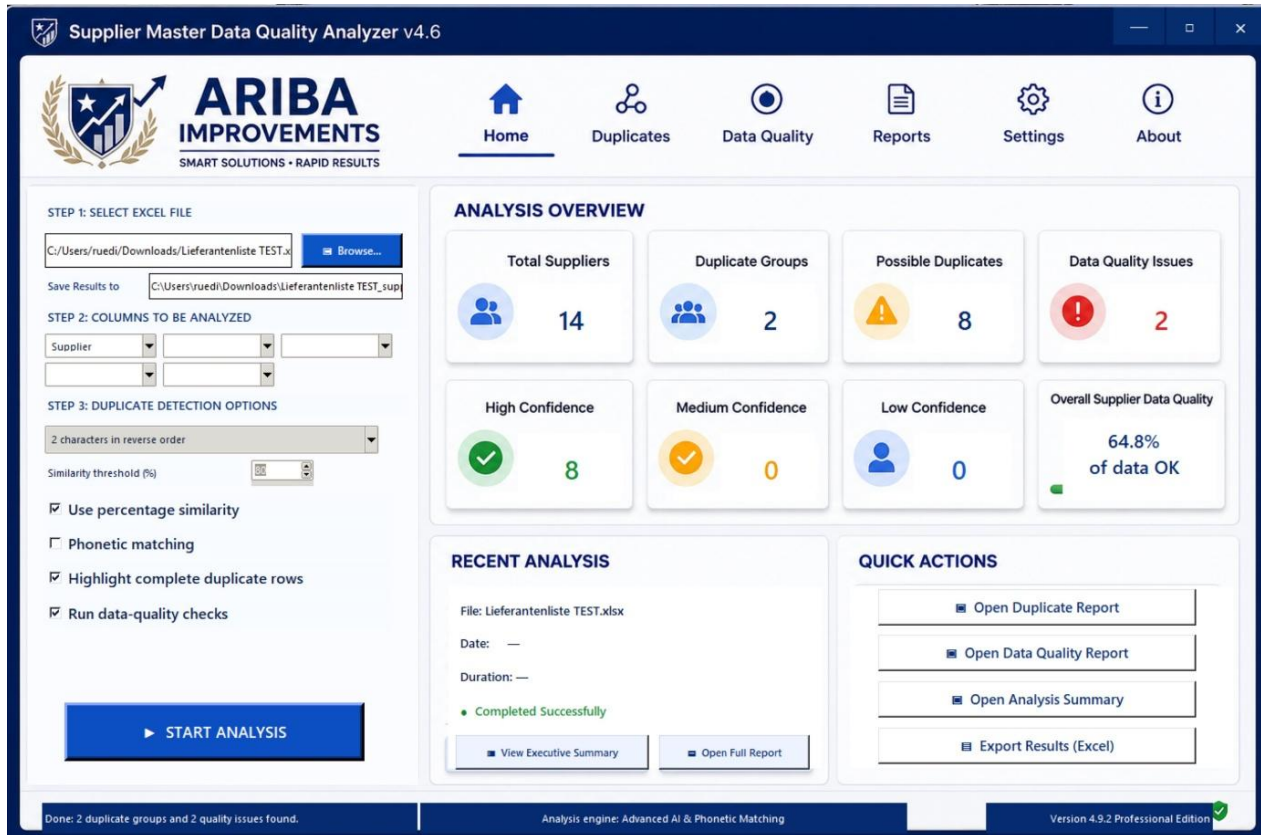


SUPPLIER DATA QUALITY

Supplier Master Data Analyzer

Find duplicates. Detect missing data. Improve supplier master data quality.



