Необходимо написать Airflow DAG для выгрузки данных, их трансформации и сохранения в Clickhouse.

Выгрузка данных

Данные, которые нужно выгружать - отзывы на английском и русском языках в Google Play для официального приложения Telegram (org.telegram.messenger). Способ выгрузки отзывов не очень важен - это может быть скраппинг данных с сайта Google Play или готовая библиотека.

Трансформация

Выгруженные данные необходимо трансформировать в следующий вид:

event_date - date
language - str
reviews_count - int
min_score - float
avg_score - float
max_score - float
insert_date - date
insert_datetime - datetime

Группировка по event_date, language score - это оценка приложения

Сохранение

Данные необходимо сохранять в локальный инстанс clickhouse

Запуск настроить ежедневно