From Messy to Mission-Ready: Mastering Data Migration (Tools, Methods & Governance)

Mathtech Best Practices

May 2026

MATHTECH

Messy to Mission-Ready

We are going to Master Data Migration and Data Governance through Data Modernization

- Data Migration & Data Modernization
- The Migration Process
- Tools
- Data Governance

Cleaning Out the Garage Before You Move

You don’t move everything blindly…

- Sort what matters
- Discard what’s broken or too old
- Label what’s important
- Clean up what’s dirty

Modernizing DMV Data Is The Same

Why Data Migration Is Difficult & High Risk

20 to 40 Years of Legacy Data with a few likely complications…

- Multiple sources and database
- Business rules embedded in code
- Duplicate and conflicting records
- Undocumented fields and interfaces
- Public safety and regulatory implications

Data Migration vs. Data Modernization

Successfully migrating data from a legacy system to a modern system is so much more than "CUT & PASTE". There are many governance decision along the way…

Data Migration

Minimally…

- **Extract** – Pull data from the legacy system
- **Transform** – Reformat to fit the new system
- **Load** – Upload the data into the new system

Data Modernization

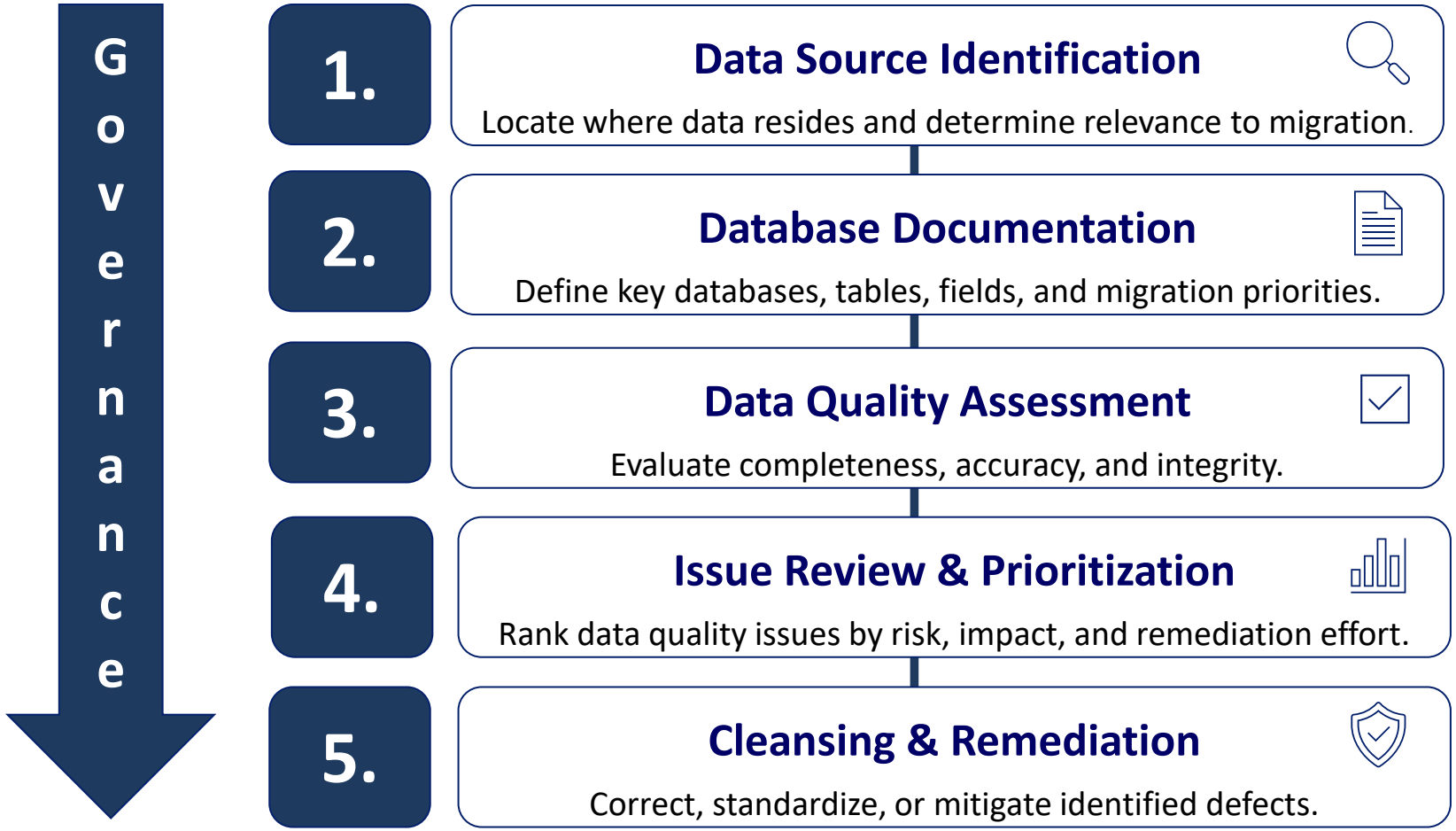
Holistically…

- Identify
- Document
- Assess
- Govern
- Cleanse
- Validate
- Improve
- Migrate

Data Modernization Process

These five steps are critical to a successful cleanup, modernization, and migration. They need to be done in order and the earlier steps are mandatory.

Data Governance is necessary to follow this path successfully and make many critical decisions.



1. Data Source Identification

