

Ataccama DQ Analyzer



Ataccama DQ Analyzer is a unique and truly complete tool for data quality analysis. DQ Analyzer can rapidly reveal basic quality attributes of your data and easily proceed to advanced features for deeper and more accurate analysis. As its core, it uses Ataccama Data Quality Center, an advanced and proven tool for Data Quality Management which is used by a number of multi-national organizations for solving data quality problems.

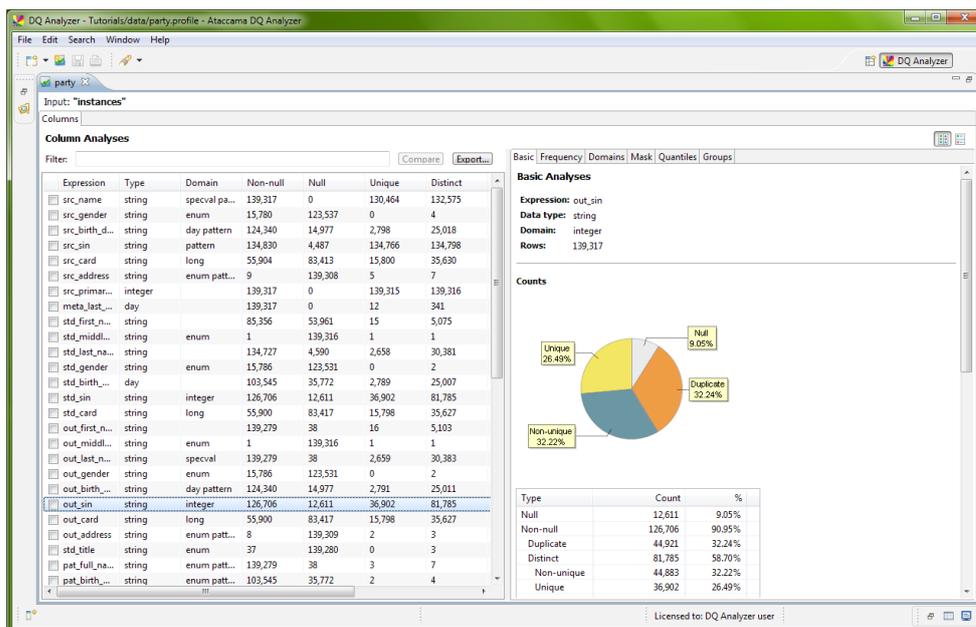
Problems with Data

Problems with data quality are pervasive and cause organizations numerous problems and incur many costs. They can result in correspondence being returned due to poor quality addresses, duplicate records in databases due to incorrect registration of names of clients or companies or the inability to contact a client due to errors in telephone numbers. These problems are common; however, organizations are frequently not able to accurately describe and quantify these errors.

financial resources and promote their implementation. In addition, data quality projects are very costly (in particular, the purchase of data cleansing tools), without the ability to foresee the benefits of such a project and the tools required for cleaning the data. In all these cases it is possible to perform a Data Quality Assessment, which enables very precise and rapid detection and quantification of errors in the client data and at the same time specifies the benefits of data quality improvement technology like Ataccama Data Quality Center (DQC) in the analyzed data.

Business Benefits

- Provides an objective picture of data quality - instantly.
- Gives an estimate of the benefits of data cleansing projects.
- Facilitates and accelerates data quality initiatives.
- Prevents making risky decisions based on inaccurate data.
- Completely free for commercial use.



Technical Advantages

- Basic analysis on data in a database or CSV file performed in a few minutes.
- DQ Analyzer is built on the Eclipse platform and its optimized user interface is designed to be intuitive and navigable.
- Simple problem solving. Plans created in DQ Analyzer can be used directly in Data Quality Center and Master Data Center without any modification.
- Large volumes of data. DQ Analyzer can handle tables with hundreds of millions of records and more.

Easy Upgrade to DQC

With DQ Analyzer, you can easily upgrade to Ataccama Data Quality Center which has all the functionality you need to cleanse, standardize, de-duplicate, code, fix and otherwise enrich your data - and since it uses the same platform, language, and advanced user interfaces, it will feel like home to you. Also, if you are into SOA, note that every Data Quality Center process can be published as a service.

Consequently, it's not possible to specify not only the effects of poor data, but also the cause, making it impossible to take effective measures to improve the quality of the data in corporate information systems. Also, the data quality projects which should solve the problem with data in organizations encounter difficulties while quantifying the impact of insufficient data quality and interpreting the financial effects. For these reasons, it is often difficult to get these projects the required

This makes it possible to quantify the financial impact of insufficient data quality and make a strong case for initiating a data quality project in the organization. At the same time it facilitates the subsequent project to clean the data since the data analysis uses the same procedures as for data cleansing. This accelerates the speed of the project and reduces its cost.

DQ Analyzer Algorithms



Text File Reader

Reads data from a delimited text file, such as CSV.



Condition

Splits the data stream based on a condition.



Column Assigner

A powerful and robust business rules engine that allows the use of complex expressions.



Fixed Width File Reader

Reads data from a text file with fixed width columns.



Union Same

Combines data streams that have the same format.



Trash

Ends the data flow.



JDBC Reader

Reads data directly from a database such as Oracle, MS SQL, Sybase, etc.



