THE UNIVERSITY OF AUCKLAND
FACULTY OF SCIENCE
Department of Statistics

Department of Statistics, University of Auckland

Summer Scholarship 2013-2014

Assessing Linkage Bias in the 1981-2006 Longitudinal Census Cohort

Name: Vera Puti Puti Clarkson

Supervisor: Barry Milne, Alan Lee

Degree: BCom/BSc, majoring in Statistics, Maths and Finance

Disclaimer

This report represents the views of the author. It does not necessarily represent the views of Statistics NZ and does not imply commitment by Statistics NZ to adopt any findings, methodologies, or recommendations. Any data analysis was carried out under the security and confidentiality provisions of the Statistics Act 1977.

Summer scholarship experience

The last 10 weeks has been a very significant aspect of my educational journey at The University of Auckland. Over this period, I got the chance to work with complex and messy real-life datasets containing as many as 7 million observations. In working with these datasets, I have developed many skills that would assist me when entering postgraduate study and the working world in the future.

This project required me to work independently, and think statistically and critically when analysing and interpreting data. It was crucial that I checked over my working thoroughly and cautiously, make educated and critical judgements on whether the results I have found were appropriate and valid.

Furthermore, this project has given me a strong taste and hands-on experience of what postgraduate study is like. It also has given me the confidence to believe that statistics is what I want to continue doing in the future.

Summary

Recently, records in New Zealand censuses from 1981 to 2006 have been linked, but not all were successfully linked. This incomplete linkage can result in a biased sample if the records linked differ in some key way from those unable to be linked. The aim of this project is to identify and investigate these differences, i.e. the factors that strongly determine linkage across NZ censuses.

This work is useful as researchers can use these factors to adjust for this bias in analyses. Once adjusting for this bias, the linked Census data would become a more accurate representation of the whole population. This would benefit researchers who want to use this linked census data to analyse, for example, longitudinal influences on mortality.

I have found that the strongest factors that determine linkage were whether or not the individual lived at the same address five years ago, their NZ deprivation score, whether or not they own their residence, their ethnicity, their relationship status, and their sex.

Abstract

The aim of this project is to identify and investigate the factors that strongly determine linkage across New Zealand Censuses. Out of the 15 cohorts created by linking censuses from 1981 to 2006, I have analysed six – all five two-census cohorts and one three-census cohort.

To identify and find the extent to which the factors determine linkage for each of the cohorts, I have found the partial correlations between linkage and each of the census variables using SAS Enterprise Guide. In conclusion, I have discovered that the factors that strongly determine linkage across these six cohorts are very similar. I have found that the most strongly associated factors were (1) whether or not the individual lived at the same address five years ago, (2) their NZ deprivation score, (3) whether or not they own their residence, (4) their ethnicity, (5) their relationship status, and (6) their sex.

Table of contents

1.0 Introduction

Recently, records have been linked between New Zealand Censuses from 1981 to 2006. This means that eligible individuals in the 2006, 2001, 1996, 1991, and 1986 censuses have been identified and 'linked' back to the previous census. To be eligible to be linked back one census, the individual must (a) be 5+ years old, and (b) have been living in the country for 5+ years.

However, not all records were able to be successfully linked. This is because individuals may not have filled in correct and consistent information in the current or previous censuses, or there may not be enough information given so that the individual can be identified in previous censuses. Another reason that makes the linking process imperfect is that personal identifiers such as the individual's name and address are removed from the dataset to keep anonymity amongst the data. Consequently, other variables such as date of birth, sex and area unit had to be used to link records across censuses.

The linkage between the five different NZ Censuses from 1981to 2006 has created 15 cohorts – shown in the diagram below. The shaded box in each row of the diagram represents a single cohort, and the number displayed within that box is the number of individuals that were successfully linked out of the eligible population. For example, the 060196 cohort shows that there are 1.592 million individuals who were linked from 2006 to 2001 and 2001 to 1996, as indicated by the shaded box.

Name of cohort	1981	1986	1991	1996	2001	2006	Number of census points in linked dataset
0601					2,311,000		**2**
0196				2,171,000			**2**
9691			2,174,000				**2**
9186		2,220,000					**2**
8681	2,078,000						**2**
060196					1,592,000		**3**
019691				1,571,000			**3**
969186			1,603,000				**3**
918681		1,581,000					**3**
06019691				1,173,000			**4**
01969186			1,177,000				**4**
96918681		1,154,000					**4**
0601969186				882,000			**5**
0196918681			850,000				**5**
060196918681				647,000			**6**

A point that should be recognized is that as the cohorts get wider in the diagram, the number of individuals successfully linked decreases. This is because we're less likely to

find individuals that have consistent census information over a longer period of time, and also because the requirements are more stringent to meet the eligibility criteria. For example, the requirements to be eligible for linkage between 2006 and 2001 is that the individual must be 5+ years old and have lived in the country for 5+ years (as of 2006), whereas the requirements to be eligible for linkage between 2006 and 1981 is that the individual must be 25+ years and have lived in the country for 25+ years (as of 2006).

Another point that should be recognized is that the wider cohorts in the diagram are a subset of some of the narrower cohorts. For example, the 060196 cohort is a subset of the 0601 cohort because those individuals that are linked from 2006 to 1996 must be linked from 2006 to 2001.

As mentioned earlier, not all records were able to be linked. Between successive censuses, approximately 70% of eligible records were able to be linked. However, researchers may want to use this linked census data to analyse, for example, job change over time or longitudinal influences on mortality. This incomplete linkage can result in a biased sample of the population if the records linked differ in some key way from those unable to be linked. The aim of this project is to identify and investigate these differences, i.e. the factors that strongly determine linkage across NZ censuses.

In my project I have analysed six of the 15 different cohorts. The cohorts I have analysed are all two-census cohorts, i.e. 0601, 0196, 9691, 9186, 8681, and also one of the three-census cohorts, 060196.

2.0 Methods

To analyse these cohorts, I used SAS Enterprise Guide.

Starting with the 2006-2001 cohort, I first identified all the variables from the 'individual_linkind0601_2006' dataset using the 2006 Data Dictionary. This dataset contained majority of the information about each individual (that was eligible to be linked back one census) in the 2006 census, and whether or not they were linked back to the 2001 census. There were also other important variables, such as Sex, NZDep and Ethnicity, that weren't located in this dataset, but in other datasets. By merging these datasets together, I created a new dataset which contained all these variables that I may have used in my analysis.

Secondly, I picked out and kept the variables which could possibly be causing bias between those records that are linked and those that are not linked. I then edited these variables – changing them all to numeric variables, simplifying the values they can take, and changing those who answered 'not specified' as 9's. I also rearranged the categorical variables into an order so that we could maximize the correlation between that variable and the linkage variable. Thirdly, I created two way frequency tables of all these variables with linkage. These tables allowed us to see the extent of the linkage bias we aim to search for.

Next, I created two correlation matrices. In the first matrix, I included all variables that I had planned to later on analyse with linkage. I used this matrix to identify all redundant variables and remove them from my analysis. I identified a redundant variable as one which was strongly correlated (i.e. greater than |0.8|) to another variable. In the second matrix, I included all variables including the linkage variable. After doing all this, I was then able to find the partial correlations between each of the variables and linkage.

Finally, I created a new dataset for this cohort, which I named 'final_dataset0601'. This dataset contains all the variables that I had created, plus the original age and years since arrival to NZ variable.

I repeated this process with the other five cohorts I have analysed. However with the three-census cohort, 2006-2001-1996 I had to have a few extra steps at the start of the process. I had to create a dataset which contained all those individuals that were eligible to be linked back from 2006 to 1996, as this was not given to us like it was in the two census cohorts. To do this I merged together those individuals that were eligible to be linked back from 2006-2001 and those eligible to be linked back from 2001-1996. I then narrowed the population down to only those individuals who were 10+ years old and have been in the country for 10+ years. This was because those who are younger than 10 or have lived in the country for less than 10 years could not possibly have been linked back 2 censuses ago. I then created a variable 'linkind060196' which indicated if

an individual was linked from 2006 to 1996 or not. After these extra steps, the same process as for all other cohorts was followed.

3.0 Results

For each cohort I have analysed, I have listed the ten most strongly correlated variables in determining linkage in the tables below. The variables are listed in order of strength based on their absolute partial correlation coefficient I have found. I ordered them by absolute value so I could focus solely on the strength of the relationship between the variable and linkage, rather than both strength and direction.

In each of these tables I have included the name of the variable, the bivariate correlation with the linkage variable, the absolute bivariate correlation with the linkage variable, the partial correlation with the linkage variable, and lastly the absolute partial correlation with the linkage variable.

I have called the linkage variable 'linkindxxxx' in my analysis. This is a binary variable which indicates whether or not the individual was linked back to the previous census or censuses. The 'xxxx' in the variable name represents the years which the linkage is between (i.e. 'xxxx' in {0601,0196,9691,9186,8681,060196}). For example, 'linkind0601' is the name of the binary variable which states whether or not the individual was linked from 06 to 01.

