Sport Participation in Canada Chapter 16

16.0 Methodology

16.1 The General Social Survey

Statistics Canada, through a sport supplement to the General Social Survey (GSS), surveyed approximately 10,000 Canadians (aged 15 years and older) about the extent and nature of their participation in sport during the previous 12 months. This survey was conducted in 1992 and again in 1998. These two sport supplements are the primary source of data for this study, specifically Cycles 7 and 12 entitled Time Use of the GSS.

The sport questions were developed by the Culture Statistics Program, Statistics Canada based on the information needs of federal and provincial/territorial partners having an interest in sport information. The questions were sponsored by a consortium of federal and provincial data users with the objective of obtaining a more comprehensive look at the extent of sport participation in Canada.

Involvement in sport was determined by asking the question:
"Did you regularly participate in any sports during the past 12 months?"
'Regularly' meant that the respondent participated in a sport at least once a week during the season or during a certain period of time. For example, although bowling is not a seasonal sport, it should be included if they bowled on a regular basis during a period of the year.

The types of sport within scope of this survey were determined using a list of sports provided by Sport Canada (see Appendix). "Sport" was defined by Sport Canada as an activity having the following characteristics:

- involves two or more participants who engage for the purpose of competitively evaluating their personal performance,
- involves formal rules and procedures,
- requires tactics and strategies,
- requires specialized neuromuscular skills which can be taught and learned,
- requires a significant involvement of large muscular groups,

- involves a high degree of difficulty, risk or effort in reproduction of movement or forms,
- its competitive mode implies the development of trained coaching personnel,
- its primary activity involves physical interaction of the participant with the environment,
- does not include activities in which the performance of a motorized vehicle is the primary determinant of the competitive outcome. Where mechanized vehicles or conveyances are used, the activity must entail significant physical effort in propelling the vehicle or conveyance.

Based on these general guidelines defining sport, several physical and leisure activities were excluded such as aerobics / dancercize / aquafit, bicycling for recreation or transportation, jogging and walking, etc.

Participation during the past 12 months was the reference period for both survey cycles (1992 and 1998). It is recognized that a reference period of one year may incur the problem of recall. As well, the risk of over-statement may be at play as respondents may be reluctant to admit, particularly in these health-conscious days, that they had not been physically active during the past year. It is not possible from this questionnaire to measure the presence or extent of either the under or over-reporting of sport participation.

The GSS, originating in 1985, conducts telephone surveys across the 10 provinces. The GSS is recognized for its continual collection of data that allows for trend analysis. The history of GSS data collection topics is outlined below.

GSS Cycle Topics	**Date of GSS Survey (cycle #)**		
	1st series	2nd series	3rd series
Health	1985 (1)	1991 (6)	
Time Use	1986 (2)	1992 (7)	1998 (12)
Victimization	1988 (3)	1993 (8)	1999 (13)
Education, Work & Retirement	1989 (4)	1994 (9)	
Family & Friends		1990 (5)	1995 (10)
Social Support	1985 (1)	1990 (5)	1996 (11)
Access & Use of Information Communication Technology	2000 (14)		

Each survey contains a core topic as well as a standard set of socio-demographic questions. The target population for the GSS is all persons 15 years of age and over residing in Canada. Residents of the territories and full-time residents of institutions are excluded from the sample. Computer Assisted Telephone Interviewing (CATI) is used to collect data for the GSS. Households without telephones were excluded. Persons living in households without telephones represent less than 2% of the target population. Survey estimates have been adjusted (i.e., weighted) to account for this population. Collecting data via telephone is attractive because of lower collection costs, as well as considerable flexibility with respect to sample design. Nevertheless, telephone interviewing does have some drawbacks: non-coverage of households, while low, is concentrated in population groups with low educational attainment or low income; response rates tend to be lower than for face-to-face interviews, and there are limitations on the amount and type of data which can be collected over the telephone.

Until 1998, the sample size was approximately 10,000 persons. This was increased to 25,000 in 1999. With this increase in sample size, more detailed results will be collected at both the national and provincial levels.

The sample design will not allow for estimates of Aboriginal peoples.

