

Appendix D. Data Quality Tables

Table DQ.1: Age distribution of household population

Single-year age distribution of household population by sex (weighted), Thailand, 2005-2006

Age	Males		Females		Age	Males		Females	
	Number	Percent	Number	Percent		Number	Percent	Number	Percent
0	512,449	1.6	466,822	1.4	43	470,418	1.5	532,075	1.6
1	503,998	1.6	476,728	1.4	44	438,677	1.4	508,829	1.5
2	478,375	1.5	474,726	1.4	45	579,790	1.8	588,619	1.8
3	481,875	1.5	499,775	1.5	46	440,936	1.4	462,940	1.4
4	495,923	1.6	466,623	1.4	47	362,961	1.1	473,843	1.4
5	408,691	1.3	390,169	1.2	48	449,538	1.4	539,363	1.6
6	490,804	1.5	445,910	1.3	49	428,603	1.3	376,668	1.1
7	452,268	1.4	440,124	1.3	50	424,185	1.3	466,501	1.4
8	544,392	1.7	560,229	1.7	51	309,115	1.0	345,344	1.0
9	554,321	1.7	494,645	1.5	52	373,161	1.2	398,611	1.2
10	551,230	1.7	514,213	1.6	53	358,766	1.1	403,282	1.2
11	535,912	1.7	504,397	1.5	54	329,969	1.0	336,080	1.0
12	530,072	1.7	505,671	1.5	55	295,088	0.9	350,464	1.1
13	515,467	1.6	515,111	1.6	56	326,245	1.0	321,323	1.0
14	529,824	1.7	507,171	1.5	57	252,675	0.8	263,301	0.8
15	622,495	1.9	573,863	1.7	58	274,113	0.9	307,201	0.9
16	529,100	1.7	534,624	1.6	59	200,817	0.6	233,316	0.7
17	525,595	1.6	541,212	1.6	60	287,088	0.9	326,190	1.0
18	588,959	1.8	495,222	1.5	61	166,289	0.5	182,726	0.6
19	411,616	1.3	421,606	1.3	62	203,355	0.6	197,365	0.6
20	551,801	1.7	501,073	1.5	63	219,413	0.7	256,033	0.8
21	507,360	1.6	479,511	1.4	64	157,122	0.5	205,298	0.6
22	533,110	1.7	524,415	1.6	65	208,366	0.7	267,356	0.8
23	587,533	1.8	533,469	1.6	66	157,575	0.5	195,239	0.6
24	519,493	1.6	585,321	1.8	67	175,224	0.5	188,965	0.6
25	589,467	1.8	567,342	1.7	68	156,822	0.5	175,358	0.5
26	494,925	1.5	526,702	1.6	69	127,971	0.4	154,242	0.5
27	460,353	1.4	504,896	1.5	70	125,581	0.4	164,910	0.5
28	576,462	1.8	511,646	1.5	71	92,297	0.3	116,820	0.4
29	575,854	1.8	551,166	1.7	72	140,494	0.4	157,785	0.5
30	612,825	1.9	557,644	1.7	73	108,146	0.3	143,242	0.4
31	470,908	1.5	588,811	1.8	74	101,445	0.3	125,181	0.4
32	513,802	1.6	568,384	1.7	75	78,314	0.2	123,856	0.4
33	511,952	1.6	529,129	1.6	76	76,383	0.2	84,856	0.3
34	567,570	1.8	552,776	1.7	77	54,901	0.2	82,747	0.2
35	543,581	1.7	560,505	1.7	78	82,643	0.3	99,518	0.3
36	546,085	1.7	614,668	1.9	79	43,419	0.1	52,681	0.2
37	539,350	1.7	584,068	1.8	80+	274,068	0.9	445,461	1.3
38	530,655	1.7	589,906	1.8	DK/Missing				
39	491,022	1.5	518,062	1.6					
40	597,915	1.9	582,469	1.8	Total	31,951,196	100.0	33,112,873	100.0
41	484,105	1.5	576,561	1.7					
42	529,728	1.7	521,921	1.6					

Typical data quality issues: Heaping on ages with digits ending with 0 and 5. If age reporting is good, the curve to be produced from these numbers should be smooth. The table should also provide insights into overreporting-underreporting at certain age groups or intervals, and the extent of missing information on age.

