



DATA VALIDATION-II

Consistency check



Objectives

- Ensuring all records meet with editing rules
- Ensuring there is no unusual/unexpected values
- Checking for whether the data makes sense when variables are tabulated or cross-tabulated



How to evaluate validation

- ❑ Cross - tabulation between variables
 - for preliminary analysis of census results, the list of tables should be prepared based on:
 - ❑ Editing rules and relation between variables
 - Tables should present all possible conditions in data without eliminating any category to verify the results



Main relation between census topics

- ❑ **Relationship to the head of households**
 - Sex and marital status of couples
 - Age interval between father/mother and children
 - ❑ At least 12-14 years and at most 65 for males, 50 for females
 - Age interval between grand parents and grand children
 - ❑ At least 25 years



Main relation between census topics

- Distribution of household size**
 - Accuracy of household size –particularly for 8, 16, ... members of households
 - There might be errors in combining the questionnaires belonging to same household



Main relation between census topics

❑ **Marital status**

- Minimum age at first marriage

❑ **CED, CS and CD**

- Sex and marital status (?) of person
- Relation between number of children ever-born, number of living children and number of dead children
- Relation between age and number of children ever born



CEB – quality assessment

Fe

| INDONESIA, 1990 Census | | | | | | |
|------------------------|--|---------|-----------|-----------|-----------|--|
| Children ever-born | Ever-married women by number of children ever-born | | | | | |
| | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | |
| 0 | 25,086 | 927,848 | 1,105,996 | 557,567 | 288,589 | |
| 1 | 6,474 | 678,524 | 2,633,719 | 1,763,409 | 702,870 | |
| 2 | 1,813 | 87,742 | 1,283,748 | 2,389,432 | 1,437,921 | |
| 3 | 897 | 10,626 | 347,446 | 1,463,715 | 1,503,832 | |
| 4 | 387 | 2,882 | 86,764 | 682,439 | 1,113,948 | |
| 5 | 159 | 1,341 | 20,739 | 258,423 | 674,712 | |
| 6 | 19 | 897 | 6,783 | 88,086 | 358,934 | |
| 7 | - | 610 | 2,965 | 27,355 | 162,477 | |
| 8 | - | 294 | 1,271 | 9,726 | 68,276 | |
| 9 | - | 64 | 799 | 3,082 | 27,204 | |

Parities wrong ?

Source, UN Statistics Division Database



CEB – quality assessment

Mongolia, 1989 Census (Source: IPUMS)

| Parity | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 |
|---------|---------|--------|--------|--------|-------|-------|-------|
| 0 | 105,548 | 43,676 | 9,824 | 2,711 | 987 | 865 | 726 |
| 1 | 4,827 | 30,834 | 15,350 | 5,432 | 2,185 | 1,302 | 1,488 |
| 2 | 896 | 17,309 | 23,960 | 10,659 | 4,479 | 2,217 | 2,053 |
| 3 | 834 | 5,382 | 19,279 | 11,159 | 4,923 | 2,663 | 1,950 |
| 4 | 199 | 1,828 | 11,831 | 11,922 | 6,974 | 3,525 | 2,658 |
| 5 | 68 | 477 | 5,730 | 11,189 | 7,426 | 4,933 | 3,379 |
| 6 | 0 | 53 | 2,161 | 7,568 | 6,348 | 4,442 | 3,619 |
| 7 | 0 | 25 | 707 | 3,737 | 4,551 | 3,638 | 2,977 |
| 8 | 15 | 0 | 263 | 2,355 | 3,879 | 3,986 | 3,706 |
| 9 | 61 | 0 | 119 | 746 | 2,190 | 2,747 | 3,059 |
| 10 | 0 | 0 | 0 | 419 | 1,300 | 2,433 | 3,253 |
| 11 | 0 | 0 | 0 | 147 | 743 | 1,183 | 1,667 |
| 12 | 22 | 38 | 11 | 53 | 262 | 845 | 1,299 |
| 13 | 0 | 0 | 0 | 19 | 161 | 403 | 898 |
| 14 | 0 | 0 | 0 | 20 | 82 | 242 | 392 |
| 15+ | 0 | 0 | 0 | 0 | 72 | 235 | 629 |
| Unknown | 218 | 0 | 65 | 58 | 35 | 35 | 20 |

Parities
wrong ?



