

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное
образовательное учреждение высшего образования
«Самарский национальный исследовательский университет
имени академика С.П. Королева»
(Самарский университет)

Институт информатики и кибернетики
Кафедра технической кибернетики

ЛАБОРАТОРНАЯ РАБОТА № 1

по курсу
Инженерия данных

Группа 6233-010402D

Студент _____

В.Ю. Мишагина

Самара 2025

1) Описание пайплайна

Заданием на лабораторную работу является реализация пайплайна.

В качестве источника данных предлагается использовать Free Weather API.

(Extract) Получить прогноз на завтра по переменным: температура, осадки, скорость и направление ветра для городов Самара и Москва. Сырые ответы API сохранить в объектном хранилище.

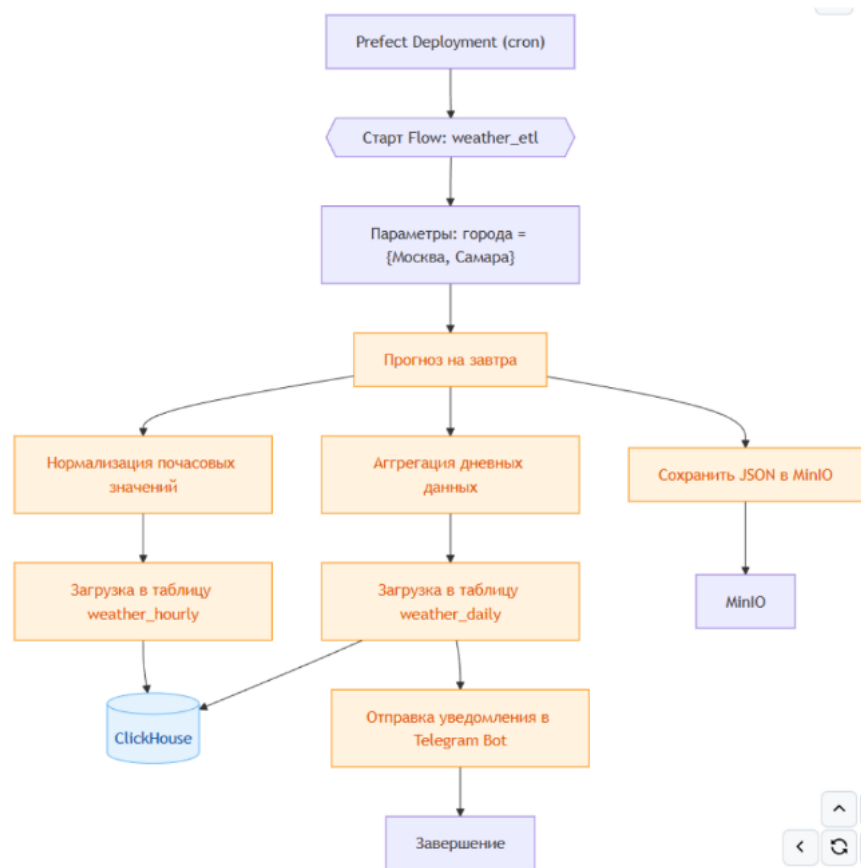
(Transform) Извлечь почасовые значения и нормализовать для таблицы `weather_hourly`. Посчитать дневную статистику (min, max, avg температуру и количество осадков) и подготовить для сохранения в таблице `weather_daily`

(Load) Загрузить преобразованные данные в соответствующие таблицы ClickHouse.

Автоматически отправить уведомления в Telegram с кратким прогнозом на завтра и предупреждать о сильном ветре/осадках.

2) Ход работы

Архитектура пайплайна:



Описание используемых инструментов.

Инструмент	Назначение
Prefect	Оркестрация ETL-пайплайна
ClickHouse	Хранение агрегированных данных
MinIO	Объектное хранилище для сырых данных
Open-Meteo	Публичный API прогноза погоды

Extract. Получаем прогноз на завтра по переменным для городов Самара и Москва и сохраняем сырые данные (полученный json) в объектном хранилище.

В качестве источника данных был использован Free Weather API <https://www.weatherapi.com/docs/>. Используемое API – http://api.weatherapi.com/v1/forecast.json?q={city}&days={forecast_days}&aqi=no&alerts={alerts}&key={api_key}.

Параметры:

- city - город для прогноза
- forecast_days – количество дней прогноза погоды
- alerts – включение/выключение оповещений о предупреждениях
- api_key – ключ для аутентификации

Ключ для аутентификации получаем на сайте при регистрации и выборе стандартного тарифного плана использования апи.

