

*BIDW/IIRM Data
Quality Analysis
Presents*

**Informatica - Data Explorer
Reports & Metadata
By: Jim Stewart, Sr. Data Analyst
AT&T Services, Inc.
925-867-7295**

Pains

- **Previous tools used were SAS, UNIX AWK & SED programs**
 - Turnaround time of data analysis requests use to be about two to three weeks
 - Modification of programs were labor intensive and mistake prone
 - Initial hours of analysis decreased approx. 50%, turnaround of requests went from two to three weeks to less than one week
- **Current reasons for uses**
 - Not able to see real data, source or target
 - Not having availability of data during project meetings
 - Not having a view of data while in development process

Who uses Data Explorer Reports?

- Intended Users
 - Business Analysts
 - Data Analysts
 - Data Modelers
 - Database Administrators
 - Physical Database Designers
 - Project Managers
 - Requirements Analysts
 - System Administrators
 - Technical Business Analysts

What does Data Explorer do?

- Data Explorer is a database profiler and mapping tool that uses production data during the profiling and design processes:
 - Process database columns or flat files for analysis
 - Understanding legacy data before transforming it
 - Decreases the chance that the final design will encounter performance or data integrity problems
 - Decreases time lapse of problem recovery in quality control of new releases

What else does Data Explorer do?

- Data Explorer builds a complete and accurate profile of the existing data, clarifying:
 - Data elements
 - Data redundancies
 - Data quality issues

What are the Reports from Data Explorer?

- Attribute Summary Report: biggest benefit – Provide a view of the data at a glance, shows distinct counts from total records, lowest value and highest value. Easy to find constant values.
- Column Profile Report: biggest benefit – Provides 25 most used and 25 least used values for all columns. Infers a data type and alternate data types based on the values processed. Provides Pattern Analysis that has been inferred by the data.
- Redundancy Analysis Report: biggest benefit – Provide a comparison between different columns data values and provides overlap percentages.

