

# Improve Data Quality

Wednesday, October 14, 2015 3:05 PM

1. Assess data quality
2. Export data
3. Cleanse and standardize data
4. Enforce data quality
5. Push clean data back into Salesforce
6. Describe the capabilities of Data.com

Improving data quality is an iterative process:

Assess - Analyze data, Identify problems, Determine next steps

Cleanse - Standardize, Clean, De-duplicate

Maintain - Educate, Enforce

Tools for Processing Data - Connect for Office, Excel Connector, Workbench

- Pull Salesforce reports directly into an Excel worksheet or a Word document.
- Excel connector allows you to query, update, insert, and delete Salesforce data from within excel
  - You can download at [developer.force.com](http://developer.force.com)
- Workbench (web-based tool that allows admins + developers to interact with their Salesforce organization using the Force.com APIs) allows you to query, update, insert, and delete Salesforce data from a website.
  - It can be used to import and export data that needs to be cleansed and standardized.
  - Available at [developer.force.com](http://developer.force.com)

Data Quality Dashboards can be used to measure data quality

Specify which critical fields should be used to measure data quality.

'Data Quality Score' formula field calculates a score based on whether or not the specified fields contain data.

Assess a score for each critical field. It's a good Idea to display a text showing successful data.

## Making Fields Required

Method	Description
Select the Required checkbox when defining a custom field.	Defines the field as universally required, which means the field must have a value when a record is saved within Salesforce, web-to-lead/web-to-case, or the Force.com API.
Set the field to Required on page layout.	Requires that the field must have a value when a record is saved from the edit page for that page layout.
Validation rule	Verifies that the data in one or more fields meets the specified criteria before the record is saved within Salesforce, web-to-lead/web-to-case, or the Force.com API.

Picklist and dependent picklist fields-Force users to enter a permissible value.  
 Custom lookup fields and lookup filters-Select a subset of records and enforce data consistency.  
 Custom formula fields-Derive values from other fields and prevent redundant data.  
 Record types and page layouts-Display only relevant fields and reduce unnecessary data.  
 Workflow field updates-Automatically update field values and reduce manual entry.

### Inconsistent Data Formatting

Enforcing Postal Code Formatting to make Postal Codes sure postal codes are correct. You can do this using REGEXes.

#### Lead Validation Rule

Error Condition Formula:

```
( Country = "United States" ) &&
NOT( ISBLANK( PostalCode)) &&
NOT( REGEX( PostalCode, "[0-9]{5}(-[0-9]{4})?" ) )
```

Error Message:

Postal code must be in 99999 or 99999-9999 format.

#### Function Guide:

REGEX(Text, RegEx\_Text) = Compares a text field to a regular expression and returns TRUE if there is a match.



A regular expression is a string used to describe a format of a string according to certain syntax rules.

Regular Expression Constructs	What Does It Match?
[ab]	A single character, either a or b
[a-zA-Z]	A single uppercase or lowercase character
\\d	A digit
\\s	A whitespace character
X?	Zero or one occurrence of X
X*	Zero or more occurrences of X
X+	One or more occurrences of X
X{n}	Exactly n occurrences of X



Regular expression syntax is based on Java Platform SE 6 syntax. However, backslash (\) must be changed to double backslash (\\).

### Duplicate Records

You can **merge duplicates** using the **Merge Wizard**. You can merge up to three accounts, contacts, or leads.

