

## Documentation for Insurance Data Set

These data are taken from the *Medical Expenditure Panel Survey* conducted in 1996. These data were provided by Professor Harvey Rosen of Princeton University and were used in his paper with Craig Perry: “The Self-Employed Are Less Likely Than Wage-Earners to Have Health Insurance. So What?” in Douglas Holtz-Eakin and Harvey S. Rosen, eds., *Entrepreneurship and Public Policy*, MIT Press 2004. The data in **Insurance** are a subset of the data used in the Perry-Rosen paper.

### Series in Data Set

Variable	Description
healthy	self reported health status binary variable 1 = Healthy
anylim	“any limitation” binary variable 1 = Limitation
insured	health insurance binary variable 1 = Insured
selfemp	employment binary variable 1 = Self Employed
age	age in years
deg_nd	education binary variable 1 = No degree
deg_ged	education binary variable 1 = GED (High School Equivalent)
deg_hs	education binary variable 1 = High School
deg_ba	education binary variable 1 = Bachelor
deg_ma	education binary variable 1 = Masters
deg_phd	education binary variable 1 = Ph.D.
deg_oth	education binary variable 1 = Other
familysz	family size
race_bl	race binary variable 1 = Black
race_wh	race binary variable 1 = White
race_ot	race binary variable 1 = Other than Black or White
reg_ne	region binary variable 1 = NorthEast
reg_mw	region binary variable 1 = MidWest
reg_so	region binary variable 1 = South
reg_we	region binary variable 1 = West
male	sex binary variable 1 = Male
married	married binary variable 1 = Married