Пример запроса:

```
curl -X 'GET'
```

```
'https://api.weatherapi.com/v1/forecast.json?q=Moscow&days=2&alerts=yes&key=49a375767078452083e93755252611' -H 'accept: application/json'
```

Transform. Преобразование данных в формат для сохранения в таблицы. Для почасовых данных извлекаем информацию о температуре, скорости и направлении ветра, вероятности дождя и снега, а также о количестве осадков. Для дневной статистики – о минимальной, максимальной и средней температуре за день и об общем количестве осадков.

Load. Сохранение преобразованных данных в ClickHouse. Схемы

таблиц:

```
CREATE TABLE IF NOT EXISTS db1.weather_hourly
(
```

```
    city varchar,
    temp_c decimal32(3),
    wind_speed_kph decimal32(3),
    wind_direction varchar,
    chance_of_rain Int32,
    chance_of_snow Int32,
    precipitation_mm decimal32(3),
    time datetime
```

```
)
```

```
ENGINE = ReplacingMergeTree
```

```
ORDER BY (city, time);
```

```
CREATE TABLE IF NOT EXISTS db1.weather_daily
```

```
(
```

```
    city varchar,
    min_temp_c decimal32(3),
    max_temp_c decimal32(3),
    avg_temp_c decimal32(3),
    precipitation_mm decimal32(3),
    date date
```

```
)
```

```
ENGINE = ReplacingMergeTree
```

```
ORDER BY (city, date);
```

```
CREATE TABLE IF NOT EXISTS db1.alerts
```

```
(
```

```
    city varchar,
    headline varchar,
    message varchar,
    category varchar,
    instruction varchar,
    date date
```

```
)
```

```
ENGINE = ReplacingMergeTree
```

```
ORDER BY (city, date)
```

В результате заполняется 3 таблицы:

- weather_hourly – для хранения почасовых данных
- weather_daily – для хранения дневной статистики
- alerts – для хранения предупреждений о непогоде

Логи Prefect

```
18:29:56.893 | INFO   | Task run 'extract_weather_data-4b6' - Finished in state Completed()
```

```
18:29:56.944 | INFO   | Task run 'save_raw_data_to_s3-52a' - Finished in state Completed()
```

```
18:29:57.024 | INFO   | Task run 'transform_weather_hourly-60d' - Finished in state Completed()
```

18:29:57.064 | INFO | Task run 'transform_weather_daily-486' - Finished in state Completed()
 18:29:57.122 | INFO | Task run 'transform_weather_alerts-293' - Finished in state Completed()
 18:29:59.653 | INFO | Task run 'load_hourly_data-a4b' - Finished in state Completed()
 18:29:59.779 | INFO | Task run 'load_daily_data-a75' - Finished in state Completed()
 18:29:59.886 | INFO | Task run 'load_alert_data-e13' - Finished in state Completed()
 18:30:00.521 | INFO | Task run 'send_message-49c' - Finished in state Completed()
 18:30:01.078 | INFO | Task run 'extract_weather_data-3bf' - Finished in state Completed()
 18:30:01.111 | INFO | Task run 'save_raw_data_to_s3-d0d' - Finished in state Completed()
 18:30:01.158 | INFO | Task run 'transform_weather_hourly-54c' - Finished in state Completed()
 18:30:01.210 | INFO | Task run 'transform_weather_daily-773' - Finished in state Completed()
 18:30:01.259 | INFO | Task run 'transform_weather_alerts-ff3' - Finished in state Completed()
 18:30:03.651 | INFO | Task run 'load_hourly_data-766' - Finished in state Completed()
 18:30:03.758 | INFO | Task run 'load_daily_data-099' - Finished in state Completed()
 18:30:03.871 | INFO | Task run 'load_alert_data-a99' - Finished in state Completed()
 18:30:04.410 | INFO | Task run 'send_message-483' - Finished in state Completed()
 18:30:04.480 | INFO | Flow run 'honored-starling' - Finished in state Completed()

Содержимое бакета minio:

Object Browser

data-engineering
Created on: Sun, Dec 14 2025 18:26:44 (GMT+4) Access: PRIVATE 69.5 KiB - 2 Objects

Name	Last Modified	Size
Moscow_20251214_182956.json	Today, 18:29	33.9 KiB
Samara_20251214_183001.json	Today, 18:30	35.5 KiB

Запросы к ClickHouse:

select * from weather_daily

Результат 1

city	min_temp_c	max_temp_c	avg_temp_c	precipitation_mm	date
Moscow	-7,3	-5,3	-6,5	0	2025-12-15 GMT+
Samara	-10,5	-9,1	-9,9	1,14	2025-12-15 GMT+

select * from weather_hourly

Результат 1

select * from weather_hourly Введите SQL выражение чтобы отфильтровать результаты

