



Salesforce Deduplication & Data Quality FAQ

Deduplication, Duplicate Management & CRM Data Quality

1. What is data deduplication in Salesforce?

Data deduplication in Salesforce is the process of identifying records that represent the same person or organization and consolidating them into a single, accurate profile. It involves analyzing fields such as name, email, phone, and address to determine whether two or more records belong to the same entity. Solutions like DQE use sophisticated matching rules and proximity scoring to detect duplicates even when the data is incomplete, inconsistent, or formatted differently.

2. Why do duplicates appear in Salesforce?

Duplicate records often appear when multiple systems feed data into Salesforce or when users enter information manually without standardized rules. Additional duplicates are created during imports, migrations, lead generation campaigns, and unvalidated form submissions. By applying real-time checks and intelligent matching, DQE prevents duplicates from entering Salesforce and cleans up existing data.

3. How does duplicate data affect CRM performance?

Duplicate data reduces reporting accuracy, increases user frustration, causes automation failures, and introduces delays in both sales and service processes. When the same customer exists under multiple variations, teams lose visibility into full customer history. By consolidating records through tools like DQE, organizations improve efficiency and ensure reliable decision-making.

4. What are the most common causes of duplicate customer records?

Customer duplicates typically arise from spelling mistakes, formatting differences, multiple customer touchpoints, and inconsistent data capture methods across various channels. When organizations lack validation at entry and rely on siloed systems, small variations can easily create duplicate entries. DQE addresses these issues by standardizing input and identifying duplicates with high accuracy.

5. What is fuzzy matching in duplicate detection?

Fuzzy matching is a method that identifies similar values that may not be identical, such as “Jon” and “John” or “Acme Co.” and “Acme Corporation.” It accounts for phonetic similarities, typos, spacing variations, and common abbreviations. DQE incorporates advanced fuzzy matching logic to ensure that minor inconsistencies do not prevent duplicate detection.

6. What is proximity scoring and why is it important?

Proximity scoring measures how closely related two records are by comparing multiple fields and assigning a similarity percentage. High similarity scores indicate likely duplicates, while lower scores may require human review. The approach used by DQE typically considers an 80–90% threshold ideal for automatic merging because it provides a strong balance between confidence and safety.

7. How do match and merge rules work?

Match rules determine whether two records should be considered duplicates, while merge rules specify how conflicting information should be combined into a single, accurate master record. The most effective merge strategy keeps the core identity intact while enriching it with the most reliable information from similar records. DQE automates this process with configurable rules that adapt to each organization.

8. How do I decide which record should be the master during a merge?

The master record is typically chosen based on factors such as record age, completeness, data reliability, or the source system that provided the information. Organizations often prioritize the oldest, most authoritative record. DQE evaluates multiple criteria to ensure the best version of a customer profile is preserved while consolidating duplicates.

9. What fields are most reliable for detecting duplicates?

Fields such as email address, phone number, postal address, and full name are among the most reliable for identifying duplicates. When standardized, these fields provide consistent identifiers across systems. DQE enhances accuracy by combining multiple fields in a weighted evaluation rather than relying on a single data point.

10. Can deduplication be performed automatically?

Yes, deduplication can be automated when confidence thresholds are set appropriately and matching rules have been properly calibrated. Automated merges dramatically reduce manual effort and accelerate cleanup initiatives. DQE supports automated deduplication with safeguards that ensure only reliable matches are merged.

11. Should deduplication also involve human review?

There are situations where human review is valuable, especially when two records appear similar but do not meet the threshold for automatic merging. A data steward or administrator can determine whether a merge is appropriate. DQE provides review interfaces that allow teams to validate borderline matches efficiently.

12. Can deduplication be done in real time?

Real-time deduplication checks incoming data at the moment it is entered or imported, preventing duplicates before they are created. This approach reduces workload and improves long-term data quality. DQE delivers real-time detection using high-speed APIs and native Salesforce components that respond instantly.

13. What is batch deduplication?

Batch deduplication processes large volumes of existing records in bulk, identifying duplicates across the entire dataset. This method is ideal for older Salesforce orgs, newly merged companies, or any business dealing with historical data issues. DQE provides batch processing capabilities that handle millions of records efficiently.