GSS Response and Non-response Rates (%)

Cycle	Topic	Survey year	Sample Size	Response Rate
1	Health	1985	11,200	83.4
2	Time Use	1986	16,390	78.9
3	Personal Risk	1988	9,870	82.4
4	Education & Work	1989	9,338	80.7
5	Family & Friends	1990	13,495	75.8
6	Health	1991	11,924	80.2
7	Time Use	1992	9,815	76.8
8	Personal Risk	1993	10,385	81.6
9	Education, Work & Retirement	1994	11,876	81.2
10	Family	1995	10,749	81.4
11	Social Support	1996	12,756	85.3
12	Time Use	1998	10,749	77.6
13	Victimization	1999	25,000 (e)	not available
14	Access to & Use of Information Communication Technology	2000	25,000 (e)	not available

Data for Cycle 12 of the GSS were collected from February 1998 through to January 1999. The sample was evenly distributed over the 12 months to represent any seasonal variation in the data.

Weighting

When a probability sample is used, as was the case for the GSS, the principle behind estimation is that each person selected in the sample 'represents' (in addition to himself/herself) several other persons not in the sample. For example, in a simple random sample of 2% of the population, each person in the sample represents 50 persons in the population.

When analyzing GSS Cycle 12 data, it is therefore necessary to use either the weighting factor WGHTFIN on the Main File or WGHTEPI on the Time Use Episode File.

Users are cautioned against releasing unweighted tables or performing any analysis based on unweighted survey results. Sampling rates as well as non-response rates vary significantly from province to province and by various demographic characteristics. For example, it is known that non-respondents are more likely to be males and more likely to be younger. In the responding sample, 3.3% were males between the ages of 15 and 19, while in the overall population, approximately 4.3% were males between 15 and 19. Therefore, it is clear that unweighted sample counts cannot be considered to be representative of the survey target population.

The survey weights must be used when producing estimates or performing analyses in order to account as much as possible for the geographic over- and under-representation and for the under- or over- representation of age-sex groups, months of the year, or days of the week in the unweighted file.

Over a large number of observations, randomly occurring errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Errors not related to sampling may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the questionnaire and errors may be introduced in the processing and tabulation of the data. These are all examples of non-sampling errors.

Considerable time and effort was made to reduce non-sampling errors in the survey. Quality assurance measures were implemented at each step of the data collection and processing cycle to monitor the quality of the data. These measures included the use of highly skilled interviewers, extensive training of interviewers with respect to the survey procedures and questionnaire, observation of interviewers to detect problems of questionnaire design or misunderstanding of instructions, procedures to ensure that data capture errors were minimized, and coding and edit quality checks to verify the processing logic.

A major source of non-sampling errors in surveys is the effect of non-response on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or a few questions) to total non-response. Total non-response occurred because the interviewer was either unable to contact the respondent, no member of the household was able to provide the information, or the respondent refused to participate in the survey. Total non-response was handled by adjusting the weight of households who responded to the survey.

Coefficient of Variation (C.V.)

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error, sound statistical practice calls for researchers to provide users with some indication of the magnitude of this sampling error.

Because of the large variety of estimates that can be produced from a survey, the standard deviation is usually expressed relative to the estimate to which it pertains. The resulting measure, known as the coefficient of variation (c.v.) of an estimate is obtained by dividing the standard error of the estimate by the estimate itself and is expressed as a percentage of the estimate. **Before releasing and/or publishing any estimates from the microdata file, users should consider whether or not to release the estimate based on the following guidelines.**

Coefficient of Variation		Policy Statement
0.0 to 16.5%	Moderate sampling variability	Releasable
16.6% to 33.3%	High sampling variability	Releasable with cautionary note
33.4% or over	Very High sampling variability	**Not releasable**

Note: The sampling variability policy should be applied to rounded estimates.

In order that estimates produced from the General Social Survey microdata files correspond to those produced by Statistics Canada, users are urged to round the estimates to the nearest thousand using the normal rounding technique. It may be misleading to release unrounded estimates, as they imply greater precision than actually exists. In instances where, due to technical or other limitations, a different rounding technique is used resulting in estimates different from Statistics Canada estimates, users are encouraged to note the reason for such differences in the released document.

T-test

A t-test was applied to the sport participation rates to determine whether or not the difference in rates is statistically significant.

A couple of assumptions are made: that the distribution is normal,
A two-tailed test will be used at a confidence level of 95%.
The critical value = 1.96

Hypothesis: If there is little or no change between the 1992 and 1998 participation rates, that is if the difference is between 1.96 and –1.96, then it can be concluded that the change is not statistically significant. But, should the t-test result in a figure greater than + or - 1.96, the change in the participation rates between 1992 and 1998 are statistically significant.

The t-test result = -13.2, well beyond the critical value of + or – 1.96, thus the difference in participation rates between 1992 and 1998 is statistically significant at a confidence level of 95%.