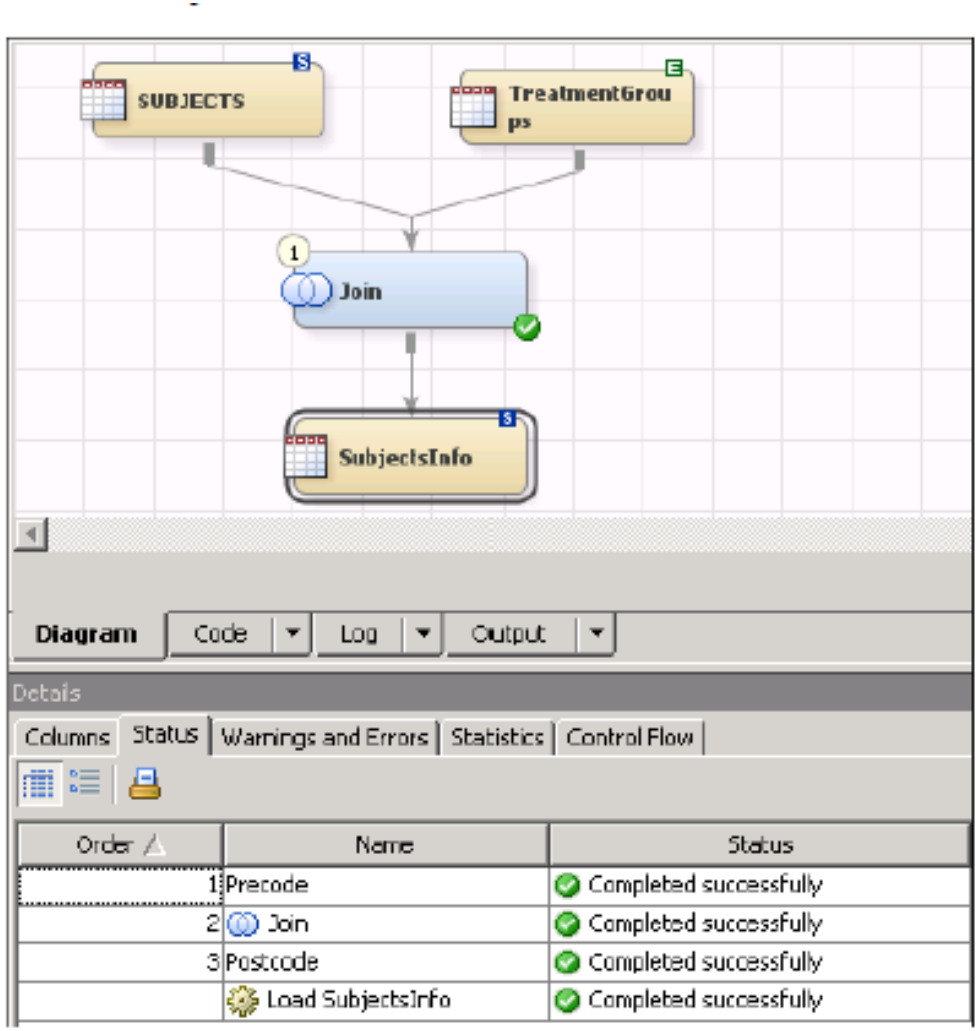
These attributes are fully editable in this location

Jobs



The Mappings tab provides control for how columns from source tables are propagated and mapped to target tables

