* 1. **Data Firewall**

A **Data Firewall** is a **data quality control mechanism** that prevents **bad or non-compliant data** from entering systems, just like a network firewall blocks harmful traffic.

**What is a Data Firewall?**

A **Data Firewall** is a tool or system that:

* **Validates data** in real-time or batch before it enters a data store or pipeline.
* **Applies business rules** to check if the data meets required standards.
* **Stops or flags** data that does not meet quality requirements.
* **Logs issues** for review, correction, or auditing.

**Key Functions**

1. **Enforces data quality rules**
   * Checks for completeness, accuracy, consistency, validity, and uniqueness.
2. **Applies business logic**
   * Ensures data aligns with business definitions and policies.
3. **Monitors and alerts**
   * Sends notifications when bad data is detected.
4. **Blocks or routes data**
   * Rejects invalid data or routes it to error-handling processes.

**Example**

If a customer phone number is required to be 10 digits and numeric:

* A data firewall will **check incoming data**.
* If the phone number is "abc123", the firewall **blocks it** or sends it to a quarantine area.
* This prevents **corrupt data** from entering the system.

**Why It Matters**

Using a Data Firewall:

* **Prevents poor quality data** from entering your systems.
* **Reduces cleanup effort** later.
* **Protects analytics and reporting** from bad inputs.
* **Ensures compliance** with regulations and business standards.

**Summary**

A **Data Firewall** is a proactive tool to **enforce data quality at the point of entry**, making sure that only clean, trusted, and compliant data is allowed into your systems.