

# Field Validation: Why & How

FMUG

March 7, 2008

# What is Field Validation?

- In computer science, **data validation** is the process of ensuring that a program operates on clean, correct and useful data. It uses routines, often called validation rules, that check for correctness or meaningfulness of data that are input to the system.
- Data validation checks that the data are valid and sensible/reasonable before they are processed.

From: [http://en.wikipedia.org/wiki/Data\\_validation](http://en.wikipedia.org/wiki/Data_validation)

# Common Types of Validation

- **Data type checks**

- Check the data type of the input and give an error message if the input data does not match with the chosen data type, e.g., In an input box accepting numeric data, if the letter 'O' was typed instead of the number zero, an error message would appear.

- **Range check**

- Checks that the data lie within a specified range of values, e.g., the month of a person's date of birth should lie between 1 and 12.

- **Limit check**

- Unlike range checks, data is checked for one limit only, upper OR lower.

- **Presence check**

- Checks that important data are actually present and have not been missed out, e.g., customers may be required to have their telephone numbers listed.

From: [http://en.wikipedia.org/wiki/Data\\_validation](http://en.wikipedia.org/wiki/Data_validation)

# Consequences of not validating data

May include:

- Data that must be cleaned before use, e.g.:
  - Zip codes for mailing labels, e.g., too few/many digits
- Inaccurate reporting, e.g.,:
  - Summary reports are not grouped properly
  - Finds won't find mistyped entries
- Failed relationships, e.g.:
  - Not all related records will appear in portals
  - Incorrectly matched records may appear related

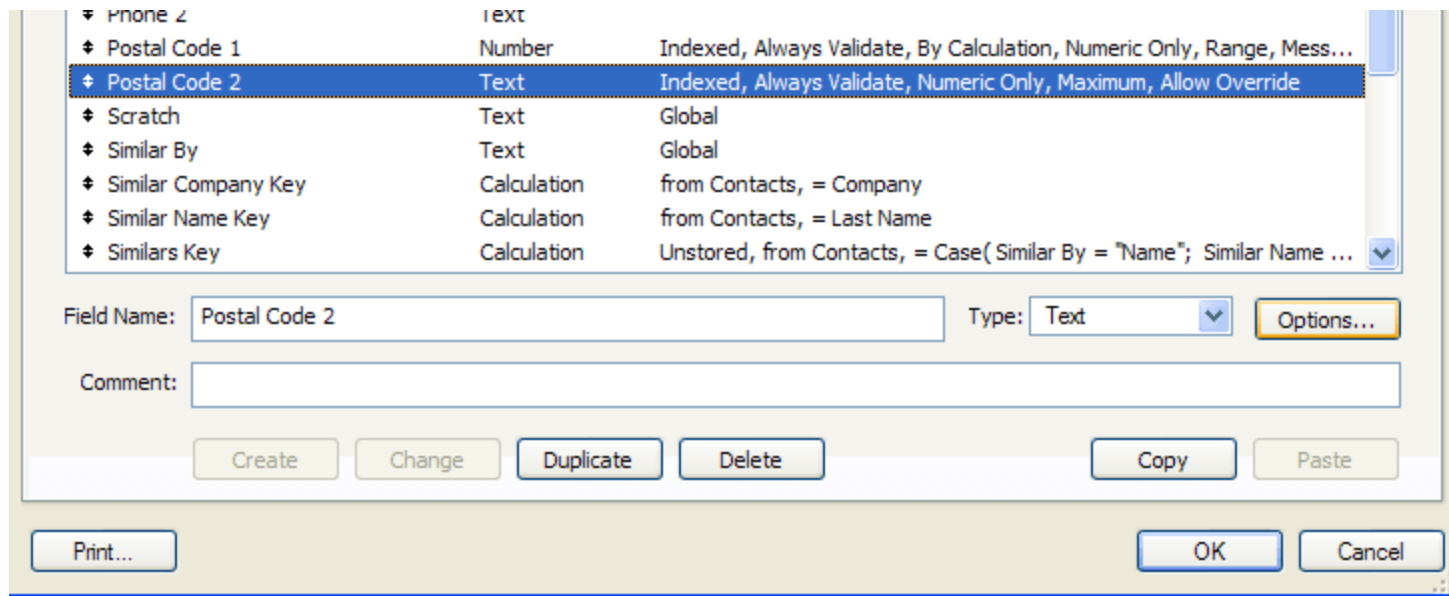
# What Validation is *not*

- An Input Mask
  - Field validation will ensure that incorrect data is not entered, but it will not ensure that correct data is entered in the format you want, e.g. phone: (617) 253-1000 vs. 617.253.1000
  - To achieve uniform formatting, you can simulate an input mask using calculated fields that use functions to manipulate the entered data