Before assessing quality or remediating issues, teams need a complete inventory of sources and a shared understanding of what the data means and what matters most.

Places to look include:

- Core systems and online portals
- Reporting and archive systems
- Vendor-hosted applications
- Interfaces and replicated datasets

Step 1. Data Source Identification

Purpose: Locate where data resides and confirm what's in scope for migration.

- **Inventory systems, databases, and source feeds that support business operations**
- **Identify data owners and key integrations/interfaces**
- **Confirm which data is in scope for conversion, reporting, and downstream use**

Deliverable:

Data source inventory + migration scope

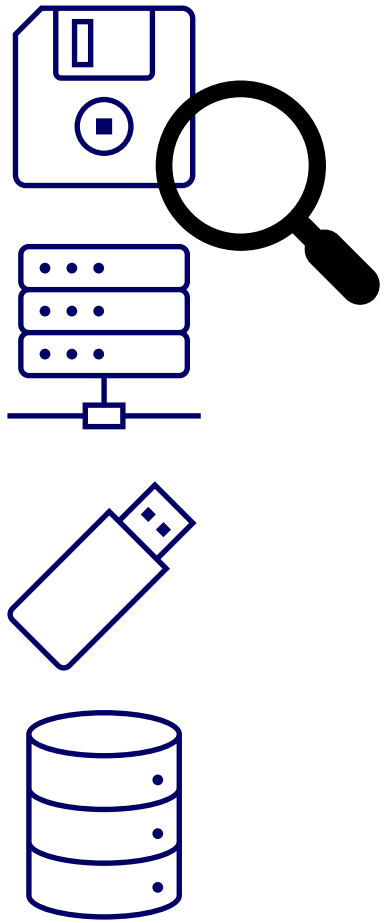
1. Data Source Identification… Where to Look

System of Record & Table of Record

- When systems disagree, who wins?
  - What is authoritative?
  - Who decides conflicts?
  - May be an IT decision… May be a Governance decision

Unofficial and Rogue Systems

- Where else is important data hiding?
  - Operational Excel files
  - Access databases
  - Legacy extracts
  - Department-built utilities



2. Database Documentation

Every database must be fully documented

- Identify the tables in each database
- Review and document each table's purpose
- Categorize the business area of each table
- Categorize the type of table
- Categorize likelihood of migration

How to Categorize?

Business Area
- AAMVA & Interfaces
- Administrative
- Customer
- Driver Lic. Issuance
- Driving Records
- Finance
- Vehicles

Table Type
- Configuration
- Lookup
- Transactional
- Performance

Step 2. Database Documentation

Purpose: Define what the data is, how it's structured, and what matters most.

