

/*This SAS program is based on paper:

Quan H. et al. Development and validation of a surname list to define Chinese ethnicity. Medical Care 2006*/

/*Two variables will be defined:

Variable names:

Chinese_1186 for the list of 1186: Yes/No

Chinese_416 for the list of 416: Yes/No*/

/*This list contains 1186 surnames. Among these,416 surnames are selected as high predictive list.

Using 1186 surnames define Chinese has a higher risk of classifying non-Chinese into Chinese compared with

the subset of 416 surnames. Therefore, the list of 416 could be used to sensitivity analysis through identifying

highly likely Chinese but will misclassify many Chinese into Non-Chinese.*/

Data a1;

set a2; /*Must change*/

if surname in ("ACHANG",

"AU YANG",

"AU YEUNG",

"AU YONG",

"AU YOUNG",

"BAGAMAN",

"BIH",

"CAI",

"CEN",

"CHAN",

"CHAO",

"CHARNG",

"CHEAZ",

"CHEE",
"CHEN",
"CHENG",
"CHENNG",
"CHEOK",
"CHERNG",
"CHEU",
"CHEUCK",
"CHEUK",
"CHEUNG",
"CHIANG",
"CHIAO",
"CHIAU",
"CHIEH",
"CHIEN",
"CHIH",
"CHIN",
"CHING",
"CHIOU",
"CHIU",
"CHIUNG",
"CHOIS",
"CHOU",
"CHOW",
"CHOY",
"CHU",
"CHUAN",
"CHUANG",
"CHUEH",

"CHUEK",
"CHUEN",
"CHUENG",
"CHUI",
"CHUO",
"CHUU",
"CHWANG",
"CHYAN",
"CHYN",
"CI",
"DENG",
"DEP",
"DER",
"DJUNG",
"DOONG",
"ENG",
"EOYANG",
"FA",
"FAN",
"FANG",
"FAT",
"FENG",
"FERNG",
"FO",
"FOK",
"FONG",
"FOO",
"FOOK",
"FOON",

"FOT",
"FOUN",
"FOUNG",
"FU",
"FUAN",
"FUH",
"FUI",
"FUN",
"FUNG",
"FUT",
"GAO",
"GET",
"GEW",
"GIN",
"GOEY",
"GOK",
"GOM",
"GONG",
"GOONG",
"GU",
"GUAN",
"GUH",
"GUNG",
"GUO",
"GWONG",
"GYATSO",
"HAUNG",
"HE",
"HEONG",

"HEUNG",
"HING",
"HIP",
"HOM",
"HOO",
"HOU",
"HSE",
"HSEI",
"HSI",
"HSIA",
"HSIAN",
"HSIANG",
"HSIAO",
"HSICH",
"HSIE",
"HSIEH",
"HSIEN",
"HSIN",
"HSING",
"HSIU",
"HSIUNG",
"HSU",
"HSUAN",
"HSUE",
"HSUEH",
"HSUI",
"HSUN",
"HU",
"HUANA",

"HUANG",

"HUEI",

"HUENG",

"HUI",

"HUNAG",

"HUNG",

"HUO",

"HWEI",

"HWONG",

"HWU",

"IEONG",

"IP",

"ISAO",

"IU",

"JAING",

"JAIR",

"JEANG",

"JEN",

"JENG",

"JER",

"JEU",

"JEUNG",

"JEW",

"JHANDA",

"JIA",

"JIAN",

"JIANG",

"JIH",

"JIM",

"JING",
"JIU",
"JOW",
"JUANG",
"JUE",
"KAO",
"KEUNG",
"KIANG",
"KIU",
"KOU",
"KOW",
"KUAI",
"KUAN",
"KUANG",
"KUEI",
"KUEY",
"KUI",
"KUNG",
"KUO",
"KWAI",
"KWAN",
"KWEI",
"KWO",
"KWOCK",
"KWOH",
"KWOK",
"KWONG",
"LAI",
"LEG",

