

ЗАДАНИЕ (КЛАСТЕРИЗАЦИЯ)

1. Перевести категориальные признаки в количественные
2. Перевести значения веса и роста в привычные нам (см, кг)
3. Разбить датасет на 2 датафрейма: в одном – мужчины, во втором – женщины. Определить для каждого поля каждого датафрейма средние значения веса и роста, значения квантилей распределения, максимальное и минимальное значение, стандартное отклонение, количество значений.
4. Построить столбчатую диаграмму распределения роста и веса для каждого датафрейма (у вас их сейчас должно быть 3 – общий, мужчины, женщины)
5. Кластеризовать общий датафрейм двумя методами (к-средних и иерархическая) на 2 кластера.
6. Определить точность работы моделей.