Duplicates

Exact + fuzzy matching

Missing data

Completeness checks

Excel output

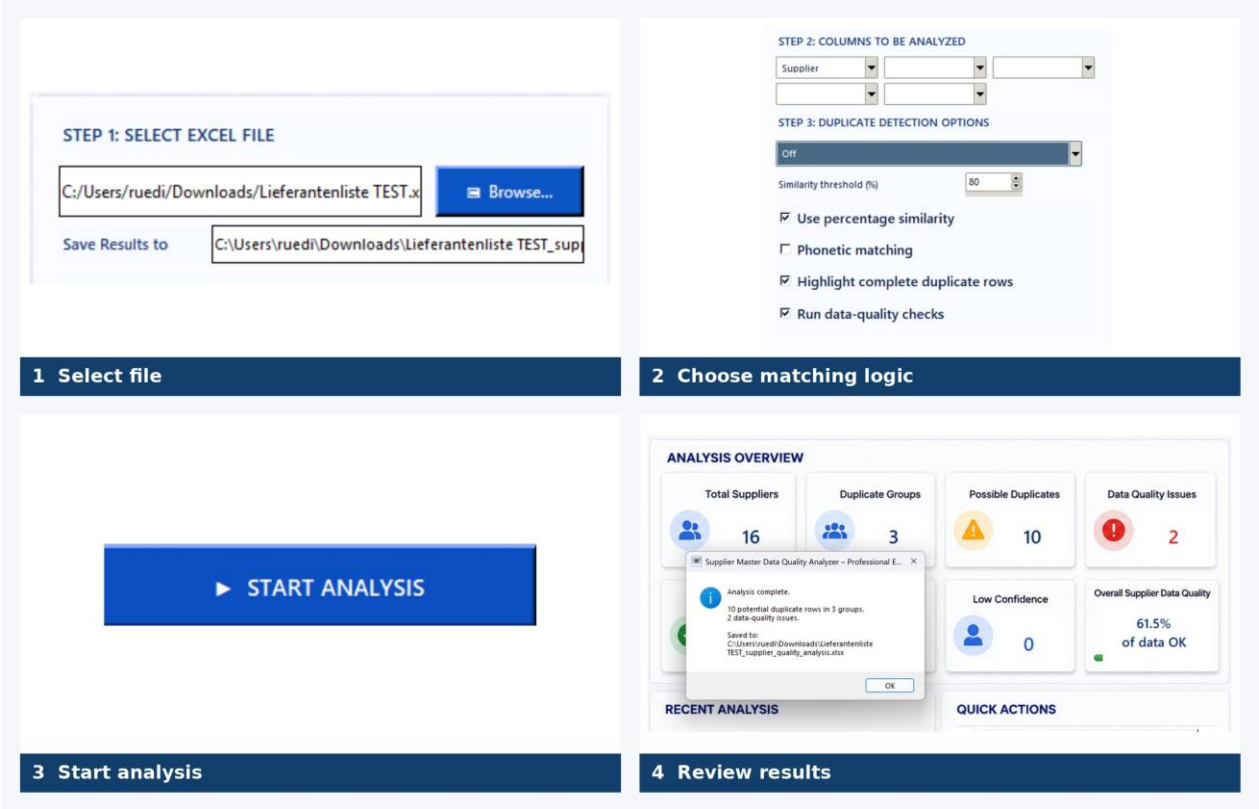
Actionable reports

One file. One analysis. See where supplier data needs attention before it becomes a project problem.

- Fast analysis of Excel-based supplier lists
- Multiple duplicate-detection methods
- Structured management and detail reports
- Runs locally on your PC

From Excel file to results in minutes

A simple workflow: select the file, choose your matching logic, run the analysis and review the findings.



1 SELECT Choose the source file and where the result should be saved.

2 CONFIGURE Select the supplier columns and the duplicate-detection method.

3 ANALYZE Start the SMDQA analysis. Duplicate patterns and data-quality issues are evaluated.

4 REVIEW Open the full Excel report with grouped candidates, missing-data issues and management summaries.

80% similarity: catch the near matches

Find suppliers that are similar enough to deserve review, even when the names are not identical.

STEP 2: COLUMNS TO BE ANALYZED

Supplier

STEP 3: DUPLICATE DETECTION OPTIONS

Off

Similarity threshold (%)

80

☒ Use percentage similarity

☐ Phonetic matching

☒ Highlight complete duplicate rows

☒ Run data-quality checks

Example configuration: Supplier column + 80% similarity threshold

ANALYSIS OVERVIEW

Total Suppliers

16

Duplicate Groups

3

Possible Duplicates

10

Data Quality Issues

2

Low Confidence

0

Overall Supplier Data Quality

61.5% of data OK

RECENT ANALYSIS

QUICK ACTIONS

Supplier Master Data Quality Analyzer – Professional E...