ATTRIBUTE SUMMARY REPORT

	Relation	Attribute	# Distinct	# Records	% Unique	# Nulls	% Null	Inferred Pattern	% Match to Pattern	Inferred Data Type	Inferred Null Rule	Inferred Minimum Value	Inferred Maximum Value
31	TATS209_EQUIPMENT	ADDITIONAL_LINE_IND	2	253625490	0	384484	0.15	U	100	CHAR (1)	Null OK	N	Y
13	TATS209_EQUIPMENT	AGREEMENT_TYPE_CD	633	253625490	0	0	0	bb	97.33	CHAR (2)	Not Null		ZZ
4	TATS209_EQUIPMENT	ATN_CUSTOMER_CD	1000	253625490	0	0	0	999	100	NUMBER (3,0)	Not Null	0	999
3	TATS209_EQUIPMENT	ATN_LINE_NBR	10000	253625490	0	0	0	9999	100	NUMBER (4,0)	Not Null	0	9999
1	TATS209_EQUIPMENT	ATN_NPA_CD	40	253625490	0	0	0	999	100	NUMBER (3,0)	Not Null	216	989
2	TATS209_EQUIPMENT	ATN_NXX_CD	1180	253625490	0	0	0	999	98.57	CHAR (3)	Not Null	151	Z99
23	TATS209_EQUIPMENT	BILLABLE_IND	2	253625490	0	0	0	U	100	CHAR (1)	Not Null	N	Y
10	TATS209_EQUIPMENT	CIRCUIT_FORMAT_CD	4	253625490	0	384484	0.15	bbb	97.83	CHAR (3)	Null OK		CLS
6	TATS209_EQUIPMENT	CIRCUIT_ID	21698774	253625490	8.56	384484	0.15	9(10)	97.68	VARCHAR2 (33)	Null OK		ZVNZ9690.001MB
7	TATS209_EQUIPMENT	CIRCUIT_ID_FORMATTED_TXT	21892527	253625490	8.63	384484	0.15	9(10)	96.51	VARCHAR2 (32)	Null OK		ZVNZ9690001MB
9	TATS209_EQUIPMENT	CIRCUIT_IDENTIFIER_CD	4	253625490	0	384484	0.15	b	97.83	CHAR (1)	Null OK		3
8	TATS209_EQUIPMENT	CIRCUIT_LEG_NBR	399	253625490	0	384484	0.15	9999	100	NUMBER (4,0)	Null OK	0	987
33	TATS209_EQUIPMENT	CLASS_OF_SERVICE_CD	532	253625490	0	384484	0.15	UUUbb	41.55	CHAR (5)	Null OK		ZZGUX
25	TATS209_EQUIPMENT	CONTRACT_NBR	3821	253625490	0	0	0	b(6)	98.04	CHAR (6)	Not Null		900000
17	TATS209_EQUIPMENT	CONTRACT_START_DT	3270	253625490	0	247473924	97.57	9999-99-99	100	DATE	Null OK	1929-11-29	2022-10-17
12	TATS209_EQUIPMENT	DATA_TERMINAL_NBR	334726	253625490	0.13	384484	0.15	b(10)	99.73	CHAR (10)	Null OK		9899842100
34	TATS209_EQUIPMENT	DW_LOAD_DT	1	253625490	0	0	0	9999-99-99	100	DATE	Not Null	2003-03-11	2003-03-11
24	TATS209_EQUIPMENT	IN_SERVICE_IND	2	253625490	0	0	0	U	100	CHAR (1)	Not Null	N	Y
19	TATS209_EQUIPMENT	MONTH_EXTENSION_DUR	8	253625490	0	0	0	9	99.99	NUMBER (2,0)	Not Null	0	84
11	TATS209_EQUIPMENT	SECONDARY_PLID_TXT	1585	253625490	0	384484	0.15	bbbb	98.83	CHAR (4)	Null OK		9014
30	TATS209_EQUIPMENT	STATE_CD	5	253625490	0	0	0	UU	100	CHAR (2)	Not Null	IL	WI
14	TATS209_EQUIPMENT	SUB_CLASS_OF_SERVICE_CD	792	253625490	0	0	0	b(5)	97.49	CHAR (5)	Not Null		ZZOCL
18	TATS209_EQUIPMENT	TERM_AGREEMENT_MONTH_DUR	84	253625490	0	0	0	9	97.58	NUMBER (3,0)	Not Null	0	999
27	TATS209_EQUIPMENT	UNIT_RATE_AMT	8961	253625490	0	0	0	9.9999	91.46	NUMBER (10,4)	Not Null	-300000.0000	218781.6600
5	TATS209_EQUIPMENT	USOC	7713	253625490	0	0	0	UUUbb	37.36	CHAR (5)	Not Null	112	ZZZ06
28	TATS209_EQUIPMENT	USOC_AMT	17809	253625490	0.01	0	0	9.99	92.88	NUMBER (8,2)	Not Null	-300000.00	550335.42
15	TATS209_EQUIPMENT	USOC_EFFECTIVE_BILL_DT	3597	253625490	0	384484	0.15	9999-99-99	100	DATE	Null OK	1959-12-31	2022-03-22
26	TATS209_EQUIPMENT	USOC_IN_SERVICE_QTY	2183	253625490	0	253241006	99.85	9	50.45	NUMBER (5,0)	Null OK	1	85440
16	TATS209_EQUIPMENT	USOC_LAST_ACTIVITY_DT	2978	253625490	0	384544	0.15	9999-99-99	100	DATE	Null OK	1993-08-03	2003-03-03
29	TATS209_EQUIPMENT	USOC_QTY	2361	253625490	0	0	0	9	99.84	NUMBER (5,0)	Not Null	0	85440
32	TATS209_EQUIPMENT	WHOLESALE_RETAIL_CD	2	253625490	0	0	0	U	100	CHAR (1)	Not Null	R	W
22	TATS209_EQUIPMENT	WTN_LINE_NBR	10001	253625490	0	384484	0.15	9999	98.09	CHAR (4)	Null OK		9999
20	TATS209_EQUIPMENT	WTN_NPA_CD	161	253625490	0	384484	0.15	999	98.09	CHAR (3)	Null OK		999
21	TATS209_EQUIPMENT	WTN_NXX_CD	1254	253625490	0	384484	0.15	999	97.94	CHAR (3)	Null OK		Z99

Sample: Column Profile Report

```

Column Profile:  ACCOUNT_NBR
Distinct Values: 2388
Records:        2388
Axio Populated: 2388
Nulls:          0
Null Rule:      Nulls Not Allowed
Data Type:      INTEGER( 10, 6 )
Matches:        2388
Minimum Values:
    237137602
    237137705
    237137710
Maximum Values:
    9981676406
    9981665325
    9981663217
Mean:           6191121703
Std Dev:        2176438714
Mismatches:     0
Alternate Data Type:
Data Type:      CHAR( 10 )
Matches:        2388
Minimum:        0237137602
Maximum:        9981676406
Mismatches:     0
Patterns:       1 of 1
                (2388) 9(10)
Most Frequent Values:
    1 0237137602
    1 1233057376
    1 1312515900
    1 1312516734
    1 1312517510
    1 1312517729
    1 1312517900
    1 1312518442
    1 1312518518
    1 1312518652
Least Frequent Values:
    1 9608322016
    1 9981582594
    1 9981593336
    1 9981594312
    1 9981594613
    1 9981594746
    1 9981604214
    1 9981604303
    1 9981676393
    1 9981676406
    
```