Table DQ.2: Age distribution of eligible and interviewed women

Household population of women age 10-54, interviewed women age 15-49, and percentage of eligible women who were interviewed (weighted), by five-year age group, Thailand, 2005-2006

Age	Household population of women age 10-54	Interviewed women age 15-49		Percentage of eligible women interviewed
	Number	Number	Percent	
10-14	2,546,564	na	na	na
15-19	2,566,527	2,542,192	13.7	99.1
20-24	2,623,789	2,598,520	14.0	99.0
25-29	2,661,753	2,639,148	14.2	99.2
30-34	2,796,744	2,788,662	15.0	99.7
35-39	2,867,209	2,842,828	15.3	99.1
40-44	2,721,855	2,707,544	14.6	99.5
45-49	2,441,433	2,423,234	13.1	99.3
50-54	1,949,818	na	na	na
15-49	18,679,308	18,542,128	100.0	99.3

Typical data quality issues: In countries with growing populations, the percentages in each age group should decline with age (Columns 2 and 4). The last column shows whether the survey was less effective in interviewing certain age groups - typically, some surveys fail to interview the younger women, sometimes because of problems in sample implementation, sometimes because of interviewers' reluctance to interview young women. These figures should be high, preferably over 95 percent, or at least 90 percent, and should not vary much by age.

Note: Weights for both household population of women and interviewed women are population weights. Age is based on the household schedule. Table should be run unweighted if major problems are identified.

Table DQ.3: Age distribution of eligible and interviewed under-5s

Household population of children age 0-4, children whose mothers/caretakers were interviewed, and percentage of under-5 children whose mothers/caretakers were interviewed (unweighted), by five-year age group, Thailand, 2005-2006

Age	<u>Household population of children age 0-7</u>	<u>Interviewed children age 0-4</u>		Percentage of eligible children interviewed
	Number	Number	Percent	
0	1,880	1,878	20.0	99.9
1	1,951	1,944	20.7	99.6
2	1,862	1,854	19.7	99.6
3	1,921	1,914	20.3	99.6
4	1,830	1,819	19.3	99.4
5	1,849	.	.	.
6	2,135	.	.	.
7	2,090	.	.	.
0-4	9,444	9,409	100.0	99.6

Typical data quality issues: The table is intended to provide information on the efficiency of the survey in collecting information on under-5s. Distribution of children by age in the household questionnaire should be smooth, with little or no heaping on age 5, which could mean out-transference of children age 0-4 to outside the eligibility range. Percentages in the last column (completion rates) should be over 90, preferably over 95.

Note: Weights for both household population of children and interviewed children are household weights. Age is based on the household schedule. Table should be run unweighted if major problems are identified.

Table DQ.3: Age distribution of eligible and interviewed under-5s

Household population of children age 0-4, children whose mothers/caretakers were interviewed, and percentage of under-5 children whose mothers/caretakers were interviewed (unweighted), by five-year age group, Thailand, 2005-2006

Age	<u>Household population of children age 0-7</u>	<u>Interviewed children age 0-4</u>		Percentage of eligible children interviewed
	Number	Number	Percent	
0	1,880	1,878	20.0	99.9
1	1,951	1,944	20.7	99.6
2	1,862	1,854	19.7	99.6
3	1,921	1,914	20.3	99.6
4	1,830	1,819	19.3	99.4
5	1,849	.	.	.
6	2,135	.	.	.
7	2,090	.	.	.
0-4	9,444	9,409	100.0	99.6

Typical data quality issues: The table is intended to provide information on the efficiency of the survey in collecting information on under-5s. Distribution of children by age in the household questionnaire should be smooth, with little or no heaping on age 5, which could mean out-transference of children age 0-4 to outside the eligibility range. Percentages in the last column (completion rates) should be over 90, preferably over 95.

Note: Weights for both household population of children and interviewed children are household weights. Age is based on the household schedule. Table should be run unweighted if major problems are identified.