Table 1. Cohort 0601

		corr with linkind0601	abs(corr)	partial corr with linkind0601	abs(partial corr)
1	*Same address 5 years ago*	0.34471	0.34471	0.2462	0.2462
2	*NZ Deprivation score*	-0.12907	0.12907	-0.04387	0.04387
3	*Own residence*	0.22449	0.22449	0.04351	0.04351
4	*Marital status*	-0.17742	0.17742	-0.04348	0.04348
5	*Maori descent*	-0.15344	0.15344	-0.0422	0.0422
6	*European ethnicity*	0.11352	0.11352	0.04183	0.04183
7	*Years lived in NZ*	0.10308	0.10308	0.0418	0.0418
8	*Live with partner*	0.09392	0.09392	0.04165	0.04165
9	*Sex*	0.04189	0.04189	0.03963	0.03963
10	*Other ethnicity* [1]	0.04238	0.04238	0.0357	0.0357

Table 2. Cohort 0196

		corr with linkind0196	abs(corr)	partial corr with linkind0196	abs(partial corr)
1	*Same address 5 years ago*	0.28444	0.28444	0.13189	0.13189
2	*Years lived at address*	0.26773	0.26773	0.06051	0.06051
3	*Own residence*	0.22318	0.22318	0.0488	0.0488
4	*Live with partner*	0.10137	0.10137	0.04744	0.04744

5	*European Ethnicity*	0.15887	0.15887	0.04685	0.04685
6	*NZ Languages spoken*^{2}	-0.14144	0.14144	-0.04505	0.04505
7	*NZ Deprivation score*	-0.12192	0.12192	-0.04185	0.04185
8	*Maori descent*	-0.1326	0.1326	-0.04022	0.04022
9	*Sex*	0.04109	0.04109	0.03979	0.03979
10	*Marital status*	-0.18408	0.18408	-0.03833	0.03833

Table 3. Cohort 9691

		corr with linkind9691	abs(corr)	partial corr with linkind9691	abs(partial corr)
1	*Same address 5 years ago*	0.32355	0.32355	0.22453	0.22453
2	*European Ethnicity*	0.15994	0.15994	0.05063	0.05063
3	*Live with partner*	0.11668	0.11668	0.04844	0.04844
4	*Marital status*	-0.17242	0.17242	-0.04802	0.04802
5	*NZ Deprivation score*	-0.11319	0.11319	-0.04511	0.04511
6	*Years lived in NZ*	0.09624	0.09624	0.03824	0.03824
7	*Maori descent*	-0.13421	0.13421	-0.03755	0.03755
8	*Smoke*	-0.1341	0.1341	-0.03713	0.03713
9	*Benefit income*	-0.10471	0.10471	-0.03621	0.03621
10	*Sex*	0.03911	0.03911	0.03493	0.03493

Table 4. Cohort 9186

		corr with linkind9186	abs(corr)	Partial corr with linkind9186	abs(partial corr)
1	*Years lived at address*	0.24583	0.24583	0.21459	0.21459
2	*Same address 5 years ago*	-0.09539	0.09539	-0.13025	0.13025
3	*European Ethnicity*	0.18745	0.18745	0.08953	0.08953
4	*Live with partner*	0.14525	0.14525	0.06961	0.06961
5	*NZ Deprivation score*	-0.12809	0.12809	-0.05387	0.05387
6	*School qualification*	0.07461	0.07461	0.03789	0.03789
7	*Live with children*	0.05578	0.05578	0.03344	0.03344
8	*Live with parents*	0.02643	0.02643	0.03255	0.03255
9	*Maori descent*	-0.08897	0.08897	-0.03095	0.03095
10	*Sex*	0.03084	0.03084	0.02686	0.02686

Table 5. Cohort 8681

		corr with linkind8681	abs(corr)	partial corr with linkind8681	abs(partial corr)
1	*Same address 5 years ago*	0.32389	0.32389	0.199	0.199

2	*European Ethnicity*	0.16544	0.16544	0.05502	0.05502
3	*Years lived at address*	0.26859	0.26859	0.05378	0.05378
4	*Born in NZ*	0.04354	0.04354	0.04404	0.04404
5	*Years lived in NZ*	0.02489	0.02489	0.04304	0.04304
6	*Maori Ethnicity*	-0.12183	0.12183	-0.03667	0.03667
7	*Pacific Ethnicity*	-0.06629	0.06629	-0.02716	0.02716
8	*Defacto status*	-0.0855	0.0855	-0.02656	0.02656
9	*Work and labour force status*	0.03662	0.03662	0.02383	0.02383
10	*Religious*	0.05885	0.05885	0.02069	0.02069

Table 6. Cohort 060196

		corr with linkind060196	abs(corr)	Partial corr with linkind060196	abs(partial corr)
1	*Npairs sameaddress*[3]	0.50011	0.50011	0.45097	0.45097
2	*Years lived at address from 06*	0.26636	0.26636	-0.15048	0.15048
3	*Years lived at address from 01*	0.25418	0.25418	0.10848	0.10848
4	*Sex in 06*	0.0467	0.0467	0.04723	0.04723
5	*Marital status in 06*	-0.17619	0.17619	-0.04123	0.04123
6	*NZ Deprivation score in 06*	-0.12282	0.12282	-0.0409	0.0409
7	*Marital status in 01*	-0.13672	0.13672	-0.04065	0.04065
8	*European Ethnicity in 06*	0.10673	0.10673	0.03864	0.03864
9	*Live with partner in 06*	0.0838	0.0838	0.03609	0.03609
10	*Maori descent in 06*	-0.14622	0.14622	-0.03422	0.03422

[1] Whether or not the individual specified an ethnicity other than European, Maori, Pacific, Asian, or MELAA
[2] Whether or not the individual speaks English and/or Maori
[3] The number of times out of two censuses (06 and 01) the individual specified that they lived in at the same address five years ago

Table 1, 2, and 3 shows similar characteristics to each other. These tables indicate that the strongest variable that determines linkage in the 0601, 0196, and 9691 cohorts is *Same address 5 years ago*, scoring the highest absolute partial correlation ranging between 0.132 and 0.246. This variable also has the highest absolute bivariate correlation in each of the tables ranging from 0.284 to 0.345. The other 9 variables on each of these tables have relatively small absolute partials, ranging from 0.035 to 0.060. Another similarity is the variables that appear. All three tables include many of the same variables, such as *Same address 5 years ago*, *NZ Deprivation score*, *Marital status*, *Maori descent*, *European Ethnicity*, *Live with partner*, and *Sex*.

Table 4 varies slightly from the previous three tables in the sense that there were two, rather than one dominant variable that determines linkage in the 9186 cohort. In this table, the variables *Years lived at address* and *Same address 5 years ago* were at the top, scoring absolute partial correlations of 0.215 and 0.130 respectively. Thus, unlike the other two-census cohorts, the variable *Years lived at address* is the strongest variable that determines linkage in the 9196 cohort, closely followed by *Same address 5 years ago*. The other 8 variables on the table have relatively small partial correlations, ranging from 0.0269 to 0.090.
The variables displayed in table 4 are similar to those in tables 1, 2, and 3. For example, the variables *Same address 5 years ago*, *European Ethnicity, Live with partner, NZ Deprivation score, Maori descent,* and *sex* were variables that had made it to this list of top ten in all four cohorts – 0601, 0196, 9691, and 9186.

Table 5 was also similar to tables 1, 2 and 3 in the sense that the strongest variable that determines linkage in the 8681 cohort is *Same address 5 years ago.* This variable scored the highest absolute partial correlation of 0.199, as well as the highest absolute bivariate correlation of 0.324. The other 9 variables on this table have relatively small absolute partial correlations, ranging from 0.021 to 0.055.
However a large proportion of variables in this table differ to those in other tables. For example, the variables *Born in NZ, Maori Ethnicity, Pacific Ethnicity, Defacto Status, work and labour force status, and religious* appeared in table 5, but did not appear in any of the other tables in my results.

Table 6 indicates that the strongest variables in determining whether an individual was linked back from 2006 to 1996 was *Npairs sameaddress*, followed by *Years lived at address from 06* and *Years lived at address from 01*. These three variables were dominant, scoring absolute partial correlations of 0.451, 0.150 and 0.109 respectively. The other 7 variables in this table had relatively small absolute partial correlations, ranging from 0.034 to 0.047.

4.0 Discussion

Our results show that the factors which strongly determine linkage vary across the different two-census cohorts. From looking at my results, variables that quite often appear in the top 10 factors across all the two –census cohorts were:

1. *Same address 5 years ago*
2. *NZ Deprivation score*
3. *Marital status*
4. *Own residence*
5. *European Ethnicity*
6. *Maori descent*
7. *Live with partner*
8. *Sex*.

My results show that the strongest variable that determines linkage across two censuses is *Same address 5 years ago*, which, out of all variables, had the strongest partial correlation with linkage in four out of the five two-census cohorts.

My results also show that the strongest factors that determine linkage across the three censuses, 2006-2001-1996, are predominantly 2006 variables.
The strongest variable in this three census cohort is the variable *Npairs sameaddress*. I had created this variable using the variables *Same address 5 years ago* from the 2006 and 2001 census datasets. The *Npairs sameaddress* variable counts the number of times the individual has lived at the same address five years ago, and takes values 0,1 and 2. However I have established that this variable is superficially strong, because those not linked from 06 to 01 cannot have 01 data, and therefore cannot have valid data for *Same address 5 years ago* in 01. Therefore, those not linked from 06 to 01 cannot achieve a score of 2 on the *Npairs sameaddress* variable. I.e., linkage 0601 partly determines the score on this variable, so it is not suprising that the *Npairs sameaddress* is strongly correlated with linkage 060196.