Regex Matching

Allows flexible parsing and matching using regular expressions.



Profiling

Flexible, fully configurable data profiler.

The screenshot shows the DQ Analyzer interface with a workflow diagram on the left and a data table on the right. The workflow includes steps like 'Text File Reader', 'Match emails in src_email', 'Extract valid emails', 'Match emails in src_address', 'Extract valid emails 2', and 'Join the flows'. The data table shows columns for 'src_primary_key', 'src_name', 'src_email', and 'src_address' with corresponding data rows.

src_primary_key	src_name	src_email	src_address	
1	90	Tomasas Tingvald	SYSTEM MIGRATION	mailto:tingvald-tomasas@smo.lvivis.com
2	94	Beatrice Kornhauser	SYSTEM MIGRATION	<kornhauserbeatrice@fastweb.com>
3	97	Vogli Cristobal	SYSTEM MIGRATION	vogli.cristobal@ecoustics.com
4	98	Debby Greenen	SYSTEM MIGRATION	greenen@greenmuseum.org
5	101	Malett Jacqueline	SYSTEM MIGRATION	jacquemalett@formula1.com
6	108	Starko Cand	SYSTEM MIGRATION	<starko.cand@edmunds.com>
7	110	Waszkiewicz Charise	SYSTEM MIGRATION	<waszkiewicz.charise@usset.com>
8	111	Louwie Ardith	SYSTEM MIGRATION	<lowieardith@multinationalmonitor.org>
9	112	Carline McPherson	SYSTEM MIGRATION	malto:mcpherson.carline@adventurelandthemapark.co
10	132	Toepke Erika	SYSTEM MIGRATION	<erika@break.com>
11	133	Lina Turdo	SYSTEM MIGRATION	<lina.turdo@cookingpastory.com>
12	136	Pollom Lyman	SYSTEM MIGRATION	lymanpollom@askasia.org
13	146	Wimbrow Mee	SYSTEM MIGRATION	<wimbrow-mee@freep.com>
14	147	Nick Teets	SYSTEM MIGRATION	<nick.teets@house.gov>
15	155	Aleta Alicandro	SYSTEM MIGRATION	<alicaandro-aleta@soncommerce.com>
16	158	Yong Mokkerow	SYSTEM MIGRATION	mokkerow_yong@medicare.gov
17	161	Eiland Genaro	SYSTEM MIGRATION	<eilandgenaro@bdsnews.com>
18	168	Osborn Kendal	SYSTEM MIGRATION	<kendalosborn@ford.com>
19	172	Dibenedict Eileen	SYSTEM MIGRATION	<eileen-dibenedict@pedigest.com>
20	176	Feduccia Milagros	SYSTEM MIGRATION	milagros-feduccia@tech.yahoo.com
21	183	Rechart Carlotta	SYSTEM MIGRATION	carlotta-rechart@2morrow.com
22	191	Anaya He	SYSTEM MIGRATION	<anaya.he@3.org>
23	195	Corda Pansy	SYSTEM MIGRATION	pansycorda@bigjailbook.com
24	203	Kinnbrew Niesha	SYSTEM MIGRATION	niesha.kinnbrew@expedia.com
25	214	Wilda Moniak	SYSTEM MIGRATION	moniak-wilda@finance-commerce.com
26	233	Knick Kathryn	SYSTEM MIGRATION	<knick.kathryn@ft.com>
27	233	Una World	SYSTEM MIGRATION	unaworld@wilkinvest.com
28	242	Kira Beagleley	SYSTEM MIGRATION	beagleley-kira@yanuar.com
29	244	Joette Smolic	SYSTEM MIGRATION	<smolic-joette@netstodirect.com>
30	257	Trevor Purdue	SYSTEM MIGRATION	mailto:trevor.purdue@medicare.gov
31	273	Fyffe Lulu	SYSTEM MIGRATION	<lulu@latimes.com>
32	276	Malchom Madine	SYSTEM MIGRATION	malchommadine@smenonrev.com

Supported Analyses

Statistical Analysis

An overview of summary values, such as extremes, distribution and frequency analysis.

Domain Analysis

A configurable analysis of data types.

Mask and Group Analysis

An overview of value formats, groups and dimensions.

Business Rules

An analysis of the results of user-defined business rules.

Foreign Key and Dependency Analyses

An inside look into complex connections in the data.

Drill Through

The option to display individual records that correspond to aggregated results.

The screenshot shows the DQ Analyzer interface with a 'Foreign Keys' analysis diagram. The diagram displays a Venn diagram and a bar chart comparing 'person' and 'address' tables. The Venn diagram shows 48438 (52.84%) matched records, 43227 (47.16%) not matched records, and 71808 (64.43%) matched records. The bar chart shows 48438 (52.84%) matched records, 43227 (47.16%) not matched records, and 37228 (33.40%) null records.

Ataccama Corporation is an international software company that prides itself in delivering cutting edge technology used for Data Quality and Master Data Management. Leading financial, commercial and government organizations choose Ataccama technologies for its proven ability to execute and deliver real business benefits.



ataccama

Ataccama Corporation

One Stamford Plaza
263 Tresser Blvd., 9th Floor
Stamford, CT 06901, USA
Tel: +1 (203) 564-1488
www.ataccama.com