- **Document key databases, tables, fields, and business definitions**
- **Identify critical data elements and data quality rules/constraints**
- **Prioritize objects for migration (high-risk / high-volume / high-value)**

Deliverable:
Data dictionary + prioritized migration dataset

3. Data Quality Assessment

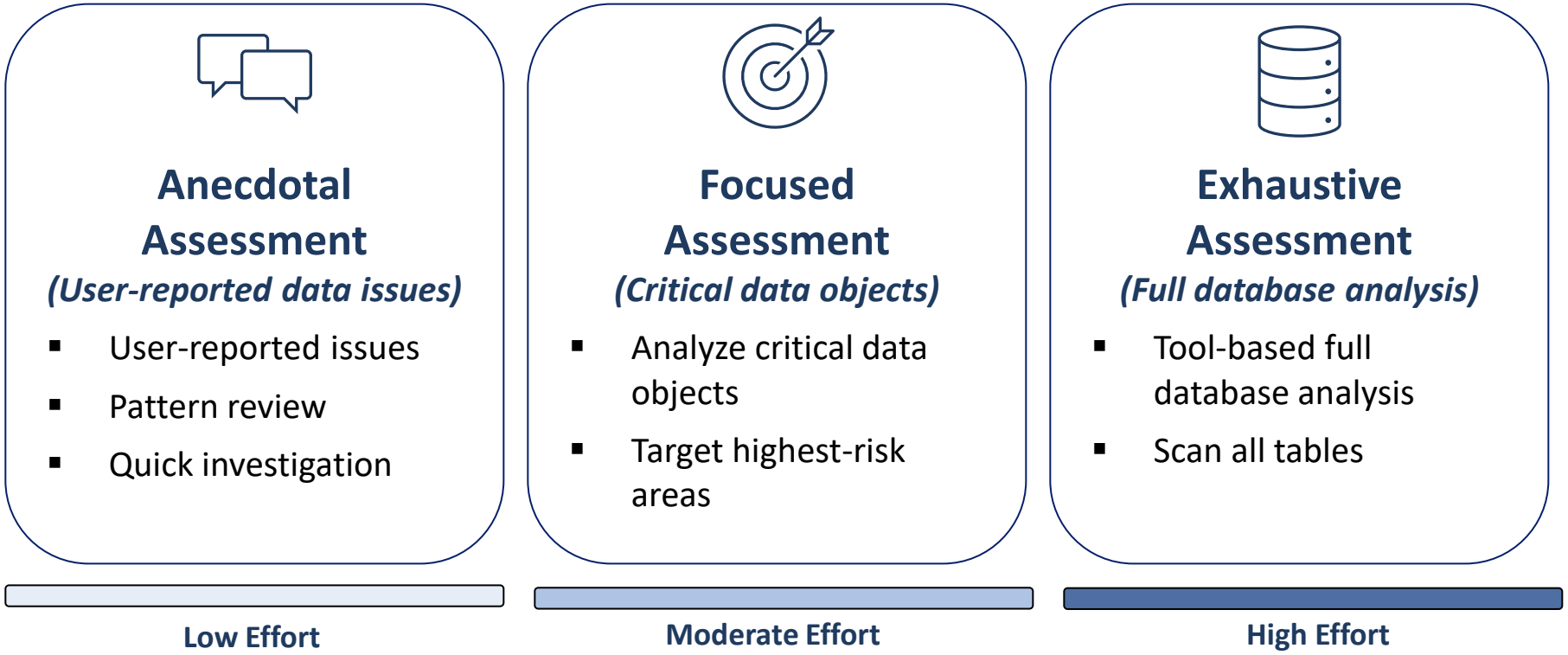
Once the databases are documented, data quality can be assessed.

Three primary strategies can be used to identify data quality issues, each with varying levels of effort.

Warning!