1) Each Record has a unique Account_Nbr

2) Info on Account_Nbr Field and populated in Metadata

3) "1" is the number of times each Account_Nbr appears on the files

4) "0237137602" is an Account_Nbr

7) Shows data using different Data Formats. All 2388 records fit into each of these Data Formats.

```

Column Profile:  CUST_CD
Distinct Values: 900
Records:        2388
Axio Populated: 900
Nulls:          0
Null Rule:      Nulls Not Allowed
Data Type:      2388 INTEGER( 2 )
Matches:        2388
Minimum Values:
    1
    2
    3
Maximum Values:
    997
    998
    999
Mean:           489
Std Dev:        290
Mismatches:     0
Alternate Data Type:
Data Type:      CHAR( 3 )
Matches:        2388
Minimum:        001
Maximum:        999
Mismatches:     0
Patterns:       1 of 1
                (2388) 999
Most Frequent Values:
    20 001
    20 777
    16 555
    8 085
    8 494
    8 617
    8 711
    8 880
    7 255
    7 256
Least Frequent Values:
    1 962
    1 964
    1 966
    1 979
    1 980
    1 988
    1 994
    1 995
    1 997
    1 998
    
```

5) Cust_Cd has 900 Unique values

6) From the "Minimum" & "Maximum", Cust_Cd's range can be determined: 001-999

8) Acct-Cust-Code of "001" and "777" are "Top Values" because they occur the most: 20 times

9) Many Acct-Cust-Codes appear just once. The list is sorted from low to high values. Here are the last ten

```

Column Profile:  ADEN_TYPE
Distinct Values: 1
Records:        2388
Axio Populated: 2
Nulls:          1355
Null Rule:      Null OK
Data Type:      CHAR( 1 )
Matches:        1033
Minimum:        C
Maximum:        C
Mismatches:     0
Patterns:       1 of 1
                (1033) U
Most Frequent Values:
    1033 C
Least Frequent Values:
    1033 C
    
```

10) Aden-Type has 1 value and only lists one, "C". If the number of "Distinct Values" is 20 or less, "Top & Bottom 10 Values" will show the entire domain set.

11) Notice Aden_Type contains "Nulls" in 1355 records

12) Axio Populated 2 domain values into the catalog when the data was processed. This includes the Nulls

Pattern Analysis Characters:

Character	Matches in pattern
'B' or 'b' or ' '	blank
'C' or 'c'	any character
'L' or 'l'	any lower case alphabetic character
'P' or 'p'	a left parenthesis - '('
'Q' or 'q'	a right parenthesis - ')'
'T' or 't'	a tab
'U' or 'u'	any upper case alphabetic character
'X' or 'x'	any upper or lower case alphabetic character
'9'	any numeric character
'(' and ')'	used to indicate repetition
Any other character	itself

