

APPENDIX B

THE NYSIIS NAME CODING PROCEDURE

1. If the first letters of the name are:  
'MAC' then change these letters to 'MCC'  
'KN' then change these letters to 'NN'  
'K' then change this letter to 'C'  
'PH' then change these letters to 'FF'  
'PF' then change these letters to 'FF'  
'SCH' then change these letters to 'SSS'
2. If the last letters of the name are:  
'EE' then change these letters to 'YØ'  
'IE' then change these letters to 'YØ'  
'DT' or 'RT' or 'RD' or 'NT' or 'ND' then change these letters to 'DØ'
3. The first character of the NYSIIS code is the first character of the name.
4. In the following rules, a scan is performed on the characters of the name. This is described in terms of a program loop. A pointer is used to point to the current position under consideration in the name. This step begins the loop and sets this pointer to point to the second character of the name.
5. Considering the position of the pointer, only one of the following statements can be executed.

If blank, then go to rule 7.

If the current position is a vowel (AEIOU) then if equal to 'EV' then change to 'AF', otherwise, change current position to 'A'.

If the current position is the letter:

'Q' then change the letter to 'G'

'Z' then change the letter to 'S'

'M' then change the letter to 'N'

If the current position is the letter 'K', then if the next letter is 'N' then replace the current position by 'N' otherwise, replace the current position by 'C'.

If the current position points to the letter string

'SCH' then replace the string with 'SSS'

'PH' then replace the string with 'FF'

If the current position is the letter 'H' and either preceding or following letter is not a vowel (AEIOU) then replace the current position with the preceding letter.

If the current position is the letter 'W' and the preceding letter is a vowel, then replace the current position with the preceding position.

If none of these rules applies, then retain the current position letter value.

6. If the current position letter is equal to the last letter placed in the code, then set the pointer to point to the next letter and go to step 5.

The next character of the NYSIIS code is the current position letter.

Increment the pointer to point at the next letter.

Go to step 5.

7. If the last character of the NYSIIS code is the letter 'S', then remove it.
8. If the last two characters of the NYSIIS code are the letters 'AY', then replace them with the single character 'Y'.
9. If the last character of the NYSIIS code is the letter 'A', then remove this letter.

# THE MODIFIED NYSIIS NAME CODING PROCEDURE

1. If the first letters of the name are:
  - 'MAC' then change these letters to 'MCC'
  - 'KN' then change these letters to 'NN'
  - 'K' then change this letter to 'C'
  - 'PH' then change these letters to 'FF'
  - 'PF' then change these letters to 'FF'
  - 'SCH' then change these letters to 'SSS'
  - \*'WR' then change these letters to 'RR'
  - \*'RH' then change these letters to 'RR'
  - \*'DG' then change these letters to 'GG'
  - \*'A,E,I,O,U' then change these letters to 'A'
- \*2. Drop terminal S or Z from all names before coding begins.
3. If the last letters of the names are:
  - 'EE' then change these letters to 'Y'
  - 'IE' then change these letters to 'Y'
  - \*'YE' then change these letters to 'Y'
  - 'DT' or 'RT' or 'RD' then change these letters to 'D'
  - \*'NT' or 'ND' then change these letters to 'N'
  - \*'IX' then change these letters to 'ICK'
  - \*'EX' then change these letters to 'ECK'
  - \*'JR' or 'SR' then call this name an error and include it in table 2 of error output.
4. The first character of the NYSIIS code is the first character of the name.
5. In the following rules, a scan is performed on the character of the name. This is described in terms of a program loop. A pointer is used to point out the current position under consideration in the name. This step begins the loop and sets this pointer to point to the second character of the name.
6. Considering the position of the pointer, only one of the following statements can be executed.
  - (a) If blank, go to rule 7.
  - (b) If the current position is a vowel (AEIOU) then if equal to 'EV' then change to 'AF', otherwise, change current position to 'A'.
  - \* (c) If the current position is a Y and it is not the last letter of the name, then change the current position to an 'A'.
  - (d) If the current position of the letter is:
    - 'Q' then change the letter to 'G'
    - 'Z' then change the letter to 'S'
    - 'M' then change the letter to 'N'
  - (e) If the current position is the letter 'K', then if the next letter is 'N' then replace the current position by 'N' otherwise, replace the current position by 'C'.

