Exercise Three: Analysis

Students have two choices of dependent variables: the Health Utility Index (HUI) or the general health question. HUI can be considered a continuous variable that can take on values between 0 and 1. (There are some cases for which a value less than zero is in fact possible, but these cases will not apply in this example.) A multiple linear regression can be used to assess the relationship between HUI and the covariates. The general health question is an ordinal variable. A log-linear analysis can be used to assess the relationship between self-assessed health and the covariates.

Hints for the regression:
You can use age as a categorical variable and decide to re-group it into a smaller number of categories (think about degrees of freedom). Since the age categories are small you could as well re-code each value to the middle of the category and treat it as a continuous variable. If so, since age is limited to the 20-65 year olds, it may be convenient to use a transformed variable (i.e. newage = age-20) to avoid having a negative intercept. DVBMI94, the derived BMI variable has a curvilinear relationship to health status. A transformation is required to use that variable in a linear regression. A healthy BMI is defined as between 20 and 24 (inclusive). One possible transformation is newbmi = absolute value (DVBMI94-22). Try to use transformations that can be interpreted, e.g. newbmi measures the difference between the values for a particular individual and an average healthy one.

Background: Survey Weights
Suppose we have a finite population U, of size $N = 100$ individuals. We are interested in estimating a total, mean, or other variable of interest from this population. In a simple random sample s of size $n = 20$ (in a simple random sample each individual has the same probability to be selected in the sample) we observe $y_1, y_2, \ldots, y_{20}$. How can we estimate the population total $Y = \sum_{i=1}^{100} y_i$? Since the population size is $N = 100$, each individual in the sample represents 5 individuals in the population and is assigned a sampling weight of 5. If w_i denotes the sampling weight of individual i in the sample, in this example, $w_1 = 5$ for $i = 1,\ldots,20$, the estimator of the total $\hat{Y}$ is

$$\hat{Y} = \sum_{i=1}^{20} w_i y_i$$

In the previous example all the individuals had the same sampling weight. In surveys it is common to select a sample with unequal probability of selection and hence unequal weights. In the data set for this case study, each individual has an associated sampling weight, but they are not all equal, since the survey was not a simple random sample. Using these in the analysis will

help the results reflect the survey population, not just the survey sample.

Background: Quality of Life

Quantity of life has been measured historically by life expectancies, mortality rates, preventable deaths and potential years of life lost. These indicators are mainly derived from vital statistics and population censuses, and can be obtained from most of the OECD countries. Life expectancy has been a very useful measure of population health. We know for instance that, with very few exceptions, life expectancy is increasing in the developed world (Monier, 1998). Social progress (especially in the field of public sanitation) and medical breakthroughs, including vaccines and other health-related technological advances over the past century, have reduced the impact of infectious and parasitic diseases to a mere fraction of what it once was.

For example, at the beginning of this century, a Canadian male could expect to live an average of 47 years and a Canadian female, 50 years. At that time, barely 38% of males and 44% of females reached the age of 65 years, and they could only expect to live roughly another decade thereafter (11 years for males and 12 for females). (Bourbeau, 1997) In 1997, life expectancy at birth in Canada reached 75.8 years for males and 81.4 years for females. More than 80% of males and almost 90% of females will celebrate their 65th birthday and they may expect to live an additional 16 and 20 years respectively thereafter (Martel, 1999).

The simple increase in the quantity of life, however, does not give any indication of whether there is an accompanying increase, decrease, or maintenance of the health-related quality of life. It provides an incomplete "snapshot" of the health of a population. Significant health intervention resources are currently being targeted not just to increase life lived, but also to reduce morbidity. This is especially true of ageing populations among which life expectancies are rising dramatically. The question of "how healthy a population is" cannot be answered by disease incidence reporting and mortality trends alone but rather must include a quantifiable estimate of health status that incorporates Canadian notions of health-related quality of life. The quality of life should also be monitored.

The Réseau Espérance de Vie en Santé/International Network on Health Expectancy (REVES), which represents university researchers and national health-related or statistical agencies across 15 countries, has promoted work in decomposing total life expectancy into non-overlapping categories, either based on disability, diseases or perceived health states. These indicators are used to identify the **expected** number of years to be lived by a population in specific health states and allow quantifying of the expected amount of time to be spent in less than full health.

These indicators have been extended to evaluate the impact of specific diseases on disability free life expectancy. For example, Nusselder et al (1996) have calculated the impact on life expectancy and disability free life expectancy of eliminating selected diseases, including arthritis, cancers and heart disease. These estimators allow the decomposition of expected length of life into specific health states (Mathers, 1992).

A second group of indicators focuses on deriving weighted life expectancy over a set of health states (however defined). These estimators attempt to factor explicit valuation of health states, using a health index, to provide estimates of "fully healthy life expectancy". Typically, weights for health states range from zero (dead or equivalent to dead) to unity (fully healthy). This group of indicators takes into account not only the expected quantity of life lived but also the health-related quality. However, much controversy exists about which health index to use. The main competitors are the EQA-5D, SF-36, the **Health Utility Index** and the Perceived-health/ADL index.

Health Utilities Index

The Health Utilities Index (HUI^{TM}) is a generic, preference-scored, comprehensive system for measuring health status, health-related quality of life, and producing utility scores.

The Health Utilities Group (HUG) focuses on preference-based measures of health-related quality of life for describing treatment process and outcomes in clinical studies, for population health studies, and economic evaluations of health care services.

For more information see http://www-fhs.mcmaster.ca/hug/