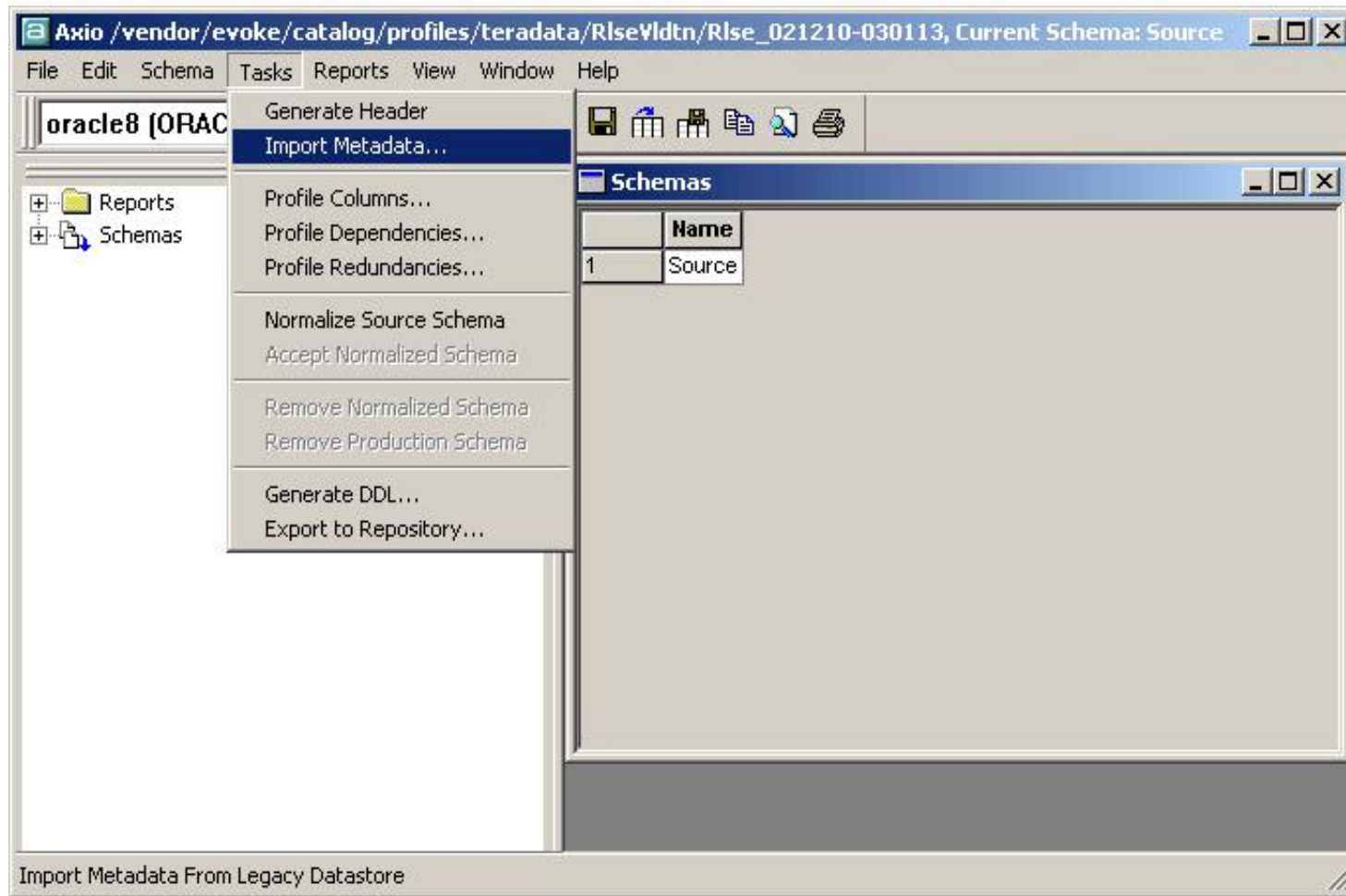
REDUNDANCY PROFILE REPORT

Estimated % Overlap	Relation 1	Attribute 1	Attribute 1 # Distinct	Relation 2	Attribute 2	Attribute 2 # Distinct	
1	100	TATS209_EQUIPMENT_EOM_CURR	ATN_LINE_NBR	10000	TCCS322_EQUIPMENT	LINE	10000
2	100	TATS209_EQUIPMENT_EOM_CURR	ATN_CUSTOMER_CD	1000	TCCS322_EQUIPMENT	CUSTCODE	1000
3	100	TCCS322_EQUIPMENT	WTN_LRN_XREF_NPA	191	TCCS322_EQUIPMENT	WTN_NPA	191
4	100	TCCS322_EQUIPMENT	WTN_LRN_XREF_NXX	801	TCCS322_EQUIPMENT	WTN_NXX	801
5	100	TCCS322_EQUIPMENT	WTN_LINE	10001	TCCS322_EQUIPMENT	WTN_LRN_XREF_LINE	10001
6	100	TCCS322_EQUIPMENT	EQUIPMENT_ACTIVITY_DATE	1030	TCCS322_EQUIPMENT	LOCATION_LIFE	1030
7	100	TCTN108_TRUNK_GROUP	TRAFFIC_CLASS_CODE	8	TCTN109_TRUNK_GROUP_FORECAST	TRAFFIC_CLASS_CODE	8
8	100	TCTN108_TRUNK_GROUP	TRAFFIC_CLASS_CODE	8	TCTN110_TRUNK_GROUP_USAGE	TRAFFIC_CLASS_CODE	8
9	100	TCTN109_TRUNK_GROUP_FORECAST	TRAFFIC_CLASS_CODE	8	TCTN110_TRUNK_GROUP_USAGE	TRAFFIC_CLASS_CODE	8
10	100	TZTF055_RPTG_ORGANIZATIO_RECST	SBC_OPERATING_UNIT_CD	15	TZTF055_RPTG_ORGANIZATION_ASGN	SBC_OPERATING_UNIT_CD	15
11	100	TZTF055_RPTG_ORGANIZATIO_RECST	RPTG_ORG_ASSIGNMENT_CLASS_CD	4	TZTF055_RPTG_ORGANIZATION_ASGN	RPTG_ORG_ASSIGNMENT_CLASS_CD	4
12	99	TATS209_EQUIPMENT_EOM_CURR	ATN_LINE_NBR	10000	TATS209_EQUIPMENT_EOM_CURR	WTN_LINE_NBR	10001
13	99	TATS209_EQUIPMENT_EOM_CURR	ATN_LINE_NBR	10000	TCCS322_EQUIPMENT	WTN_LINE	10001
14	99	TATS209_EQUIPMENT_EOM_CURR	ATN_LINE_NBR	10000	TCCS322_EQUIPMENT	WTN_LRN_XREF_LINE	10001
15	99	TATN605_NPA_NXX	NXX_CD	999	TATS209_EQUIPMENT_EOM_CURR	ATN_CUSTOMER_CD	1000
16	99	TCCS322_EQUIPMENT	LINE	10000	TATS209_EQUIPMENT_EOM_CURR	WTN_LINE_NBR	10001
17	99	TCCS322_EQUIPMENT	WTN_LINE	10001	TATS209_EQUIPMENT_EOM_CURR	WTN_LINE_NBR	10001
18	99	TCCS322_EQUIPMENT	WTN_LRN_XREF_LINE	10001	TATS209_EQUIPMENT_EOM_CURR	WTN_LINE_NBR	10001
19	99	TCCS322_EQUIPMENT	WTN_NXX	801	TCCS322_EQUIPMENT	NXX	806