	AZ city	123 temp_c	123 wind_speed_kph	AZ wind_direction	123 chance_of_rain	123 chance_of_snow	123 precipitation_mm	time
2	Moscow	-6,7	11,5	WNW	0	0	0	2025-12-15 01:00:00
3	Moscow	-6,7	10,8	WNW	0	0	0	2025-12-15 02:00:00
4	Moscow	-6,7	9,7	WNW	0	0	0	2025-12-15 03:00:00
5	Moscow	-6,6	9	WNW	0	0	0	2025-12-15 04:00:00
6	Moscow	-6,6	6,5	WNW	0	0	0	2025-12-15 05:00:00
7	Moscow	-6,6	5,4	WNW	0	0	0	2025-12-15 06:00:00
8	Moscow	-6,3	4	WNW	0	0	0	2025-12-15 07:00:00
9	Moscow	-6,4	2,2	N	0	0	0	2025-12-15 08:00:00
10	Moscow	-6,3	3,2	ENE	0	0	0	2025-12-15 09:00:00
11	Moscow	-6	3,2	NE	0	0	0	2025-12-15 10:00:00
12	Moscow	-5,6	3,6	NE	0	0	0	2025-12-15 11:00:00
13	Moscow	-5,4	4,7	NE	0	0	0	2025-12-15 12:00:00
14	Moscow	-5,3	5	NNE	0	0	0	2025-12-15 13:00:00
15	Moscow	-5,5	6,8	N	0	0	0	2025-12-15 14:00:00
16	Moscow	-6,1	6,8	NNE	0	0	0	2025-12-15 15:00:00
17	Moscow	-6,4	7,9	N	0	0	0	2025-12-15 16:00:00
18	Moscow	-6,8	6,8	NNE	0	0	0	2025-12-15 17:00:00
19	Moscow	-7	7,2	N	0	0	0	2025-12-15 18:00:00
20	Moscow	-7,1	8,3	NNE	0	0	0	2025-12-15 19:00:00
21	Moscow	-7,3	3,6	N	0	0	0	2025-12-15 20:00:00
22	Moscow	-7,3	5	N	0	0	0	2025-12-15 21:00:00
23	Moscow	-7,3	4	NW	0	0	0	2025-12-15 22:00:00
24	Moscow	-7,2	2,9	NW	0	0	0	2025-12-15 23:00:00
25	Samara	-10,4	22,7	WNW	0	80	0,05	2025-12-15 00:00:00
26	Samara	-10,4	22,3	WNW	100	75	0,06	2025-12-15 01:00:00

select * from alerts

Результат 1

select * from alerts Введите SQL выражение чтобы отфильтровать результаты

	AZ city	AZ headline	AZ message	AZ category	AZ instruction	date
	Moscow	Прочие опасности	На дорогах гололедица.	Прочие опасности		2025-12-15 GMT+
	Samara	Ветер	14 декабря сильный северо-западный ветер 17-22 м/с	Ветер		2025-12-15 GMT+

Пример уведомлений в Telegram:

December 14

Прогноз погоды на 2025-12-15 для города Moscow

🌡️ Min temp: -7.3°C

🌡️ Max temp: -5.3°C

☁️ Осадки: 0.0 мм

Внимание!

⚠️ Прочие опасности. На дорогах гололедица. 18:30

Прогноз погоды на 2025-12-15 для города Samara

🌡️ Min temp: -10.5°C

🌡️ Max temp: -9.1°C

☁️ Осадки: 1.1 мм

Внимание!

⚠️ Ветер. 14 декабря сильный северо-западный ветер 17-22 м/с 18:30

December 13 0:47

⚠️ Прочие опасности. на дорогах снежные заносы

Прогноз погоды на 2025-12-14 для города Moscow

🌡️ Min temp: -6.7°C

🌡️ Max temp: -5.4°C

☁️ Осадки: 0.0 мм

Внимание!

⚠️ Гололедица. На дорогах гололедица.

⚠️ Ветер. Местами усиление северо-западного ветра с порывами до 17 м/с.

⚠️ Гололедно - изморозевое отложение. На дорогах области сохранится гололедица.

⚠️ Прочие опасности. На дорогах гололедица. 18:32

Прогноз погоды на 2025-12-14 для города Samara

🌡️ Min temp: -10.1°C

🌡️ Max temp: -8.4°C

☁️ Осадки: 2.5 мм

Внимание!

⚠️ Прочие опасности. на дорогах снежные заносы

⚠️ Ветер. местами порывы 17-22 м/с, метель при ухудшении видимости 500-1000 м

⚠️ Снег. Местами очень сильный снег 18:33

3) Выводы

Сложным показалось настройка и запуск Prefect в контейнере Docker-контейнере, а также сбор и написание Docker-compose файла.

Я бы добавила обработку ошибок для предотвращения возможных точек сбоя.