

Car Data

Рассматриваются данные о 74 типах автомобилей. Каждый автомобиль (наблюдение) характеризуется 13 переменными. Данные за 1977-1978 годы.

Данные

Описание переменных

ЦЕНА	цена
РАССТ	расстояние в милях на галлон
РЕМОНТ78	насколько автомобиль требовал ремонт в 78 г в баллах по 5-бальной шкале; 5 – лучший, 1 – худший
РЕМОНТ77	насколько автомобиль требовал ремонт в 77 г в баллах по 5-бальной шкале; 5 – лучший, 1 – худший
ВЫСОТА	высота салона (дюймы)
ЗАДН_СИД	расстояние от переднего до заднего сиденья в дюймах
ОБЪ_САЛО	объем салона в кубических футах
ВЕС	вес в фунтах
ДЛИНА	длина в дюймах
ДИАМ_ПОВ	диаметр разворота в футах
ДВИГАТЕЛ	объем двигателя в кубических дюймах
G	Gear ratio для самой высокой передачи
КОМПАНИЯ	расположение руководства компании (1 = США, 2 = Япония, 3 = Европа).

Описание данных на английском

The car data set consists of 13 variables measured for 74 car types. The abbreviations 3 are as follows:

- X1: P Price,
- X2: M Mileage (in miles per gallone),
- X3: R78 Repair record 1978
(rated on a 5-point scale; 5 best, 1 worst),
- X4: R77 Repair record 1977 (scale as before),
- X5: H Headroom (in inches),
- X6: R Rear seat clearance
(distance from front seat back to rear seat, in inches),
- X7: Tr Trunk space (in cubic feet),
- X8: W Weight (in pound),
- X9: L Length (in inches),
- X10: T Turning diameter (clearance required to make a U-turn, in feet),
- X11: D Displacement (in cubic inches),
- X12: G Gear ratio for high gear,
- X13: C Company headquarter

(1 for U.S., 2 for Japan, 3 for Europe).

Данные впервые рассматривались в книге Chambers, Cleveland, Kleiner and Tukey, 1983

Данные обсуждаются в книге

Hardle, Simar Applied Multivariate Statistical Analysis

Задача

Постройте правило классификации, которое по показателям определяет, какая это машина: японская, американская или европейская.

Исследуйте, можно ли уменьшить число используемых показателей без потери качества классификации. Приведите полученную модель.