

# Трек #2 – кодбук

В общей сложности в рамках данного трека вы получите пять архивов: основной и по одному архиву на каждую дополнительную выгрузку данных.

В основном архиве находятся следующие датасеты:

- `train.csv` — основной датасет с обучающей выборкой по гидрологическим данным, содержащий целевую переменную `delta_stage_max`;
- `meteo_3hours.csv` — восьмисрочные (через каждые 3 часа) наблюдения на метеостанциях;
- `meteo_1day.csv` — ежедневные наблюдения на метеостанциях;
- `meteo_1month.csv` — значения метеопараметров за месяц;
- `hydro_coord.csv` — географические координаты гидропостов;
- `meteo_coord.csv` — географические координаты метеостанций;
- `reference_water_codes.csv` — справочник по кодам режимных групп с расшифровкой, а также группировкой экспертов;
- `reference_horiz_visib.csv` — справочник по кодам дальности горизонтальной видимости;
- `ice_saw.csv` — информация по мероприятиям по ослаблению прочности льда.

Кроме того, последовательно будут выгружены ещё 4 архива. Архив номер  $n$  включает в себя тестовую выборку для периода от 21 апреля +  $2 * n$  недели, для которой и необходимо сделать предсказание. Поскольку задача о предсказании стоит только на **1 неделю вперед**, из каждого из 9 тестовых лет в тестовую выборку попала только одна неделя из двух. Если попавшая неделя четная, то предшествующая ей нечетная выдается с уже проставленными значениями, которые можно использовать в прогнозе; если попавшая неделя нечетная, то четная будет выдана в следующем архиве.

- `test.csv` — датасет с такими же идентификаторами, как и `train.csv`, с пропущенными значениями целевой переменной. Содержит всегда ровно 9 новых недель (по одной из каждого тестового года). Каждая новая тестовая выборка *содержит* предыдущую;
- `extra_train.csv` — дополнительные данные, предшествующие тестовым. В первом случае содержит ровно 4 недели (для 4 лет, в которых в первый тестсет попали четные недели), для оставшихся — также 9 новых недель. Новые тренировочные датасеты *не содержат* предыдущие;
- `extra_meteo_3hours.csv` — восьмисрочные наблюдения на метеостанциях за период, покрытый `extra_train.csv`;
- `extra_meteo_1day.csv` — ежедневные наблюдения на метеостанциях за период, покрытый `extra_train.csv`;
- `extra_meteo_1month.csv` — значения метеопараметров за месяц за период, покрытый `extra_train.csv`;
- `forecast_meteo_3hours.csv` — восьмисрочный прогноз («идеальный прогноз», фактически – реальные данные) четырех метеопараметров (`air_temperature`,

precipitation, wind\_speed\_aver, wind\_direction) на метеостанциях за период, покрытый `test.csv`. Новые прогнозные датасеты *не содержат* предыдущие.

Основные данные к хакатону делятся на две группы: гидрологические наблюдения, характеризующие состояние водных объектов вблизи гидропостов, и метеонаблюдения, описывающие погоду и климатические явления в окрестности метеостанций. Данные покрывают временной период с 1 января 1985 года; географически “работаем” [в этом регионе](#).

Из вспомогательных данных у нас есть географическая привязка пунктов наблюдения и справочники для расшифровки кодов категориальных признаков.

К важным (пусть и немногочисленным) данным также относятся сведения о мероприятиях по ослаблению льда — это противозаторная мера, антропогенный фактор в (не)возникновении заторов. Эти данные нельзя использовать в качестве предиктора (наблюдения в тестовые годы отсутствуют), однако они дают представление о способах использования будущей модели.

## **train.csv — наблюдения на гидропостах**

- `station_id` (INT) — общероссийский номер гидрологического пункта. Соответствует полю `station_id` в таблице `hydro_coord.csv`. Примеры: 3019, 3021, 3024.
- `date` (DATE) — дата наблюдения в формате `yyyy-mm-dd`. Пример: 2000-04-17.
- `stage_avg` (INT) — среднесуточный уровень воды, см. Пример: 17.
- `stage_min` (INT) — минимальный уровень воды за сутки, см. Пример: 16.
- `stage_max` (INT) — максимальный уровень воды за сутки, см. Пример: 18.
- `temp` (FLOAT) — среднесуточная температура воды, °C. Пример: 20.3.
- `water_code` (TEXT) — код режимной группы. Поле может содержать несколько кодов, которые разделены запятой. Режимная группа характеризует состояние водного объекта. Информация о режимных группах приведена в таблице `reference_water_codes.csv`. Примеры: '27', '5, 4, 1', '9,26,25,27'.
- `ice_thickness` (INT) — толщина льда, см. Пример: 33.
- `snow_height` (INT) — высота снега на льду, см. Пример: 56.
- `place` (TEXT) — код места ледовых измерений. В датасете атрибут принимает одно из трех значений:
  - '1' — на середине реки;
  - '2' — у берега;
  - '13' — на середине реки, на плесе.
- `discharge` (FLOAT) — среднесуточный расход воды, м³/с. Пример: 76.6.
- `delta_stage_max` (INT) — целевая переменная, обозначающая изменение максимального уровня воды между вчерашним днём и сегодняшним.4

## **hydro\_coord.csv — географическая привязка гидропостов**

- `station_id` (INT) — общероссийский номер гидрологического пункта. Соответствует полю `station_id` в таблицах `train.csv` и `test.csv`. Пример: 3019.

- `name` (TEXT) — название гидропоста и водного объекта, на котором он расположен. Примеры: 'КИРЕНСК - Р.ЛЕНА', 'БАТАМАЙ (ГП ПРИ БАТАМАЙ М) - Р.АЛДАН'.
- `lat` (FLOAT) — широта, градусы. Пример: 62.013708.
- `lon` (FLOAT) — долгота, градусы. Пример: 129.668088.
- `distance_from_source` (FLOAT) — расстояние от истока, км. Пример: 728.0.
- `drainage_area` (INT) — площадь водосбора, км<sup>2</sup>. Это площадь территории земной поверхности, с которой все поверхностные и грунтовые воды стекают в данный водоём или водоток, включая различные его притоки. Пример: 912000.
- `z_null` (FLOAT) — высотная отметка нуля графика на гидропосте, м. Дана по Байкальской системе высот. Пример: 175.51.