Table DQ.4: Age distribution of under-5 children

Age distribution of under-5 children by 3-month groups (weighted), Thailand, 2005-2006

	Males		Females		Total	
	Number	Percent	Number	Percent	Number	Percent
Age in months						
0-2	91,640	3.7	104,708	4.4	196,348	4.1
3-5	142,198	5.8	114,343	4.8	256,541	5.3
6-8	147,171	6.0	113,781	4.8	260,953	5.4
9-11	122,824	5.0	120,614	5.1	243,437	5.0
12-14	125,275	5.1	122,045	5.1	247,321	5.1
15-17	132,465	5.4	127,673	5.4	260,138	5.4
18-20	113,534	4.6	126,199	5.3	239,733	5.0
21-23	122,818	5.0	104,853	4.4	227,670	4.7
24-26	127,322	5.2	114,787	4.8	242,109	5.0
27-29	124,029	5.0	127,515	5.4	251,544	5.2
30-32	131,746	5.3	126,820	5.3	258,566	5.3
33-35	104,529	4.2	104,370	4.4	208,899	4.3
36-38	110,917	4.5	120,821	5.1	231,738	4.8
39-41	132,824	5.4	126,972	5.3	259,796	5.4
42-44	113,531	4.6	111,247	4.7	224,778	4.6
45-47	122,105	5.0	137,061	5.8	259,165	5.4
48-50	124,370	5.0	118,792	5.0	243,162	5.0
51-53	121,730	4.9	117,305	4.9	239,036	4.9
54-56	121,113	4.9	118,453	5.0	239,566	5.0
57-59	130,730	5.3	116,451	4.9	247,181	5.1
Total	2,462,868	100.0	2,374,812	100.0	4,837,680	100.0

Typical data quality issues: The table is intended to provide information on the quality of age reporting for under-5s. In fact, the information is collected by asking the date of birth of children in the under-5 questionnaire, which is later converted into ages during data processing and analysis. The distribution should be smooth. Poor interviewing will reveal itself in heaping on certain ages.

Table DQ.5: Heaping on ages and periods

Age and period ratios at boundaries of eligibility by type of information collected (weighted), Thailand, 2005-2006

	Age and period ratios*			Eligibility boundary (lower-upper)	Module or questionnaire
	Males	Females	Total		
Age in household questionnaire					
1	1.01	1.01	1.01		
2	0.98	0.98	0.98	Lower	Child discipline and child disability
3	0.99	1.04	1.02		
4	1.07	1.03	1.05	Upper	Under-5 questionnaire
5	0.88	0.90	0.89	Lower	Child labour and education
6	1.09	1.05	1.07		
8	1.05	1.12	1.09		
9	1.01	0.95	0.98	Upper	Child disability
10	1.01	1.02	1.01		
13	0.98	1.01	1.00		
14	0.95	0.95	0.95	Upper	Child labour and child discipline
15	1.11	1.07	1.09	Lower	Women's questionnaire
16	0.95	0.97	0.96		
17	0.96	1.03	1.00	Upper	Orphaned and vulnerable children
18	1.03	1.11	1.07		
23	1.07	0.97	1.02		
24	0.92	1.04	0.98	Upper	Education
25	1.10	1.01	1.06		
48	1.09	1.16	1.13		
49	0.99	0.82	0.90	Upper	Women's questionnaire
50	1.10	1.18	1.14		
Age in women's questionnaire					
23	na	0.98	na		
24	na	1.04	na	Upper	Sexual behaviour
25	na	1.01	na		(This module not include in questionnaire)
Months since last birth in women's questionnaire					
6-11	na	1.00	na		
12-17	na	1.05	na		
18-23	na	0.91	na	Upper	Tetanus toxoid and maternal and child health
24-29	na	1.09	na		
30-35	na	0.92	na		

* Age or period ratios are calculated as $x / ((x_{n-1} + x_n + x_{n+1}) / 3)$, where x is age or period.

Typical data quality issues: Age and period ratios in the table are calculated for two purposes: To check for evidence of heaping on certain periods or ages, particularly on those at the boundaries of eligibility, and to check if interviewers had transferred cases out of eligibility intervals. The table is indicative of the quality of fieldwork. Interviewers sometimes "transfer out" cases so as to avoid extra work - for instance, interviewers may transfer the age a 15 year-old woman to 14 to avoid an individual interview, in which case the age ratio on age 15 will be depressed (a deficit of females at age 15) and the age ratio on age 14 significantly higher than 1.00.