Analysis complete.

10 potential duplicate rows in 3 groups.
2 data-quality issues.

Saved to:
C:\Users\vuedi\Downloads\Lieferantenliste
TEST_supplier_quality_analysis.xlsx

OK

The analysis overview immediately shows duplicate groups, possible duplicates and quality issues.

What you get: Potential duplicate rows are grouped for review. The same group can be followed through the Excel output.

80%	3	10
Similarity threshold	Duplicate groups	Potential rows

Exact matches first

Start with the easiest wins: identical or 100%-similar supplier records.

STEP 2: COLUMNS TO BE ANALYZED

Supplier

STEP 3: DUPLICATE DETECTION OPTIONS

Off

Similarity threshold (%)

100

☒ Use percentage similarity

☐ Phonetic matching

☒ Highlight complete duplicate rows

☒ Run data-quality checks

100% similarity

	A	B	C	D	E	F	G	H
	Supplier	Street	No	City	Phone	Duplicate Group	Duplicate Confidence	Possible Duplicate Of
1								
2	Krone KG	Regenstrasse		Passau	999123			
3	Krones AG	Peterstrasse		1 Neutraubli	110333			
4	Saarstahl AG	Röchlingstras		1 Saarbrücke	456789			
5	Sea Men Ltd	Main Street		London	123544			
6	Siemens Nixdorf AG	Hauptstrasse	21	München	111333			
7	Siemen AG	Hauptstrasse		München	111233			
8	Siemens Nixdorf AG	Hauptstrasse	21	München	111333	1	99.5%	Siemens Nixdorf AG
9	Siemens Nixdorf AG	Hauptstrasse	33	München	223544	1	99.5%	Siemens Nixdorf AG
10	Siemens AG	Orthografieweg		Hamburg	222444			
11	Siemens Ltd.	Siemensstrasse		Regensburg	987654			
12	Stahlsaar AG	Völklinger Weg		Homburg	888321			
13	Weimens GmbH	Hauptstrasse	21	München	122444			
14	BASF							
15	SFBA							
16	Werner GmbH	Waldstrasse	3	Hamburg	30456	2	99.5%	Werner's GmbH
17	Werner's GmbH	Feldstrasse	4	Hamburg	30477	2	99.5%	Werner GmbH
18	Identical Ltd	Main Street		1 Manchester		3	99.5%	Identical Ltd
19	Identical Ltd	Main Street		1 Manchester		3	99.5%	Identical Ltd
20								
21								

Exact matches grouped in the original list

Why start here? Exact duplicates are usually the fastest records to validate and clean. Removing them improves reporting and reduces avoidable master-data noise.

Typical use

- Identify repeated supplier records with identical names.
- Group exact matches for quick validation.
- Create a clean baseline before reviewing less obvious duplicate candidates.

Catch typing errors and letter swaps

Small differences can still represent the same supplier record.