Data Quality Assessments and Remediation can go on forever…

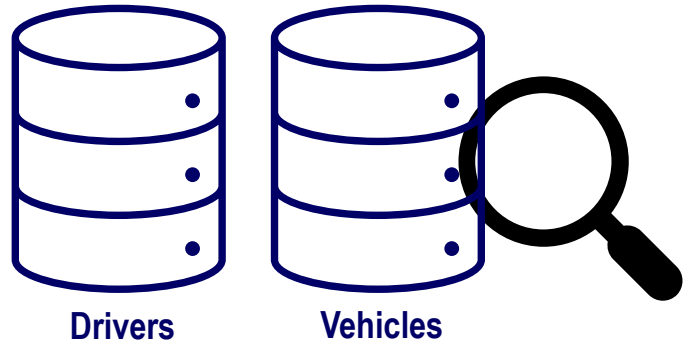
Start early, know when to focus and when to be done.



3. Data Quality Assessment… Anecdotal Assessments

Where to start…?

- Survey frontline staff
- Identify recurring issues
- Validate patterns through analysis



Your team knows where the big data quality issues are…

- **Duplicate Driver Records** caused by third parties – Driving Citations, County Data Entry Errors…
- **Duplicate Vehicle Owner Records** caused by lack of a unique key or business rules…
- **Misspelled Vehicle Manufacturer Names** caused by free form text fields…
- **Duplicate Insurance Company and Bank Records** caused multiple branch locations…
- **Inaccurate Address Information** caused by free form text fields and limited validation…
- **Unclear Lienholder Information** caused by bank mergers…

3. Data Quality Assessment… Focused Assessment

There are key data objects that are critical to successful operations.

The following are 12 high priority items that should be the focus of a detailed data quality assessment.

Individual Customer Record
- Customer
- Demographic Data
- Address
- Contact Information
- Identity Documents
- Photo & Signature
- Transaction & History Records

Business Customer Record
- Business Info
- Owner Info
- Address
- Contact Information
- Transaction & History Records

Driver License Record
- Current License
- Endorsements
- Restrictions
- Sanctions
- Transaction & History Records

Driver Sanction Records
- Violations & Transactions
- Suspensions
- Restoration
- General Status

Test Records
- Knowledge Tests
- Skill Tests

Vehicle Record
- Vehicle Data

Titles
- Title Data
- Link to Vehicle
- Link to Customers
- Transaction & History Records

Registrations
- Registration Data
- Link to Vehicle
- Link to Customers
- Transaction & History Records

Liens
- Lien Data
- Link to Title
- Link to Customers
- Transaction & History Records