- \*(f) If the current position is the letter 'S' and the next letter 'CH' then change to 'SSA' if end of the word or change to 'SSS' if not end of word.
  - \*(g) If the current position is the letter 'S' and the next letter 'H' then change to 'SA' if end of the word or change to 'SS' if not end of word.
  - (h) If the current position is the letter 'P' and the next letter 'H' then change 'PH' to 'FF'.
  - \*(i) If the current position is the letter 'G' and the next two letters are 'HT', then change 'GHT' to 'TTT'.
  - \*(j) If the current position is the letter 'D' and the next letter is 'G', then change 'DG' to 'GG'.
  - \*(k) If the current position is the letter 'W' and the next letter is 'R', then change 'WR' to 'RR'.
  - (l) If the current position is the letter 'H' and either preceding or following letter is not a vowel then replace the current position with the preceding letter.
  - (m) If the current position is the letter 'W' and the preceding letter is a vowel then replace the current position ('W') with the preceding position.
  - (n) If none of these rules apply, then retain the current position letter value.
7. If the current position letter is equal to the last letter placed in the code, then set the pointer to point to the next letter and go to step 6. The next character of the NYSIIS code is the current position letter. Increment the pointer to point at the next letter. Go to step 6.
  8. If the last character of the NYSIIS code is the letter 'S', then remove it.
  9. If the last two characters of the NYSIIS code are the letters 'AY', then replace the letters 'AY' with the single character 'Y'.
  10. If the last character of the NYSIIS code is the letter 'A', then remove this letter.
  - \*11. If the first character of the NYSIIS code is either 'A' or space, then replace it with the first letter of the original name.

\*Modifications made to the original NYSIIS coding technique.

THE CENSUS MODIFIED STATISTICS CANADA  
NAME CODING PROCEDURE

1. Insert first character of name in first code position.
2. Examine remaining characters of name deleting all vowels and the letter 'Y'.
3. Make all multiple adjacent letters occurrence single.
4. Compress the name removing all embedded blanks.
5. Truncate to four character. If the procedures yield a code of less than four characters, blanks to the right are valid and do not need change.

THE LEIN NAME CODING PROCEDURE

1. Insert first character of name word in first code position.
2. Examine the remaining letters of the name words removing all vowels and the letters 'Y', 'W', and 'H'.
3. Make all multiple adjacent letters single and truncate to four characters.
4. Code the 2nd thru 4th characters with the table below padding with 0's to the right if needed to make four characters.

NOTE: In step 3 and 2, you would compress the name removing all embedded blanks before continuing.

Table for Lein Name Coding Method

<u>Letters</u>	<u>Code Number</u>
D, T	1
M, N	2
L, R	3
B, F, P, V	4
C, J, K, G, Q, S,	5
X, Z	

# THE ROGER ROOT NAME CODING PROCEDURE

**\*\*The phonic code consists of five numeric digits.**

Example: BROWNER (09424)  
STANLEY (00125)

**\*\*The first letter or combination of letters are coded from the '1st Letter' table. The remainder of the letters are coded from the 'Basic' table. When vowels and the letters H, Y, and W appear other than as first letters, they are not coded.**

Example: CHALMAN (06532) would be coded as follows-  
CH - 06 (as shown in '1st Letter' table)  
A - not coded  
L - 5  
M - 3  
A - not coded  
N - 2

**\*\*If a fully coded name results in less than five digits, pad with zeros.**

Example: CHING (06270)

**\*\*If a name is too long for the five-digit code, code as many letters as possible and ignore remainder.**

Example: ANDERSON (12140)  
OVERSTREET (18401)

**\*\*When two letters with the same numerical value are together, they are considered as one letter.**

Example: HECKEL (27500)  
WYSZYNSKI (40207)

**\*\*Consonants separated by a vowel or by the letters H, Y, or W are coded separately and carry their individual values.**

Example: WHITTED (41100)  
ONGOOO (12770)