# Field Validation in FileMaker

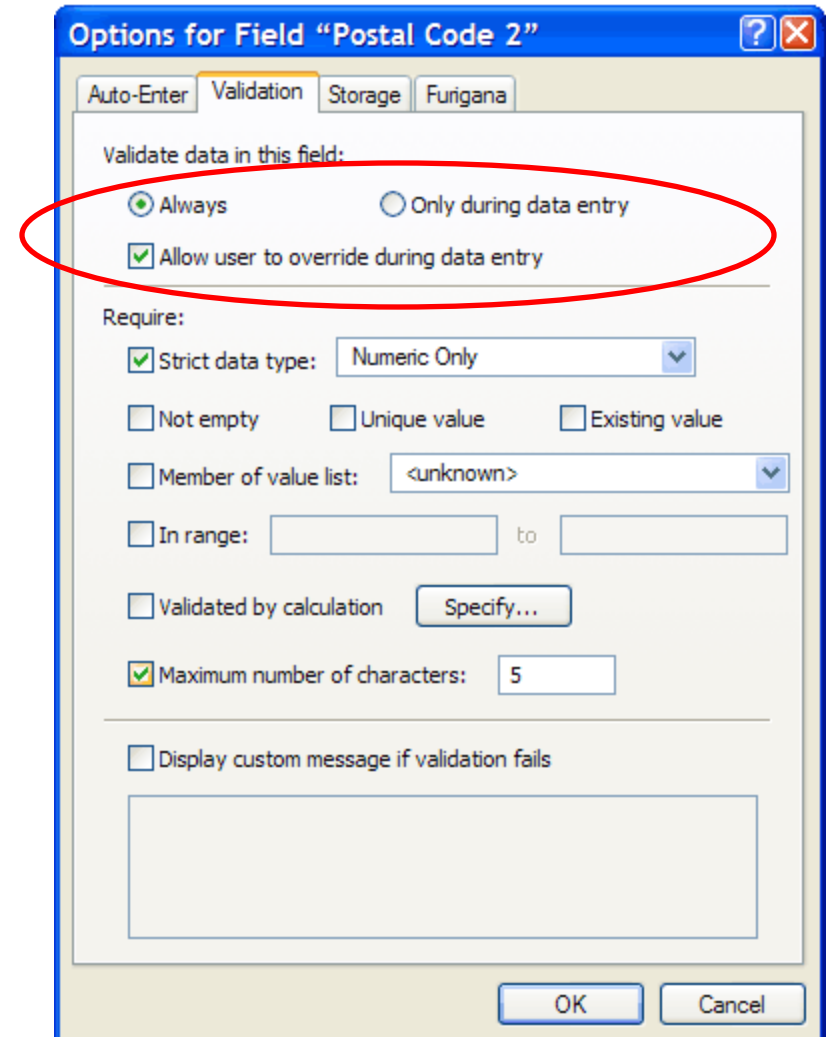
## How to implement rules:

- File > Manage > Database
- On the Fields tab, create a new field or select an existing field
  - double-click
  - or click the Options button



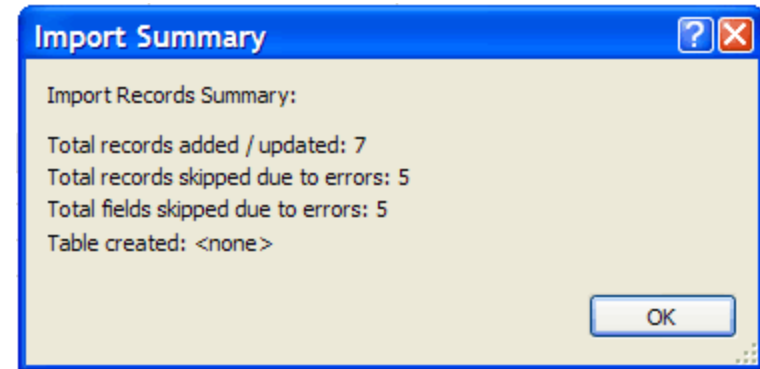
# Rules: When to validate

- If you receive data via some kind of feed or import, you should check “Always”
- If any imported data fails your validation rules
  - You will see an error, but the import may still occur
  - You’ll need to fix the errors



# More About Validation on Import

- How the record is validated
  - The entire record is added/updated or rejected
    - If it's a calculated rule
    - If "Not empty" is checked off
  - Only the field in question is rejected
    - If the record in question fails any other of the validation options.
    - The rest of the record is updated/added
- The summary won't tell you which records or fields were skipped





# Rules: Data Type

- First, know your field type, e.g.,
  - Text: may contain numbers and carriage returns
  - Number: may contain letters
- Data type validation options:
  - Numeric Only
  - 4-digit year date
  - Time of Day
- Think twice about Numbers vs. Text
  - Sorting text example: Task ID