Vendor Records
- Company Data
- Transaction & History Records

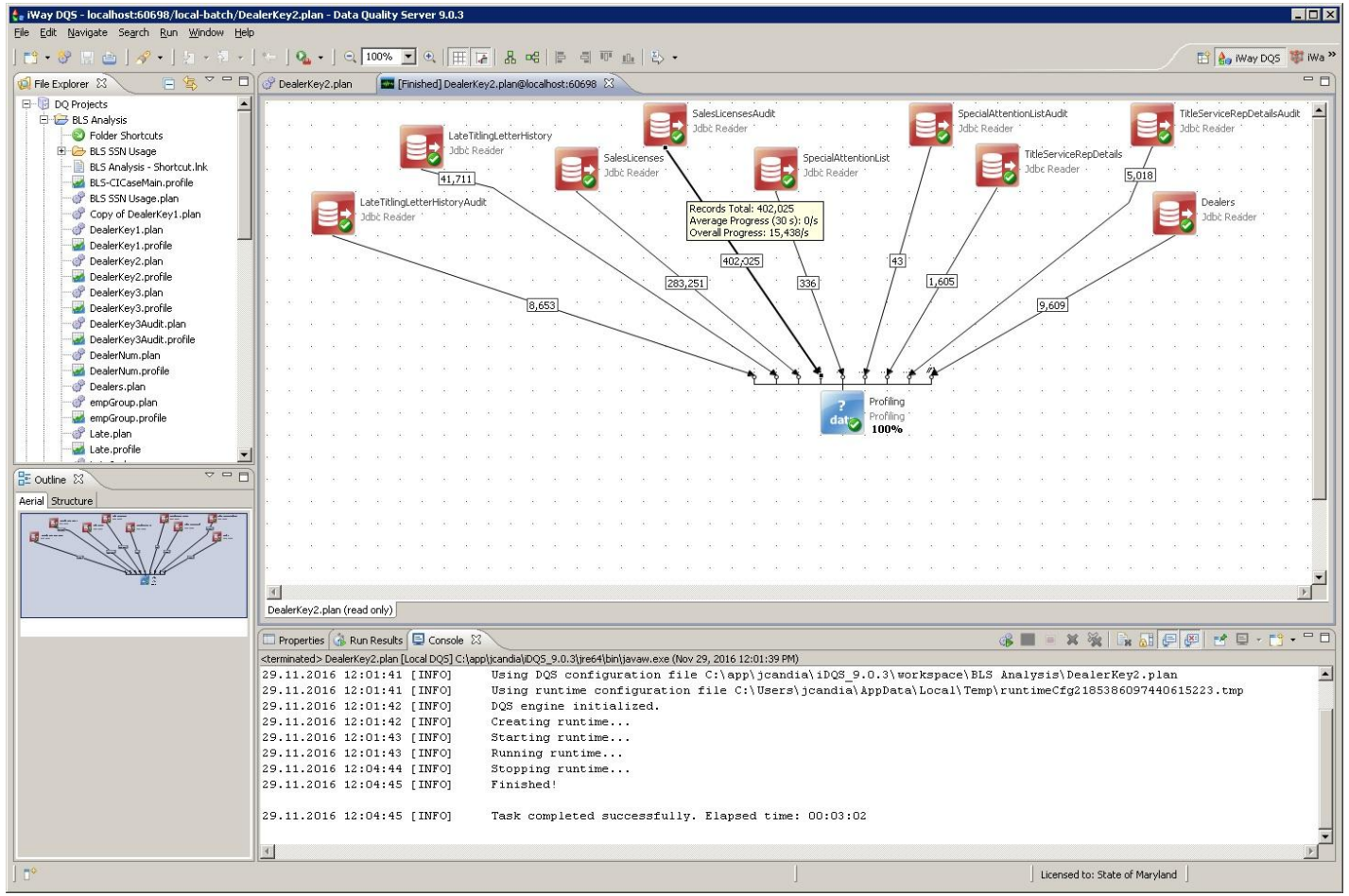
Dealer Records
- Company Data
- Transaction & History Records

Lender Records
- Company Data
- Transaction & History Records

3. Data Quality Assessment… Exhaustive Assessment

Data profiling tools provide an option to performing an exhaustive assessment which may be valuable when minimal documentation is available.

- Pros
  - Can identify a full range of issues
  - Can find errors across multiple tables
  - Can help with documentation
- Cons
  - Could be expensive
  - Requires specialized training
  - Requires copy of production database
  - May not add practical value over a knowledgeable IT team



4. Data Quality Issue Review & Prioritization

Now that you found data errors, what do you do?

Three steps guide teams in evaluating and prioritizing identified data quality issues.

ASSESS BUSINESS IMPACT

Does the quality issue affect DMV operations or the customer experience?

- **Process or Operations –** Impact on staff workflows or reporting
- **Fraud Risk –** Confusion or conditions that could enable fraud
- **Customer Service –** Difficulty completing transactions or obtaining assistance

ASSESS ABILITY TO REPAIR

Can the issue realistically be corrected?

- **Permissible –** Database corrections will not conflict with signed documents or official records
- **Possible –** The missing or incorrect data is available or can be reconstructed
- **Realistic** – The repair can be automated or completed with reasonable effort