### **meteo\_3hours.csv — восьмисрочные наблюдения на метеостанциях**

- `station_id` (INT) — синоптический индекс метеостанции; соответствует полю `station_id` в таблице `meteo_coord.csv`. Примеры: 24449, 30393.
- `year_GMT` (INT) — год по Гринвичу. Пример: 1984.
- `month_GMT` (INT) — месяц по Гринвичу. Пример: 12.
- `day_GMT` (INT) — день по Гринвичу. Пример: 31.
- `synop_hour_GMT` (INT) — срок по Гринвичу. Наблюдения проводились в стандартные синоптические сроки с интервалом в 3 часа. Пример: 15.
- `year_local` (INT) — год по местному времени. Пример: 1985.
- `month_local` (INT) — месяц по местному времени. Пример: 1.
- `day_local` (INT) — день по местному времени. Пример: 1.
- `synop_hour_local` (INT) — срок по местному времени. Сроки наблюдений соответствуют стандартным синоптическим срокам: 0, 3, 6, 9, 12, 15, 18, 21. Таким образом, шаг между наблюдениями составляет 3 часа. Пример: 18.
- `synop_hour_order_PDZV` (INT) — номер срока в сутках по поясному декретному зимнему времени. Пример: 1.
- `time_local` (INT) — час по местному времени. Пример: 0.
- `timezone` (INT) — номер часового пояса. Пример: 8.
- `meteo_day_start_PDZV` (INT) — начало метеорологических суток по поясному декретному зимнему времени. Пример: 21.
- `source_index` (INT) — индекс архива-источника. В датасете принимает значение 4.
- `horizontal_visibility` (INT) — горизонтальная дальность видимости; наибольшее расстояние, с которого в светлое время суток перестает быть видимым абсолютно черный объект размером более 15', проектирующийся на фон неба у горизонта. Дальность видимости является показателем оптического состояния атмосферы. На метеорологических станциях измерение производится с помощью приборов, а в их отсутствие — визуально с помощью специально выбранных ориентиров. Горизонтальная дальность видимости приводится в цифрах кода. При инструментальном способе измерения используются цифры от 0 до 89, за исключением 51–55, а при визуальном — от 90 до 99. Коды обозначают следующее (также продублировано в `reference_horiz_visib.csv`):

<b>КОД</b>	<b>ДАЛЬНОСТЬ ВИДИМОСТИ, КМ</b>
0	< 0.1
1	0.1
2	0.2
3	0.3
4	0.4
5	0.5
6	0.6
7	0.7
8	0.8
9	0.9
10	1
11	1.1
12	1.2
13	1.3
14	1.4
15	1.5
16	1.6
17	1.7
18	1.8
19	1.9
20	2
21	2.1
22	2.2
23	2.3
24	2.4
25	2.5
26	2.6
27	2.7

28	2.8
29	2.9
30	3
31	3.1
<b>код</b>	<b>дальность видимости, км</b>
32	3.2
33	3.3
34	3.4
35	3.5
36	3.6
37	3.7
38	3.8
39	3.9
40	4
41	4.1
42	4.2
43	4.3
44	4.4
45	4.5
46	4.6
47	4.7
48	4.8
49	4.9
50	5
56	6
57	7
58	8
59	9
60	10

61	11
62	12
63	13
64	14
65	15
66	16
67	17
68	18
<b>код</b>	<b>дальность видимости, км</b>
69	19
70	20
71	21
72	22
73	23
74	24
75	25
76	26
77	27
78	28
79	29
80	30
81	35
82	40
83	45
84	50
85	55
86	60
87	65
88	70

89	> 70
90	< 0.05
91	0.05
92	0.2
93	0.5
94	1
95	2
96	4
97	10
98	20
99	> 50

- `horizontal_visibility_qual` (INT) — признак качества атрибута `horizontal_visibility`. Расшифровка кодов качества приведена после этого списка.
- `horizontal_visibility_qual_1` (INT) — признак наличия знака “>” у значений `horizontal_visibility`: 0 — знак “>” отсутствует; 1 — знак “>” присутствует.
- `cloud_amount_total` (INT) — общее количество облачности. Оценивается визуально как степень покрытия небосвода облаками по 13-бальной шкале. Кодировается в баллах от 0 до 13. Величина 0 означает полное отсутствие облаков или покрытие облаками менее 1/10 небосвода, а значение 10 означает, что небосвод полностью покрыт облаками. 11 обозначает наличие следов облаков; 12 — 10 баллов с просветами; 13 — облака невозможно определить. Просветы между индивидуальными облачными элементами, типичные для некоторых форм облаков (*Altostratus*, *Stratocumulus*) не включаются в общее количество облачности, т.е. они считаются чистым небом. Значение 99 означает отсутствие наблюдений.
- `cloud_amount_total_qual` (INT) — признак качества атрибута `cloud_amount_total`. Расшифровка кодов качества приведена после этого списка.
- `cloud_amount_low_level` (INT) — количество облачности нижнего яруса. Оценивается визуально как степень покрытия небосвода облаками по 13-бальной шкале. Кодировается в баллах от 0 до 13. Величина 0 означает полное отсутствие облаков или покрытие облаками менее 1/10 небосвода, а значение 10 означает, что небосвод полностью покрыт облаками. 11 обозначает наличие следов облаков; 12 — 10 баллов с просветами; 13 — облака невозможно определить. Просветы между индивидуальными облачными элементами, типичные для некоторых форм облаков (*Altostratus*, *Stratocumulus*) не включаются в общее количество облачности, т.е. они считаются чистым небом. Значение 99 означает отсутствие наблюдений.
- `cloud_amount_low_level_qual` (INT) — признак качества атрибута `cloud_amount_low_level`. Расшифровка кодов качества приведена после этого списка.