The screenshot shows a dialog box titled "Options for Field 'Postal Code 2'". It has four tabs: "Auto-Enter", "Validation", "Storage", and "Furigana". The "Validation" tab is selected. Under "Validate data in this field:", there are two radio buttons: "Always" (unselected) and "Only during data entry" (selected). Below that is a checked checkbox "Allow user to override during data entry". Under the "Require:" section, there are several options: "Strict data type:" (checked) with a dropdown menu set to "Numeric Only" (circled in red), "Not empty" (unchecked), "Unique value" (unchecked), and "Existing value" (unchecked). Other options include "Member of value list:" (unchecked) with a dropdown set to "<unknown>", "In range:" (unchecked) with two empty input fields, "Validated by calculation" (unchecked) with a "Specify..." button, and "Maximum number of characters:" (unchecked) with an input field containing "1". At the bottom, there is an unchecked checkbox "Display custom message if validation fails" and a large empty text area. "OK" and "Cancel" buttons are at the bottom right.

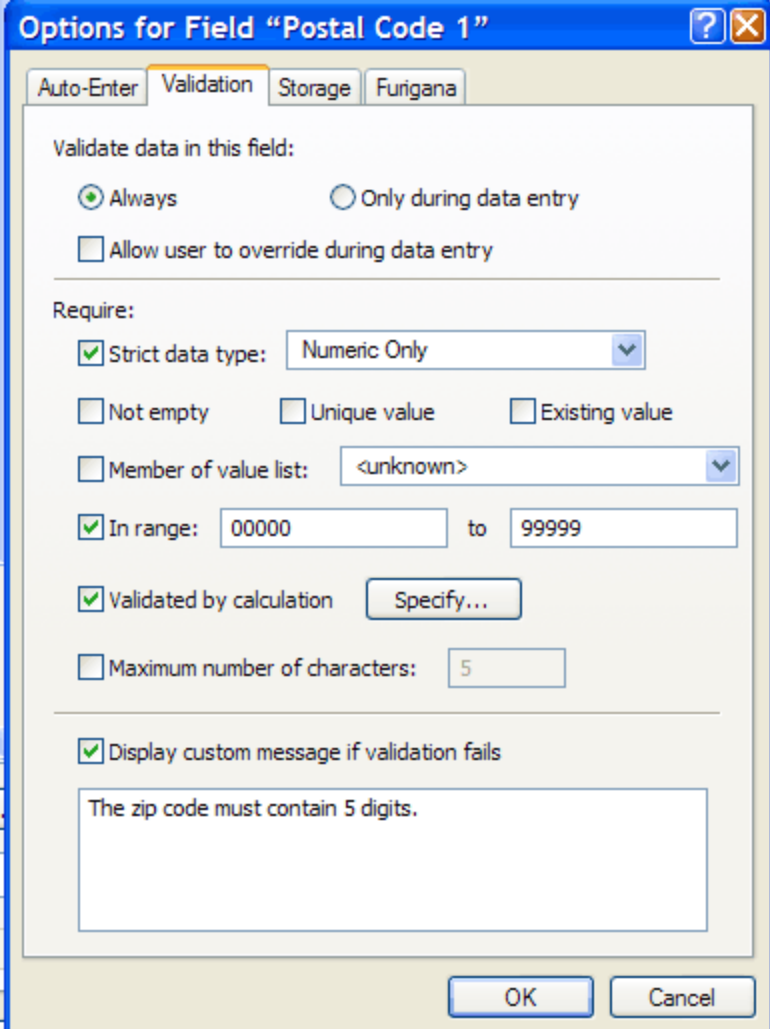
# Rules: More about Data Type

- You can define a text field that will contain numbers
  - Will preserve the leading and trailing zeroes
  - But could still cause integration issues
- Incompatibility with other applications, e.g., zip codes
  - If you export the field contents
    - Use the text file format because...
    - Open/Import it into Excel using menu commands, not by double-clicking
      - Opening by double-clicking automatically strips out leading zeros from number fields, whether the file was saved as .xls or .csv.
      - Opening by double-clicking automatically strips out leading zeros from text fields if the file was saved as a .csv
      - Opening by double-clicking automatically preserves leading zeros from text fields if the file was saved as .xls, but displays an error in Excel
    - Opening .txt/.tab file will force you through the import process
      - *Designate zip as a text field during import to Excel to preserve leading zeros*

# Rules: More about the Values

More about numbers and text:

- You can define a number field, and then
  - Set a strict Numeric Only data type because the Number data type allows non-numeric characters
  - Give it a range
  - Specify a max number of characters or, better yet, create a calculation to ensure 5 characters
  - Display a custom error message



Options for Field "Postal Code 1"

Auto-Enter Validation Storage Furigana

Validate data in this field:

Always  Only during data entry

Allow user to override during data entry

---

Require:

Strict data type: Numeric Only

Not empty  Unique value  Existing value

Member of value list: <unknown>

In range: 00000 to 99999

Validated by calculation Specify...