I faced a few difficulties when analyzing cohorts with more than 2 censuses. In the 2006-2001-1996 cohort, the first difficulty I had come across was that I had all the 2006 as well as 2001 variables to deal with here – twice the number of variables than in the two census cohorts. It only became more difficult when analysing cohorts across four or five censuses. At an attempt to solve this first difficulty, I encountered another as I tried to create new variables that covered the same variables over successive censuses, for example the *Npairs sameaddress* variable. This variable took into account both the 2006 and 2001 *Same address 5 years ago* variable so they became redundant in the analysis and I had less variables to deal with. However I had established that this had resulted in a superficially strong variable, concluding that merging variables across censuses like so was not appropriate.

I also faced the difficulty of not having much information from the 1991 and 1986 censuses which made it difficult to analyse the 9186 and 8681 cohorts. There were no data dictionaries created for these years so it made it difficult to identify variables in the data set, especially the 1986 census data set. Consequently, there were only a handful of variables that I could analyse in these cohorts. This explains why the top 10 variables included in the table for the 8681 cohort differed significantly from the tables – I could only use variables that I was able to identify in the data set.

A disadvantage of using correlations to analyse the strength of these variables in determining linkage is that the correlation coefficients only measure the linear relationship between the two variables. The variable 'age' wasn't included in any of our results because it had a very weak partial correlation with linkage across all cohorts. This is because the relationship between age and linkage is non-linear – the lowest linkage rate occurs around the age of 25, where only about 50% of eligible records are matched, whereas the linkage rate peaks to about 80% around the age of 70. This is due to the life-stage transitions and the mobility levels of these age groups (Statistics NZ, 2013). There may also be other variables in the analysis that have a non-linear relationship with linkage and it would be more appropriate to analyse these variables differently.

There are also other methods, other than partial correlations, that I could have used to identify and rank the factors most strongly determine linkage. For example, Thompson (2009) compared six methods for ranking the strength of associations: *standardized coefficients*, *p-value of Wald chi-square test*, *adequacy*, *c-statistic*, and *information value*, which would also have been appropriate to use in my analysis. These methods were found to not be equivalent – some yielded very different results to others. Furthermore, each method had its own advantages and disadvantages which should be taken into consideration when used. For example, the *p-value of Wald chi-square test* indicated the strength of evidence that the two groups – in terms of our analysis, those linked and those non-linked - differed from each other. However a disadvantage of using p-values is that it does not indicate the magnitude that the two groups differ by. Overall, it is unclear as to which method is the best method. The most sensible solution to make my analysis more reliable would be to use multiple methods to identify and rank the most important factors.

After analysing all fifteen cohorts and finding the extent to which each variable determines linkage, the next step would be to determine how to adjust for this bias in analyses. To do this, weights would have to be created for each variable combination which is calculated as the inverse of the linkage probability. The linked census population would be a more accurate representation of the whole population once these weights are applied.

5.0 References

Statistics New Zealand (2013). Developing a historic longitudinal dataset in New Zealand: A feasibility study. Available from www.stats.govt.nz

Thompson D. (2009). Ranking Predictors in Logistic Regression. Paper D10-2009 http://www.mwsug.org/proceedings/2009/stats/MWSUG-2009-D10.pdf

Appendices

A. Full lists of bivariate and partial correlations (with original variable names) for each cohort. A description of each of these variables is in appendix B.

	Corr with LinkInd0601	abs(corr)
same_addr_5yrs_ago	0.34471	0.34471
yrs_at_addr	0.23216	0.23216
own_residence	0.22449	0.22449
marital_status_legal	-0.17742	0.17742
income_support	-0.15636	0.15636
maori_descent	-0.15344	0.15344
language_indicator	-0.14453	0.14453
smoke	-0.1428	0.1428
income_source_count	0.14116	0.14116
NZDep2006	-0.12907	0.12907
ttl_personal_income	0.11835	0.11835
Mao06	-0.11677	0.11677
Eur06	0.11352	0.11352
unpaid_acty_count	0.10681	0.10681
yrs_in_NZ	0.10308	0.10308
difficulty_acty_count	-0.10039	0.10039
highest_qual	0.10027	0.10027
live_with_partner	0.09392	0.09392
live_with_flatmates	-0.08808	0.08808
disability_ind	-0.08801	0.08801
benefit_income	-0.08575	0.08575
iwi_count	-0.08563	0.08563
children_born	0.0852	0.0852
religious	0.08367	0.08367
age	0.07602	0.07602
Pac06	-0.07451	0.07451
hrs_work_mainjob	-0.07055	0.07055
EthNS	-0.06299	0.06299
travel_work	-0.06036	0.06036
work_labour_force_status	0.05197	0.05197
live_with_children	0.04356	0.04356
Asian06	-0.04279	0.04279
Other06	0.04238	0.04238
sex_female	0.04189	0.04189
adult	0.02732	0.02732
MELAA06	-0.02313	0.02313
live_with_other	-0.02159	0.02159
live_with_siblings	0.00666	0.00666
live_alone	0.00646	0.00646
Language_count	0.00522	0.00522
live_with_parents	0.00466	0.00466

	partial corr with linkind0601	abs(partial corr)
same_addr_5yrs_ago	0.2462	0.2462
nzdep2006	-0.04387	0.04387
own_residence	0.04351	0.04351
marital_status_legal	-0.04348	0.04348
maori_descent	-0.0422	0.0422
Eur06	0.04183	0.04183
yrs_in_NZ	0.0418	0.0418
live_with_partner	0.04165	0.04165
sex_female	0.03963	0.03963
Other06	0.0357	0.0357
income_source_count	0.03556	0.03556
religious	0.03424	0.03424
language_indicator	-0.03345	0.03345
smoke	-0.03328	0.03328
Pac06	-0.02705	0.02705
live_with_children	0.02625	0.02625
benefit_income	-0.02611	0.02611
highest_qual	0.02419	0.02419
live_with_siblings	0.02285	0.02285
difficulty_acty_count	-0.02261	0.02261
live_with_parents	0.02226	0.02226
unpaid_acty_count	0.02017	0.02017
age	0.01824	0.01824
ttl_personal_income	0.01722	0.01722
hrs_work_mainjob	-0.0147	0.0147
Mao06	-0.0143	0.0143
iwi_count	-0.01184	0.01184
live_alone	0.01127	0.01127
live_with_flatmates	-0.01119	0.01119
disability_ind	-0.0072	0.0072
income_support	0.00608	0.00608
adult	-0.00562	0.00562
MELAA06	-0.00492	0.00492
live_with_other	0.00329	0.00329
Language_count	-0.00298	0.00298
EthNS	-0.00206	0.00206
Asian06	0.00098	0.00098
yrs_at_addr	-0.00079	0.00079
travel_work	-0.00079	0.00079

	corr with LinkInd0196	abs(corr)
same_addr_5yrs_ago	0.28444	0.28444
yrs_at_addr	0.26773	0.26773
own_residence	0.22318	0.22318
marital_status_legal	-0.18408	0.18408
income_support	-0.16246	0.16246
Eur01	0.15887	0.15887
language_indicator	-0.14144	0.14144
maori_descent	-0.1326	0.1326
nzdep2001	-0.12192	0.12192
currently_separated	-0.12083	0.12083
income_source_count	0.11996	0.11996
Mao01	-0.11977	0.11977
ttl_personal_income	0.10574	0.10574
benefit_income	-0.10363	0.10363
live_with_partner	0.10137	0.10137
live_with_flatmates	-0.09411	0.09411
unpaid_acty_count	0.08858	0.08858
difficulty_acty_count	-0.08812	0.08812
iwi_count	-0.08679	0.08679
iwi_ind	0.08206	0.08206
highest_qual	0.08203	0.08203
religious	0.08169	0.08169
age	0.07777	0.07777
disability_ind	-0.07719	0.07719
EthNS	-0.07647	0.07647
Pac01	-0.0731	0.0731
disability	-0.06722	0.06722
work_labour_force_status	0.05742	0.05742
work_at_home	0.05492	0.05492
travel_work	-0.05149	0.05149
NZ_born	0.04821	0.04821
sex_female	0.04109	0.04109
live_with_other	-0.04108	0.04108
live_with_children	0.04058	0.04058
hrs_work_mainjob	0.03486	0.03486
yrs_in_NZ	0.02729	0.02729
Asian01	-0.02542	0.02542
adult	0.02531	0.02531
MELAA01	-0.01505	0.01505
live_with_parents	0.00509	0.00509
live_with_siblings	0.00357	0.00357
live_alone	0.00351	0.00351
Other01	-0.00242	0.00242
Language_count	0.00075	0.00075

	partial corr with linkind0196	abs(partial corr)
same_addr_5yrs_ago	0.13189	0.13189
yrs_at_addr	0.06051	0.06051
own_residence	0.0488	0.0488
live_with_partner	0.04744	0.04744
Eur01	0.04685	0.04685
language_indicator	-0.04505	0.04505
nzdep2001	-0.04185	0.04185
maori_descent	-0.04022	0.04022
sex_female	0.03979	0.03979
marital_status_legal	-0.03833	0.03833
religious	0.03319	0.03319
income_source_count	0.03268	0.03268
highest_qual	0.03008	0.03008
unpaid_acty_count	0.02955	0.02955
live_with_children	0.02678	0.02678
live_with_parents	0.0242	0.0242
Pac01	-0.0212	0.0212
live_with_siblings	0.02119	0.02119
ttl_personal_income	0.01889	0.01889
difficulty_acty_count	-0.01805	0.01805
live_alone	0.01588	0.01588
benefit_income	-0.01547	0.01547
iwi_ind	0.01295	0.01295
work_labour_force_status	0.01156	0.01156
travel_work	0.0115	0.0115
NZ_born	0.00949	0.00949
age	0.00901	0.00901
Mao01	-0.00887	0.00887
EthNS	-0.00844	0.00844
Asian01	0.00753	0.00753
disability	-0.00627	0.00627
work_at_home	-0.00601	0.00601
live_with_flatmates	-0.00599	0.00599
income_support	-0.00553	0.00553
iwi_count	0.0049	0.0049
adult	-0.00465	0.00465
hrs_work_mainjob	-0.00448	0.00448
disability_ind	-0.00349	0.00349
currently_separated	-0.0029	0.0029
live_with_other	-0.00235	0.00235
MELAA01	-0.00194	0.00194
Language_count	-0.00166	0.00166
Other01	0.00038	0.00038
yrs_in_NZ	0.00003	0.00003

