

Семинар 11.

1. Набор данных `healthcare.dta` представляет собой несбалансированную панель из 7293 человек. Всего насчитывается 27 326 наблюдений. Количество наблюдений колеблется от 1 до 7. В набор данных включены следующие показатели:

- ID = идентификационный номер
- YEAR = календарный год наблюдения
- DOCVIS = количество посещений врача за последние три месяца
- AGE = возраст (в годах)
- EDUC = годы обучения в школе
- FEMALE = женщина = 1; мужчина = 0
- HHNINC = номинальный ежемесячный чистый доход домохозяйства в немецких марках/10000
- HANDDUM = наличие инвалидности = 1; в противном случае = 0
- HADPER = степень инвалидности в процентах (0 – 100)
- HSAT = удовлетворенность здоровьем от 0 (низкая) – 10 (высокая)
- PUBLIC = застрахован в системе государственного медицинского страхования = 1; в противном случае = 0
- WORKING = занятый = 1; в противном случае = 0
- HHKIDS = наличие детей в возрасте до 16 лет = 1; в противном случае = 0
- MARRIED = женат = 1; в противном случае = 0
- BLUEC = синие воротнички (рабочий класс) = 1; в противном случае = 0
- WHITEC = белые воротнички (наемные работники, занимающиеся умственным трудом) = 1; в противном случае = 0
- SELF = самозанятый = 1; в противном случае = 0
- BEAMT = государственный служащий = 1; в противном случае = 0
- HOSPVIS = количество посещений больниц за последний календарный год
- ADDON = наличие дополнительного страхования = 1; в противном случае = 0
- NUMOBS = количество наблюдений за этим человеком
- NEWHSAT = перекодированное значение HSAT с исправленными ошибками кодирования.