Table DQ.6: Completeness of reporting

Percentage of observations missing information for selected questions and indicators (weighted), Thailand, 2005-2006

Questionnaire and Subject	Reference group	Percent with missing information*	Number of cases
Age			
Salt testing	All households surveyed	0.0	18,031,070
Women			
Date of Birth	All women age 15-49		
Month only		4.9	18,542,128
Month and year missing		-	18,542,128
Date of first birth	All women age 15-49 with at least one live birth		
Month only		2.3	11,950,256
Month and year missing		1.6	11,950,256
Completed years since first birth	All women age 15-49 with at least one live birth	-	242,669
Date of last birth	All women age 15-49 with at least one live birth		
Month only		-	11,950,256
Month and year missing		-	11,950,256
Date of first marriage/union	All ever married women age 15-49		
Month only		18.7	13,544,028
Month and year missing		27.4	13,544,028
Age at first marriage/union	All ever married women age 15-49	0.3	13,544,028
Under-5			
Date of Birth	All under five children surveyed		
Month only		-	4,837,680
Month and year missing		-	4,837,680
Anthropometry	All under five children surveyed		
Height			
Weight		2.5	4,837,680
Height or Weight		2.9	4,837,680
		2.9	4,837,680

* Includes "Don't know" responses

Typical data quality issues: Surveys always have cases with missing information. The extent of missing information is important, because it can result in biased results if such proportions are high. Particularly informative is the extent of missing information on measurements, ages, dates of events.

Table DQ.7: Presence of mother in the household and the person interviewed for the under-5 questionnaire

Distribution of children under five by whether the mother lives in the same household, and the person interviewed for the under-5 questionnaire (weighted), Thailand, 2005-2006

	Mother in the household				Mother not in the household				Total	Number of children aged 0-4 years
	Mother interviewed	Father interviewed	Other adult female interviewed	Other adult male interviewed	Father interviewed	Other adult female interviewed	Other adult male interviewed	Child(<15) interviewed		
Age										
0	84.6	0.0	-	-	0.1	15.2	0.1	0.0	100.0	979,271
1	78.9	0.0	-	-	0.1	20.4	0.5	0.0	100.0	980,725
2	75.0	0.0	-	-	0.5	23.8	0.6	0.0	100.0	953,101
3	76.5	0.0	-	-	1.2	21.8	0.6	0.0	100.0	981,650
4	77.1	0.0	-	-	1.3	20.7	0.9	0.0	100.0	962,546
			-	-						
Total	78.5	0.0			0.6	20.4	0.5	0.0	100.0	4,857,293

Typical data quality issues: The under-5 questionnaire should be administered to the mother, if the mother was in the household. The table is informative on how the questionnaire was administered during the fieldwork. Not all information will have been collected from mothers, but cases where the mother is in the household but somebody else was interviewed can be problematic.

Table DQ.8: School attendance by single age

Distribution of household population age 5-24 by educational level and grade attended in the current year (weighted), Thailand, 2005-2006