"LEI",
"LEING",
"LEONG",
"LEUNG",
"LI",
"LIANG",
"LIAO",
"LIAU",
"LIAW",
"LIEH",
"LIGH",
"LII",
"LIN",
"LIOU",
"LIU",
"LIY",
"LO",
"LOO",
"LOONG",
"LOU",
"LOUI",
"LOUIE",
"LU",
"LUEN",
"LUENG",
"LUI",
"LUK",
"LUM",
"LUO",

"MA",
"MAH",
"MAR",
"MIAO",
"MIH",
"MOY",
"MUI",
"NG",
"NGAI",
"NGAR",
"NGEE",
"NGEW",
"NGI",
"NGIP",
"NGOK",
"NGONG",
"NGOOK",
"NGUEY",
"NGUN",
"NI",
"NIEH",
"NING",
"NIP",
"NIU",
"NUEN",
"NUI",
"NUM",
"OONG",
"OUYANG",

"OW",
"OWYANG",
"OWYOUNG",
"OY",
"OYANG",
"OYOUNG",
"PAN",
"PEI",
"PENG",
"PERNG",
"PHOON",
"PON",
"POON",
"POY",
"PU",
"PUI",
"PUN",
"QI",
"QIAN",
"QIN",
"QIU",
"QU",
"QUING",
"QUN",
"QUOCK",
"QUON",
"QUONG",
"RU",
"RUAN",

"RUMJAHN",
"SEETO",
"SEETOO",
"SETO",
"SEW",
"SEZTO",
"SHANG",
"SHANGKUAN",
"SHAO",
"SHE",
"SHEE",
"SHEN",
"SHENG",
"SHEONG",
"SHEU",
"SHEUNG",
"SHIANG",
"SHIAO",
"SHIAU",
"SHIEH",
"SHIH",
"SHING",
"SHIT",
"SHIU",
"SHIUE",
"SHIUNG",
"SHU",
"SHUAI",
"SHUEH",

"SHUEN",
"SHUI",
"SHUN",
"SHUNG",
"SHYONG",
"SHYR",
"SHYU",
"SHYY",
"SIH",
"SIOU",
"SIT",
"SITU",
"SIU",
"SOOHOO",
"SOONG",
"SOW",
"SUEN",
"SUI",
"SUJO",
"SUN",
"SWEI",
"SZE",
"SZETO",
"SZTO",
"SZU",
"SZUTU",
"TAAM",
"TAI",
"TAM",

"TAO",
"TARNG",
"TAUR",
"TEONG",
"TI",
"TIEH",
"TING",
"TOM",
"TOO",
"TSA",
"TSAI",
"TSANG",
"TSAO",
"TSAU",
"TSAY",
"TSE",
"TSEN",
"TSENG",
"TSERNG",
"TSEU",
"TSEUNG",
"TSI",
"TSIA",
"TSIANG",
"TSIAO",
"TSIEN",
"TSIM",
"TSING",
"TSIU",

"TSO",
"TSOI",
"TSON",
"TSONG",
"TSOU",
"TSOW",
"TSU",
"TSUAN",
"TSUE",
"TSUEI",
"TSUI",
"TSUN",
"TSUNG",
"TUNG",
"TWU",
"TZE",
"TZEN",
"TZENG",
"TZO",
"TZOU",
"TZU",
"UANG",
"UENG",
"WAH",
"WAI",
"WAN",
"WANG",
"WAT",
"WAUNG",

"WEI",
"WEN",
"WO",
"WOEY",
"WONG",
"WOT",
"WU",
"WUN",
"WUNG",
"WUU",
"XI",
"XIA",
"XIAN",
"XIANG",
"XIAO",
"XIE",
"XIN",
"XING",
"XU",
"XUE",
"YAI",
"YAN",
"YAO",
"YAT",
"YAU",
"YEE",
"YEH",
"YEI",
"YEIN",

"YEN",
"YENG",
"YEP",
"YET",
"YEUNG",
"YICK",
"YIEH",
"YIEN",
"YIH",
"YIN",
"YING",
"YIP",
"YIU",
"YOK",
"YOT",
"YUAM",
"YUAN",
"YUANG",
"YUE",
"YUEH",
"YUEN",
"YUENG",
"YUET",
"YUEY",
"ZAO",
"ZAU",
"ZENG",
"ZHAI",
"ZHAN",

