

**Duplicate FHA Single-Family Mortgage  
Records and Related Problems**

**(PLAYING WITH MATCHES)**

THOMAS N. HERZOG, PhD, ASA  
CHIEF ACTUARY  
FHA/HUD

---

---

---

---

---

---

---

---

**DATA QUALITY AND RECORD  
LINKAGE TECHNIQUES**

THOMAS HERZOG  
FRITZ SCHEUREN  
WILLIAM WINKLER

SPRINGER -- 2007

---

---

---

---

---

---

---

---

**PART ONE  
DATA QUALITY: WHAT IT IS, WHY  
IT IS IMPORTANT, AND HOW TO  
ACHIEVE IT**

1. INTRODUCTION
2. WHAT IS DATA QUALITY AND WHY SHOULD WE CARE?
3. EXAMPLES OF COMPANIES USING DATA TO THEIR ADVANTAGE
4. PROPERTIES OF DATA QUALITY AND METRICS FOR MEASURING IT
5. BASIC DATA QUALITY TOOLS

---

---

---

---

---

---

---

---

**PART TWO  
SPECIALIZED TOOLS FOR DATABASE  
IMPROVEMENT**

- 6. MATHEMATICAL PRELIMINARIES FOR SPECIALIZED DATA QUALITY TECHNIQUES
- 7. AUTOMATIC EDITING AND IMPUTATION OF SAMPLE SURVEY DATA
- 8. RECORD LINKAGE -- METHODOLOGY
- 9. ESTIMATING THE PARAMETERS OF THE FELLEGI-SUNTER RECORD LINKAGE MODEL

---

---

---

---

---

---

---

**PART THREE  
TOOLS FOR LINKING RECORDS  
BY NAME AND ADDRESS**

- 10. STANDARDIZATION AND PARSING
- 11. PHONETIC CODING SYSTEMS FOR NAMES
- 12. BLOCKING
- 13. STRING COMPARATOR METRICS FOR TYPOGRAPHICAL ERROR

---

---

---

---

---

---

---

**PART FOUR  
RECORD LINKAGE CASE STUDIES**

- 14. FHA SINGLE-FAMILY MORTGAGE RECORDS
- 15. RECORD LINKAGE CASE STUDIES IN THE MEDICAL, BIOMEDICAL, AND HIGHWAY SAFETY AREAS
- 16. CONSTRUCTING LIST FRAMES AND ADMINISTRATIVE LISTS
- 17. SOCIAL SECURITY AND RELATED TOPICS

---

---

---

---

---

---

---

**PART FIVE  
OTHER TOPICS**

- 18. CONFIDENTIALITY: MAXIMIZING PUBLIC ACCESS TO MICRODATA WHILE PROTECTING RESPONDENT PRIVACY
- 19. REVIEW OF RECORD LINKAGE SOFTWARE
- 20. LEADERSHIP/MANAGEMENT
- 21. CONCLUDING THOUGHTS ON RECORD LINKAGE TECHNIQUES

---

---

---

---

---

---

---

**MOTIVATION:**

CLASSIC TEXTS WERE WRITTEN MANY YEARS AGO AND ARE NOW OUT-OF-PRINT

- NAUS [1975]  
– *DATA QUALITY CONTROL AND EDITING*
- NEWCOMBE [1988]  
– *HANDBOOK OF RECORD LINKAGE: METHODS FOR HEALTH AND STATISTICAL STUDIES, ADMINISTRATION, AND BUSINESS*

---

---

---

---

---

---

---

**BASIC CONCEPTS**

- Most techniques to improve data quality are a type of internal consistency check.
- Record linkage procedures can also be used to improve data quality as well as to enhance the usefulness of existing databases by adding new data elements or improving the completeness/accuracy of existing data elements.

---

---

---

---

---

---

---

## APPLICATION

ENHANCED VERSION OF  
CASE STUDY OF  
FHA-INSURED SINGLE-FAMILY  
MORTGAGE RECORDS  
(Chapter 14)

---

---

---

---

---

---

---

---

## ABOUT FHA

- FHA insures mortgages against the risk of foreclosure (i.e., that the borrower(s) will be unable/unwilling to continuing making payments to his/her/their mortgage).
- FHA has insured over 34 million single-family mortgages since its inception in 1934.

---

---

---

---

---

---

---

---

## Numbering Scheme for FHA Case Numbers

- SSO-NNNNNNNC
  - SS = State Code
  - O = Office Code within State
  - NNNNNN = Serial Portion of Case Number
  - C = Check Digit (to prevent typographical error)

---

---

---

---

---

---

---

---

### Examples of FHA Case Number

- 411-1451573
  - 41 = Ohio
  - 411 = Cincinnati Office within Ohio
  - 145157 = Serial Portion of Case Number
  - 3 = Check Digit
- 541-2760279
  - 54 = Virginia
  - 541 = Richmond
  - 276027 = Serial Portion of Case Number
  - 9 = Check Digit

---

---

---

---

---

---

---

---

### FHA CASE RECORD FAILING INTERNAL CONSISTENCY CHECK

FHA CASE NUMBER	STREET ADDRESS	CITY	STATE	ZIP CODE	NAME(S) OF BORROWER
441-1451573	714 GRAND AVENUE	CINCINNATI	PA	45232	HOLLINS, LARRY & JUANITA