Maximum number of characters: 5

---

Display custom message if validation fails

The zip code must contain 5 digits.

OK Cancel

# Help Yourself

- You can use conditional formatting to tip you off to validation errors before you get the error message
  - All rules except “Not Empty” and “Validate by Calculation” occur when you commit the field, not the record
    - So, calculated validation will leave bad data in place as long as you are still performing data entry on the record
  - Set some conditional formatting to trigger on validation errors, so that you can fix them after leaving the field but before leaving the record

# Rules: List Membership

- Require field entry to be selected from a value list
  - Results in cleaner statistics when running summary reports
  - Eliminates typos
  - Turn on auto-complete for long lists
    - In Layout Mode:
      - Format > Field/Control > Setup...
      - Check off “Auto-complete using value list”

The screenshot shows the 'Options for Field "Status"' dialog box with the 'Validation' tab selected. The 'Validate data in this field:' section has 'Only during data entry' selected. The 'Require:' section has 'Member of value list' checked and set to 'Status', which is circled in red. Other options like 'Strict data type', 'Not empty', 'Unique value', 'Existing value', 'In range', 'Validated by calculation', 'Maximum number of characters', and 'Display custom message if validation fails' are unchecked.

# Rules: List Membership Override

- Decide whether you want to allow override
  - Allowing override will add non-list values to be entered in the field
  - Will not add new values to the list
    - Unless your list is defined by the field itself and not by custom values
  - Disallowing override is overridden if you check off “Include Edit...” in Field Setup (go figure)

The screenshot shows the 'Options for Field "Status"' dialog box with the 'Validation' tab selected. The 'Validate data in this field:' section has 'Only during data entry' selected. The 'Require:' section includes 'Member of value list' checked and set to 'Status'. The 'Allow user to override during data entry' checkbox is unchecked and highlighted by a purple arrow. Other options like 'Strict data type' (set to 'Numeric Only'), 'Not empty', 'Unique value', 'Existing value', 'In range', 'Validated by calculation', 'Maximum number of characters' (set to 1), and 'Display custom message if validation fails' are also visible.

# Rules: Allowing nulls

- Check off “Not Empty” for data that must be collected
  - Don’t go nuts with this!
  - If your business process doesn’t present the necessary data at the time of record creation, you’ll just frustrate your users

The screenshot shows the 'Options for Field "Contact ID"' dialog box with the 'Validation' tab selected. The 'Require:' section is circled in red, highlighting the 'Not empty' checkbox, which is currently unchecked. Other options include 'Strict data type' (set to 'Numeric Only'), 'Unique value', and 'Existing value'. The 'Allow user to override during data entry' checkbox is checked. The 'Display custom message if validation fails' checkbox is unchecked.

# Rules: Unique Identifiers

- Unique identifier as a primary key
  - “Not Empty” *and* “Unique Value”
  - Use this for relationship joins
- When bad things happen to good data
  - Duplicate values in join fields result in mismatched or “hidden” related records
  - Nulls in join fields result in hidden related records

Options for Field "Contact ID"

Auto-Enter Validation Storage Furigana

Validate data in this field:

Always  Only during data entry

Allow user to override during data entry

---

Require:

Strict data type: Numeric Only

Not empty  Unique value  Existing value

Member of value list: <unknown>

In range: to

Validated by calculation Specify...

Maximum number of characters: 1

---

Display custom message if validation fails

OK Cancel



# More About Primary Keys

- Support your validation rule with Auto-Enter values
- Do *not* allow users to modify primary keys
  - This will break relationships

Options for Field "Contact ID"

Auto-Enter Validation Storage Furigana

Automatically enter the following data into this field:

Creation Date

Modification Date

Serial number

Generate:  On creation  On commit

next value 16 increment by 1

value from last visited record

Data:

Calculated value Specify...

Do not replace existing value for field (if any)

Looked-up value Specify...

Prohibit modification of value during data entry

OK Cancel

# But wait! There's more!

- Serial vs. “meaningful” identifiers
  - Better to use serial for internal record ids (primary keys)
    - You do not want to rely on any value that could
      - Change, e.g., department codes
      - Result in duplicates, e.g., calculations based on first and last name
      - Be mistyped and later revised, e.g., email address
      - Possibly be null for a particular record, e.g., kerberos for external people
  - You may need meaningful identifiers for integrating with external data
    - So, you will probably still store values like MIT ID, email, kerb
    - Avoid SSN like the plague!!!!

# Questions?