	corr with LinkInd9691	Abs(corr)
same_addr_5yrs_ago	0.32355	0.32355
yrs_at_addr	0.24187	0.24187
marital_status_legal	-0.17242	0.17242
Eur96	0.15994	0.15994
maori_ancestry	-0.13421	0.13421
smoke	-0.1341	0.1341
live_with_partner	0.11668	0.11668
NZdep1996	-0.11319	0.11319
Language_count	-0.11123	0.11123
Mao96	-0.11013	0.11013
benefit_income	-0.10471	0.10471
income_source_count	0.1009	0.1009
ttl_personal_income	0.09754	0.09754
yrs_in_NZ	0.09624	0.09624
EthNS	-0.0911	0.0911
children_born	0.08734	0.08734
health_problems	-0.08499	0.08499
religious	0.083	0.083
unpaid_work	0.08049	0.08049
disability	-0.07909	0.07909
live_with_children	0.07846	0.07846
live_with_parents	-0.07223	0.07223
longterm_disability	-0.07163	0.07163
Pac96	-0.06996	0.06996
unpaid_acty_count	0.0691	0.0691
highest_qualification	0.06884	0.06884
age	0.06581	0.06581
NZ_born	0.06202	0.06202
work_labour_force_status	0.05415	0.05415
live_with_siblings	-0.04297	0.04297
sex_female	0.03911	0.03911
language_indicator	0.03699	0.03699
Asian96	-0.02621	0.02621
adult	0.01565	0.01565
hrs_work_mainjob	-0.01398	0.01398
MELAA96	-0.00772	0.00772
Other96	-0.0022	0.0022

	partial corr with linkind9691	abs(partial corr)
same_addr_5yrs_ago	0.22453	0.22453
Eur96	0.05063	0.05063
live_with_partner	0.04844	0.04844
marital_status_legal	-0.04802	0.04802
NZdep1996	-0.04511	0.04511
yrs_in_NZ	0.03824	0.03824
maori_ancestry	-0.03755	0.03755
smoke	-0.03713	0.03713
benefit_income	-0.03621	0.03621
sex_female	0.03493	0.03493
hrs_work_mainjob	-0.03481	0.03481
religious	0.03433	0.03433
income_source_count	0.03408	0.03408
live_with_children	0.0272	0.0272
ttl_personal_income	0.02567	0.02567
highest_qualification	0.01969	0.01969
health_problems	-0.01914	0.01914
Language_count	-0.01879	0.01879
Pac96	-0.01877	0.01877
Mao96	-0.01737	0.01737
work_labour_force_status	0.0151	0.0151
yrs_at_addr	0.01375	0.01375
unpaid_work	0.01319	0.01319
age	-0.01196	0.01196
Asian96	0.01178	0.01178
adult	-0.01134	0.01134
longterm_disability	-0.01062	0.01062
EthNS	-0.00728	0.00728
live_with_siblings	0.00719	0.00719
language_indicator	0.00602	0.00602
live_with_parents	-0.005	0.005
disability	-0.00274	0.00274
unpaid_acty_count	0.00239	0.00239
MELAA96	0.00122	0.00122
Other96	0.00026	0.00026

	corr with LinkInd9186	abs(corr)
yrs_at_addr	0.24583	0.24583
Eur91	0.18745	0.18745
live_with_spouse	0.14525	0.14525
Mao91	-0.1339	0.1339
NZdep1991	-0.12809	0.12809
Pac91	-0.09745	0.09745
live_with_other_rel	-0.09564	0.09564
same_addr_5yrs_ago	-0.09539	0.09539
maori_ancestry	-0.08897	0.08897
school_qual	0.07461	0.07461
work_labour_force_status	0.07101	0.07101
age	0.06502	0.06502
EthNS	-0.06087	0.06087
religious	-0.06078	0.06078
live_with_children	0.05578	0.05578
NZ_born	0.04188	0.04188
sex_female	0.03084	0.03084
ttl_work_hrs	0.0296	0.0296
hrs_work_mainjob	0.0288	0.0288
Asian91	-0.0274	0.0274
live_with_parents	0.02643	0.02643
live_with_siblings	0.02603	0.02603
adult	0.00789	0.00789
MELAA91	-0.0055	0.0055
Other91	-0.001	0.001

	Partial corr with linkind9186	abs(partial corr)
yrs_at_addr	0.21459	0.21459
same_addr_5yrs_ago	-0.13025	0.13025
Eur91	0.08953	0.08953
live_with_spouse	0.06961	0.06961
NZdep1991	-0.05387	0.05387
school_qual	0.03789	0.03789
live_with_children	0.03344	0.03344
live_with_parents	0.03255	0.03255
maori_ancestry	-0.03095	0.03095
sex_female	0.02686	0.02686
EthNS	0.02679	0.02679
work_labour_force_stat	0.02658	0.02658
live_with_siblings	0.02316	0.02316
age	0.02024	0.02024
adult	-0.01433	0.01433
NZ_born	0.01257	0.01257
religious	-0.0092	0.0092
Asian91	0.00885	0.00885
live_with_other_rel	-0.00874	0.00874
MELAA91	0.00258	0.00258
hrs_work_mainjob	-0.00256	0.00256
Mao91	-0.00117	0.00117
Other91	0.00087	0.00087

	corr with LinkInd8681	abs(corr)
same_addr_5yrs_ago	0.32389	0.32389
yrs_at_addr	0.26859	0.26859
Eur86	0.16544	0.16544
Mao86	-0.12183	0.12183
defacto_status	-0.0855	0.0855
EthNS	-0.08094	0.08094
Pac86	-0.06629	0.06629
Religious	0.05885	0.05885
age	0.053	0.053
NZ_Born	0.04354	0.04354
work_labour_force_status	0.03662	0.03662
yrs_in_NZ	0.02489	0.02489
sex_female	0.01979	0.01979
Asian86	-0.01647	0.01647
labour_force_dummy	0.01223	0.01223
ttl_work_hrs	0.00741	0.00741
hrs_work_mainjob	0.00531	0.00531
Other86	-0.00425	0.00425
MELAA86	-0.00292	0.00292
adult	0.00148	0.00148

	partial corr with linkind8681	abs(partial corr)
same_addr_5yrs_ago	0.199	0.199
Eur86	0.05502	0.05502
yrs_at_addr	0.05378	0.05378
NZ_Born	0.04404	0.04404
yrs_in_NZ	0.04304	0.04304
Mao86	-0.03667	0.03667
Pac86	-0.02716	0.02716
defacto_status	-0.02656	0.02656
work_labour_force_status	0.02383	0.02383
Religious	0.02069	0.02069
age	-0.01865	0.01865
sex_female	0.01772	0.01772
Asian86	0.00715	0.00715
ttl_work_hrs	0.00645	0.00645
hrs_work_mainjob	-0.00345	0.00345
MELAA86	0.00223	0.00223
EthNS	0.00171	0.00171
adult	0.00157	0.00157
Other86	0.00014	0.00014

	corr with linkind060196	abs(corr)
linkind060196	1	1
LinkInd0196	1	1
LinkInd0601	0.69894	0.69894
npairs_sameaddress	0.50011	0.50011
same_addr_5yrs_ago06	0.30823	0.30823
same_addr_5yrs_ago01	0.29018	0.29018
yrs_at_addr06	0.26636	0.26636
yrs_at_addr01	0.25418	0.25418
own_residence06	0.21478	0.21478
marital_status_legal06	-0.17619	0.17619
own_residence01	0.15709	0.15709
income_support06	-0.1473	0.1473
maori_descent06	-0.14622	0.14622
smoke06	-0.13929	0.13929
marital_status_legal01	-0.13672	0.13672
Mao06	-0.13435	0.13435
age_code_num_06	0.13422	0.13422
income_source_count06	0.13093	0.13093
nzdep2006	-0.12282	0.12282
language_indicator06	-0.12184	0.12184
Eur06	0.10673	0.10673
Eur01	0.10615	0.10615
age_code_num_01	0.10517	0.10517
benefit_income06	-0.10015	0.10015
iwi_count06	-0.09872	0.09872
income_support01	-0.09852	0.09852
age06	0.0939	0.0939
Mao01	-0.09325	0.09325
religious06	0.08907	0.08907
ttl_personal_income06	0.08781	0.08781
nzdep2001	-0.0877	0.0877
live_with_flatmates01	-0.08733	0.08733
benefit_income01	-0.08633	0.08633
live_with_partner06	0.0838	0.0838
live_with_flatmates06	-0.08141	0.08141
maori_descent01	-0.07954	0.07954
difficulty_acty_count06	-0.07895	0.07895
Pac06	-0.0759	0.0759
highest_qual06	0.0755	0.0755
language_indicator01	-0.07413	0.07413
age01	0.07347	0.07347
iwi_count01	-0.06977	0.06977
yrs_in_NZ06	0.06836	0.06836
iwi_ind01	0.06817	0.06817
currently_separated01	-0.06764	0.06764
income_source_count01	0.06553	0.06553
disability_ind06	-0.06433	0.06433

