Инструкция для парсера

Для работы парсера необходим установленный пайтон. <https://www.python.org/ftp/python/3.11.4/python-3.11.4-amd64.exe>

При установке обходимо установить чек-бокс



Так же на компьютере необходим установленный браузер Chrome

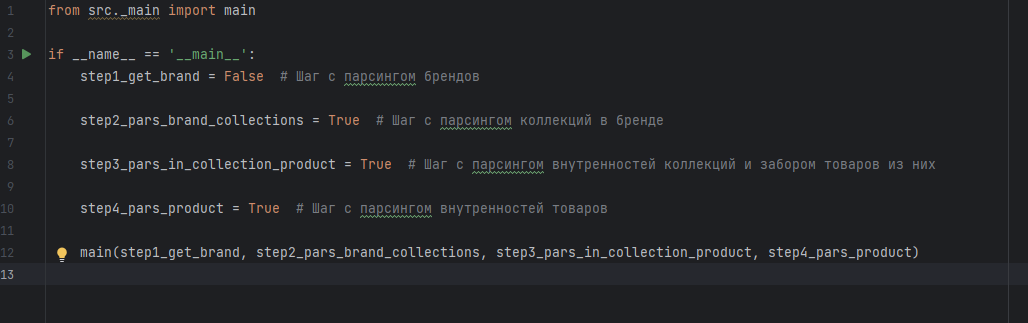
После установки пайтона необходимо распаковать архив в папку, в пути недолжно быть кириллицы

Для установки, необходимо один раз запустить **install.bat**

Для запуска настроенного парсер запускать **start.bat**

**Настройка парсера**

Что бы настроить парсер необходимо открыть файл main.py блокнотом, желательно notepad ++



Для того что бы парсер собрал все бренды с сайта то надо установить step1\_get\_brand в True

Т.е.

step1\_get\_brand = True

После выполнения данного шага все бренды сохранятся в файле brands.txt в каталоге с программой

Что бы собрать только бренды и более ни чего-то остальные шаги необходимо установить в False

Если нужно спарсить определенный бренд то данный шаг необходимо выключить т.е.

step1\_get\_brand = False

step2\_pars\_brand\_collections данный шаг отвечает за сбор коллекций по бренду, без товаров. Необходимо держать включенным.

step3\_pars\_in\_collection\_product данный шаг отвечает за парсинг внутренностей коллекций и сборку релевантного товара

step4\_pars\_product отвечает за шаг парсинга товара.

Если необходимо пропарсить 1 или несколько брендов (не все что есть)

То не обходимо выставить сборбу бренда в False, остальные шаги в True

Если нужно пропарсить все бренды то все шаги в True

!!! После парсинга, если нужно собрать новую информацию то папку files необходимо удалять т.к. там содержится «кеш» коллекций и товаров. Так же необходимо удалить базу данных rusplitka.db

В итоге каждый новый запуск необходимо удалить папку и базу данных.

**\*Результаты работы сохраняются в папку files/result/tovar**