STEP 2: COLUMNS TO BE ANALYZED

Supplier

STEP 3: DUPLICATE DETECTION OPTIONS

Maximum 2 characters difference

Similarity threshold (%) 100

☐ Use percentage similarity

☐ Phonetic matching

☒ Highlight complete duplicate rows

☒ Run data-quality checks

Max. 2 characters difference

STEP 2: COLUMNS TO BE ANALYZED

Supplier

STEP 3: DUPLICATE DETECTION OPTIONS

2 characters in reverse order

Similarity threshold (%) 100

☐ Use percentage similarity

☐ Phonetic matching

☒ Highlight complete duplicate rows

☒ Run data-quality checks

2 characters reversed

Group	Item 1	Supplier 1	Item 2	Supplier 2	Confidence	Similarity	Character Difference
1	2	Siemens AG	2	Siemens AG	100.0%	100.0%	0
2	5	Siemens AG	7	Siemens AG	92.0%	76.0%	2
3	7	Siemens AG	11	Siemens AG	92.0%	87.1%	2
4	11	Siemens AG	13	Siemens AG	92.0%	83.0%	2
5	9	Siemens AG	9	Siemens AG	99.0%	100.0%	0
6	9	Siemens AG	9	Siemens AG	99.0%	95.2%	2
7	9	Siemens AG	9	Siemens AG	99.0%	95.2%	2
8	19	Siemens AG	17	Siemens AG	99.0%	100.0%	0
9	19	Siemens AG	17	Siemens AG	99.0%	100.0%	0
10	19	Siemens AG	17	Siemens AG	99.0%	100.0%	0
11	19	Siemens AG	17	Siemens AG	99.0%	100.0%	0

Result: grouped candidates

Group	Item 1	Supplier 1	Item 2	Supplier 2	Confidence	Similarity	Character Difference
1	9	Siemens AG	9	Siemens AG	99.0%	95.2%	2
2	9	Siemens AG	9	Siemens AG	99.0%	95.2%	2

Result: transposed characters

Maximum two characters difference: Useful where the supplier name differs by one or two characters only.

Two characters reversed: Useful for transposed letters - for example, “Schneider” versus “Scheidner”.

Why it matters

These are exactly the kinds of errors that are easy to miss in a long supplier list but can create duplicate master records.

Phonetic matching

Find supplier names that sound alike but are spelled differently.

STEP 2: COLUMNS TO BE ANALYZED

Supplier

STEP 3: DUPLICATE DETECTION OPTIONS

Off

Similarity threshold (%)

☐ Use percentage similarity

☒ Phonetic matching

☒ Highlight complete duplicate rows

☒ Run data-quality checks

Phonetic matching settings

Group	Row 1	Supplier 1	Row 2	Supplier 2	Confidence	Similarity	Character Difference	Reason
1	5	Sea Men Ltd	7	Siemen AG	92.0%	76.9%	2	Only two characters differ; Names have the same phonetic pattern
2	8	Siemens Nixdorf AG	9	Siemens Nixdorf AG	99.5%	100.0%	0	Similar normalized names (100.0%); Names have the same phonetic pattern
3	2	6 Siemens Nixdorf AG	8	Siemens Nixdorf AG	96.0%	93.3%	2	Adjacent letters appear transposed; Names have the same phonetic pattern
4	2	6 Siemens Nixdorf AG	9	Siemens Nixdorf AG	96.0%	93.3%	2	Adjacent letters appear transposed; Names have the same phonetic pattern
5	3	18 Identical Ltd	19	Identical Ltd	99.5%	100.0%	0	Similar normalized names (100.0%); Names have the same phonetic pattern

Detected phonetic candidates

Important: Phonetic matching works together with a similarity threshold to keep the result set meaningful.

Useful when supplier data contains:

- regional spelling variants
- names entered by different users
- multilingual supplier records
- minor spelling differences that sound alike

Missing data is a quality issue too

The SMDQA does more than detect duplicates. It also exposes incomplete supplier records.

Supplier	Email	VAT ID	Country	Result
Alpha Components	info@alpha.de	DE123456789	Germany	Complete
Beta Industrial AG	—	CH234567890	Switzerland	Missing email
Gamma Supply Ltd	sales@gamma.com	—	—	Missing VAT ID + country

	A	B	C	D
1	Row	Supplier	Field	Issue
2		14 BASF	Supplier	No recognizable legal form in supplier name
3		15 SFBA	Supplier	No recognizable legal form in supplier name
4				
5				

