1. **Анализ звуков, издаваемых различными видами лягушек**

*Описание:*

Датасет " Frogs\_MFCCs " представляет собой набор данных о записях звуков, издаваемых различными видами лягушек (anurans). Вот краткое описание переменных:

1. MFCCs\_1 - MFCCs\_22: Эти поля содержат значения 22 мел-частотных кепстральных коэффициентов (Mel-Frequency Cepstral Coefficients - MFCCs). MFCCs - это набор функций, широко используемых для представления звуковых сигналов.

2. Family: семейство, к которому принадлежит данная запись.

3. Genus: род, к которому принадлежит данная запись.

4. Species: вид, к которому принадлежит данная запись

5. RecordID: уникальный идентификатор.

1. **Предобработка данных:**  
   • Загрузите датасет и проверьте наличие пропущенных значений. Обработайте их  
   • Преобразуйте категориальные переменные, если они есть

2. **Выбор переменных для кластеризации:**  
  
3. **Анализ данных:**  
   • Постройте визуализации (гистограммы, диаграммы рассеяния) для понимания распределения выбранных переменных.  
   • Рассчитайте основные статистики (средние значения, медианы) для количественных переменных  
  
4. **Нормализация данных:**  
   • Примените нормализацию или стандартизацию к количественным переменным чтобы избежать влияния масштаба на кластеризацию.  
  
5. **Определение числа кластеров:**  
   • Используйте метод локтя или силуэтный метод для определения оптимального количества кластеров для K-means или другого алгоритма кластеризации.  
  
6. **Кластеризация:**  
   • Примените алгоритм K-means или иерархическую кластеризацию к подготовленным данным.

7. **Визуализация кластеров:**  
   • Постройте визуализации для представления результатов кластеризации (например, графики с центрами кластеров и их распределением по различным переменным).

8. Удалить из датасета признак(признаки), являющийся целевым и повторить кластеризацию. Сравнить результаты