- `cloud_form_high_level` (INT) — форма облаков верхнего яруса. К таким облакам относятся облака, нижняя граница которых находится выше 6000 м, а именно: *cirrus* (Ci), *cirrocumulus* (Cc), *cirrostratus* (Cs). Данная характеристика кодируется следующим образом:
  - 0 — облака отсутствуют;
  - 1 — Ci;
  - 2 — Cc;
  - 3 — Cs;
  - 4 — Ci и Cc;
  - 5 — Ci и Cs;
  - 6 — Cc и Cs;
  - 7 — Ci, Cc и Cs;
  - 8 — туман или форму облаков невозможно определить;
  - 9 — форму облаков невозможно определить из-за темноты или атмосферных явлений.
- `cloud_form_high_level_qual` (INT) — признак качества атрибута `cloud_form_high_level`. Расшифровка кодов качества приведена после этого списка.
- `cloud_form_middle_level` (INT) — форма облаков среднего яруса. К облакам среднего яруса относятся облака, нижняя граница которых находится в пределах от 2000 до 6000 м, а именно: *altocumulus* (Ac), *altostratus* (As). Кодируется следующим образом:
  - 0 — облака отсутствуют;
  - 1 — Ac;
  - 2 — As;
  - 3 — не используется;
  - 4 — Ac и As;
  - 5–7 — не используются;
  - 8 — туман или форму облаков невозможно определить;
  - 9 — форму облаков невозможно определить из-за темноты или атмосферных явлений.
- `cloud_form_middle_level_qual` (INT) — признак качества атрибута `cloud_form_middle_level`. Расшифровка кодов качества приведена после этого списка.
- `cloud_form_vertical_develop` (INT) — форма облаков вертикального развития. Эти облака — *cumulus* (Cu) и *cumulonimbus* (Cb) — относятся к облакам нижнего яруса, хотя по высоте занимают несколько ярусов, но их нижняя граница находится в нижнем ярусе, т.е. ниже 2000 м. Кодируется следующим образом:
  - 0 — облака отсутствуют;
  - 1 — Cu;
  - 2 — Cb;
  - 3 — не используется;
  - 4 — Cu и Cb;
  - 5–7 — не используются;
  - 8 — туман или форму облаков невозможно определить;



- 9 — форму облаков невозможно определить из-за темноты или атмосферных явлений.
- cloud\_form\_vertical\_develop\_qual (INT) — признак качества атрибута cloud\_form\_vertical\_develop. Расшифровка кодов качества приведена после этого списка.
- cloud\_form\_strat\_stratocum (INT) — форма слоистых и слоисто-кучевых облаков. Эта группа облаков, которая включает stratus (St) и stratocumulus (Sc), также относится к облакам нижнего яруса. Кодировается следующим образом:
  - 0 — облака отсутствуют;
  - 1 — St;
  - 2 — Sc;
  - 3 — не используется;
  - 4 — St и Sc;
  - 5–7 — не используются;
  - 8 — туман или форму облаков невозможно определить;
  - 9 — форму облаков невозможно определить из-за темноты или атмосферных явлений.
- cloud\_form\_strat\_stratocum\_qual (INT) — признак качества атрибута cloud\_form\_strat\_stratocum. Расшифровка кодов качества приведена после этого списка.
- cloud\_form\_strat\_rain (INT) — форма слоисто-дождевых и разорвано-дождевых облаков. Последняя группа облаков нижнего яруса, которая состоит из nimbostratus (Ns) и fractonimbus (Frnb), кодируется следующим образом:
  - 0 — облака отсутствуют;
  - 1 — не используется;
  - 2 — Ns;
  - 3 — Frnb;
  - 4–5 — не используются;
  - 6 — Ns и Frnb;
  - 7 — не используется;
  - 8 — туман или форму облаков невозможно определить;
  - 9 — форму облаков невозможно определить из-за темноты или атмосферных явлений.
- cloud\_form\_strat\_rain\_qual (INT) — признак качества атрибута cloud\_form\_strat\_rain. Расшифровка кодов качества приведена после этого списка.
- cloud\_base\_altitude (INT) — высота нижней границы облаков, м. При тумане высота нижней границы облаков кодируется одной цифрой 0.
- cloud\_base\_altitude\_qual (INT) — признак качества атрибута cloud\_base\_altitude. Расшифровка кодов качества приведена после этого списка.
- cloud\_base\_altitude\_qual\_1 (INT) — признак способа определения высоты нижней границы облачности cloud\_base\_altitude. Если высота нижней границы облаков определялась визуально, то признак способа определения высоты принимает значение 0. При инструментальном определении высоты нижней границы облаков этот признак равен 9.

- `cloud_below_station` (INT) — признак наличия облачности ниже уровня станции. Используется только на высокогорных станциях и только тогда, когда наблюдалась облачность ниже уровня станции. Он может принимать значения:
  - 1 — окрестность станции ниже ее уровня частично покрыта облаками;
  - 2 — ниже уровня станции наблюдается сплошная облачность.
- `cloud_below_station_qual` (INT) — признак качества атрибута `cloud_below_station`. Расшифровка кодов качества приведена после этого списка.
- `soil_surface_condition` (INT) — состояние поверхности почвы. В зависимости от времени года кодируется следующим образом:

лето		зима	
код	описание	код	описание
0	Поверхность почвы сухая (без трещин, заметного количества пыли и сыпучего песка)	0	Лед, в основном покрывающий поверхность земли
1	Поверхность почвы влажная (без луж)	1	Слежавшийся или мокрый снег (со льдом или без него), покрывающий менее половины поверхности земли
2	Поверхность почвы сырая (вода застаивается на поверхности и образует малые и большие лужи)	2	Слежавшийся или мокрый снег (со льдом или без него), покрывающий половину или более поверхности земли (но не полностью)
3	Поверхность почвы затоплена водой	3	Равномерный слой слежавшегося или мокрого снега, покрывающий поверхность земли полностью
4	Поверхность почвы замерзшая	4	Неравномерный слой слежавшегося или мокрого снега, покрывающий поверхность земли полностью
5	Поверхность почвы покрыта коркой льда, но без снега или таящего снега	5	Сухой рассыпчатый снег, покрывающий менее половины поверхности земли
6	Поверхность почвы частично покрыта сухой пылью или сыпучим песком	6	Сухой рассыпчатый снег, покрывающий половину или более половины поверхности земли (но не полностью)
7	Поверхность почвы полностью покрыта тонким слоем сухой пыли или сыпучего песка	7	Равномерный слой сухого рассыпчатого снега, покрывающий поверхность земли полностью

8	Поверхность почвы полностью покрыта умеренным или толстым слоем сухой пыли или сыпучего песка	8	Неравномерный слой сухого рассыпчатого снега, покрывающий поверхность земли полностью
9	Поверхность почвы чрезвычайно сухая с трещинами	9	Снег с глубокими сугробами, заносами, покрывающий поверхность земли полностью

- `soil_surface_condition_qual` (INT) — признак качества атрибута `soil_surface_condition`. Расшифровка кодов качества приведена после этого списка.
- `weather_before` (INT) — погода в течение трех часов, предшествующих сроку наблюдения. Кодировка следующим образом:
  - 0 — ясно или облачность не более 5 баллов;
  - 1 — меняющаяся облачность: в течение рассматриваемого периода облачность была временами более 5 баллов, а временами 5 баллов и менее;
  - 2 — пасмурно или облачность более 5 баллов;
  - 3 — песчаная или пыльная буря; поземок или низовая метель;
  - 4 — туман или ледяной туман; сильная мгла;
  - 5 — морось;
  - 6 — дождь;
  - 7 — снег или дождь со снегом;
  - 8 — ливневые осадки;
  - 9 — гроза с осадками или без них.
- `weather_before_qual` (INT) — признак качества атрибута `weather_before`. Расшифровка кодов качества приведена после этого списка.
- `weather` (INT) — погода в срок наблюдения или в течение последнего часа перед сроком наблюдения. Приводится в цифрах кода от 0 до 99. Цифры кода позволяют закодировать 100 различных характеристик погоды. Эти характеристики разделены на десятки и на две большие группы — без осадков на станции в срок наблюдения и с осадками. Кодировка следующим образом:

код	описание
А. Без осадков на станции в срок наблюдения	
0–19	погода без осадков, тумана, ледяного тумана (за исключением 11–12), пыльной или песчаной бури, низовой метели или поземка на станции в срок наблюдения и (за исключением 9 и 17) в течение последнего часа
0	условия развития облаков неизвестны
1	облака в целом рассеивались
2	состояние неба в целом не изменилось
3	облака образовались или развивались

4	видимость ухудшена из-за дыма или вулканического пепла
5	мгла
6	пыль в срок наблюдения, взвешенная в воздухе на обширном пространстве, но не поднятая ветром на станции или вблизи нее
7	пыль или песок, поднятые ветром на станции, но без развития песчаных вихрей или пыльной бури
8	хорошо развитый пыльный или песчаный вихрь, но никакой пыльной или песчаной бури не наблюдается
9	пыльная или песчаная буря в поле зрения
10	дымка (видимость 1000 м или более)
11	поземный туман или поземный ледяной туман клочками
12	поземный туман или поземный ледяной туман более или менее сплошным слоем
13	зарница
14	осадки в поле зрения, но не достигающие поверхности земли
15	осадки в поле зрения, достигающие поверхности земли на расстоянии более 5 км от станции
16	осадки в поле зрения, достигающие поверхности земли вблизи, но не на станции
17	гроза в срок наблюдения, но без осадков
18	шквал
19	смерч
20– 29	осадки, туман, ледяной туман или гроза на станции в течение последнего часа, но не в срок наблюдения
20	морось или снежные зерна
21	дождь
22	снег
23	дождь со снегом или ледяной дождь
24	морось или дождь с образованием гололеда
25	ливневый дождь
26	ливневый снег или ливневый дождь со снегом
27	град, ледяная или снежная крупа с дождем или без дождя
28	туман или ледяной туман (видимость менее 1000 м)

29	гроза с осадками или без них
30– 39	пыльная или песчаная буря, поземок или низовая метель в срок наблюдения
30	слабая или умеренная пыльная или песчаная буря ослабела в течение последнего часа
31	слабая или умеренная пыльная или песчаная буря без заметного изменения интенсивности в течение последнего часа
32	слабая или умеренная пыльная или песчаная буря началась или усилилась в течение последнего часа
33	сильная пыльная или песчаная буря ослабела в течение последнего часа
34	сильная пыльная или песчаная буря без заметного изменения интенсивности в течение последнего часа
35	сильная пыльная или песчаная буря началась или усилилась в течение последнего часа
36	слабый или умеренный поземок, при котором перенос снега происходит ниже уровня глаз наблюдателя
37	сильный поземок
38	слабая или умеренная низовая метель
39	сильная низовая метель
40– 49	туман или ледяной туман в срок наблюдения
40	туман или ледяной туман в окрестностях станции
41	туман или ледяной туман местами
42	туман или ледяной туман ослаб в течение последнего часа, небо видно
43	туман или ледяной туман ослаб в течение последнего часа, небо не видно
44	туман или ледяной туман без заметного изменения интенсивности в течение последнего часа, небо видно
45	туман или ледяной туман без заметного изменения интенсивности в течение последнего часа, небо не видно
46	туман или ледяной туман начался или усилился в течение последнего часа, небо видно
47	туман или ледяной туман начался или усилился в течение последнего часа, небо не видно