Age	Primary school										Secondary school						Total	Number	
	Preschool	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Higher	Non-standard curriculum	Don't know			Not attending school
5	98.7	0.8	0.1	-	0.0	0.1	-	-	-	-	-	-	-	-	0.1	-	0.2	100.0	798,860
6	80.4	18.1	1.0	0.2	0.2	-	-	-	-	-	-	-	-	-	-	-	0.0	100.0	936,714
7	8.4	70.3	19.6	1.1	0.4	-	-	-	-	-	-	-	-	-	0.0	-	0.1	100.0	892,392
8	0.6	7.3	67.5	21.6	2.2	0.7	-	-	-	-	-	-	-	-	0.0	-	0.1	100.0	1,104,622
9	0.1	0.4	7.7	69.7	20.9	1.0	0.2	-	-	-	-	-	-	-	0.0	-	0.1	100.0	1,048,966
10	0.0	0.2	0.5	7.9	67.2	23.1	1.0	-	-	-	-	-	-	-	-	-	0.1	100.0	1,065,444
11	0.1	0.2	0.4	1.2	8.9	67.9	20.2	0.8	0.2	0	-	0.0	-	-	0.0	-	0.2	100.0	1,040,309
12	-	0.0	0.1	0.3	1.8	10.0	65.8	18.4	2.8	0	0.0	-	-	-	0.0	-	0.6	100.0	1,035,743
13	-	0.0	0.1	0.0	0.3	1.6	9.6	62.6	22.6	1	0.0	-	-	-	0.0	-	1.9	100.0	1,030,578
14	-	0.0	0.2	0.3	-	0.2	1.3	7.5	62.7	20	1.1	0.1	-	-	-	-	6.3	100.0	1,036,996
15	-	-	-	-	-	-	0.2	0.8	6.7	66	14.2	0.6	0.3	0.1	-	-	10.9	100.0	1,196,358
16	-	-	-	-	0.1	0.1	0.4	0.4	2.5	11	51.1	11.7	2.3	0.3	-	-	20.7	100.0	1,063,724
17	-	-	-	0.0	-	-	-	0.1	0.3	3	9.3	45.0	12.1	0.8	-	-	30.0	100.0	1,066,807
18	-	-	0.1	-	-	-	-	0.1	0.1	1	3.5	15.9	38.2	6.9	-	-	34.4	100.0	1,084,181
19	-	-	-	-	-	0.0	0.1	0.1	0.1	1	1.9	3.5	10.3	25.7	-	-	57.5	100.0	833,222
20	-	-	-	-	-	-	-	-	0.0	0	0.7	1.1	2.3	27.7	-	-	67.7	100.0	1,052,874
21	-	-	-	-	-	-	-	0.0	0.1	0	0.9	0.3	1.8	27.1	-	-	69.5	100.0	986,871
22	-	-	-	-	-	-	-	0.1	0.1	0	0.8	0.3	0.7	16.1	-	-	81.6	100.0	1,057,525
23	-	-	-	-	-	-	-	0.1	0.1	0	0.3	0.0	0.2	6.0	-	-	93.2	100.0	1,121,002
24	-	-	-	-	-	-	0.1	-	-	0	0.1	0.1	0.3	5.5	-	-	93.7	100.0	1,104,815

Typical data quality issues: The table could be used to look at the outliers. Data entry programs do not check age versus grade. If data has been collected and entered correctly, one should see cases concentrated over the diagonal, and should not expect such cases as 22 year old persons at primary school grades, very young people at grade 6 of secondary school etc. Before running the table grades should be adapted to the system in the country.

Note: Levels and grades refer to the most recent school year if data collection was completed between school years

Table DQ.9: Sex ratio at birth among children ever born and living

Sex ratio at birth among children ever born, children living, and deceased children, by age of women (weighted), Thailand, 2005-2006

Age	Children Ever Born			Children Living			Children deceased			Number of women
	Number of sons ever born	Number of daughters ever born	Sex ratio	Number of sons living	Number of daughters living	Sex ratio	Number of deceased sons	Number of deceased daughters	Sex ratio	
15-19	114,548.1	85,888.5	1.3	113,054	85,888	1.3	1,494	-	.	2,542,192
20-24	640,643.0	655,798.8	1.0	630,850	654,319	1.0	9,793	1,479	7	2,598,520
25-29	1,309,274.5	1,179,727.1	1.1	1,301,082	1,159,349	1.1	8,192	20,378	0	2,639,148
30-34	2,074,026.1	1,956,630.6	1.1	2,035,725	1,936,896	1.1	38,301	19,734	2	2,788,662
35-39	2,589,721.7	2,480,728.2	1.0	2,527,934	2,449,024	1.0	61,787	31,704	2	2,842,828
40-44	2,812,055.0	2,685,799.9	1.0	2,710,172	2,628,372	1.0	101,883	57,428	2	2,707,544
45-49	2,830,224.5	2,621,132.3	1.1	2,671,757	2,525,168	1.1	158,467	95,964	2	2,423,234
Total	12,370,492.8	11,665,705.4	1.1	11,990,575	11,439,018	1.0	379,918	226,687	2	18,542,128

Typical data quality issues: Universally, the sex ratio among live births is around 1.05, typically ranging from 1.03 to 1.07 in sizeable populations (with the exception of populations where sex-selective abortions is practiced). The values in column 3 should be within these ranges. However, since sample surveys are influenced by chance fluctuations, one should be looking for systematically low or high ratios (in several countries, very young daughters may not be reported, or deaths of males may not be reported). In most populations, death rates at early ages are higher for males than females - hence, the sex ratios among deceased children (Column 6) should also be above 1.