Main relation between census topics

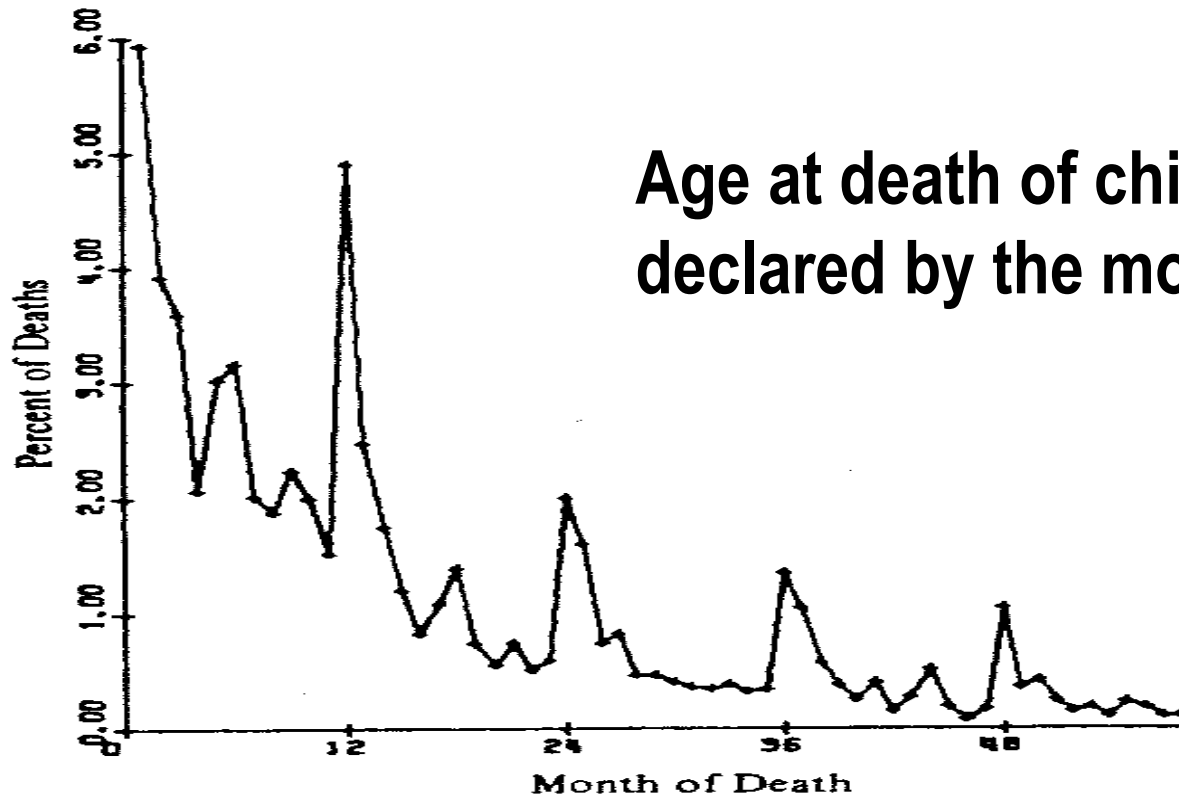
❑ **Date of last live birth**

- Distribution of last live births by months and years
 - ❑ Misreporting of the date
 - ❑ Errors in data capture
- Consistency with age of mother –minimum and maximum age for giving a birth
- Consistency with children ever born-at least one child
- Consistency with number of living children- if the last live birth is alive, at least one living children



Main relation between census topics

- Household deaths
 - Sex and age of death for maternal mortality
 - Consistency with the last live birth if the infant is dead



Age at death of children (in month) declared by the mother, Nepal 1975



Main relation between census topics

Education

Literacy

- Relation with the highest completed level of education

School attendance

- Relation between Q20 and Q21
- Relation with age -minimum age for entering school and possibility of giving maximum age??
- Consistency between school attendance and student category of working status



