

Description of Smoking Data Set (rates)

The dataset contains four variables:

1. *smoke_yrs*: Years of smoking (categorical) {15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59}.
2. *cigar_day*: Cigarettes per day (numerical).
3. *person_yrs*: Person-years at risk of lung cancer (numerical).
4. *case*: Number of lung cancer cases (count).