"ZHANG",

"ZHAO",

"ZHEN",

"ZHENG",

"ZHONG",

"ZHOU",

"ZHU",

"ZHUANG",

"ZHUO",

"ZI",

"ZIANG",

"ZOU",

"ZUO") then Chinese_416="Yes";else Chinese_416="No";

if Chinese_416="Yes" or surname in

("AI",

"AN",

"ANG",

"AO",

"AU",

"AYE",

"BA",

"BAI",

"BAK",

"BAO",

"BAT",

"BAW",

"BEI",

"BEN",

"BI",

"BIAN",
"BIAO",
"BING",
"BIU",
"BO",
"BOK",
"BONG",
"BOU",
"BU",
"BUK",
"BUY",
"CAN",
"CANG",
"CAO",
"CENG",
"CHA",
"CHAI",
"CHAIM",
"CHAING",
"CHAK",
"CHANG",
"CHANGCHIEN",
"CHAT",
"CHAU",
"CHAUN",
"CHAUNG",
"CHAW",
"CHE",
"CHEAH",

"CHEANG",
"CHEAR",
"CHEI",
"CHEING",
"CHENCHANG",
"CHEO",
"CHEONG",
"CHERN",
"CHEUG",
"CHEUICHEUK",
"CHEUN",
"CHEUNJ",
"CHEW",
"CHHAY",
"CHHOR",
"CHI",
"CHIA",
"CHIAN",
"CHIEU",
"CHIEW",
"CHIK",
"CHIM",
"CHINN",
"CHIONG",
"CHITING",
"CHIUZ",
"CHIW",
"CHO",
"CHOCK",

"CHOENG",

"CHOH",

"CHONG",

"CHOO",

"CHOONG",

"CHOUNN",

"CHOUW",

"CHOWN",

"CHUA",

"CHUE",

"CHUH",

"CHUK",

"CHUN",

"CHUNE",

"CHUNG",

"CHUONG",

"CHUP",

"CHUY",

"CHYE",

"CHYI",

"CHYOU",

"CHYU",

"CHZUNG",

"CING",

"CIU",

"CO",

"COI",

"CONG",

"COU",

"COY",
"CU",
"CUAN",
"CUI",
"CYU",
"CYUN",
"DAI",
"DAN",
"DAU",
"DEA",
"DEE",
"DEFU",
"DEONG",
"DERE",
"DI",
"DIAN",
"DIAO",
"DIEC",
"DIEP",
"DIK",
"DIN",
"DING",
"DIU",
"DJANG",
"DJENG",
"DO",
"DOI",
"DONG",
"DOO",

"DOOH",
"DOU",
"DOUNG",
"DOW",
"DU",
"DUAN",
"DUANMU",
"DUEN",
"DUGU",
"DUNG",
"DUO",
"DUONG",
"EAR",
"ERH",
"FAH",
"FAI",
"FANCHIANG",
"FEI",
"FI",
"FING",
"FOCK",
"FOONG",
"FOU",
"FUEN",
"FUK",
"GA",
"GAA",
"GAI",
"GAM",

"GAN",
"GANG",
"GANN",
"GAU",
"GAUK",
"GAW",
"GE",
"GEI",
"GEN",
"GENE",
"GENG",
"GEOI",
"GEUNG",
"GIANG",
"GIM",
"GING",
"GIT",
"GIU",
"GO",
"GOENG",
"GOH",
"GON",
"GOO",
"GOOEY",
"GOOI",
"GOON",
"GOT",
"GOU",
"GOY",

"GUANG",

"GUEN",

"GUEY",

"GUI",

"GUK",

"GUM",

"GUN",

"GUU",

"GWAI",

"GWANG",

"GWIK",

"GWOCK",

"GWOK",

"GWON",

"GYUN",

"HA",

"HAHM",

"HAI",

"HAK",

"HAN",

"HANG",

"HAO",

"HAP",

"HAR",

"HAU",

"HECK",

"HEE",

"HEI",

"HENG",

"HEOI",
"HEP",
"HEU",
"HEW",
"HG",
"HIM",
"HIU",
"HO",
"HOI",
"HON",
"HONG",
"HOP",
"HOR",
"HORNG",
"HOTUNG",
"HOUNG",
"HSA",
"HSANG",
"HSIAD",
"HSIGH",
"HSIH",
"HSIUN",
"HSUEN",
"HSUNG",
"HUA",
"HUAI",
"HUAN",
"HUE",
"HUEN",