Jobs

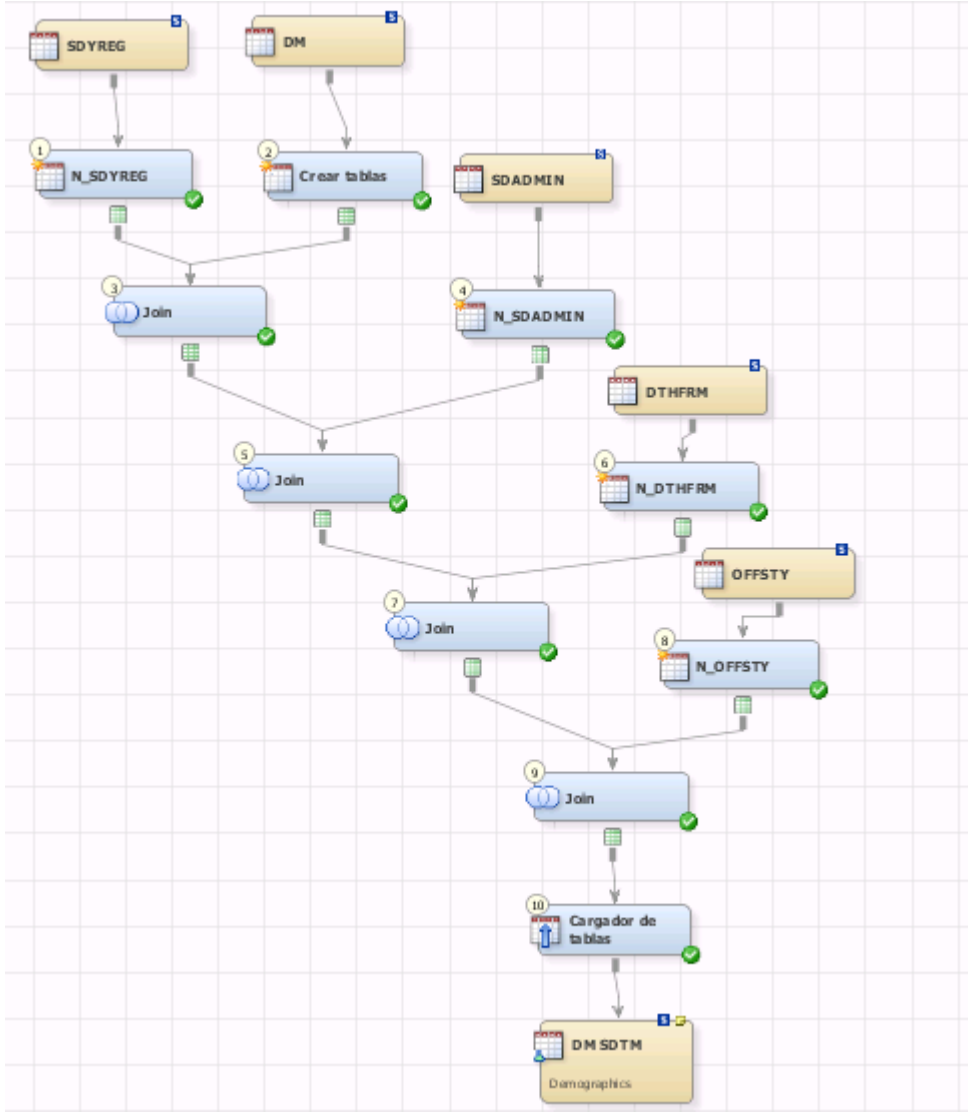


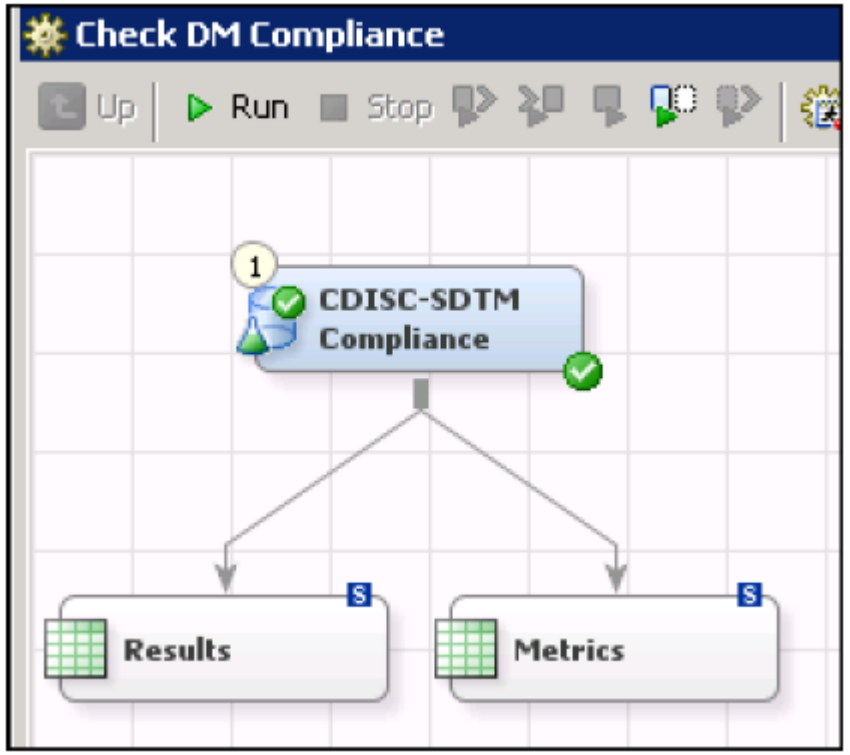
After run the job by clicking “RUN”, the Status tab indicate that the job completed without generating any errors or warnings

Code tab shows the code generated by the transformation

Log tab shows the log

DM domain mapping example






Thank you