48	туман с отложением изморози, небо видно
49	туман с отложением изморози, небо не видно
Б. Осадки на станции в срок наблюдения	
50– 59	морось
50	морось с перерывами, слабая
51	морось непрерывная, слабая
52	морось с перерывами, умеренная
53	морось непрерывная, умеренная
54	морось с перерывами, сильная
55	морось непрерывная, сильная
56	морось слабая, образующая гололед
57	морось умеренная и сильная, образующая гололед
58	морось слабая с дождем
59	морось умеренная и сильная с дождем
60– 69	дождь
60	дождь с перерывами, слабый
61	дождь непрерывный, слабый
62	дождь с перерывами, умеренный
63	дождь непрерывный, умеренный
64	дождь с перерывами, сильный
65	дождь непрерывный, сильный
66	дождь слабый, образующий гололед
67	дождь умеренный или сильный, образующий гололед
68	дождь или морось со снегом, слабые
69	дождь или морось со снегом, умеренные или сильные
70– 79	твердые осадки, не ливневые
70	снег с перерывами, слабый

71	снег непрерывный, слабый
72	снег с перерывами, умеренный
73	снег непрерывный, умеренный
74	снег с перерывами, сильный
75	снег непрерывный, сильный
76	ледяные иглы
77	снежные зерна
78	отдельные снежные кристаллы, похожие на звездочки
79	ледяной дождь
80– 89	ливневые осадки без грозы
80	ливневый дождь слабый
81	ливневый дождь умеренный или сильный
82	ливневый дождь очень сильный
83	ливневый дождь со снегом, слабый
84	ливневый дождь со снегом, умеренный или сильный
85	ливневый снег, слабый
86	ливневый снег, умеренный или сильный
87	ледяная или снежная крупа слабая, с дождем, со снегом и дождем или без них
88	ледяная или снежная крупа умеренная или сильная, с дождем, со снегом и дождем или без них
89	град слабый с дождем, со снегом и дождем или без них
90	град умеренный или сильный с дождем, со снегом и дождем или без них
91– 99	гроза в срок наблюдения или в течение последнего часа
91	дождь слабый, гроза в течение последнего часа
92	дождь умеренный или сильный, гроза в течение последнего часа
93	снег или снег с дождем, град или крупа, слабые, гроза в течение последнего часа
94	снег или снег с дождем, град или крупа, умеренные или сильные, гроза в течение последнего часа

95	гроза слабая или умеренная с дождем и/или снегом в срок наблюдения
96	гроза слабая или умеренная с градом или крупой в срок наблюдения
97	гроза сильная с дождем или снегом
98	гроза вместе с песчаной или пыльной бурей в срок наблюдения
99	гроза сильная с градом или крупой

- `weather_qual` (INT) — признак качества атрибута `weather`. Расшифровка кодов качества приведена после этого списка.
- `wind_direction` (INT) — направление ветра. Дается в градусах. Штиль кодируется одной цифрой 0, а переменное направление — 999.
- `wind_direction_qual` (INT) — признак качества атрибута `wind_direction`. Расшифровка кодов качества приведена после этого списка.
- `wind_speed_aver` (INT) — средняя скорость ветра. Скорость ветра измеряется в м/с на высоте 10–12 м. При штиле скорость кодируется цифрой 0.
- `wind_speed_aver_qual` (INT) — признак качества атрибута `wind_speed_aver`. Расшифровка кодов качества приведена после этого списка.
- `wind_speed_aver_qual_1` (INT) — признак наличия знака “>” у значений `wind_speed_aver`: 0 — знак “>” отсутствует; 1 — знак “>” присутствует.
- `wind_speed_max` (INT) — максимальная скорость ветра за 3 часа, включая порывы. Кодировка по тем же правилам, что и средняя скорость ветра `wind_speed_aver`.
- `wind_speed_max_qual` (INT) — признак качества атрибута `wind_speed_max`. Расшифровка кодов качества приведена после этого списка.
- `wind_speed_max_qual_1` (INT) — признак наличия знака “>” у значений `wind_speed_max`: 0 — знак “>” отсутствует; 1 — знак “>” присутствует.
- `precipitation` (FLOAT) — сумма осадков за период между сроками, когда измеряются осадки, в мм с точностью до десятых долей. Согласно «Наставлению гидрометеорологическим станциям и постам» (вып. 3, ч. 1, 1985), измерение количества осадков, выпавших за ночную и дневную половины суток, на метеостанциях бывшего СССР производится в сроки, ближайшие к 8 и 20 часам поясного декретного (зимнего) времени. С 1966 года в каждое измерение осадков непосредственно на станции вносится поправка на смачивание, равная 0.1 мм для твердых осадков и 0.2 мм — для жидких. Значение суммы осадков равное 0 означает отсутствие осадков, если признак качества равен 5, и наличие следов осадков, если признак качества равен 0.
- `precipitation_qual` (INT) — признак качества атрибута `precipitation`. Расшифровка кодов качества приведена после этого списка.
- `soil_surface_temperature` (FLOAT) — температура поверхности почвы. Значение температуры поверхности почвы по срочному термометру в градусах с точностью до десятых долей. Температура поверхности измеряется на оголенной от растительности поверхности почвы или поверхности снежного покрова.
- `soil_surface_temperature_qual` (INT) — признак качества атрибута `soil_surface_temperature`. Расшифровка кодов качества приведена после этого списка.