Lead  
**Steve Shermer**



[Customize Page](#) | [Edit Layout](#) | [Printable View](#) | [Help for this Page](#)

[Show Feed](#) [Follow](#)

[Back to List: Leads](#)

[Open Activities \(0\)](#) | [Activity History \(0\)](#) | [HTML Email Status \(0\)](#) | [Campaign History \(3\)](#)

**Lead Detail** [Edit](#) [Delete](#) [Convert](#) [Clone](#) [Sharing](#) [Find Duplicates](#)

Lead Owner	 <a href="#">Lorenzo Costa</a> <a href="#">Change</a>	Lead Status	Open
Name	Steve Shermer	Phone	(408) 555-6673
Company	Carl Financial Group	Email	<a href="mailto:sshermer@trainingorg.com">sshermer@trainingorg.com</a>
Title	Manager	Rating	Hot
Region	US		
Employee Referral			

Lead  
**Steve Shermer - Search for Duplicates**

[Help for this Page](#)

[Search](#) [Convert Lead](#) [Cancel](#)

Name	<input type="checkbox"/> Steve Shermer
Last Name	<input checked="" type="checkbox"/> Shermer
Company	<input type="checkbox"/> Carl Financial Group
Email	<input checked="" type="checkbox"/> sshermer@trainingorg.com
Email Domain	<input type="checkbox"/> @trainingorg.com
Phone	<input checked="" type="checkbox"/> (408) 555-6673

[Search](#) [Convert Lead](#) [Cancel](#)

**Matching Leads** [Merge Leads](#)

<input type="checkbox"/>	Name	Title	Phone	Company	Email	Lead Status	Owner Alias
<input type="checkbox"/>	<a href="#">Steve Shermer</a>	Manager	(408) 555-6673	<a href="#">Carl Financial Group</a>	<a href="mailto:sshermer@trainingorg.com">sshermer@trainingorg.com</a>	Open	<a href="#">lcost</a>
<input type="checkbox"/>	<a href="#">Stephen Shermer</a>	Manager	(408) 555-6673	<a href="#">Watkins Group</a>		Open	<a href="#">msato</a>
<input type="checkbox"/>	<a href="#">Elizabeth Decker</a>	CIO	(408) 555-6673	<a href="#">Carl Financial Group</a>	<a href="mailto:elizabeth.decker@trainingorg-carlfinancialgroup.com">elizabeth.decker@trainingorg-carlfinancialgroup.com</a>	Open	<a href="#">cbrow</a>

## Merge Leads

[Help for this Page](#) 

[Previous](#) **Merge** [Cancel](#)

Select the values that you want to retain in the merged record. Highlighted rows indicate fields that contain conflicting data. The Master Record selected will retain read-only and hidden field values. The oldest Created By date and user will be retained in the merged record. Note: All related records including any campaigns and activities will be associated with the new merged record.

	Steve Shermer <a href="#">[Select All]</a>	Stephen Shermer <a href="#">[Select All]</a>
Master Record	<input checked="" type="radio"/>	<input type="radio"/>
Lead Owner	<input checked="" type="radio"/> Lorenzo Costa	<input type="radio"/> Mimi Sato
Name	<input checked="" type="radio"/> Steve Shermer	<input type="radio"/> Stephen Shermer
Company	<input checked="" type="radio"/> Carl Financial Group	<input type="radio"/> Watkins Group
Title	Manager	Manager
Lead Currency	USD	USD
Region	US	US
Employee Referral		
Lead Status	Open	Open
Phone	(408) 555-6673	(408) 555-6673
Email	<input checked="" type="radio"/> sshermer@trainingorg.com	<input type="radio"/>
Rating	<input checked="" type="radio"/> Hot	<input type="radio"/> Warm

Use a custom field and a workflow field update to enforce uniqueness of a standard field, such as lead email address, and thus prevent duplicates.

1. Create a unique custom field to shadow a standard field.

**Lead Custom Fields & Relationships**
[New](#)
[Map Lead Fields](#)
[Field Dependencies](#)

Action	Field Label	API Name	Data Type
<a href="#">Edit</a>   <a href="#">Del</a>	<u>Email Unique</u>	Email_Unique__c	Text(255) (Unique Case Insensitive)

2. Use a workflow field update to set its value to the field you are shadowing.

**Workflow Rule Detail**
[Edit](#)
[Clone](#)
[Deactivate](#)

Rule Name	Lead Email Unique	Object	Lead
Active	<input checked="" type="checkbox"/>	Evaluation Criteria	Evaluate the rule when a record is created, and every time it's edited
Description		Field to Update	Lead: Email Unique
Rule Criteria	true	Field Data Type	Text
		Re-evaluate Workflow Rules after Field Change	<input type="checkbox"/>
		Formula Value	Email

3. Use Data Loader to set the Email Unique field for existing leads.

- A developer can build an Apex trigger to prevent a duplicate record from being saved when a lead is created or updated.
- An Apex trigger is code that executes before or after specific system events.