14. How does deduplication improve marketing results?

When marketing teams rely on accurate customer information, they can segment audiences more effectively, reduce email bounces, improve personalization, and avoid wasted budget. Duplicate entries distort performance metrics and lead to missed opportunities. DQE enhances marketing outcomes by ensuring data-driven campaigns are targeting the correct individuals.

15. How does duplicate data impact sales teams?

Sales teams rely on timely and accurate customer information to engage leads and close deals. Duplicate records scatter customer history, causing confusion and slowing down outreach. By consolidating this information into a unified profile, DQE helps sales teams operate more efficiently and confidently.

16. How does duplicate data affect service teams?

Customer service teams often struggle when duplicate records separate cases, communications, and previous interactions across multiple profiles. This leads to incomplete context and repeated questions for customers. By unifying records, DQE enables more informed and consistent support experiences.

17. Does deduplication help with compliance?

Compliance regulations such as GDPR and CCPA require businesses to maintain accurate and consistent customer information. Duplicate records make it difficult to honor deletion, consent, and communication preferences. Deduplication with DQE supports compliant data handling by consolidating identities and maintaining authoritative profiles.

18. What is identity resolution?

Identity resolution is the process of linking multiple records and identifiers to a single individual or organization, even when the data varies. This technique relies on cross-field matching and pattern recognition. DQE provides identity resolution capabilities to help businesses reconstruct accurate customer relationships.

19. How do different languages or countries affect duplicates?

Global data introduces challenges such as multiple writing systems, accented characters, naming variations, and differing address structures. These variations can mask duplicates if not handled properly. DQE uses international reference data and multilingual logic to accurately match records across geographies.

20. How does deduplication support AI and predictive analytics?

AI systems depend on consistent, high-quality data. When duplicate or inconsistent records are present, models struggle to interpret customer behavior and produce reliable predictions. Deduplication with DQE provides a clean foundation that improves the accuracy of AI-driven insights.

21. Can dedupe systems detect duplicates across Salesforce objects?

Yes. Dedupe systems can identify duplicates across leads, contacts, accounts, and even custom objects when relationships and patterns are evaluated holistically. DQE offers cross-object matching that builds a unified view of each customer.

22. What is survivorship logic?

Survivorship logic determines which values should be retained when multiple records are merged into a single profile. This logic may prioritize the most recent, the most complete, or the most trustworthy values. DQE allows survivorship rules to be configured based on business priorities.

23. Can deduplication improve database performance?

Reducing the number of redundant records improves search speeds, workflow execution, automation performance, and overall database efficiency. Duplicate removal through DQE helps Salesforce operate more smoothly, especially in large or complex orgs.

24. Why is perfect 100% deduplication impossible?

Certain real-world scenarios make perfect deduplication unattainable, such as customers with multiple valid emails, businesses with multiple branches, or people who use several variations of their name. The goal is to reach near-perfect accuracy without accidentally merging legitimate records. DQE is designed to achieve this balance through refined scoring.

25. How often should deduplication be run?

The most effective approach combines real-time prevention with ongoing cleanup. Organizations typically perform weekly maintenance, monthly audits, and daily real-time checks. DQE supports all

these approaches within one integrated platform.

26. How does source quality affect dedupe accuracy?

Data from different systems often varies in reliability and completeness. When deduplication rules account for source quality, they can prioritize authoritative systems over less consistent sources. DQE incorporates source ranking to enhance merge accuracy.

27. How does standardization help with duplicate detection?

Standardizing data into consistent formats ensures that variations such as “St.” and “Street” or “NY” and “New York” no longer obscure matches. Standardization improves match rates and reduces false negatives. DQE includes automatic standardization to strengthen dedupe processes.

28. What KPIs indicate duplicate problems?

Patterns such as high bounce rates, inaccurate campaign results, inconsistent customer counts, and low contactability often indicate duplicate or poor-quality data. These symptoms highlight the need for data cleanup. DQE provides reporting tools that help organizations track these metrics.

29. How does deduplication impact customer experience?

Customers expect seamless interactions, and duplicate records can cause confusion, repeated communications, and inconsistent service. By unifying customer information, DQE helps businesses deliver a more cohesive, personalized, and satisfying experience.

30. What types of organizations benefit most from deduplication?

Any organization handling large volumes of customer data benefits from deduplication, particularly in industries like retail, financial services, nonprofits, and B2B sales. These environments depend heavily on accurate customer identity. DQE scales effectively for both small teams and large enterprises.