

# Определение оптимального количества кластеров

Введение в кластерный анализ

# Как определить, на каком шаге остановиться?

1. Анализ «протокола объединения».
2. Анализ дендрограммы.
3. Анализ решений, проверка устойчивости моделей с разным числом кластеров.

# Анализ протокола объединения

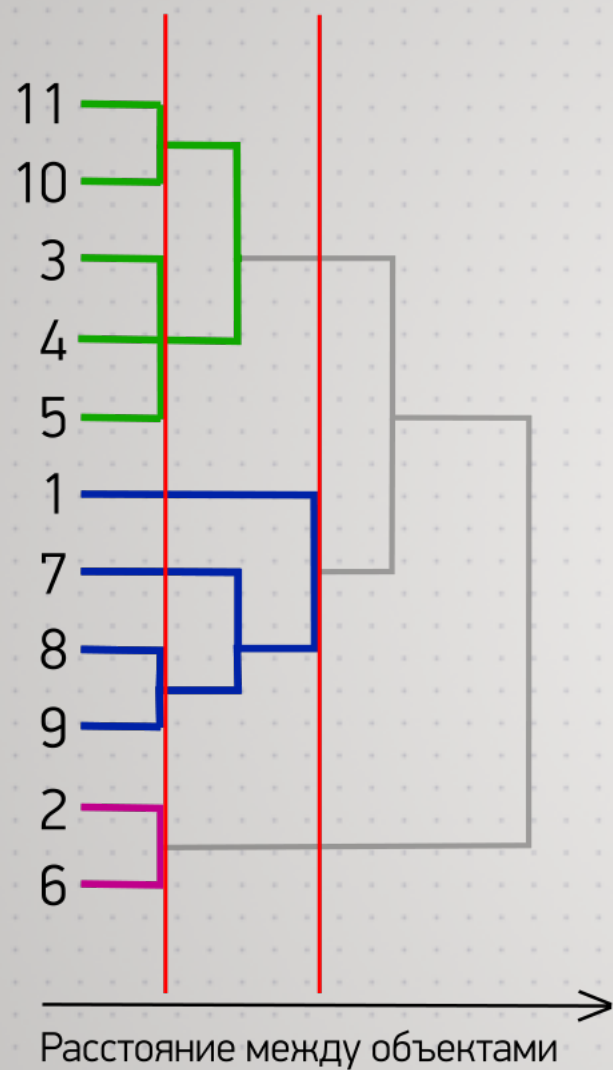
$n$  — число объектов

$N$  — число шагов ( $n - 1$ )

$k$  — номер шага, на котором произошёл скачок

$N - k$  — число кластеров

# Графическое отображение объединения объектов в кластеры: дендрограмма классификации



# Параметры для выбора количества кластеров

- Наполненность (10%).
- Сходство объектов внутри одного кластера.
- Различие объектов в разных кластерах.
- Устойчивость.
- Интерпретируемость.