---

---

---

---

---

---

---

---

### FHA CASE RECORD FAILING INTERNAL CONSISTENCY CHECK

FHA CASE NUMBER	STREET ADDRESS	CITY	STATE	ZIP CODE	NAME(S) OF BORROWER
441-1451573 <b>411-1451573</b>	714 GRAND AVENUE	CINCINNATI	<b>OH</b>	45232	HOLLINS, LARRY & JUANITA

---

---

---

---

---

---

---

---

**FHA CASE RECORD FAILING  
INTERNAL CONSISTENCY CHECK**

FHA CASE NUMBER	STREET ADDRESS	CITY	STATE	ZIP CODE	NAME(S) OF BORROWER
441-2760279	1309 BROAD STREET	RICH-MOND	PA	23231	BROWN, ROBERT

---

---

---

---

---

---

---

---

**FHA CASE RECORD FAILING  
INTERNAL CONSISTENCY CHECK**

FHA CASE NUMBER	STREET ADDRESS	CITY	STATE	ZIP CODE	NAME(S) OF BORROWER
441-2760279	1309 BROAD STREET	RICH-MOND	PA	23231	BROWN, ROBERT
541-2760279			VA		

---

---

---

---

---

---

---

---

**FHA Case Record Out of Sequence**

FHA Case Number	Initial Loan Amount	Begin Amortization Date	Contract Interest Rate	Name(s) of Borrower
131-5132066	\$70,100	Aug-1987	10.5%	Copp, Jeffrey S.
131-5132095	\$47,750	Oct-1987	11.0%	Budge, Robert W.
131-5132116	\$76,500	Aug-1987	10.5%	Woods, Sherrie D.
131-5132151	\$68,800	Jun-1984	14.5%	O'Hara Tom L Kim A
131-5132180	\$47,600	Aug-1987	10.5%	Epps, Sherri A.
131-5132197	\$43,250	Oct-1987	10.5%	Hollins, Richard M.

---

---

---

---

---

---

---

---

FHA CASE RECORDS FAILING RANGE TEST			
FHA Case Number	Begin Amortization Date	Contract Interest Rate	Name(s) of Borrower
372-0101064	Dec-1979	11.5%	Spruce, Ronald H
372-0113867	Apr-1983	12.0%	Stewart, P D Q
372-0116726	June-1983	12.0%	Landrieu W J
372-0707605	Jan-1975	9.0%	Richardson B F L
372-0707736	Jan-1975	9.5%	Bowers Newton P
372-0708494	Jan-1975	9.0%	McDonald K L

---

---

---

---

---

---

---

---

Duplicate FHA Case Records					
FHA Case Number	Property Address	Initial Loan Amount	Closing Date Of Loan	Contract Interest Rate	Name(s) of Borrower
372-1132617	109 Elm St Bethany, NY 14054	\$32,000	Feb-1983	12.5%	Brown John N
374-1132613	109 Elm St Bethany, NY 14054	\$32,000	Feb-1983	12.5%	Brown JR

---

---

---

---

---

---

---

---

FHA Case Records with Incorrect Termination Status					
FHA Case Number	Property Address	Closing Date of Loan	Status of Loan	Contract Interest Rate	Name(s) of Borrower
371-1019310	109-07 211 <sup>th</sup> Place Queens Village, NY 11429	Mar-1982	Active	16.5%	Smith John Paulette
374-4413730	109-07 211 <sup>th</sup> Place Queens Village, NY 11429	Sep-2004	Active	5.5%	Smith, Paulette

---

---

---

---

---

---

---

---

FHA Case Records with Incorrect Termination Status					
FHA Case Number	Property Address	Closing Date	Status of Loan	Contract Interest Rate	Name(s) of Borrower
372-1221854	323 Oak St Buffalo, NY 14215	Apr-1984	Active	13%	J. & K. Falkides
372-519223	323 Oak St Buffalo, NY 14215	Jan-1987	Prepaid	9%	Falkides, John P

---

---

---

---

---

---

---

---

FHA Case Records with Incorrect Termination Status					
FHA Case Number	Property Address	Closing Date	Status of Loan	Contract Interest Rate	Name(s) of Borrower
131-5109339	14104 RTE 1750 Coal Valley, IL 61240	Oct-1986	Active	9.0%	Angel, Dale J
131-7200510	14104 RTE 1750 Coal Valley, IL 61240	Jul-1994	Claim	7.5%	Angel, Dale I

---

---

---

---

---

---

---

---

**Ginnie Mae**

- Ginnie Mae = Government National Mortgage Assn.
- Ginnie Mae is a wholly-owned govt. corp.
- Ginnie Mae is a component of HUD
- Ginnie Mae's role is to guarantee the timely payment of principal and interest to owners of Ginnie Mae residential mortgage-backed securities (MBS) initially pooled by the lenders.

---

---

---

---

---

---

---

---



## Linking FHA & Ginnie Mae Records

- This was feasible because both databases had a field with the FHA Case Number.
- So exact matching could be performed via the FHA case number.

---

---

---

---

---

---

---

## Why Link FHA & Ginnie Mae Records?

- To correct termination status field within records on FHA's single-family data warehouse. If Ginnie Mae showed that mortgage was "terminated", why did FHA show it was still "active"?
- FHA's single-family data warehouse had 25,000 "active" case records without a property address. By linking we were able to use Ginnie Mae database to obtain about 5,000 such addresses.

---

---

---

---

---

---

---

**THANK  
YOU!!!**

---

---

---

---

---

---

---