"HUIE",
"HUK",
"HUM",
"HUMS",
"HUN",
"HUNE",
"HUU",
"HWA",
"HWAI",
"HWANG",
"HWE",
"HWO",
"HWUNG",
"I",
"ING",
"ISANG",
"JAIN",
"JAM",
"JANG",
"JAO",
"JAU",
"JAW",
"JEA",
"JEE",
"JEHNG",
"JEI",
"JI",
"JIAO",
"JIE",

"JIEN",
"JIN",
"JIP",
"JOENG",
"JOIE",
"JOM",
"JONE",
"JONG",
"JOO",
"JOR",
"JOU",
"JOY",
"JSE",
"JU",
"JUAN",
"JUM",
"JOO",
"JYU",
"JYUN",
"KA",
"KAI",
"KAM",
"KAMKAN",
"KAN",
"KANG",
"KAT",
"KAU",
"KAUNG",
"KAWN",

"KE",
"KEE",
"KEET",
"KEH",
"KEI",
"KEN",
"KEOI",
"KEONG",
"KER",
"KEU",
"KEW",
"KHAY",
"KHOR",
"KHU",
"KIAO",
"KIEW",
"KIN",
"KIP",
"KIT",
"KLYHN",
"KO",
"KOA",
"KOE",
"KOH",
"KOK",
"KONG",
"KOO",
"KOT",
"KOUNG",

"KOWNG",

"KOY",

"KU",

"KUEN",

"KUK",

"KUM",

"KUN",

"KUOK",

"KUON",

"KUT",

"KWA",

"KWAIN",

"KWANG",

"KWON",

"KWUNG",

"KYOK",

"LA",

"LAM",

"LAN",

"LAO",

"LAU",

"LEM",

"LENG",

"LET",

"LEU",

"LEVNG",

"LEW",

"LIAN",

"LIEU",

"LIEW",
"LIK",
"LING",
"LIP",
"LIV",
"LOENG",
"LOH",
"LOI",
"LOK",
"LOR",
"LOUEN",
"LOUNG",
"LUAN",
"LUN",
"LUNG",
"LUONG",
"LYUN",
"MAI",
"MAK",
"MAN",
"MANG",
"MAO",
"MAT",
"MAU",
"MAW",
"MEI",
"MEN",
"MENG",
"MI",

"MIAN",
"MIAU",
"MIE",
"MIN",
"MING",
"MIU",
"MO",
"MOH",
"MOI",
"MOK",
"MON",
"MONG",
"MONK",
"MOU",
"MOW",
"MU",
"MUA",
"MUK",
"MUN",
"MUQI",
"NA",
"NAN",
"NANG",
"NANN",
"NAP",
"NAU",
"NEI",
"NENG",
"NEU",

"NEW",
"NGAN",
"NGAO",
"NGAU",
"NGAY",
"NGEUN",
"NGHEM",
"NGHIEM",
"NGIM",
"NGIN",
"NGING",
"NGON",
"NGOON",
"NGOR",
"NGOW",
"NGU",
"NGUI",
"NGUY",
"NIAN",
"NIE",
"NIN",
"NIPP",
"NOE",
"NOI",
"NOM",
"NUNG",
"NYI",
"OCK",
"OEI",

"OI",
"ON",
"ONG",
"OOI",
"OR",
"OU",
"OUONG",
"OUYOUNG",
"OWN",
"PA",
"PAI",
"PAK",
"PANG",
"PAO",
"PAU",
"PHANG",
"PHUA",
"PHUNG",
"PI",
"PIAN",
"PIAU",
"PIEN",
"PIH",
"PIN",
"PING",
"PIU",
"PO",
"POK",
"PONG",