PRIORITIZE ISSUE

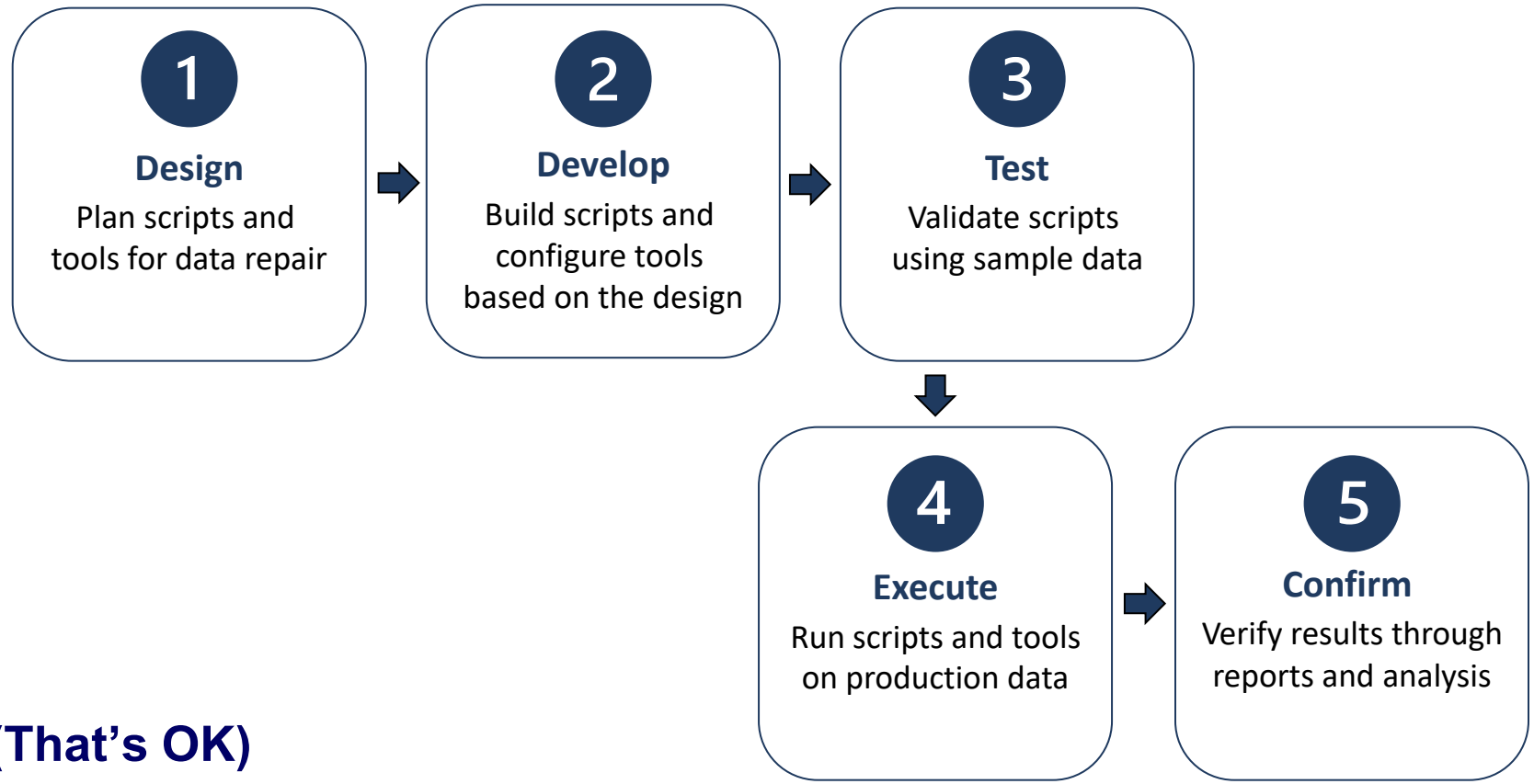
How critical is this issue relative to others?

- **Mandatory –** Must be resolved prior to migration to avoid operational or customer impacts
- **Important –** Supports successful system implementation and data quality
- **Desirable –** Beneficial but may be addressed later

5. Cleansing & Remediation

If data quality repairs require scripts or automated tools, the effort should follow a structured SDLC approach.