- `soil_surface_alco_temperature` (FLOAT) — температура поверхности почвы по спирту минимального термометра; приводится в градусах с точностью до десятых долей.
- `soil_surface_alco_temperature_qual` (INT) — признак качества атрибута `soil_surface_alco_temperature`. Расшифровка кодов качества приведена после этого списка.
- `soil_surface_temperature_min_before` (FLOAT) — минимальная температура поверхности почвы за период между сроками по штифту минимального термометра; приводится в градусах с точностью до десятых долей.
- `soil_surface_temperature_min_before_qual` (INT) — признак качества атрибута `soil_surface_temperature_min_before`. Расшифровка кодов качества приведена после этого списка.
- `soil_surface_temperature_max_before` (FLOAT) — максимальная температура поверхности почвы за период между сроками наблюдений по максимальному термометру; приводится в градусах с точностью до десятых долей.
- `soil_surface_temperature_max_before_qual` (INT) — признак качества атрибута `soil_surface_temperature_max_before`. Расшифровка кодов качества приведена после этого списка.
- `soil_surface_max_temperature` (FLOAT) — температура поверхности почвы по максимальному термометру после встряхивания; приводится в градусах с точностью до десятых долей.
- `soil_surface_max_temperature_qual` (INT) — признак качества атрибута `soil_surface_max_temperature`. Расшифровка кодов качества приведена после этого списка.
- `air_temperature` (FLOAT) — температура воздуха по сухому термометру; приводится в градусах с точностью до десятых долей. При температуре воздуха ниже -36 °C кодируется значение низкоградусного спиртового термометра, а в случае его отсутствия значение температуры определяется по столбику спирта минимального термометра.
- `air_temperature_qual` (INT) — признак качества атрибута `air_temperature`. Расшифровка кодов качества приведена после этого списка.
- `air_temperature_wet_bulb` (FLOAT) — температура воздуха по смоченному термометру; приводится в градусах с точностью до десятых долей. В холодную часть года при температуре ниже -10 °C характеристики влажности снимают с лент самописцев.
- `air_temperature_wet_bulb_qual` (INT) — признак качества атрибута `air_temperature_wet_bulb`. Расшифровка кодов качества приведена после этого списка.
- `air_temperature_wet_bulb_qual_1` (INT) — признак наличия льда на батисте смоченного термометра: 0 — лед на батисте отсутствует; 1 — лед на батисте присутствует.
- `air_temperature_min_alco_temperature` (FLOAT) — температура воздуха по спирту минимального термометра. Кодируется по тем же правилам, что и температура воздуха `air_temperature`.

- `air_temperature_min_alco_temperature_qual` (INT) — признак качества атрибута `air_temperature_min_alco_temperature`. Расшифровка кодов качества приведена после этого списка.
- `air_temperature_min_before` (FLOAT) — минимальная температура воздуха между сроками по штифту минимального термометра с учетом поправки из поверочного свидетельства, но без учета добавочной поправки. Приводится в градусах с точностью до десятых долей.
- `air_temperature_min_before_qual` (INT) — признак качества атрибута `air_temperature_min_before`. Расшифровка кодов качества приведена после этого списка.
- `air_temperature_max_before` (FLOAT) — максимальная температура воздуха между сроками наблюдений по максимальному термометру в градусах с точностью до десятых долей. При температуре воздуха ниже  $-36^{\circ}\text{C}$  выбирают с лент термографа.
- `air_temperature_max_before_qual` (INT) — признак качества атрибута `air_temperature_max_before`. Расшифровка кодов качества приведена после этого списка.
- `air_max_temperature` (FLOAT) — температура воздуха по максимальному термометру после встряхивания. Приводится в градусах с точностью до десятых долей.
- `air_max_temperature_qual` (INT) — признак качества атрибута `air_max_temperature`. Расшифровка кодов качества приведена после этого списка.
- `water_vapour_partial_pressure` (FLOAT) — парциальное давление водяного пара (упругость водяного пара). Основная характеристика влажности; представляет собой парциальное давление водяного пара, содержащегося в воздухе. Выражается в миллибарах или миллиметрах ртутного столба, как и давление воздуха. Определяется с помощью психрометрических таблиц по измерениям температуры сухого и смоченного термометров, а при температуре ниже  $-10^{\circ}\text{C}$  — по исправленным показаниям гигрометра и сухого термометра. Значение парциального давления приводится с точностью до десятых долей (при этом дополнительная характеристика `water_vapour_partial_pressure_qual_1` равна 1) или до сотых долей (дополнительная характеристика `water_vapour_partial_pressure_qual_1` равна 2).
- `water_vapour_partial_pressure_qual` (INT) — признак качества атрибута `water_vapour_partial_pressure`. Расшифровка кодов качества приведена после этого списка.
- `water_vapour_partial_pressure_qual_1` (INT) — дополнительная характеристика атрибута `water_vapour_partial_pressure`. Принимает значение 1, если парциальное давление `water_vapour_partial_pressure` приводится с точностью до десятых долей. Принимает значение 2, если парциальное давление `water_vapour_partial_pressure` приводится с точностью до сотых долей.
- `relative_humidity` (INT) — относительная влажность — это отношение фактической упругости водяного пара к упругости насыщенного воздуха при той же температуре, выраженное в процентах. Характеризует степень насыщения воздуха водяным паром. Может принимать значения от 0 до 100.