	Partial corr with linkind060196	abs(partial corr)
npairs_sameaddress	0.45097	0.45097
yrs_at_addr06	-0.15048	0.15048
yrs_at_addr01	0.10848	0.10848
sex_female06	0.04723	0.04723
marital_status_legal06	-0.04123	0.04123
nzdep2006	-0.0409	0.0409
marital_status_legal01	-0.04065	0.04065
Eur06	0.03864	0.03864
live_with_partner06	0.03609	0.03609
maori_descent06	-0.03422	0.03422
religious06	0.03382	0.03382
income_source_count06	0.03336	0.03336
Pac06	-0.03303	0.03303
Other06	0.03281	0.03281
own_residence01	0.03162	0.03162
language_indicator06	-0.03114	0.03114
smoke06	-0.02966	0.02966
live_with_children01	0.02945	0.02945
own_residence06	0.02587	0.02587
yrs_in_NZ06	0.02577	0.02577
benefit_income01	-0.02413	0.02413
income_source_count01	0.02412	0.02412
benefit_income06	-0.02404	0.02404
Eur01	0.02389	0.02389
highest_qual06	0.02257	0.02257
language_indicator01	-0.02162	0.02162
maori_descent01	-0.02122	0.02122
live_with_parents01	0.02056	0.02056
highest_qual01	0.01996	0.01996
difficulty_acty_count06	-0.01948	0.01948
live_with_siblings01	0.01904	0.01904
live_with_partner01	0.01882	0.01882
religious01	0.01878	0.01878
nzdep2001	-0.01788	0.01788
unpaid_acty_count01	0.01702	0.01702
ttl_personal_income06	0.01663	0.01663
Mao06	-0.01612	0.01612
live_with_siblings06	0.01417	0.01417
live_with_flatmates01	-0.014	0.014
unpaid_acty_count06	0.01307	0.01307
work_labour_force_status(	0.01171	0.01171
difficulty_acty_count01	-0.01148	0.01148
live_with_children06	0.01114	0.01114
hrs_work_mainjob06	-0.01075	0.01075
travel_work01	0.01073	0.01073
ttl_personal_income01	0.01052	0.01052
iwi_count06	-0.01006	0.01006

unpaid_acty_count06	0.06266	0.06266
live_with_partner01	0.05833	0.05833
Pac01	-0.05772	0.05772
religious01	0.05665	0.05665
EthNS2006	-0.04805	0.04805
sex_female06	0.0467	0.0467
difficulty_acty_count01	-0.03857	0.03857
ttl_personal_income01	0.03751	0.03751
travel_work06	-0.03463	0.03463
Other06	0.03412	0.03412
live_with_other01	-0.03235	0.03235
unpaid_acty_count01	0.03176	0.03176
disability_ind01	-0.02678	0.02678
highest_qual01	0.02677	0.02677
live_with_other06	-0.0262	0.0262
sex_female01	0.0258	0.0258
adult06	0.02518	0.02518
work_labour_force_status01	0.02499	0.02499
EthNS2001	-0.02493	0.02493
Language_count01	-0.02228	0.02228
adult01	0.02165	0.02165
Asian06	-0.02133	0.02133
years_in_nz_code_num_06	-0.02119	0.02119
disability01	-0.02034	0.02034
Asian01	-0.0189	0.0189
live_with_children01	0.0187	0.0187
yrs_in_NZ01	0.01674	0.01674
work_at_home01	0.01549	0.01549
MELAA06	-0.01361	0.01361
live_alone06	0.01179	0.01179
travel_work01	-0.00943	0.00943
hrs_work_mainjob01	-0.00845	0.00845
NZ_born01	0.00828	0.00828
MELAA01	-0.00824	0.00824
Language_count06	-0.00715	0.00715
years_in_NZ_code_num_01	-0.00598	0.00598
live_with_children06	0.00542	0.00542
live_alone01	0.00514	0.00514
hrs_work_mainjob06	0.00479	0.00479
live_with_parents06	0.00219	0.00219
live_with_parents01	-0.00172	0.00172
Other01	-0.00094	0.00094
live_with_siblings01	0.00085	0.00085
live_with_siblings06	0.00033	0.00033

iwi_ind01	0.00777	0.00777
Language_count06	-0.00767	0.00767
live_alone06	0.00659	0.00659
disability_ind06	-0.00658	0.00658
live_alone01	0.00641	0.00641
live_with_parents06	0.00632	0.00632
income_support06	0.00575	0.00575
adult01	-0.00567	0.00567
work_at_home01	-0.00566	0.00566
hrs_work_mainjob01	-0.0055	0.0055
iwi_count01	0.00534	0.00534
age06	0.00496	0.00496
live_with_other01	-0.00396	0.00396
EthNS2006	-0.00395	0.00395
live_with_other06	0.00394	0.00394
EthNS2001	-0.00301	0.00301
yrs_in_NZ01	-0.00268	0.00268
NZ_born01	0.00246	0.00246
disability01	-0.00233	0.00233
live_with_flatmates06	-0.00192	0.00192
Asian06	0.00176	0.00176
travel_work06	0.00156	0.00156
disability_ind01	-0.00097	0.00097
currently_separated01	0.00096	0.00096
Other01	0.00077	0.00077
MELAA06	-0.00038	0.00038
Language_count01	-0.0003	0.0003
MELAA01	0.0003	0.0003
adult06	-0.00008	0.00008

B. Descriptions of variables in Appendix A

Variable name	Description
LinkindXXXX	Whether or not the individual was linked back to the previous census (or censuses), where XXXX is in {0601,0196, 9691, 9186, 8681, 060196}
adult	Whether or not the individual is an adult
age	The age of the individual
AsianXX	Whether or not the individual is Asian in the year XX
benefit_income	whether or not the individual is on a benefit
children_born	the number of children the individual has given birth to
currently_separated	whether or not the individual is separated from their partner
defacto_status	the individual's defacto status
difficulty_acty_count	number of activities the individual has difficulties with
disability	whether or not the individual has a long term disability (lasting 6 months or more)
disability_ind, longterm_disability	whether or not the individual indicated any difficulties with an activity
EthNS	Whether or not the individual indicated their ethnicity
EurXX	whether or not the individual is European in the year XX
health_problems	Whether or not the individual indicated any difficulties with an activity
highest_qual, highest_qualification	the individual's highest qualification
hrs_work_mainjob	number of hours the individual works in main job
income_source_count	number of income sources the individual has
income_support	whether or not the individual is on income support
iwi_count	the number of iwi the individual has
iwi_ind	whether or not the individual specified if they have any iwi
Language_count	The number of languages the individual speaks
language_indicator	whether the individual speaks just English, just Maori, both or neither
live_alone	whether or not the individual lives alone
live_with_children	whether or not the individual lives with their children
live_with_flatmates	whether or not the individual lives with flatmates
live_with_other	whether or not the individual lives with people other than relatives and flatmates
live_with_other_rel	Whether or not the individual lives with relatives other than parents, children, siblings or partner.
live_with_parents	whether or not the individual lives with their parents
live_with_partner, live with spouse	whether or not the individual lives with their partner
live_with_siblings	whether or not the individual lives with their siblings
maori_ancestry, maori_descent	whether or not the individual is of Maori descent
MaoXX	Whether or not the individual is Maori in the year XX

marital_status_legal	The individual's marital status
MELAA01	whether or not the individual is Middle Eastern/Latin American/African in XX
NZ_born	Whether or not the individual is born in NZ
NZdepXXXX	The individual's NZ Deprivation Score in the year XXXX
OtherXX	Whether the individual specified an ethnicity other than European, Pacific, Asian, MELAA, Maori
own_residence	Whether or not the individual owns their residence
Pac01	Whether or not the individual is pacific in year XX
religious	whether or not the individual is religious
same_addr_5yrs_ago	whether or not the individual lived at the same address 5 years ago, or didn't specify
school_qual	whether or not the individual has a school qualification
sex_female	whether or not the individual is female
smoke	whether or not the individual smokes
travel_work	How the individual travels to work
ttl_personal_income	The individual's personal income
ttl_work_hrs	The total number of hours the individual worked
unpaid_acty_count	the number or unpaid activities the individual has done in the last 7 days
unpaid_work	whether or not the individual did any unpaid work over the last 7 days
work_at_home	whether or not the individual works at home
work_labour_force_status	The individual's work and labour force status
yrs_at_addr	The number of years the individual has lived at their current address