20	99	TCCS322_EQUIPMENT	WTN_LRN_XREF_NXX	801	TCCS322_EQUIPMENT	NXX	806
21	99	TCCS322_EQUIPMENT	LINE	10000	TCCS322_EQUIPMENT	WTN_LINE	10001
22	99	TCCS322_EQUIPMENT	LINE	10000	TCCS322_EQUIPMENT	WTN_LRN_XREF_LINE	10001
23	99	TATN605_NPA_NXX	NXX_CD	999	TCCS322_EQUIPMENT	CUSTCODE	1000
24	99	TATN605_NPA_NXX	RATE_CENTER_VERTICAL_COORD	9064	TATN605_NPA_NXX	RATE_CENTER_HORIZONTAL_COORD	9067
25	98	TCCS322_EQUIPMENT	WTN_NPA	191	TCCS322_EQUIPMENT	NPA	194
26	98	TCCS322_EQUIPMENT	WTN_LRN_XREF_NPA	191	TCCS322_EQUIPMENT	NPA	194
27	98	TCCS322_EQUIPMENT	RATE	192	TCCS322_EQUIPMENT	REVENUE_DERIVED	199
28	97	TCTN109_TRUNK_GROUP_FORECAST	SWTCH_CLLI_A	14416	TCTN108_TRUNK_GROUP	SWITCH_CLLI_A	15129
29	97	TCTN109_TRUNK_GROUP_FORECAST	SWTCH_CLLI_Z	14812	TCTN108_TRUNK_GROUP	SWITCH_CLLI_Z	15538
30	97	TCTN109_TRUNK_GROUP_FORECAST	MODIFICATION_CODE	5217	TCTN108_TRUNK_GROUP	MODIFICATION_CODE	5446
31	96	TATS209_EQUIPMENT_EOM_CURR	ATN_LINE_NBR	10000	TCCS322_EQUIPMENT	TERMINAL_NBR	10590
32	96	TATS209_EQUIPMENT_EOM_CURR	WTN_LINE_NBR	10001	TCCS322_EQUIPMENT	TERMINAL_NBR	10590
33	96	TCCS322_EQUIPMENT	LINE	10000	TCCS322_EQUIPMENT	TERMINAL_NBR	10590
34	96	TCCS322_EQUIPMENT	WTN_LINE	10001	TCCS322_EQUIPMENT	TERMINAL_NBR	10590
35	96	TCCS322_EQUIPMENT	WTN_LRN_XREF_LINE	10001	TCCS322_EQUIPMENT	TERMINAL_NBR	10590
36	96	TCTN110_TRUNK_GROUP_USAGE	OFFICE_CLASS_CODE	30	TCTN108_TRUNK_GROUP	OFFICE_CLASS_CODE	32
37	94	TCTN109_TRUNK_GROUP_FORECAST	TRAFFIC_USE_CODE	70	TCTN108_TRUNK_GROUP	TRAFFIC_USE_CODE	78
38	93	TCTN109_TRUNK_GROUP_FORECAST	OFFICE_CLASS_CODE	28	TCTN108_TRUNK_GROUP	OFFICE_CLASS_CODE	32
39	93	TZTF055_RPTG_ORGANIZATIO_RECST	REPORTING_CUSTOMER_SEGMENT_CD	74	TZTF055_RPTG_ORGANIZATION_ASGN	REPORTING_CUSTOMER_SEGMENT_CD	84
40	92	TCTN110_TRUNK_GROUP_USAGE	TRAFFIC_USE_CODE	67	TCTN108_TRUNK_GROUP	TRAFFIC_USE_CODE	78
41	92	TZTF055_RPTG_ORGANIZATIO_RECST	LAST_UPDATE_DT	12	TZTF055_RPTG_ORGANIZATION_ASGN	LAST_UPDATE_DT	14