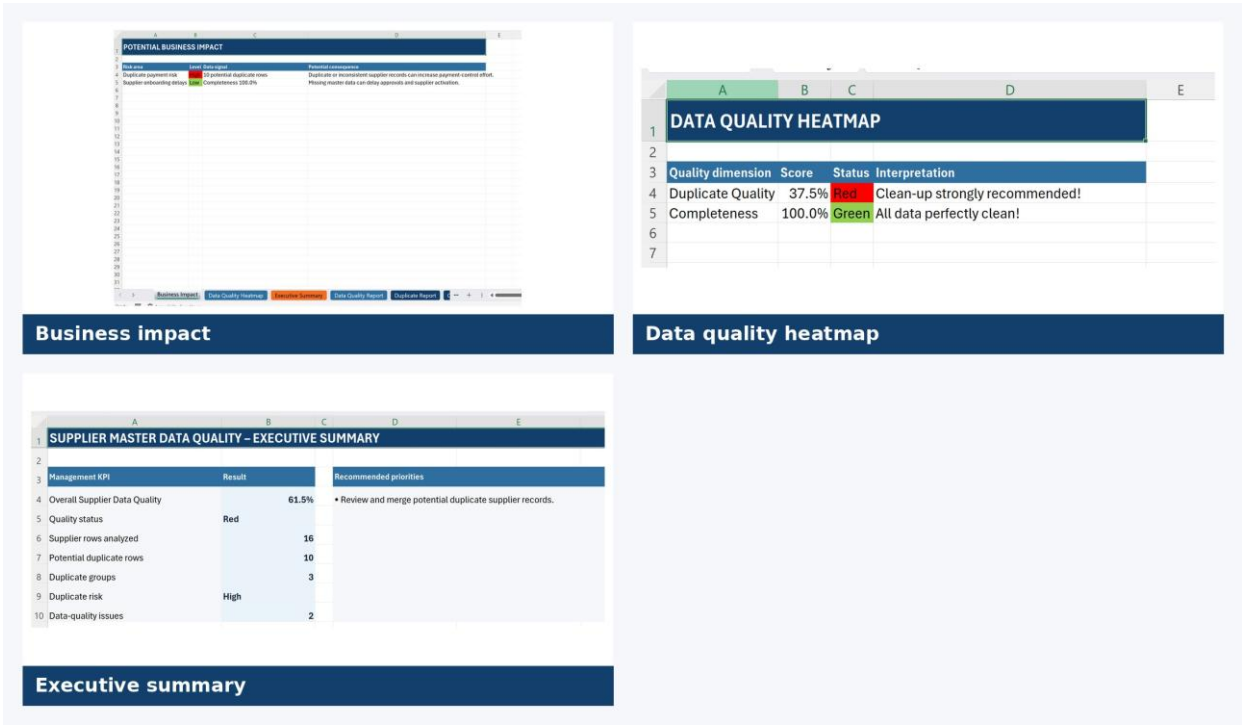
Example from the Data Quality Report: incomplete or inconsistent supplier information is called out explicitly.

Business value: Missing email, tax, country or address information can delay onboarding, approvals and supplier communication.

- Missing data is visible at record level.
- Issues can be reviewed in the detailed report.
- Completeness contributes to the overall data-quality assessment.

From raw supplier data to management insight

The analysis does not stop at a list of matches. It summarizes what the findings mean.



Executive Summary: A compact view of overall supplier data quality, duplicate risk, affected rows and recommended priorities.

Data Quality Heatmap: Shows where quality problems are concentrated and where clean-up is most urgent.

Business Impact: Connects findings with practical consequences such as duplicate-payment effort or onboarding delays.

See every finding in context

Review duplicate groups, matching reasons and data-quality issues directly in Excel.

Group	Row 1	Supplier 1	Row 2	Supplier 2	Confidence	Similarity	Character Difference	Reason
1	2	Krone KG	3	Krones AG	95.5%	95.5%		1 Only one character differs; One normalized name contains the other; Name similarity is 95.5% (threshold 80%)
2	8	Siemens Nixdorf AG	9	Siemens Nixdorf AG	99.5%	100.0%		0 Name similarity is 100.0% (threshold 80%)
3	6	Seimens Nixdorf AG	8	Siemens Nixdorf AG	96.0%	93.3%		2 Adjacent letters appear transposed; Names have the same phonetic pattern; Name similarity is 93.3% (threshold 80%)
4	6	Seimens Nixdorf AG	9	Siemens Nixdorf AG	96.0%	93.3%		2 Adjacent letters appear transposed; Names have the same phonetic pattern; Name similarity is 93.3% (threshold 80%)
5	7	Siemen AG	8	Siemens Nixdorf AG	94.0%	90.0%		8 One normalized name contains the other; Name similarity is 90.0% (threshold 80%)
6	7	Siemen AG	9	Siemens Nixdorf AG	94.0%	90.0%		8 One normalized name contains the other; Name similarity is 90.0% (threshold 80%)
7	11	Siemens Ltd.	11	Siemens Ltd.	91.0%	87.1%		2 Only two characters differ; Name similarity is 87.1% (threshold 80%)
8	2	11 Siemens Ltd.	13	Weimens GmbH	91.0%	83.9%		2 Only two characters differ; Name similarity is 83.9% (threshold 80%)
9	3	16 Werner GmbH	17	Werner's GmbH	99.5%	100.0%		1 Only one character differs; One normalized name contains the other; Name similarity is 100.0% (threshold 80%)