Apex Trigger: leadDuplicatePreventer

Lead Edit Help for this Page ?

**New Lead**

---

**Lead Edit** Save Save & New Cancel

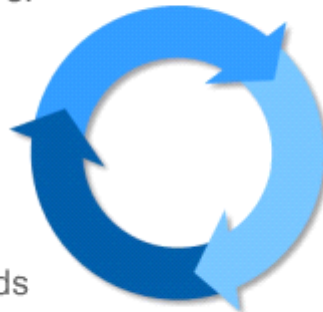
**Error: Invalid Data.**  
Review all error messages below to correct your data.

**Lead Information** = Required Information

First Name	--None-- Jon	Company	Berk Hath Inc
Last Name	Airaudi	Email	jon.airaudi@trainingorg-be

**Error: A lead with this email address already exists.**

Determine records with no activities or updates during the past two years.



Delete old records to maintain a high quality of data in the system.

Archive data using reports, data loader, or weekly data export service.

**Data.com** is an external app that allows admins to clean and enrich their Salesforce records with data from Data.com.

A Data.com License is required for use.

**Account Detail**

Edit
Delete
Get Contacts
Clean

Account Owner
Admin User [Change]
Clean Status
In Sync

Account Name
Apple Inc. [View Hierarchy]
Phone
+1.408.996.1010

Salesforce
Data.com
Select All | Clear All | Hide Identical Fields

Account Name	Apple Inc.	✓ Apple Inc.
Employees	<input type="checkbox"/> 46,600	Enrich
Annual Revenue	<input type="checkbox"/> \$65,225,000,000	
Ticker Symbol	<input type="checkbox"/> AAPL	
Website	<a href="http://www.apple.com">http://www.apple.com</a>	✓ <a href="http://www.apple.com">http://www.apple.com</a>
Industry	Electronics	✓ Electronics
Ownership	Public	✓ Public
Billing Address	1 Infinite Loop Cupertino, CA 95014-2083 United States	✓ 1 Infinite Loop Cupertino, CA 95014-2083 United States
Phone	<input type="checkbox"/> +1.408.996.1000	Clean
Fax	<input type="checkbox"/> +1.408.974.2400	
		<input type="checkbox"/> +1.408.996.1010
		<input type="checkbox"/> +1.408.974.2483

Administrators can schedule automatic clean jobs to:

- Examine Salesforce account, contact, and lead data.
- Flag or overwrite records based on selected preferences.
- Automatically fill in blank values with Data.com fields.

### Clean Jobs

**Clean Job Scheduling**

To keep your records' Clean status up to date, schedule automated jobs to compare your records with Data.com. Jobs are queued according to your scheduling settings, and they run independently for Accounts, Contacts, and Leads.

Run Clean Jobs Every...

Sun
Mon
Tue
Wed
Thu
Fri
Sat

☒
☒
☒
☒
☒
☒
☒

Preferred Start Time

10:00 AM
Pacific Daylight Time

Next Job Scheduled For
3/6/2013 10:00 AM

Full Sync With Next Job

On
Last user-run full sync: 3 days ago

Save
Undo

**Clean Jobs History**

Date +	Time Submitted	Job Type	Job Status	Records	Matched	Unmatched	Inactive
3/5/2013	10:12 AM	Account (D-U-N-S Number Lookup)	Completed	—	—	—	—