C. SAS Program coding sample: Cohort 0601

```
libname ver2 'U:\DataLab\MA\MAA2013-18 Linkage Bias Longitudinal
Census\Vera\Updated data dec 2013';
libname vernzdep 'U:\DataLab\RO\MAA2013-18 Linkage Bias Longitudinal
Census\Updated data dec 2013'; *CHANGE FOR OTHER COHORTS TOO*Use for
NZDep;
libname created 'U:\DataLab\MA\MAA2013-18 Linkage Bias Longitudinal
Census\Vera\Created data sets';
data aa;
set ver2.individual_linkind0601_2006;
run;
data datasex (keep=ID sex);
set ver2.spine_dataset;
run;
data dataeth (keep=ID eur06_06 Mao06_06 Pac06_06 Asian06_06 MELAA06_06
Other06_06 EthNS06);
set ver2.ethnicity_info;
run;


data datanzdep (keep=MB06 nzdep2006);
set vernzdep.nzdep2006_mb06;
nzdep2006=nzdep2006_mb06;
if nzdep2006=. then nzdep2006=99;
run;
/*proc freq data=datanzdep;
tables nzdep2006;
run;*/

/*data datanzdep (keep=mb06 nzdep2006); *adding this new dataset to
the work library;
set vernzdep.nzdep06_mb06;
if nzdep2006=. then nzdep2006=99; *meshblocks with no NZDep have a
NZDep score of 99;
run;*/
data datamb (keep=id mb06); *adding this new dataset to the work
library;
set ver2.geogr_linkspine_0601;
mb06=URSmb06_06;
run;
proc sort data=datamb; by mb06; run;
proc sort data=datanzdep; by mb06; run;
data data_mb_nzdep; *merging to form a dataset with ID mb06 nzdep2006;
merge datamb datanzdep;
by mb06;
run;
proc sort data=data_mb_nzdep;by ID;run;
data data2_mb_nzdep; *removing those meshblocks which no IDs live in;
```

```
set data mb_nzdep;
if id ~=.;
run;


proc sort data=aa;by ID;run;
proc sort data=datasex;by ID;run;
proc sort data=dataeth;by ID;run;
proc sort data=data2_mb_nzdep;by ID;run;


data aa2; *merging all datasets together;
merge aa datasex dataeth data2_mb_nzdep;
by ID;
run;
data aa3; *reducing dataset down to those that exist in 2006;
set aa2;
if Linkind0601~=.;
run;

*******;


proc contents data=aa3 position;
run;

proc freq data=aa3;
tables
LinkInd0601      age_code_06      individual_rec_type_code_06
      years_at_addr_code_06 addr_5years_ago_code_06
      birth_country_code_06 years_in_nz_code_06
      official_language_code_06   languages_count_code_06
      maori_descent_code_06 recode_maori_descent_code_06
      iwi_ind_code_06  iwi_count_code_06      religion1_code_06
      living_arrangement1_code_06 living_arrangement2_code_06
      living_arrangement3_code_06 living_arrangement4_code_06
      living_arrangement5_code_06 living_arrangement6_code_06
      living_arrangement7_code_06 living_arrangement8_code_06
      living_arrangement9_code_06 living_arrangement10_code_06
      living_arrangement11_code_06      family_grp_code_06
      family_role_code_06    related_family_grp_code_06
      related_family_role_code_06 recode_family_grp_code_06
      recode_family_role_code_06  difficulty_acty1_code_06
      difficulty_acty2_code_06    difficulty_acty3_code_06
      difficulty_acty4_code dsblty_code_06   dsblty_ind_code_06
      smoking_status_code_06      tenure_holder_code_06
      legal_marital_status_code_06      social_marital_status_code_06
      fertility_code_06      child_depend_code_06  highest_qual_code_06
      std_highest_qual_code_06    income_srce_count_code_06
      income_srce1_code_06  income_srce2_code_06  income_srce3_code_06
      income_srce4_code_06  income_srce5_code_06  income_srce6_code_06
      income_srce7_code_06  income_srce8_code_06  income_srce9_code_06
```

```
      income_srce10_code_06 income_srce11_code_06 income_srce12_code_06
      income_srce13_code_06 income_srce14_code_06 total_income_code_06
      income_support_count_code_06     job_ind_code_06
      emp_status_code_06    wklfs_code_06    sect_code_06
      workplace_addr_ind_code_06 work_hrs1_code_06
      travel_work_code_06   unpaid_acty1_code_06  unpaid_acty2_code_06
      unpaid_acty3_code_06  unpaid_acty4_code_06  unpaid_acty5_code_06
      unpaid_acty6_code_06  unpaid_acty7_code_06  Sex   Eur06_06
      Mao06_06   Pac06_06   Asian06_06 MELAA06_06 Other06_06 EthNS06
      nzdep2006  ;
run;

data aa4;
set aa3;
age=input(age_code_06,f4.0);
if 0<=age<=14 then age=0; if 15<=age<=19 then age=1; if 20<=age<=24
then age=2; if 25<=age<=29 then age=3; if 30<=age<=49 then age=4; if
50<=age<=64 then age=5; if 65<=age<=74 then age=6; if 75<=age<=84 then
age=7; if 85<=age<=116 then age=8;
adult=0;
if individual_rec_type_code_06='3' then adult=1;
yrs_at_addr=input(years_at_addr_code_06,F4.0);
if 1<=yrs_at_addr<=5 then yrs_at_addr=1; if 6<=yrs_at_addr<=10 then
yrs_at_addr=2; if 11<=yrs_at_addr<=20 then yrs_at_addr=3; if
21<=yrs_at_addr<=98 then yrs_at_addr=4; if yrs_at_addr in(777,999)
then yrs_at_addr=9;
same_addr_5yrs_ago=input(addr_5years_ago_code_06,f4.0);
if same_addr_5yrs_ago=2 then same_addr_5yrs_ago=0; if
same_addr_5yrs_ago in (5,99) then same_addr_5yrs_ago=9;
NZ_born=0;
if birth_country_code_06='1201' then NZ_born=1; if
birth_country_code_06='9999' then NZ_born=9;
yrs_in_NZ=input(years_in_NZ_code_06,f4.0);
if 0<=yrs_in_NZ<=9 then yrs_in_NZ=1; if 10<=yrs_in_NZ<=19 then
yrs_in_NZ=2; if 20<=yrs_in_NZ<=97 then yrs_in_NZ=3; if yrs_in_NZ=888
then yrs_in_NZ=4; if yrs_in_NZ in (777,999) then yrs_in_NZ=9;
offic_lang_dummy=input(official_language_code_06,f4.0);
if offic_lang_dummy in (2,5) then language_indicator=1; if
offic_lang_dummy in (1,3,4,6) then language_indicator=2; if
offic_lang_dummy in (7,8) then language_indicator=3; if
offic_lang_dummy in (97,98,99) then language_indicator=9;
Language_count=input(languages_count_code_06,f4.0);
if 3<=language_count<=6 then language_count=3; if 7<=language_count<=9
then language_count=9;
maori_descent= input(recode_maori_descent_code_06, f4.0);
if maori_descent=2 then maori_descent=0; if maori_descent in (4,7,9)
then maori_descent=9;
iwi_ind=input(iwi_ind_code_06,f4.0);
if iwi_ind=2 then iwi_ind=0; if iwi_ind=7 then iwi_ind=9;
iwi_count=input(iwi_count_code_06,f4.0);
if 3<=iwi_count<=5 then iwi_count=3; if iwi_count in (9) then
iwi_count=0; if iwi_count in (7,8) then iwi_count=9;
```

```
*Created the variable so that if not maori, then iwi_count=0;
religious=1;
if religion1_code_06='00000' then religious=0; if religion1_code_06 in
('94444','95555','97777','98888','99999') then religious=9;
live_with_parents=0;
if living_arrangement6_code_06='0411' then live_with_parents=1;
live_with_partner=0;
if living_arrangement1_code_06='0211' then live_with_partner=1; if
living_arrangement2_code_06='0211' then live_with_partner=1; if
living_arrangement3_code_06='0212' then live_with_partner=1; if
living_arrangement4_code_06='0211' then live_with_partner=1; if
living_arrangement5_code_06='0212' then live_with_partner=1;
live_with_siblings=0;
if living_arrangement8_code_06='0511' then live_with_siblings=1;
live_with_children=0;
if living_arrangement7_code_06='0311' then live_with_children=1;
live_alone=0;
if living_arrangement11_code_06='0111' then live_alone=1;
live_with_flatmates=0;
if living_arrangement9_code_06='0611' then live_with_flatmates=1;
live_with_other=0;
if living_arrangement10_code_06='7777' then live_with_other=1;
difficulty_acty_count=0;
if difficulty_acty1_code_06='1' then
difficulty_acty_count=difficulty_acty_count+1;
if difficulty_acty2_code_06='2' then
difficulty_acty_count=difficulty_acty_count+1;
if difficulty_acty3_code_06='3' then
difficulty_acty_count=difficulty_acty_count+1;
if difficulty_acty4_code='4' then
difficulty_acty_count=difficulty_acty_count+1;
if difficulty_acty_count in (4,0) then difficulty_acty_count=9;
if difficulty_acty4_code='4' then difficulty_acty_count=0;
disability=0;
if dsblty_code_06='1' then disability=1; if dsblty_code_06 in ('7','9')
then disability=9;
disability_ind=input(dsblty_ind_code_06,f4.0);
if disability_ind=7 then disability_ind=9;
smoke=input(smoking_Status_code_06, f4.0);
if smoke in (2,3) then smoke=0; if smoke in (7,9) then smoke=9;
own_residence=input(tenure_holder_code_06,f4.0);
if own_residence=2 then own_residence=0; if own_residence in (7,9)
then own_residence=9;
marital_status_legal=input(legal_marital_status_code_06,f4.0);
if marital_status_legal in (2) then marital_status_legal=0; if
marital_status_legal in (1,3) then marital_status_legal=3; if
marital_status_legal in (6) then marital_status_legal=1; if
marital_status_legal in (4,5) then marital_status_legal=2; if
marital_status_legal in (7,9) then marital_status_legal=9;
children_born=input(fertility_code_06,f4.0);
if 2<=children_born<=10 then children_born=2; if children_born in
(55,77,88,99) then children_born=9;
```

```
had_children=children_born;
if children_born=2 then had_children=1;
highest_qual=input(highest_qual_code_06, f4.0);
if highest_qual in (1,2,3,4) then highest_qual=1; if highest_qual in
(5,6,7,8,9,10) then highest_qual=2; if highest_qual in (11,12,13,14)
then highest_qual=3; if highest_qual in (97,99) then highest_qual=9;
income_source_count=input(income_srce_count_code_06, f4.0);
if 3<=income_source_count<=10 then income_source_count=3; if
income_source_count=99 then income_source_count=9;
benefit_income=0;
if income_srce7_code_06='07' then benefit_income=1; if
income_srce8_code_06='08' then benefit_income=1; if
income_srce9_code_06='09' then benefit_income=1; if
income_srce10_code_06='10' then benefit_income=1; if
income_srce11_code_06='11' then benefit_income=1; if
income_srce12_code_06='12' then benefit_income=1;
if adult=0 then benefit_income=.;
ttl_personal_income=input(total_income_code_06,f4.0);
if ttl_personal_income in (1,2) then ttl_personal_income=0; if
3<=ttl_personal_income<=8 then ttl_personal_income=1; if
9<=ttl_personal_income<=14 then ttl_personal_income=2;  if