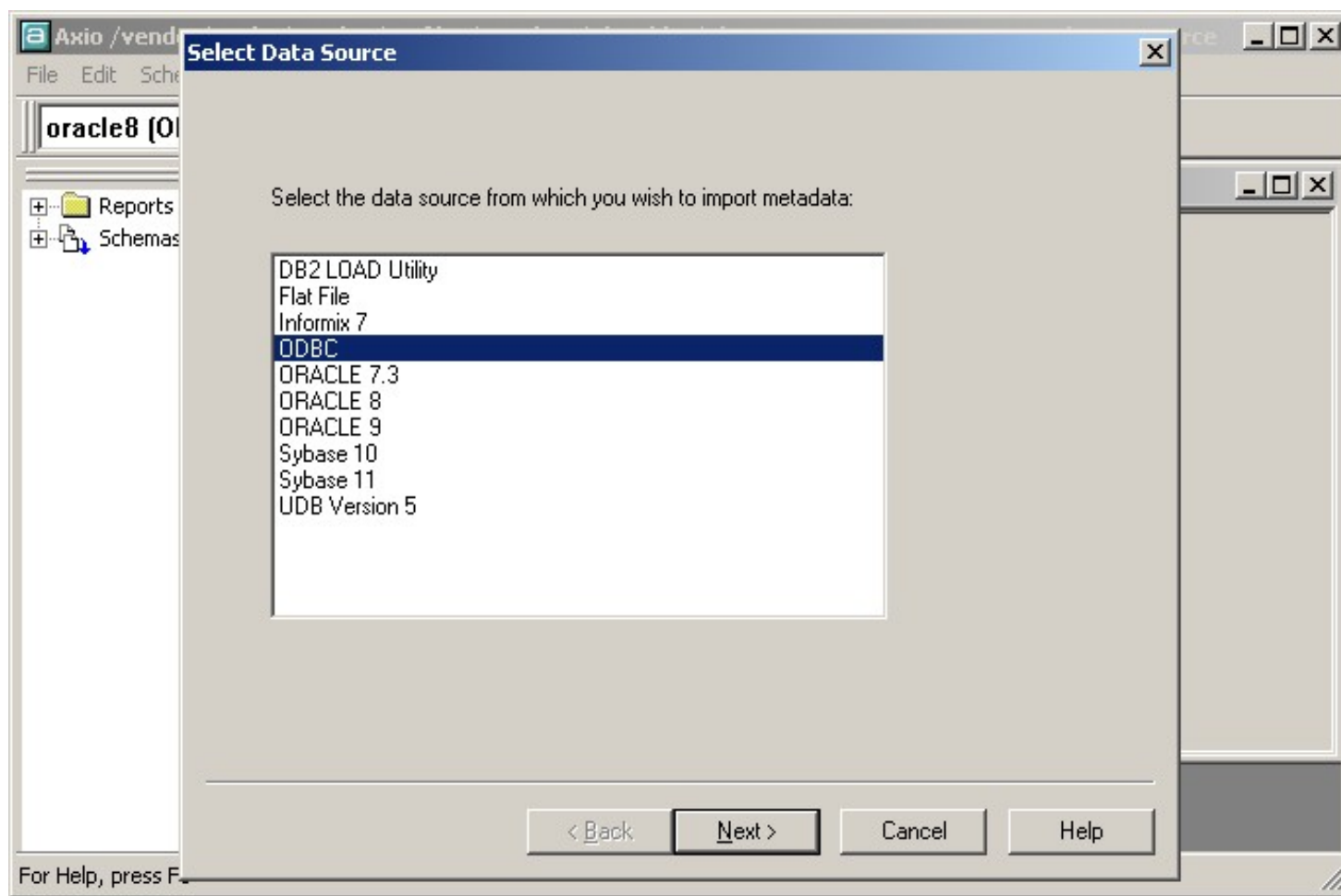
Metadata

- We update our Metadata to provide Data Explorer Analysis on each column of our EDW (Enterprise Data Warehouse).
- Our Metadata is stored in a CA Advantage Repository using MVS DB2.

Import Metadata



Select Data Source



Enter Database Access Params

The screenshot shows a dialog box titled "Enter Database Access Parameters" overlaid on a software interface. The dialog box is divided into two main sections. The top section, "Connect to ODBC Data Source as:", contains three input fields: "User Name:" with the text "EVOKEID", "Password:" with "xxxxxxx", and "Data Source:" with a dropdown menu showing "Teradata_EDW [Teradata CNEW Dr". The bottom section, "Limit Search of Tables/Views", contains two input fields: "Schema:" with "finance" and "Table Name:" which is empty. Below these fields are three radio buttons: "Include Tables", "Include Views", and "Include Tables and Views", with the last one selected. At the bottom of the dialog box are four buttons: "< Back", "Next >", "Cancel", and "Help".

Connect to ODBC Data Source as:

User Name: EVOKEID

Password: xxxxxxx

Data Source: Teradata_EDW [Teradata CNEW Dr

Limit Search of Tables/Views

Schema: finance

Table Name:

