Data Guidebook

Data Validation Rules

Data Validation Rules

Fields	**Type**	
Order No	Non-null integer (e.g.,101).	
Quantity Ordered	Positive integer (e.g.,1, 10, 100, 1000).	
Price Each	Positive decimal value, formatted as currency (e.g., 108.50).	
Orderline Number	Positive integer, sequential within each order (e.g., 1, 2, 3).	
Sales	Positive decimal calculated as Quantity Ordered × Price Each (e.g.,252.36)	

Order Date	Must be in YYYY-MM-DD format (e.g., 2024-02-04).	
Status	It must be one of predefined values (Shipped, Cancelled, On Hold, In Process).	
QTR_ID	Must be an integer between 1 and 4.	
Month_ID	Must be an integer between 1 and 12	
Year_ID	It must be a four-digit positive integer (e.g., 2024).	
Product Line	Must be a predefined category (e.g., Motorcycles, Trucks, Ships).	
MSRP	It must be a positive decimal value formatted as currency.	
Product Code	Must be a unique alphanumeric code (e.g., S10_1678).	
Customer Name	It must be a non-empty text string containing only valid characters A-Z Ä, Ö,Ü,ß (e.g., Neva Bhrana).	
Phone	Must be a valid phone number format (e.g., +1-444-123-4869).	
Address Line1	It must be a non-empty text string.	
Address Line2	Optional, can be null or text.	
City	Must be a non-empty text string containing only valid characters A-Z Ä,Ö,Ü,ß.	
State	It must be a valid state/province name or abbreviation (A-Z Ä, Ö, Ü,ß).	
Postal Code	Must match the postal code format for the corresponding country (e.g., 12345 for the US, A1B 2C3 for Canada).	
Country	Must be a valid country name should be String from a predefined list.	

Territory	The region identifier must	t be a predefined regio, whichr can be text or nul (e.g., NA, EMEA, APAC).
Contact Last Name	Must be a non-empty text string with only valid characters (A-Z Ä, Ö, Ü,ß).	
Contact First Name	Must be a non-empty text string with only valid characters(A-Z Ä, Ö, Ü,ß).	
Deal Size	Must be categorized as text string Small, Medium, or Large based on predefined sales value ranges.	
Full Name	Must be a non-empty text string with only valid characters (A-Z Ä, Ö, Ü,ß).	

Data Access Controls

Role	View Data	Edit Data	Modify Permissions	Access Level
CEO / Executives	✅ All company data	✅ Strategic, financial, HR, and legal	✅ High-level decisions	Full Access
Chief Data Officer (CDO)	✅ Enterprise-wide data	✅ Define data strategy, compliance	✅ Approve data governance policies	Enterprise Data
Data Governance Manager	✅ Data policies, compliance frameworks	✅ Modify data governance rules	✅ Assign data owners & stewards	Data Governance
Data Owner	✅ Data under their responsibility	✅ Define data access rules	✅ Approve modifications to datasets	Business Unit
Data Steward	✅ Operational data, data dictionaries	✅ Ensure data quality, correct errors	❌ Cannot assign access rights	Data Management
Data Architect	✅ Database structures, metadata	✅ Design data models, integration	✅ Define system architecture	IT & Data
Database Administrator	✅ All databases	✅ Manage database performance, backups	✅ Modify access controls	Database & IT

Cybersecurity Specialist	✅ Security logs, access control lists	✅ Monitor threats, enforce policies	✅ Manage encryption & security	Security & IT
Project Manager	✅ Project data, progress reports	✅ Update project timelines	❌ No access to financial data	Project Management
Product Manager	✅ Product development data	✅ Modify product requirement	❌ No access to HR or finance	Product Development
Data Analyst	✅ Business intelligence reports	✅ Clean, transform, and analyze data	❌ Cannot modify database structure	Analytics
IT Administrator	✅ System & security logs, infrastructure	✅ Manage user access, IT policies	✅ Full system control	System & Security
Finance Manager	✅ Budgeting, payroll, invoices	✅ Approve transactions	❌ No access to IT or HR	Finance
HR Manager	✅ Employee records, attendance	✅ Edit employee details, leave approvals	❌ Cannot modify system access	HR
Sales Manager	✅ Sales pipeline, lead data	✅ Modify deals, approve discounts	❌ No access to finance or HR	Sales
Marketing Manager	✅ Campaign performance data	✅ Modify marketing strategies	❌ No access to financials	Marketing
Intern / Contractor	✅ Only assigned project data	❌ No access to internal systems	❌ Cannot modify company records	☐ Restricted Access
Data Entry Operator	✅ View assigned datasets	✅ Input, update records	❌ Cannot modify system permissions	Data Entry

Data Model Standardization

1. Entries and Attributes

Entries	Attributes

Customers	Customer Name, Phone, Address Line1, Address Line2, City, State, Postal Code, Country, Territory Contact Last Name, Contact First Name, Full Name
Orders	Order No, Order Date, Status, QTR_ID Month_ID, Year_ID
Order Lines	Orderline Number, Quantity Ordered Price Each, Sales
Product	Product Line, MSRP, Product Code
Deals	Deal Size

2. Relationship

- Customers can place many Orders.
- Each Order is associated with one Customer.
- An Order can have multiple Order Lines.
- Each Order Line is associated with one Order and one Product.
- A Product can be part of many Order Lines.