ttl_personal_income=99 then ttl_personal_income=9;
income_support=input(income_support_count_code_06,f4.0);
if 1<=income_support<=6 then income_support=1;
hrs_work_mainjob=input(work_hrs1_code_06,f4.0);
if 1<=hrs_work_mainjob<=29 then hrs_work_mainjob=1; if
30<=hrs_work_mainjob<=49 then hrs_work_mainjob=2; if
50<=hrs_work_mainjob<=168 then hrs_work_mainjob=3; if hrs_work_mainjob
in (777) then hrs_work_mainjob=9; if hrs_work_mainjob=999 then
hrs_work_mainjob=0;
if wklfs_code_06='3' then work_labour_force_status=0; if
wklfs_code_06='4' then work_labour_force_status=1; if
wklfs_code_06='1' then work_labour_force_status=2;  if
wklfs_code_06='2' then work_labour_force_status=3; if
wklfs_code_06='9' then work_labour_force_status=9;
travel_work=input(travel_work_code_06,f4.0);
if travel_work in (1,2) then travel_work=0; if travel_work in (3,4,5,8)
then travel_work=1; if travel_work in (6,7) then travel_work=2; if
travel_work in (9,10) then travel_work=3; if travel_work in (15,77,99)
then travel_work=9;
unpaid_acty_count=0;
if unpaid_acty1_code_06='01' then
unpaid_acty_count=unpaid_acty_count+1;
if unpaid_acty2_code_06='02' then
unpaid_acty_count=unpaid_acty_count+1;
if unpaid_acty3_code_06='03' then
unpaid_acty_count=unpaid_acty_count+1;
if unpaid_acty4_code_06='04' then
unpaid_acty_count=unpaid_acty_count+1;
if unpaid_acty5_code_06='05' then
unpaid_acty_count=unpaid_acty_count+1;
```

```
if unpaid_acty6_code_06='06' then
unpaid_acty_count=unpaid_acty_count+1;
if unpaid_acty_count=0 then unpaid_acty_count=9;
if unpaid_acty7_code_06='00' then unpaid_acty_count=0;
if adult=0 then unpaid_acty_count=.;
if unpaid_acty_count in (3,4,5,6) then unpaid_acty_count=3;
sex_female=0;
if sex='2' then sex_female=1;
Eur06=input(Eur06_06,f4.0);
if Eur06=. then Eur06=0;
Mao06=input(Mao06_06,f4.0);
if Mao06=. then mao06=0;
Pac06=input(Pac06_06,f4.0);
if Pac06=. then Pac06=0;
Asian06=input(Asian06_06,f4.0);
if Asian06=. then Asian06=0;
MELAA06=input(MELAA06_06,f4.0);
if MELAA06=. then MELAA06=0;
Other06=input(Other06_06,f4.0);
if Other06=. then Other06=0;
EthNS=input(ethNS06,f4.0);
if EthNS=. then EthNS=0;
run;

proc contents data=aa4 position;
run;

proc freq data=aa4;
tables
age   adult yrs_at_addr      same_addr_5yrs_ago    NZ_born    yrs_in_NZ
      language_indicator    Language_count  maori_descent    iwi_ind
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability disability_ind   smoke own_residence
      marital_status_legal  children_born    had_children
      highest_qual     income_source_count   benefit_income
      ttl_personal_income   income_support   hrs_work_mainjob
      work_labour_force_status    travel_work      unpaid_acty_count
      sex_female eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      nzdep2006  ;
run;
proc freq data=aa4;
tables
(age  adult yrs_at_addr      same_addr_5yrs_ago    NZ_born    yrs_in_NZ
      language_indicator    Language_count  maori_descent    iwi_ind
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability disability_ind   smoke own_residence
      marital_status_legal  children_born    had_children
      highest_qual     income_source_count   benefit_income
```

```
      ttl_personal_income   income_support   hrs_work_mainjob
      work_labour_force_status    travel_work      unpaid_acty_count
      sex_female eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      nzdep2006)
*LinkInd0601;
run;
*******;
*proc freq data=aa4;
*tables age*adult;
*run;

*proc corr data=aa4;
*var disability disability_ind;
*run;

*proc corr data=aa4;
*var disability linkind0601;
*run;


***LABELS FOR data=aa2 (aa4):
Age: 0-14yo=0, 15-19yo=1, 20-24yo=2, 25-29yo=3, 30-49yo=4, 50-64yo=5,
65-75yo=6, 75-85yo=7, 85+yo=8
?individual_rec_type_cod_06: adult=3 child=4
yrs_at_addr: 0yrs=0 1-5yrs=1 6-10yrs=2 10-20yrs=3 21+yrs=4
NotSpecified=9(-1)
same_addr_5yrs_ago: yes=1 no=0 notspecified=9
NZ_born: yes=1 no=0 notspecified=9(-1)
yrs_in_nz: 0-9yrs=1 10-19yrs=2 20+yrs=3 NotSpecified=9(4)
language_indicator: english-no-maori=1 speak-maori=2 neither-maori-
nor-english=0 NotSpecified=9(4)
language_count: 0language=0 1language=1 2language=2 3+language=3
NotSpecified=9(-1)
maori_descent: yes=1 no=0 NotSpecified=9(2)
?????????iwi_ind: yes=1 no=9 not specified/not maori=9
iwi_count: None=0 1iwi=1 2iwi=2 3-5iwi=3 Notspecified=9
religious: yes=1 no=0 NotSpecified=9(-1)
live_with_parents, Live_with_children, Live_with_siblings,
live_with_partner, live_alone: yes=1 no=0 notSpecified=9
difficulty_acty_count:
disability: yes=1 no=0
disability_ind:
smoke: smoker=1 non-smoker=0 NotSpecified=9
own_residence: own(incl.partly-own)residence=1 don't-own-residence=0
NotSpecified=9
marital_status_legal: currently-married=0 single(widowed)=1
single(incl.divorced)=2 never-married=3 NotSpecified=9
children_born: none=0 1child=1 2+children=2 NotSpecified-or-male=9
child_depend_code_06: dependentchild=1 adultchild=2
highest_qual: none=0 school-qualification=1 post-school-
qualification=2 bachelor/higher-degree=3 NotSpecified=9(-1)
```

```
*income_source_count: no-source=0 1-source=1 2-sources=2 3+sources=3
NotSpecified=9(-1)
-(almost the same as below) benefit_income: yes=1 no=0 NotSpecified=9
ttl_personal_income: loss/$0=0 $1-$30,000=1 $30,001+=2
NotSpecified=9(-1)
-income_support: yes=1 no=0 NotSpecified=9
hrs_worked_mainjob: 1-29hrs=1 30-49hrs=2 50-168hrs=3 NotSpecified=9
work_labour_force_status: unemployed=0 Not-in-labour-force=1 employed-
full-time=2 employed-part-time=3  NotSpecified=9
travel_work: stay-home=0 car/motorbike=1 public-transport=2 walk=3
NotSpecified=9
unpaid_acty_count: 0unpaid-activities/Unidentifiable/NotSpecified=0
1unpaid-activity=1 2unpaid-activities=2 3+unpaid-activities=3
;

*proc freq data= aa4 (where =(adult=0));
*tables age_code_06;
*run;

data aa5;
set aa4;
if yrs_at_addr=9 then yrs_at_addr=-1;
if same_addr_5yrs_ago=9 then same_addr_5yrs_ago=-1;
if NZ_born=9 then NZ_born=-1;
if yrs_in_nz=9 then yrs_in_nz=-1; *how would changing it from 5 to -1
affect it?;
if language_indicator=9 then language_indicator=4;
if language_count=9 then Language_count=-1;
if maori_descent=9 then maori_descent=2;
if iwi_ind=9 then iwi_ind=2;
if iwi_count=9 then iwi_count=4;
if religious=9 then religious=-1;
if difficulty_acty_count=9 then difficulty_acty_count=4;
if disability=9 then disability=2;
if disability_ind=9 then disability_ind=2;
if smoke=9 then smoke=2;
if own_residence=9 then own_residence=-1;
if marital_status_legal=9 then marital_status_legal=4;
if children_born=9 then children_born=-1;
if had_children=9 then had_children=-1;
if highest_qual=9 then highest_qual=-1;
if income_source_count=9 then income_source_count=-1;
if ttl_personal_income=9 then ttl_personal_income=-1;
if income_support=9 then income_support=2;
if hrs_work_mainjob=9 then hrs_work_mainjob=4;
if travel_work=9 then travel_work=4;
if unpaid_acty_count=9 then unpaid_acty_count=-1;
if nzdep2006=99 then nzdep2006=11;
run;

*proc freq data=aa5;
```

```
*tables adult * (benefit_income unpaid_act_count disability iwi_count
live_alone live_with_children live_with_parents live_with_partner
live_with_siblings religious) ;
*run;

*correlations;
proc corr data=aa5;
var
linkInd0601
age   adult yrs_at_addr      same_addr_5yrs_ago    NZ_born    yrs_in_NZ
      language_indicator    Language_count   maori_descent    iwi_ind
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability disability_ind   smoke own_residence
      marital_status_legal  children_born    had_children
      highest_qual     income_source_count   benefit_income
      ttl_personal_income   income_support   hrs_work_mainjob
      work_labour_force_status    travel_work      unpaid_acty_count
      sex_female Eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      NZDep2006;
run;

data aa6;
set aa5 (drop= NZ_born iwi_ind had_children children_born disability
work_labour_force_status);
run;



*partial correlations;
proc corr data=aa6;
partial
      adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var age LinkInd0601;
run;
proc corr data=aa6;
partial
age         yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var adult LinkInd0601;
```

```
run;
proc corr data=aa6;
partial
age   adult      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var yrs_at_addr LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr                 yrs_in_NZ  language_indicator
      Language_count   maori_descent          iwi_count  religious
      live_with_parents     live_with_partner     live_with_siblings
      live_with_children    live_alone live_with_flatmates
      live_with_other  difficulty_acty_count       disability_ind
      sex_female Eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      NZDep2006;
var same_addr_5yrs_ago LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var yrs_in_NZ LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      Language_count   maori_descent          iwi_count  religious
      live_with_parents     live_with_partner     live_with_siblings
      live_with_children    live_alone live_with_flatmates
      live_with_other  difficulty_acty_count       disability_ind
      sex_female Eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      NZDep2006;
var language_indicator LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator          maori_descent          iwi_count
      religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
```

```
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var Language_count LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count              iwi_count
      religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var maori_descent LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var iwi_count LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count        live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var religious LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious       live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var live_with_parents LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
```

```
      iwi_count  religious  live_with_parents
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var live_with_partner LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_children    live_alone live_with_flatmates
      live_with_other  difficulty_acty_count       disability_ind
      sex_female Eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      NZDep2006;
var live_with_siblings LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings          live_alone live_with_flatmates
      live_with_other  difficulty_acty_count       disability_ind
      sex_female Eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      NZDep2006;
var live_with_children LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var live_alone LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_other  difficulty_acty_count       disability_ind
      sex_female Eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      NZDep2006;
var live_with_flatmates LinkInd0601;
run;
proc corr data=aa6;
```

```
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates         difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var live_with_other LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other             disability_ind
      sex_female Eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      NZDep2006;
var difficulty_acty_count LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      sex_female Eur06 Mao06 Pac06 Asian06    MELAA06    Other06    EthNS
      NZDep2006;
var disability_ind LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind        Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var sex_female LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female       Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
```

```
var Eur06 LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06      Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006;
var Mao06 LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06      Asian06    MELAA06
      Other06    EthNS NZDep2006;
var Pac06 LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06      MELAA06
      Other06    EthNS NZDep2006;
var Asian06 LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06
      Other06    EthNS NZDep2006;
var MELAA06 LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
```

```
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      EthNS NZDep2006;
var Other06 LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06          NZDep2006;
var EthNS LinkInd0601;
run;
proc corr data=aa6;
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS ;
var NZDep2006 LinkInd0601;
run;
*variables just for adults
smoke
own_residence
marital_status_legal


highest_qual
income_source_count
benefit_income
ttl_personal_income
income_support
hrs_work_mainjob

travel_work
unpaid_acty_count
;

*for questions answered by just adults;

proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
```

```
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006        own_residence
      marital_status_legal              highest_qual
      income_source_count   benefit_income   ttl_personal_income
      income_support   hrs_work_mainjob      travel_work
      unpaid_acty_count     ;
var smoke LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke      marital_status_legal
      highest_qual     income_source_count   benefit_income
      ttl_personal_income   income_support   hrs_work_mainjob
      travel_work      unpaid_acty_count     ;
var own_residence LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      highest_qual     income_source_count   benefit_income
      ttl_personal_income   income_support   hrs_work_mainjob
      travel_work      unpaid_acty_count     ;
var marital_status_legal LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      marital_status_legal                  income_source_count
      benefit_income   ttl_personal_income   income_support
      hrs_work_mainjob      travel_work      unpaid_acty_count     ;
var highest_qual LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
```

```
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      marital_status_legal             highest_qual
      benefit_income   ttl_personal_income   income_support
      hrs_work_mainjob      travel_work     unpaid_acty_count     ;
var income_source_count LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      marital_status_legal             highest_qual
      income_source_count         ttl_personal_income   income_support
      hrs_work_mainjob      travel_work     unpaid_acty_count     ;
var benefit_income LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      marital_status_legal             highest_qual
      income_source_count   benefit_income        income_support
      hrs_work_mainjob      travel_work     unpaid_acty_count     ;
var ttl_personal_income LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      marital_status_legal             highest_qual
```

```
      income_source_count   benefit_income   ttl_personal_income
      hrs_work_mainjob      travel_work      unpaid_acty_count     ;
var income_support LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      marital_status_legal             highest_qual
      income_source_count   benefit_income   ttl_personal_income
      income_support              travel_work      unpaid_acty_count    ;
var hrs_work_mainjob LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      marital_status_legal             highest_qual
      income_source_count   benefit_income   ttl_personal_income
      income_support   hrs_work_mainjob           unpaid_acty_count    ;
var travel_work LinkInd0601;
run;
proc corr data=aa6 (where=(adult=1));
partial
age   adult yrs_at_addr      same_addr_5yrs_ago          yrs_in_NZ
      language_indicator    Language_count   maori_descent
      iwi_count  religious  live_with_parents     live_with_partner
      live_with_siblings    live_with_children    live_alone
      live_with_flatmates   live_with_other  difficulty_acty_count
      disability_ind   sex_female Eur06 Mao06 Pac06 Asian06    MELAA06
      Other06    EthNS NZDep2006  smoke own_residence
      marital_status_legal             highest_qual
      income_source_count   benefit_income   ttl_personal_income
      income_support   hrs_work_mainjob      travel_work          ;
var unpaid_acty_count LinkInd0601;
run;

data created.final_dataset0601
(Keep=
ID    linkind0601      age06 age_code_num_06  adult06    yrs_at_addr06
      same_addr_5yrs_ago06        yrs_in_NZ06
      years_in_nz_code_num_06     language_indicator06  Language_count06
```

```
      maori_descent06       iwi_count06     religious06
      live_with_parents06   live_with_partner06   live_with_siblings06
      live_with_children06  live_alone06    live_with_flatmates06
      live_with_other06     difficulty_acty_count06    disability06
      disability_ind06 sex_female06     Eur06 Mao06 Pac06 Asian06
      MELAA06    Other06    EthNS2006  NZDep2006  smoke06
      own_residence06  marital_status_legal06          highest_qual06
      income_source_count06 benefit_income06 ttl_personal_income06
      income_support06 hrs_work_mainjob06    work_labour_force_status06
      travel_work06    unpaid_acty_count06
);
set aa6;
linkind0601      =     linkind0601      ;
age06 =     age   ;
age_code_num_06= input(age_code_06,f4.0);
adult06    =     adult ;
yrs_at_addr06    =     yrs_at_addr      ;
same_addr_5yrs_ago06  =     same_addr_5yrs_ago    ;

yrs_in_NZ06      =     yrs_in_NZ  ;
years_in_nz_code_num_06 = input(years_in_nz_code_06,f4.0);
language_indicator06  =     language_indicator    ;
Language_count06 =     Language_count   ;
maori_descent06  =     maori_descent    ;

iwi_count06      =     iwi_count  ;
religious06      =     religious  ;
live_with_parents06   =     live_with_parents     ;
live_with_partner06   =     live_with_partner     ;
live_with_siblings06  =     live_with_siblings    ;
live_with_children06  =     live_with_children    ;
live_alone06     =     live_alone ;
live_with_flatmates06 =     live_with_flatmates   ;
live_with_other06     =     live_with_other  ;
difficulty_acty_count06     =     difficulty_acty_count ;

disability_ind06 =     disability_ind   ;
sex_female06     =     sex_female ;
Eur06 =     Eur06 ;
Mao06 =     Mao06 ;
Pac06 =     Pac06 ;
Asian06    =     Asian06    ;
MELAA06    =     MELAA06    ;
Other06    =     Other06    ;
EthNS2006  =     EthNS ;
NZDep2006  =     NZDep2006  ;
smoke06    =     smoke ;
own_residence06  =     own_residence    ;
marital_status_legal06      =     marital_status_legal  ;


highest_qual06   =     highest_qual     ;
```

```
income_source_count06 =    income_source_count   ;
benefit_income06=    benefit_income   ;
ttl_personal_income06 =    ttl_personal_income   ;
income_support06=    income_support   ;
hrs_work_mainjob06    =    hrs_work_mainjob;

travel_work06    =    travel_work     ;
unpaid_acty_count06   =    unpaid_acty_count     ;

run;
```