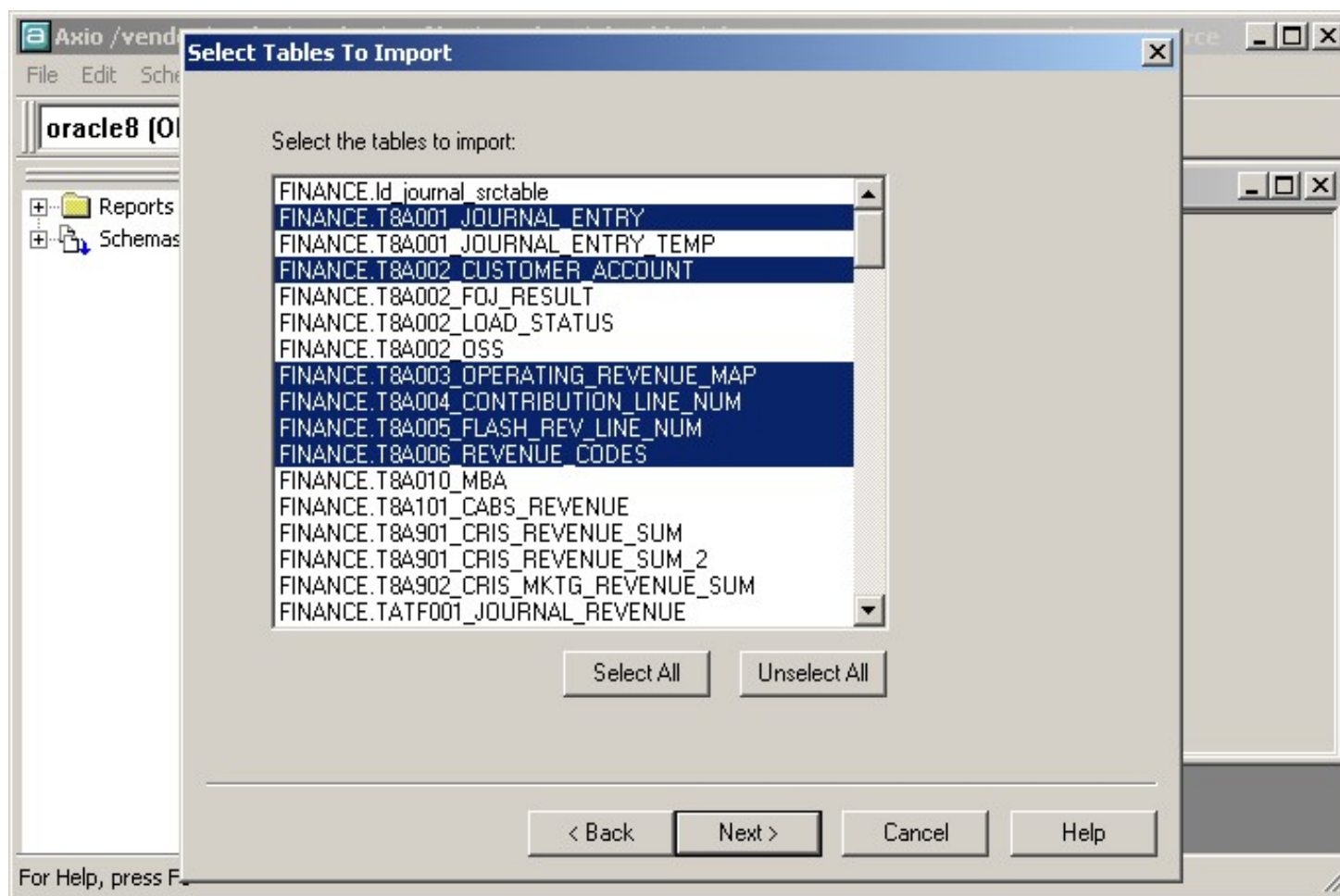
Include Tables

Include Views

Include Tables and Views

< Back Next > Cancel Help

Select Tables to Import



EDW Metadata

- Column Profile Report is parsed for upload to EDW Metadata.
- Search page for a table TATS209.

The screenshot shows the EDW Metadata Home Page in a Microsoft Internet Explorer browser window. The address bar shows the URL <http://edw.sbc.com/metadata/index.htm>. The page features a navigation menu with links for Home, What's New, Feedback, Search, Metrics, EDW Definition Update, EDW Data Naming Guidelines, and EDW Home Page. Below the navigation menu is a search form with the following fields and options:

- Select:** A dropdown menu with options: Database, Table (selected), View, and Column.
- Name:** A checked radio button.
- Contains:** An unchecked radio button.
- Search Term:** A text input field containing "tats209".
- AND:** A red text label.
- Description:** An unchecked radio button.
- Search:** A button.
- Reset:** A button.

Below the search form, there is explanatory text: "Search on various metadata objects (e.g. Database, Table, etc.). Select a Name search, Description search (or both). Name searches can be "begins with" or "contains" searches. Description searches are always "contains" searches. Only simple searches are available at this time (no Boolean operators)."

Below the explanatory text, there is a warning: "Also, be aware that Descriptions searches and some 'Name Contains' searches can take time to complete. Searching without a word (except for Databases and Stewards) is not recommended!"

At the bottom of the page, there is a footer: "Proprietary Information Not for use or disclosure outside SBC Services, Inc. and affiliates except under written permission."

EDW Metadata

- Column List for table TATS209.

