16
Constructing List Frames and Administrative Lists

Construction of a frame or administrative list of businesses or agricultural entities for an entire country or a large region of a country involves at least the following major steps. This discussion is based on Winkler [1998].)

1. **Identify existing lists that can be used to create the frame or administrative list—the main list.** It is better to concentrate on no more than 10 lists, choosing those with the most comprehensive coverage and the fewest errors. If an update list does not add enough new entries to the main list, then do not use that update list. If it is likely that an update list will cause too many duplicates to be added to the main list, then do not use that update list. If an update list is on an incompatible computer system or is in an incompatible format (e.g., microfiche), then do not use that update list.

2. **Obtain an annotated layout for each list.** The layout should identify all of the possible values or codes used for each of the fields of the records on the list. The layout should describe how name and address standardization were performed. If no annotated layout is available for a list, then do not use the list.

3. **Make sure that the entries on each update list are in a format compatible with your duplicate-detection and standardization software.** If an update list can’t be run through your duplicate-detection or standardization software, then do not use the update list. If a large portion of the entries on an update list can’t be standardized, then do not use the list or just use the records that can be standardized.

4. **If resources permit, append each update list to the main list sequentially.** Appending lists one-at-a-time and then completing a clerical review before adding another list usually enables clerical staff to identify and eliminate more duplicate records. Of course, such a process is usually more expensive because it uses more staff resources.

5. **Design the main list so that it has fields that identify the source (i.e., the update) list of each record, the date the source list was obtained, and the date each record was last updated.** Such information is often crucial to eliminating duplicates and to ensuring that current information is on the main list.
6. Frequently – at least several times per year – check the main list to identify duplicate records and to correct data entries having typographical or other errors. Because small businesses in the United States (1) go in and out of business frequently, (2) move (thereby changing their address) or (3) change names, it is difficult to keep information on business lists up-to-date. For this reason, it is often helpful with business lists to include a field that denotes the operating status ("open" or "out of business") and another field that indicates whether the entity is a corporate "subsidiary" or the "parent" company. Moreover, if the entity is a subsidiary, then it is important to know the name of the parent company. Finally, it is important to know if the subsidiary reports for itself or if the parent company is the only reporting entity. Frequently, when company “A” buys company “B”, company “B” will continue to report for itself for 2 to 3 years after the purchase.

16.1. National Address Register of Residences in Canada

An address register is a list of residential addresses. Statistics Canada has done lots of work on address registers over the years. The goal of those creating the 1991 Canadian address register was to improve the coverage of the 1991 Census of Canada – that is, to improve the proportion of residences enumerated in the 1991 Census. Specifically, Statistics Canada estimated that the use of the 1991 address register would add 34,000 occupied dwellings and 68,000 persons in the areas for which it would be constructed.

For British Columbia, the address register for the 1991 Canadian Census included all urban population centers as well as some rural areas – approximately 88% of the population. For the other Canadian provinces, the 1991 register was limited to medium and large urban centers. British Columbia was singled out because it had a high rate of undercoverage, 4.49%, in the 1986 Census of Canada.

16.1.1. Sources of Addresses

Statistics Canada decided to use, wherever possible, four types of administrative databases to obtain addresses for the 1991 address register:

- Telephone billing databases
- Municipal assessment rolls
- Hydro company billing databases
- The T1 Personal Income Tax File.

---

1 This section is based on Swain et al. [1992].