**Алгоритм определения курса наличной валюты**

Детализация:

* Периодичность: тики
* Набор: доллары, евро
* География: Россия в целом; Москва; Регионы (?)

Источники:

* [**https://t.me/DollarEuroFrank**](https://t.me/DollarEuroFrank) **(рекомендация ДЗ - только этот сейчас)**
* <https://t.me/obmenvmsk> (тут больше людей и сообщений)
* <https://t.me/obmen_valuty_usd> (комментарии)
* <https://t.me/valyuty_obmen>

Схема исходных данных:

|  |  |
| --- | --- |
| **Поле** | **Значение** |
| dt | Дата и время сообщения |
| trans\_type | Направление операции |
| currency | Валюта USD / EUR / etc. |
| rate | Курс, руб. |
| volume | Объём покупки / продажи |
| region | Регион (опционально) |

Порядок сбора исходных данных

1. Получение web app api на https://my.telegram.org
2. Запрос данных чатов с помощью библиотеки telethon: загрузка json-файлов с сообщениями, выгрузка их в csv и загрузка в базу данных Clickhouse
3. Формирование правил распознавания котировок о покупке и продаже валюты на основе регулярных выражений

Ограничения

* не парсим количество
* не парсим регион
* не собираем несколько сообщений
* не собираем несколько котировок из одного сообщения