EDW Metadata Home Page - Microsoft Internet Explorer provided by SBC Services

Address: <http://edw.sbc.com/metadata/index.htm>

EDW Shareable Metadata edw.sbc.com

Home What's New ^{NEW} Feedback Search Metrics EDW Definition Update EDW Data Naming Guidelines EDW Home Page

Select: Database Table View Column Name Begins Contains Description Contains

AND

Names which Begin "tats209"

1 Tables found

Database	Table	Description	Load Freq	Go To ...
PRODUCTS	TATS209_EQUIPMENT_EOM_CURR	Account level product information that supports the multi-tier and/or special pricing plans that a SBC Ameritech customer can have on their account. NOTE: This table contains both billable and in-service quantities. To avoid double counting you will need to specify which type of counts you're looking for (i.e. bill or in-service). The attribute BILLABLE_IND or IN_SERVICE_IND should be included in all queries against this TATS209 table. If the bill or in-service attribute is excluded double counting can/will occur when revenue (USOC_AMT) is involved.		Table Column View Metadata History List Rpt/Upd Excel File

Return to [EDW Home Page](#)

http://edw.sbc.com/metadata/metanvfm.cfm?SrcObject=Table&TgtObject=Column&tb_id=11548322&tb_name=TATS209_EQUIPMENT_EO Local intranet

EDW Metadata Column List

- Listing of all the columns for table TATS209.
- Pick from any column that has an **Evoke** hot button next to it.

Column Name	Length	Data Type	Precision	Description	Action
STATE_CD	30	CHAR	2	XX(2) of USOC Billable QTY for a given USOC. <i>This description was last updated on 2002/11/27 by jz1366 via erwin desc import</i>	Evoke
ADDITIONAL_LINE_IND	31	CHAR	1	XX(1) Two character state identification. Values are: IL = Ameritech - Illinois, MI = Ameritech - Michigan, IN = Ameritech - Indiana, OH = Ameritech - Ohio, WI = Ameritech - Wisconsin; <i>This description was last updated on 2002/11/27 by jz1366 via erwin desc import</i>	Evoke
WHOLESALE_RETAIL_CD	32	CHAR	1	XX(1) Indicates if this is an additional line account. Values are: Y = YES; N = NO; <i>This description was last updated on 2002/11/27 by jz1366 via erwin desc import</i>	Evoke
CLASS_OF_SERVICE_CD	33	CHAR	5	XX(5) A code that designates that the account is a wholesale account or retail account or both. Valid values are W for wholesale, R for retail, or B for both. For AIT data, W and R are the only valid values. <i>This description was last updated on 2002/11/27 by jz1366 via erwin desc import</i>	Evoke
DW_LOAD_DT	34	DATE	0	YYYY/MM/DD The type of service the customer has selected and is identified on the service order. It is used by the various departments to classify the service for their own processing. <i>This description was last updated on 2002/11/27 by jz1366 via erwin desc import</i>	Evoke
				Operational date on which this row was inserted into the EDW data warehouse table. <i>This description was last updated on 2002/11/27 by jz1366 via erwin desc import</i>	Evoke

EDW Metadata for Column

- Metadata for Database=Products, Table=TATS209 and Column=Additional_Line_Ind.
- Extract Date is a snapshot in time of that data.

EDW Metadata Home Page - Microsoft Internet Explorer provided by SBC Services

Address: <http://edw.sbc.com/metadata/index.htm>
[Back to EDW Metadata Search - metanvfm.cfm]

EDW
Shareable Metadata edw.sbc.com

[Home](#) [What's New](#) [Feedback](#) [Search](#) [Metrics](#) [EDW Definition Update](#) [EDW Data Naming Guidelines](#) [EDW Home Page](#)

[Data Quality Analysis](#) (Click [Here](#) for [Evoke Help](#))

DataBase: PRODUCTS
Table Name: TATS209_EQUIPMENT_EOM_CURR
Column Name: ADDITIONAL_LINE_IND

Evoke Analysis:

EVOKE EXTRACT DATE: 03/18/2003
DISTINCT VALUES: 2
RECORDS: 253625490
AXIO POPULATED: 3
NULLS: 384484
NULL RULE: NULL OK
DATA TYPE: CHAR(1)
MATCHES: 253241006
MINIMUM VALUES:
N
Y
MAXIMUM VALUES:
Y
N
PATTERNS: 1 OF 1
(253241006) U
MOST FREQUENT VALUES:
(251514324) N
(1726682) Y
LEAST FREQUENT VALUES:
(251514324) N
(1726682) Y

Local Intranet

Data Explorer Benefits

- **Project accurate timelines, estimate appropriate risk**
 - Identify risks and inaccuracies involved with working with source data
 - Example: >50% reduction on hours for analysis by using Data Explorer
 - Today we are able to process thousands of tables into our EDW Metadata without program modification
 - Overnight profiling resolved for 95% of projects
 - Addition of Pattern Analysis to Column Profiling Report enhanced format and business rule validation
 - Example: Data Modeler carried Column Profile Report of source data for 2 years during project development lifecycle
- **Integrated into web based metadata tool that is used by the Enterprise Data Warehouse community**
 - Fulfilled requests to see summary of live data at the column level. Data Explorer provides this avenue via the EDW Metadata web site
 - Immediate value to everyone with quicker and accurate response time compared to tools used previously
 - 1997 purchase of DBStar (currently Informatica Data Explorer) revealed ROI within 8 months of purchase through labor hours and accuracy

Summary

- Data Explorer can provide valuable time saving analysis that would normally take much longer when done manually or with a combination of other tools.
- Any questions?