- Leverage Tools You Know
- Use Staff Familiar With The System
- Allow Sufficient Time
- Collaborate with Implementation Team for Solutions

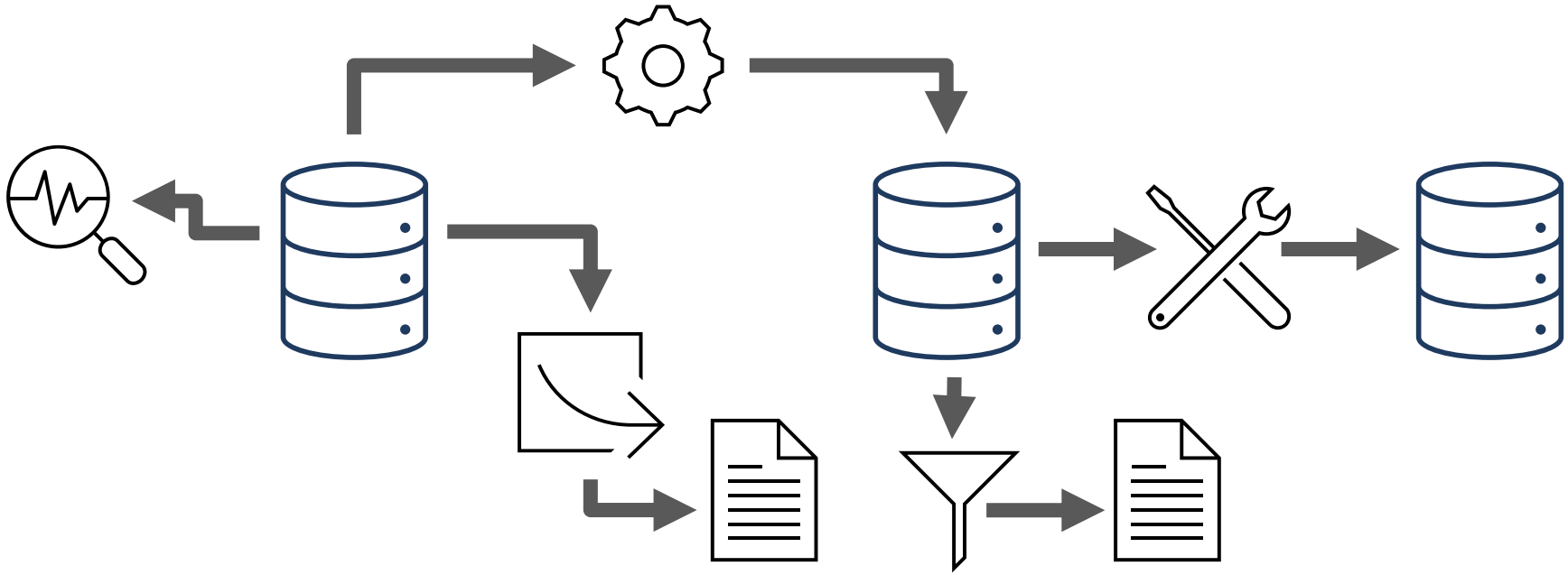


You'll Probably Never Finish (That's OK)

Supporting Tools

Every legacy system is different, but tools are available to assist in every step.

- **Staging Database** – for creating fresh repositories of data
- **Schema Tools** – for creating documentation
- **SQL & Scripting Tools** – for analysis, extraction, and corrections
- **Reporting Tools** – for analysis and reconciliation
- **ETL** – for large scale extraction, transformation and loading of data



Data Modernization & Migration Require Governance

Governance During Migration

- Who owns the data?
- Who approves cleansing?
- Who validates reconciliation?
- Who defines authoritative sources?

Essential Governance Roles

- Executive Sponsor
- Data Governance Council
- Data Owner & Steward
- Technical Custodian
- Security & Compliance

Beyond Migration

- Improved reporting
- Better fraud detection
- Stronger audit posture
- Cleaner integrations
- Improved customer service

Data Modernization as Governance Accelerator

- Establish data catalog
- Define ownership
- Implement quality processes
- Improve security posture
- Align with industry standards

Key Takeaways

Data migration is one of the highest modernization risk

- Get started now
- Documentation, cleansing, and governance is an investment
- Structured methodology reduces risk
- Governance guides you to success
- Data modernization eliminates data "debt"

Questions, Discussion

Contact Information

For more information regarding this presentation please contact:

Steven E. Young, PMP, Vice President
syoung@mathtechinc.com

Mathtech, Inc.
Mathtech, Inc.
2465 Kuser Road
Suite 200
Hamilton, NJ 08690
Phone: 609-495-2540
Fax: 609-495-2505
www.mathtechinc.com