Main relation between census topics

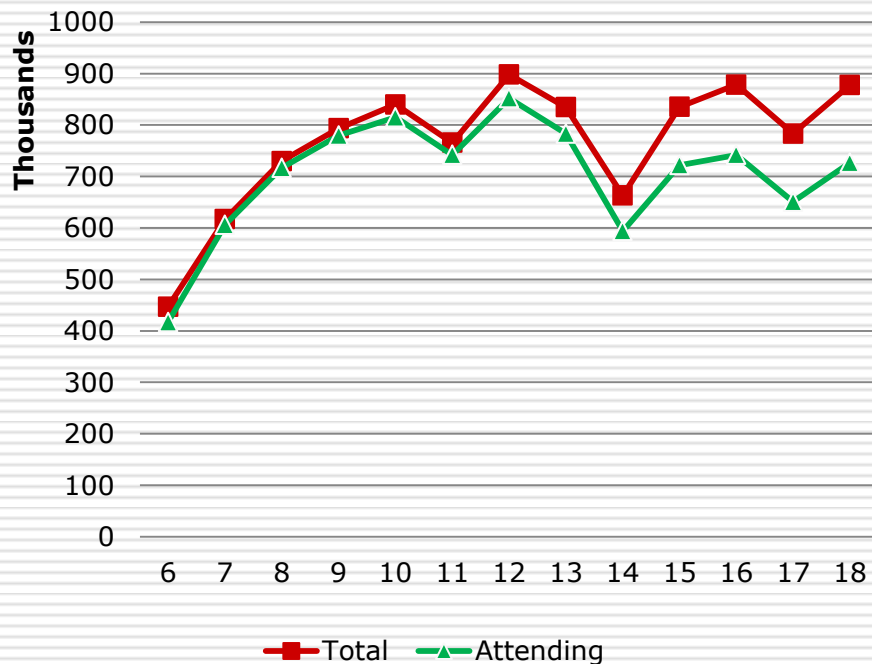
□ Education

- Educational attainment- highest level completed
 - Consistency with school attendance-previously attended or currently attending
 - Relation with age –minimum age for completing school
 - Usually it is calculated by taking minimum age for entering school plus number of years-minimum- for the certain level of school.
Example:
 - Minimum age for primary education is age 6
 - If primary education is 8 years, minimum age for completing primary school would be age 13

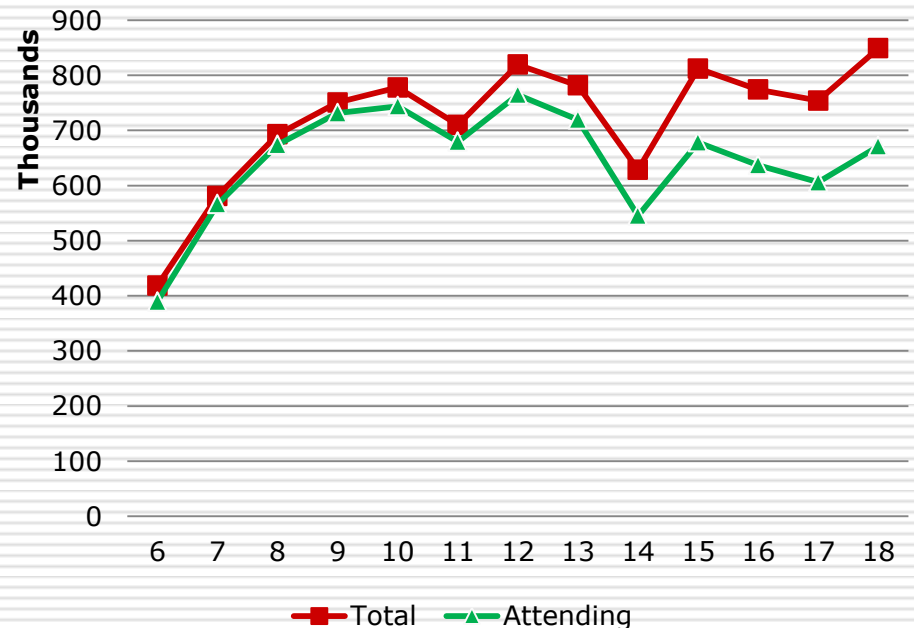


School attendance – quality assessment

School attendance, Male, Egypt, 2006 Census



School attendance, Female, Egypt, 2006 Census





Main relation between census topics

❑ **Economic activity**

- Consistency between students and school attendance
- Consistency between disabled and disability status
- Relation of age to employment –minimum age
- Relation of certain occupations to the level of education



Main relation between census topics

Migration

- Place of birth

- Place of previous residence and place of usual residence
 - If usual residence is different from place of birth, place of previous residence must be filled

- Duration of residence
 - Relation of age to duration of residence



Duration of residence– quality assessment

Population by Duration of Residence in Hong Kong, Sex, Age Group and Ethnicity, 2011

| Chinese | | | | | | |
|-----------|-----------|--------|-----------|--------|-----------|---------|
| Age group | 4-6 years | | 7-9 years | | 10+ years | |
| | Male | Female | Male | Female | Male | Female |
| 0 - 4 | 17 419 | 16 466 | - | - | - | - |
| 5 - 9 | 51 874 | 48 090 | 58 591 | 56 157 | - | - |
| 10 - 14 | 8 367 | 7 654 | 16 499 | 15 798 | 132 367 | 124 173 |
| 15 - 19 | 7 935 | 7 186 | 8 257 | 8 179 | 190 190 | 179 202 |
| 20 - 24 | 6 182 | 6 244 | 6 447 | 6 079 | 199 768 | 191 527 |
| 25 - 29 | 2 855 | 6 569 | 3 355 | 5 432 | 210 041 | 204 774 |
| 30 - 34 | 2 191 | 17 096 | 2 118 | 8 298 | 206 303 | 208 118 |
| 35 - 39 | 3 307 | 19 463 | 2 982 | 24 388 | 211 779 | 212 258 |
| 40 - 44 | 3 941 | 11 760 | 4 014 | 21 819 | 220 225 | 248 425 |
| 45 - 49 | 3 517 | 6 579 | 3 123 | 11 733 | 276 563 | 308 599 |
| 50 - 54 | 1 904 | 2 237 | 1 584 | 4 564 | 301 055 | 301 369 |



Main relation between census topics

❑ **Population living abroad**

- Relation of age to year of departure
- Consistency between completed age and relationship to the head of household