**\*\*The ten most common names on file would be coded as follows:**

JOHNSON (32020)  
WILLIAMS (45300)  
SMITH (00310)  
JONES (32000)  
BROWN (09420)  
DAVIS (01800)  
JACKSON (37020)  
WILSON (45020)  
LEE (05000)  
THOMAS (01300)



1st Letter Table

A	1----
B	09----
CE	00----
CH	06----
CI	00----
CY	00----
C	07----
DG	07----
D	01----
E	1----
F	08----
GF	08----
GM	03----
GN	02----
G	07----
H	2----
I	1----
J	3----
KN	02----
K	07----
L	05----
M	03----
N	02----
O	1----
PF	08----
PH	08----
PN	02----
P	09----
Q	07----
R	04----
SCH	06----
SH	06----
S	00----
TSCH	06----
TSH	06----
TS	00----
T	01----
U	1----
V	08----
WR	04----
W	4----
X	07----
Y	5----
Z	00----

Basic Table

B	9
CE	0
CH	6
CI	0
CY	0
C	7
DG	7
D	1
F	8
G	7
J	6
K	7
L	5
M	3
N	2
PH	8
P	9
Q	7
R	4
SCH	6
SH	6
S	0
TSCH	6
TSH	6
TS	0
T	1
V	8
X	7
Z	0

COMPOSITION OF SURNAME CODE  
FROM EACH PROCEDURE THAT CONTAINS DAVIS

Lein:

Dubose	Doubek	Debose	Dobosh
Dubs	Defigh	Daves	Dupois
Dubbs	Defazio	Dipiazza	Dufek
Dopps	Davis	Dobbs	Duffek
Doviak	Debaca	Dobak	Dupuis
Dubke	Dabbs	Dobis	Dupas
Dubus	Davies	Dobish	Devese
Dubois	Dubukey	Doepke	Devos
Duboise	Debus	Divish	Deveaux
			Devies

Roger Root:

Defouw	Davey	Devos	Tevis
Davis	Davies	Dafoe	Tiffée
Dauphi	Daves	Dove	Tivis
Defazio	Deife	Duff	Thevis
Defay	Dehoff	Duffey	Tovey
Davy	Devese	Duffie	Toeves
Defee	Devoe	Duffy	Tuffs
Dayhoff	Devee	Duyava	
Davie	Devies	Tafoya	

NYSIIS Eight Character:

Daves  
Davies  
Davis  
Devies  
Divish  
Dove  
Devese  
Devies  
Devos

Census Canada:

Daves  
Davies  
Davis  
Devese  
Devies  
Devos

CIA Dictionary:

Davis  
Davies

COMPOSITION OF SURNAME CODE FROM EACH PROCEDURE THAT CONTAINS SMITH

Lein:

Sand  
Sandau  
Sande  
Sandia  
Sando  
Sandoe  
Sandy  
Santee  
Santi  
Santo  
Send  
Sennet  
Shemoit  
Shenot  
Shumate  
Simmet  
Simot  
Sineath  
Sinnott  
Sintay  
Smead  
Smeda  
Smit

Smite  
Smith  
Smithey  
Smithy  
Smoot  
Smyth  
Snead  
Sneath  
Sneed  
Snoddy  
Sunday  
Sunanday  
Sund  
Sunda  
Sunday  
Sundy  
Swanda  
Swenda  
Swent  
Swint  
Synott

Census Canada:

Smathers  
Smith  
Smithart  
Smithbower  
Smitherman  
Smithey  
Smithgall  
Smithingall  
Smithmyer  
Smithpeter  
Smithson  
Smithwick  
Smithy  
Smotherman  
Smothers  
Smyth

Roger Root:

Samotid  
Simmet  
Simot  
Smead  
Smeda  
Smit  
Smite  
Smith  
Smithe  
Smithey  
Smithson  
Smithy  
Smoot  
Smyth  
Szmodis  
Zemaitis  
Zmuda

CIA:

Smith

NYSIIS Eight Character & Modified:

Schmit  
Schmitt  
Schmitz  
Schmoutz  
Schnitt  
Smit  
Smite  
Smith  
Smits  
Smoot  
Smuts  
Sneath  
Smyth  
Smithy  
Smithey