- `relative_humidity_qual` (INT) — признак качества атрибута `relative_humidity`. Расшифровка кодов качества приведена после этого списка.
- `vapour_pressure_deficit` (FLOAT) — дефицит насыщения водяного пара — разность между насыщающей и фактической упругостью водяного пара. Кодировается так же, как и парциальное давление водяного пара `water_vapour_partial_pressure`. Значение приводится с точностью до десятых долей (при этом дополнительная характеристика `vapour_pressure_deficit_qual_1` равна 1) или до сотых долей (дополнительная характеристика `vapour_pressure_deficit_qual_1` равна 2).
- `vapour_pressure_deficit_qual` (INT) — признак качества атрибута `vapour_pressure_deficit`. Расшифровка кодов качества приведена после этого списка.
- `vapour_pressure_deficit_qual_1` (INT) — дополнительная характеристика атрибута `vapour_pressure_deficit`. Принимает значение 1, если дефицит насыщения водяного пара `vapour_pressure_deficit` приводится с точностью до десятых долей. Принимает значение 2, если дефицит насыщения водяного пара `vapour_pressure_deficit` приводится с точностью до сотых долей.
- `dew_point_temperature` (FLOAT) — температура точки росы — это температура, при которой воздух достигает состояния насыщения при данном содержании водяного пара и неизменном давлении. При насыщении, т.е. при относительной влажности 100%, температура воздуха совпадает с температурой точки росы. Приводится в градусах с точностью до десятых долей.
- `dew_point_temperature_qual` (INT) — признак качества атрибута `dew_point_temperature`. Расшифровка кодов качества приведена после этого списка.
- `pressure` (FLOAT) — атмосферное давление на уровне станции; на метеорологических станциях измеряется с помощью станционного чашечного ртутного барометра. Приводится в гПа (мб) с точностью до десятых долей.
- `pressure_qual` (INT) — признак качества атрибута `pressure`. Расшифровка кодов качества приведена после этого списка.
- `pressure_sea_level` (FLOAT) — атмосферное давление на уровне моря. Согласно «Наставлению гидрометеорологическим станциям и постам» (часть 1, выпуск 3, 1985), вычисляется давление на уровне моря (для станций, расположенных не выше 1000 м над уровнем моря) или высота ближайшей изобарической поверхности (для станций, расположенных выше 1000 м). Приводится в гПа (мб) с точностью до десятых долей.
- `pressure_sea_level_qual` (INT) — признак качества атрибута `pressure_sea_level`. Расшифровка кодов качества приведена после этого списка.
- `barometric_tendency_characteristic` (INT) — характеристика барической тенденции; характеризует изменение атмосферного давления на станции за последние 3 часа. Определяется по записи барографа и кодируется следующим образом:
  - 0 — рост, затем падение;
  - 1 — рост, затем без изменения или более слабый рост; давление в срок наблюдения выше, чем 3 часа назад;

- 2 — рост равномерный или неравномерный; давление в срок наблюдения выше, чем 3 часа назад;
- 3 — падение, затем рост; без изменения, затем рост; рост, а затем более сильный рост; давление в срок наблюдения выше, чем 3 часа назад;
- 4 — ровный или неровный ход; давление такое же, что и 3 часа назад;
- 5 — падение, затем рост; давление такое же, что и 3 часа назад;
- 6 — падение, затем без изменений; давление ниже, чем 3 часа назад;
- 7 — равномерное или неравномерное падение; давление ниже, чем 3 часа назад;
- 8 — рост, затем падение; без изменения, затем падение; падение, затем более сильное падение; давление ниже, чем 3 часа назад.
- `barometric_tendency_characteristic_qual` (INT) — признак качества атрибута `barometric_tendency_characteristic`. Расшифровка кодов качества приведена после этого списка.
- `barometric_tendency` (FLOAT) — величина барической барической тенденции — разница между текущим значением атмосферного давления на станции и тем, что наблюдалось 3 часа назад. Приводится с точностью до десятых долей гПа (мб).
- `barometric_tendency_qual` (INT) — признак качества атрибута `barometric_tendency`. Расшифровка кодов качества приведена после этого списка.

Большинству элементов массива присвоены признаки качества (имеют название `%feature%_qual`), которые могут принимать следующие значения:

Код качества	Описание
0	значение элемента достоверно
1	значение элемента достоверно и восстановлено вручную
2	значение элемента достоверно и восстановлено автоматически
3	значение элемента сомнительно
4	значение элемента забраковано программами синтаксического и семантического контроля
5	значение элемента отсутствует, но наблюдения проводились
6	значение элемента забраковано на станции
7	значения элемента отсутствуют, т.к. наблюдения не производились

### **meteo\_1day.csv — ежедневные наблюдения на метеостанциях**

В эту таблицу вошли данные четырех тематик: маршрутные снегомерные съемки, наблюдения за снежным покровом на станции, температура почвы на глубине и температура почвы по коленчатым термометрам Савинова.

Поскольку частота наблюдений различна, в этой таблице ожидается много пропусков. Например, маршрутные снегосъемки осуществляются через каждые 10 дней в течение холодного периода (каждые пять дней в период интенсивного снеготаяния), а наблюдения за снежным покровом на станции — ежедневные.

- `station_id` (INT) — синоптический индекс метеостанции. Соответствует полю `station_id` в таблице `meteo_coord.csv`. Примеры: 24966, 30493.
- `year` (INT) — год. Пример: 1985.
- `month` (INT) — месяц. Пример: 1.
- `day` (INT) — день. Пример: 1.
- `route_type` (INT) — тип маршрута. Снегосъемки осуществляются отдельно для трех видов ландшафта, и коды типов маршрутов распределены следующим образом: 1 — поле, 2 — лес и 3 — овраги. Длина маршрута составляет 1 или 2 км (в поле и в лесу). Каждые 10 (в лесу) или 20 (в поле) метров измеряется высота снежного покрова, каждые 100 (в лесу) или 200 (в поле) метров измеряются остальные характеристики снежного покрова. В оврагах измерения проводятся только по заданию Гидрометеорологической обсерватории.
- `snow_coverage_near_station` (INT) — степень покрытия окрестности станции снегом. Значения изменяются от 0 до 10, где 10 — 100% покрытия.
- `snow_coverage_route` (INT) — степень покрытия маршрута снегом. Значения изменяются от 0 до 10, где 10 — 100% покрытия.
- `ice_crust_route` (INT) — Степень покрытия маршрута ледяной коркой. Значения изменяются от 0 до 10, где 10 — 100% покрытия.
- `snow_height_aver` (INT) — средняя высота снежного покрова на маршруте, см.
- `snow_height_max` (INT) — наибольшая высота снежного покрова на маршруте, см.
- `snow_height_min` (INT) — наименьшая высота снежного покрова на маршруте, см.
- `snow_density_aver` (FLOAT) — средняя плотность снега, г/см<sup>3</sup>.
- `ice_crust_aver` (INT) — средняя толщина ледяной корки, мм.
- `snow_saturated_thickness` (INT) — толщина слоя снега, насыщенного водой, мм.
- `water_thickness` (INT) — толщина слоя чистой воды, мм.
- `water_in_snow` (INT) — запас воды в снеге, мм.
- `water_total` (INT) — общий запас воды, мм.
- `snow_coverage_charact` (INT) — характер залегания снежного покрова. Значения кодов:
  - 0 — равномерный снежный покров на замерзшей почве без сугробов;
  - 1 — равномерный снежный покров на оттаявшей почве без сугробов;
  - 2 — равномерный снежный покров без сугробов, состояние почвы не известно;
  - 3 — неравномерный снежный покров на замерзшей почве, небольшие сугробы;
  - 4 — неравномерный снежный покров на оттаявшей почве, небольшие сугробы;
  - 5 — неравномерный снежный покров, состояние почвы не известно, небольшие сугробы;
  - 6 — очень неравномерный снежный покров на замерзшей почве, большие сугробы;

