**Программа для получения отдельных данных записей GenBank из NCBI**

Ссылка на GitHub: <https://github.com/xfluffymuffin/astrovirus.git>

Описание:

Программа использует три библиотеки: pandas, biopython, collections.

**Блок 1** – вначале с помощью biopython происходит извлечение из базы данных “NCBI Nucleotide” ID записей по запросу, задаваемому в рамках переменной *handle.*

**Блок 2** – для каждого из полученных ID затем производится загрузка соответствующих записей формата GenBank, из них выборочно извлекаются исходные данные, которые будут напечатаны в последнем блоке (source: ‘country’, ‘host’, ‘collection\_date’, ‘strain’, ‘isolate’).

**Блок 3** – с помощью pandas для каждой записи создается DataFrame: в них хранится извлеченная информация. Дополнительно подсчитывается и сохраняется в переменную *count* кол-во записей, для которых известна дата выделения образца.

**Блок 4** – создание Excel-файла с данными датафреймов посредством pandas. Необходимо, чтобы избежать замусоривания вывода ненужной в данный момент информацией и сохранить ее для дальнейшего анализа.

**Блок 5** – вывод необходимых данных на экран.