3. Logical Data Model

Customers	CustomerID (Primary Key), CustomerName, Phone, AddressLine1, AddressLine2, City State, PostalCode, Country, Territory ContactFullName
Orders	OrderID (Primary Key), CustomerID (Foreign Key referencing Customers), OrderNo, OrderDate, Status, QTR_ID Month_ID, Year_ID
OrderLines	OrderLineID (Primary Key), OrderID (Foreign Key referencing Orders), ProductCode (Foreign Key referencing Products), OrderlineNumber, QuantityOrdered PriceEach, Sales
Products	ProductCode (Primary Key), ProductLine, MSRP
Deals	DealID (Primary Key), OrderID (Foreign Key referencing Orders), DealSize

Data Steward & Data Ownership Roles & Policies

1. Data Owner

1.1 Role of a Data Owner

A Data Owner is an executive or senior manager responsible for the data assets of a specific business unit. They ensure that data is used correctly, securely stored, and accessible only to authorized personnel.

1.2 Responsibilities of a Data Owner

- Define Access Permissions → Determine who can view, edit, and delete data.
- Ensure Compliance → Ensure data follows legal, regulatory, and company policies.
- Approve Data Changes → Validate modifications to critical datasets.
- Monitor Data Usage → Ensure that employees follow data security and governance rules.
- Collaborate with Data Stewards → Work with data stewards to resolve data quality issues.

1.3 Data Ownership Policies

Policy	Description	Enforced By
Data Access Policy	Owners approve access based on business needs.	IT Security & Data Owner
Data Modification Policy	Owners must approve critical data edits.	Data Owner & Stewards
Data Security Policy	Owners ensure encryption, backups, and security measures are in place.	IT & Cybersecurity
Data Retention Policy	Data is stored based on legal and business requirements (e.g., 5-7 years for financial records).	Compliance Team
Compliance & Audit Policy	Owners ensure regulatory compliance (e.g., GDPR, HIPAA).	Legal & Data Governance

2. Data Steward

2.1 Role of a Data Steward

A Data Steward is responsible for maintaining data accuracy, consistency, and integrity. They work closely with Data Owners to ensure that data governance policies are enforced.

2.2 Responsibilities of a Data Steward

- Monitor Data Quality → Identify and correct errors in data.
- Ensure Data Consistency → Standardize formats, definitions, and metadata.
- Implement Data Governance Rules → Follow and enforce company policies.
- Assist in Data Integration → Ensure smooth data sharing between systems.
- Support Data Users → Train employees on best practices for data handling.

2.3 Data Stewardship Policies

Policy	Description	Enforced By

Data Access Policy	Owners approve access based on business needs.	IT Security & Data Owner
Data Modification Policy	Owners must approve critical data edits.	Data Owner & Stewards
Data Security Policy	Owners ensure encryption, backups, and security measures are in place.	IT & Cybersecurity
Data Retention Policy	Data is stored based on legal and business requirements (e.g., 5-7 years for financial records).	Compliance Team
Compliance & Audit Policy	Owners ensure regulatory compliance (e.g., GDPR, HIPAA).	Legal & Data Governance

3. Collaboration

Task	Data Owner	Data Steward
Define who can access data	Approve	Enforce & monitor
Ensure data security	Set policies	Implement security measures
Maintain data quality	Establish rules	Clean & validate data
Resolve data issues	Make decisions	Identify & escalate issues
Data integration & reporting	Approve changes	Ensure consistency

Data Archiving Policies

This policy establishes guidelines for the retention, storage, and secure archiving of company data to ensure compliance, efficiency, and data integrity. This applies to all employees, contractors, and third-party vendors handling company data, including structured (databases) and unstructured (emails, documents) data, Customer records, Employee records, Financial transactions, Project documentation, IT logs, and emails.