Duplicate report

Supplier	Street	City	Phone	Duplicate Group	Duplicate Confidence	Possible Duplicate Of	Duplicate Reason	Data Quality Issues
2	Krone KG	Regenstrasse	Passau	999123	1	95.5%	Krones AG	Only one character differs; One normalized name contains the other; Name similarity is 95.5% (threshold 80%)
3	Krones AG	Peterstrasse	Neutraubling	110333	1	95.5%	Krone KG	Only one character differs; One normalized name contains the other; Name similarity is 95.5% (threshold 80%)
4	Saarstahl AG	Rödingstrasse	Saarbrücken	456789				
5	Sea Main Ltd	Main Street	London	123544				
6	Siemens Nixdorf AG	Hauptstrasse	21 München	111333	2	96.0%	Siemens Nixdorf AG	Adjacent letters appear transposed; Names have the same phonetic pattern; Name similarity is 93.3% (threshold 80%)
7	Siemen AG	Hauptstrasse	München	111233	2	94.0%	Siemens Nixdorf AG	One normalized name contains the other; Name similarity is 90.0% (threshold 80%)
8	Siemens Nixdorf AG	Hauptstrasse	21 München	111333	2	99.5%	Siemens Nixdorf AG	Name similarity is 100.0% (threshold 80%)
9	Siemens Nixdorf AG	Hauptstrasse	33 München	223544	2	99.5%	Siemens Nixdorf AG	Name similarity is 100.0% (threshold 80%)
10	Siemens AG	Ordnungsleweg	Hamburg	223444				
11	Siemens Ltd.	Siemstrasse	Regensburg	987654	2	91.0%	Siemen AG	Only two characters differ; Name similarity is 87.1% (threshold 80%)
12	Stahlwerk AG	Völklinger Weg	Homburg	888321				
13	Weimens GmbH	Hauptstrasse	21 München	122444	2	91.0%	Siemens Ltd.	Only two characters differ; Name similarity is 83.9% (threshold 80%)
14	BASF							No recognizable legal form in supplier name
15	OFBA							No recognizable legal form in supplier name
16	Werner GmbH	Waldstrasse	3 Hamburg	30456	3	99.5%	Werner's GmbH	Only one character differs; One normalized name contains the other; Name similarity is 100.0% (threshold 80%)
17	Werner's GmbH	Feldstrasse	4 Hamburg	30477	3	99.5%	Werner GmbH	Only one character differs; One normalized name contains the other; Name similarity is 100.0% (threshold 80%)

Original list

Duplicate Report: Supplier pairs are grouped and color-coded so likely duplicates can be reviewed quickly.

Original List: Every supplier stays in the context of the original records, enriched with duplicate group, reason and quality information.

Designed for follow-up, not just diagnosis.

The Excel output gives procurement or master-data teams a practical basis for validation, cleansing and next actions.

See what is hiding in your supplier master

A practical first step before an SAP Ariba implementation, migration, supplier onboarding initiative or data-quality program.

Find near and exact duplicates	Detect missing supplier data
Similarity, character and phonetic matching	Completeness issues surfaced clearly
Understand the impact	Keep data local
Executive summary, heatmap and business impact	Analysis runs on your own PC

DOWNLOAD THE SUPPLIER MASTER DATA ANALYZER

Runs locally on your PC • Structured Excel output • Supplier data is not uploaded