Table DQ.10: Distribution of women by time since last birth

Distribution of women aged 15-49 with at least one live birth, by months since last birth (weighted), Thailand, 2005-2006

Months since last birth					
Age	Number	Percent	Age	Number	Percent
0	54,496	2.1	18	74,080	2.8
1	74,449	2.9	19	70,734	2.7
2	105,540	4.1	20	50,110	1.9
3	97,519	3.8	21	49,867	1.9
4	83,988	3.2	22	81,433	3.1
5	72,748	2.8	23	60,636	2.3
6	87,106	3.3	24	80,230	3.1
7	85,152	3.3	25	58,790	2.3
8	82,755	3.2	26	59,775	2.3
9	76,197	2.9	27	76,432	2.9
10	83,331	3.2	28	66,415	2.6
11	65,957	2.5	29	79,622	3.1
12	75,962	2.9	30	64,003	2.5
13	72,219	2.8	31	66,687	2.6
14	90,380	3.5	32	77,142	3.0
15	69,839	2.7	33	53,982	2.1
16	85,281	3.3	34	53,643	2.1
17	73,022	2.8	35	39,962	1.5
			Total	2,600,338	100.0

Typical data quality issues: Months since last birth may be heaped on periods of 6 months, 12 months, 24 months etc. In particular, the heaping on 24 months is problematic, since some women had a birth in the last 2 years, but did not declare them so.

Figure	Description
1	Scatterplot of weight (Y-axis) by height (x-axis), unweighted
2	Scatterplot of weights of children by age in months
3	Scatterplot of heights of children by age in months
	<p>Figures 1-3 are intended to provide a visual insight into the quality of anthropometric measurements. The data points should be concentrated along a diagonal. Outliers can be easily spotted visually. Remember that data problems may be due to poor reporting of age, or poor measurement of heights or weights, or any combination of the three.</p>
4	Number of male household population (Y-axis) by single ages (X-axis) (Line graph) (unweighted and weighted)
5	Number of female household population (Y-axis) by single ages (X-axis) (Line graph) (unweighted and weighted)
	<p>Figures 4-5 are based on Table DQ.1, and are intended to provide information on the extent of age heaping, deficits of household population at certain ages or age intervals. Both unweighted and weighted distributions are shown.</p>
6	Population pyramid, Thailand, 2005-2006

Figure 1. Scatterplot of weight (Y-axis) by height (X-axis) (unweighted), Thailand, 2005-2006

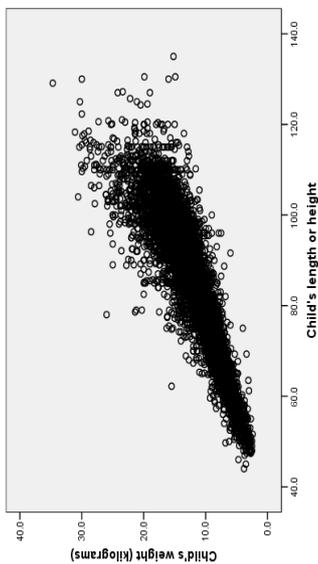


Figure 2. Scatterplot of weights of children by age in months (unweighted), Thailand, 2005-2006

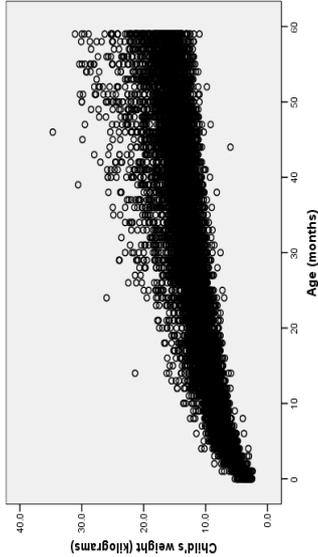


Figure 3. Scatterplot of heights of children by age in months (unweighted), Thailand, 2005-2006