"POO",
"POU",
"PUEN",
"QIANG",
"QIAO",
"QING",
"QUAN",
"QUE",
"QUELCH",
"QUEN",
"QUEY",
"QUNG",
"QUOORK",
"QWONG",
"RAN",
"REN",
"RONG",
"ROUGH",
"RUI",
"RUN",
"RUO",
"SA",
"SAEBAE",
"SAEUNG",
"SAH",
"SAI",
"SAM",
"SAN",
"SAT",

"SAU",
"SE",
"SEAH",
"SEC",
"SEEN",
"SEID",
"SEK",
"SEN",
"SENG",
"SEON",
"SEOW",
"SET",
"SHA",
"SHAM",
"SHAN",
"SHANGGUAN",
"SHAU",
"SHAZHA",
"SHEIH",
"SHEK",
"SHENH",
"SHEW",
"SHI",
"SHIC",
"SHINN",
"SHION",
"SHO",
"SHOON",
"SHOU",

"SHUANG",

"SHUE",

"SHUM",

"SI",

"SIAO",

"SIAUW",

"SID",

"SIE",

"SIEU",

"SIEW",

"SIKONG",

"SIMA",

"SIN",

"SING",

"SITOU",

"SO",

"SOK",

"SONG",

"SOO",

"SOP",

"SOU",

"SU",

"SUE",

"SUEY",

"SUK",

"SUM",

"SUNG",

"SUO",

"Surname",

"SUTU",
"SY",
"SYU",
"SYUN",
"SZETU",
"TA",
"TAN",
"TANG",
"TAT",
"TAY",
"TCHAO",
"TCHENG",
"TCHOU",
"TEANG",
"TED",
"TEH",
"TENG",
"TEO",
"TEOH",
"TEU",
"TEUNG",
"THEAN",
"THEUNG",
"THIAN",
"THOO",
"THOON",
"THUI",
"TIAN",
"TIAO",

"TIEN",
"TIEO",
"TIGHT",
"TIN",
"TIONG",
"TIP",
"TIU",
"TO",
"TOA",
"TOI",
"TON",
"TONG",
"TORNG",
"TOU",
"TOV",
"TOW",
"TOWE",
"TOY",
"TRI",
"TROUNG",
"TSAN",
"TSAUR",
"TSIN",
"TSOY",
"TSUANG",
"TSZ",
"TU",
"TUAN",
"TUANMU",

"TUEN",
"TUI",
"TUIN",
"TWAN",
"TZONG",
"UEN",
"UNG",
"VANG",
"VHANG",
"VONG",
"VOO",
"VUON",
"VUONGVY",
"WA",
"WE",
"WEE",
"WENG",
"WEY",
"WHENG",
"WHO",
"WHU",
"WING",
"WIP",
"WN",
"WNG",
"WOH",
"WON",
"WONE",
"WOO",

"WOON",
"WOU",
"WUE",
"WUENG",
"WUI",
"WY",
"XHO",
"XIONG",
"XIU",
"XUAN",
"XUN",
"YAM",
"YANG",
"YAP",
"YE",
"YEAP",
"YEEN",
"YEL",
"YEOH",
"YEONG",
"YEU",
"YEUONG",
"YEW",
"YI",
"YIANG",
"YIE",
"YII",
"YIK",
"YIM",

"YIT",
"YN",
"YO",
"YOEN",
"YONG",
"YOOK",
"YOON",
"YOR",
"YOU",
"YU",
"YUCHI",
"YUCK",
"YUI",
"YUIN",
"YUK",
"YUM",
"YUN",
"YUNG",
"YUP",
"YUT",
"YUU",
"ZAI",
"ZAM",
"ZAN",
"ZANG",
"ZE",
"ZEE",
"ZEON",
"ZHA",

"ZHAD",

"ZHE",

"ZHI",

"ZHON",

"ZHUGE",

"ZHUONG",

"ZIK",

"ZIU",

"ZOENG",

"ZONG",

"ZOO",

"ZU",

"ZUK",

"ZUOREN",

"ZYU")

then Chinese_1186="Yes";else Chiese_1186="No";

run;