1. Data Retention Periods

Data Type	Examples	Period	Compliance
Legal & Financial Records	Invoices, Tax Filings, Contracts	7 years	IRS, GDPR
Customer & Sales Data	Order history, Contact details	5 years	GDPR
Employee Records	Payroll, Performance reports	6 years	GDPR, Local Labor Laws
Marketing Data	Campaign results, Leads	3 years	GDPR

IT System Logs	Access logs, Error logs	1 year	Security Best Practices
Emails & Communications	Internal & External emails	3 years	GDPR, Business Needs
Healthcare or PHI Data (if applicable)	Medical records, Insurance details	6 years	HIPAA

2. Data Storage and Security Measures

- Archived data will be stored in encrypted cloud storage and offline backup servers.
- Physical documents will be digitized and stored securely before disposal.
- All archived data will be encrypted using AES-256 encryption.
- Access to archived data is role-based and requires multi-factor authentication (MFA).
- An audit log will track all access to archived records.

3. Data Retention and Deletion Process

- Identification: Data is reviewed for archiving based on the retention policy.
- Transfer: Data is moved to secure, long-term storage with limited access.
- Monitoring: Archived data is periodically reviewed to ensure compliance.
- Once data reaches the end of its retention period, it will be securely deleted using the following methods: Digital Data: Permanent deletion using data-wiping software (DoD 5220.22-M standard), Physical Documents: Shredding or incineration.
- A Data Deletion Certificate will be generated for compliance records.

4. Compliance with Legal and Regulatory Frameworks

- GDPR (General Data Protection Regulation) for customer data privacy.
- HIPAA (if handling healthcare data) for patient record security
- Local tax laws for financial data retention.
- Industry standards for secure data management.

5. Data Restoration and Accessibility

- Authorized personnel may request access to archived data via the IT Helpdesk.
- Data retrieval is logged for security auditing.
- Critical archived data is backed up every 6 months to prevent loss.

6. Review and Policy Update

This policy will be reviewed annually by the IT & Compliance team to ensure it meets current legal and business requirements.

Approved by: CEO
Effective Date: 01.01.2025
Next Review Date: 01.01.2026

File Conversions

The company is committed to ensuring efficient, secure, and standardized file conversions to maintain data integrity, optimize storage, and improve compatibility across systems. This policy defines guidelines for converting files while ensuring compliance with legal, security, and operational requirements.

This policy applies to all employees, contractors, and third-party vendors handling file conversions within DigiFlow Solutions. It covers:

- Document conversions (e.g., Word to PDF, Excel to CSV).
- Image & media conversions (e.g., PNG to JPG, MP4 to WebP).
- Data format conversions (e.g., XML to JSON, CSV to JSON).
- Compliance and security considerations.

1. File Conversion Methods

Document File Conversions

File Type	Conversion Format	Conversion Reasons
Word (.docx) → PDF (.pdf)	Standardized read-only format	Microsoft Word, Adobe Acrobat
Excel (.xlsx) → CSV (.csv)	Data processing & integration	Microsoft Excel, Google Sheets
PDF (.pdf) → Word (.docx)	Editing & modifications	Adobe Acrobat, Smallpdf
TXT (.txt) → Markdown (.md)	Documentation & readability	Notepad++, Pandoc

Image and Media File Conversion

File Type	Conversion Format	Conversion Reason
PNG (.png) → JPG (.jpg)	Reduce file size for web use	Photoshop, GIMP
JPG (.jpg) → WebP (.webp)	Optimize images for fast loading	Squoosh, TinyPNG
MP4 (.mp4) → WebM (.webm)	Web-friendly video format	HandBrake, VLC

Data Format Conversion

File type	Conversion Format	Conversion Reason
JSON (.json) → XML (.xml)	Data interchange format	Online tools, Python scripts
CSV (.csv) → SQL (.sql)	Database migration	MySQL Workbench, Pandas

2. File Conversion Policy

Standardization & Format Guidelines

- All official documents should be converted to PDF before external sharing.
- Large images should be converted to WebP for web usage to reduce bandwidth.
- CSV should be used when exporting structured data for compatibility.
- Lossless formats should be used for long-term storage (e.g., PNG instead of JPG for critical graphics).

Security & Compliance

- Confidential files (contracts, HR records) must be encrypted before conversion.
- No third-party online tools should be used for sensitive file conversions without IT approval.

- File integrity should be checked post-conversion to prevent data corruption.

Automation & Efficiency

- Batch conversions should be used for large datasets (e.g., scripts for bulk CSV to JSON).
- Automation tools (Power Automate, Python) should be used where possible.
- Duplicate copies should be deleted after conversion to save storage.

File Retention & Archiving

- Original files should be retained for at least 6 months before deletion.
- Converted files should be named properly (e.g., Reportdoc_2024_final.pdf).
- Older formats (e.g., legacy .xls files) should be converted to modern formats for compatibility.

3. Review and Policy Updates

This policy will be reviewed annually to align with business and compliance requirements.

Approved by: CEO
Effective Date: 01.01.2025
Next Review Date: 01.01.2026