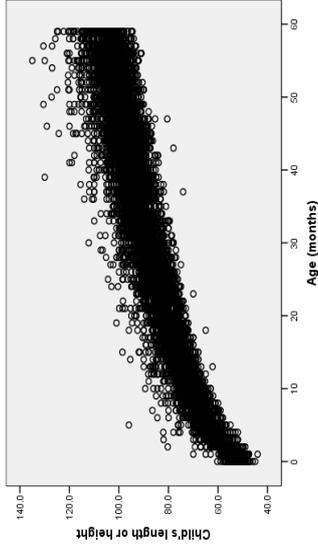


Figure 4. Number of male household population (Y-axis) by single ages (X-axis) (unweighted), Thailand, 2005-2006

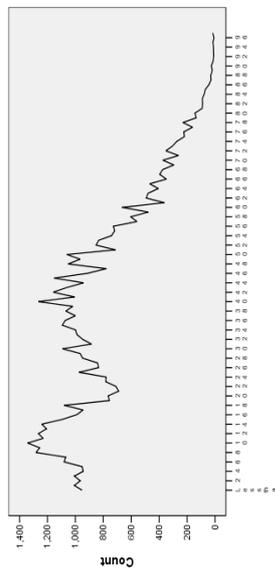


Figure 5. Number of female household population (Y-axis) by single ages (X-axis) (unweighted), Thailand, 2005-2006

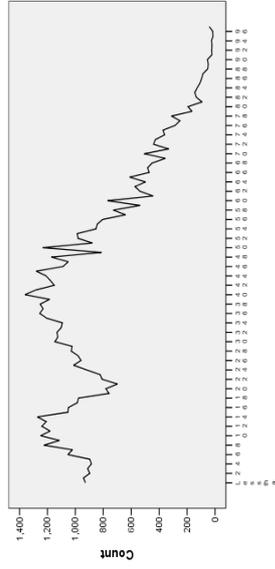


Figure 4. Number of male household population (Y-axis) by single ages (X-axis) (weighted), Thailand, 2005-2006

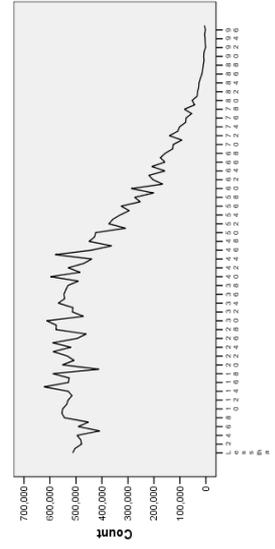


Figure 5. Number of female household population (Y-axis) by single ages (X-axis) (weighted), Thailand, 2005-2006

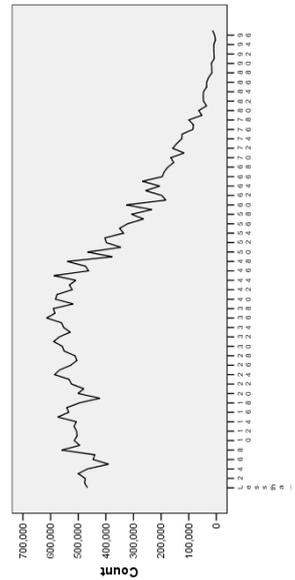
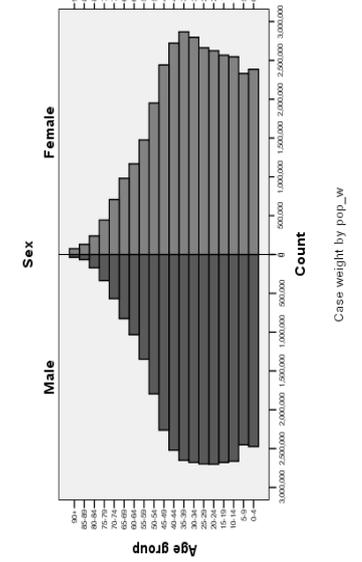


Figure 6. Population pyramid, Thailand, 2005-2006



Cases weighted by pop_w

Cases weighted by pop_w