

Exercises for Use with Large Data Sets Available on the Following Website:
www.wiley.com/college/daniel

Refer to the dataset NC BIRTH800. The North Carolina State Center for Health Statistics and Howard W. Odum Institute for Research in Social Science at the University of North Carolina at Chapel Hill (A-20) make publicly available birth and infant death data for all children born in the state of North Carolina. These data can be accessed at www.irss.unc.edu/ncvital/bfd1down.html. Records on birth data go back to 1968. This comprehensive data set for the births in 2001 contains 120,300 records. The data represents a random sample of 800 of those births and selected variables. The variables are as follows:

Variable Label	Description
PLURALITY	Number of children born of the pregnancy
SEX	Sex of child (1 = male, 2 = female)
MAGE	Age of mother (years)
WEEKS	Completed weeks of gestation (weeks)
MARITAL	Marital status (1 = married, 2 = not married)
RACEMOM	Race of mother (0 = other non-White, 1 = White, 2 = Black 3 = American Indian, 4 = Chinese, 5 = Japanese, 6 = Hawaiian, 7 = Filipino, 8 = Other Asian or Pacific Islander)
HISPMOM	Mother of Hispanic origin (C = Cuban, M = Mexican, N = Non-Hispanic, O = other and unknown Hispanic, P = Puerto Rican, S = Central/South American, U = not classifiable)
GAINED	Weight gained during pregnancy (pounds)
SMOKE	0 = mother did not smoke during pregnancy 1 = mother did smoke during pregnancy
DRINK	0 = mother did not consume alcohol during pregnancy 1 = mother did consume alcohol during pregnancy
TOUNCES	Weight of child (ounces)
TGRAMS	Weight of child (grams)
LOW	0 = infant was not low birth weight 1 = infant was low birth weight
PREMIE	0 = infant was not premature 1 = infant was premature Premature defined at 36 weeks or sooner