- 7 — очень неравномерный снежный покров на оттаявшей почве, большие сугробы;
  - 8 — очень неравномерный снежный покров, состояние почвы не известно, большие сугробы;
  - 9 — снежный покров с проталинами.
- snow\_charact (INT) — характер снежного покрова. Значения кодов:
  - 0 — свежий снег, пылевидный;
  - 1 — свежий снег, пушистый;
  - 2 — свежий снег, липкий;
  - 3 — старый снег, рассыпчатый;
  - 4 — старый снег, плотный;
  - 5 — старый снег, влажный;
  - 6 — снежная корка, не связанная со снегом под ней;
  - 7 — плотный снег с коркой на поверхности;
  - 8 — влажный снег с коркой на поверхности;
  - 9 — снег, насыщенный водой.
- snow\_height (INT) — высота снега, см. Значение 9999 означает, что значение высоты снега забраковано или наблюдения не проводились.
- snow\_coverage\_station (INT) — степень покрытия окрестности станции снегом. Значения изменяются от 0 до 10, где 10 — 100% покрытия. 99 означает, что значение степени покрытия окрестности станции снегом забраковано или наблюдения не проводились.
- snow\_height\_q1 (INT) — дополнительная информация о высоте снежного покрова. Коды обозначают:
  - 0 — данные о высоте снежного покрова верные;
  - 1 — отсутствие снега;
  - 2 — снежный покров отсутствует на станции, однако в окрестностях станции снег есть;
  - 3 — высота снега меньше 0.5 см;
  - 9 — наблюдения не проводились или значение высота снега забраковано.
- snow\_height\_q2 (INT) — признак качества по высоте снежного покрова. Осуществлено сравнение *последовательных* значений высоты снежного покрова: snow\_height\_q2 присваивается значение 1, если разница перепадов последовательных значений высоты снега составляет 30 см или более, причем высота снега соседних наблюдений различается на 10 и более см; во всех остальных случаях snow\_height\_q2 присваивается значение 0.
- snow\_height\_q3 (INT) — дополнительная информация с учетом температуры воздуха. Осуществлен анализ значений высоты снежного покрова при различных значениях среднесуточной и минимальной за сутки температуры воздуха. snow\_height\_q3 присваивается значение 1, если при положительной минимальной температуре и средней за сутки температуре воздуха более 5 °C высота снежного покрова больше 0. Во всех остальных случаях snow\_height\_q3 присваивается значение 0.
- temperature\_20cm (INT) — температура почвы на глубине 20 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.

- `temperature_20cm_qual (INT)` — признак качества признака `temperature_20cm`.  
Расшифровка кодов качества приведена после этого списка в первой таблице.
- `temperature_40cm (INT)` — температура почвы на глубине 40 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_40cm_qual (INT)` — признак качества признака `temperature_40cm`.  
Расшифровка кодов качества приведена после этого списка в первой таблице.
- `temperature_80cm (INT)` — температура почвы на глубине 80 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_80cm_qual (INT)` — признак качества признака `temperature_80cm`.  
Расшифровка кодов качества приведена после этого списка в первой таблице.
- `temperature_120cm (INT)` — температура почвы на глубине 120 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_120cm_qual (INT)` — признак качества признака `temperature_120cm`.  
Расшифровка кодов качества приведена после этого списка в первой таблице.
- `temperature_160cm (INT)` — температура почвы на глубине 160 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_160cm_qual (INT)` — признак качества признака `temperature_160cm`.  
Расшифровка кодов качества приведена после этого списка в первой таблице.
- `temperature_240cm (INT)` — температура почвы на глубине 240 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_240cm_qual (INT)` — признак качества признака `temperature_240cm`.  
Расшифровка кодов качества приведена после этого списка в первой таблице.
- `temperature_320cm (INT)` — температура почвы на глубине 320 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_320cm_qual (INT)` — признак качества признака `temperature_320cm`.  
Расшифровка кодов качества приведена после этого списка в первой таблице.
- `temperature_ks_5cm (INT)` — температура почвы по коленчатым термометрам Савинова на глубине 5 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_ks_5cm_qual (INT)` — признак качества признака `temperature_ks_5cm`.  
Расшифровка кодов качества приведена после этого списка во второй таблице.
- `temperature_ks_10cm (INT)` — температура почвы по коленчатым термометрам Савинова на глубине 10 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_ks_10cm_qual (INT)` — признак качества признака `temperature_ks_10cm`.  
Расшифровка кодов качества приведена после этого списка во второй таблице.
- `temperature_ks_15cm (INT)` — температура почвы по коленчатым термометрам Савинова на глубине 15 см, увеличенная в 10 раз, °C. Константой отсутствия данных является 9999.
- `temperature_ks_15cm_qual (INT)` — признак качества признака `temperature_ks_15cm`.  
Расшифровка кодов качества приведена после этого списка во второй таблице.

- temperature\_ks\_20cm (INT) — температура почвы по коленчатым термометрам Савинова на глубине 20 см, увеличенная в 10 раз, °С. Константой отсутствия данных является 9999.
- temperature\_ks\_20cm\_qual (INT) — признак качества признака temperature\_ks\_20cm. Расшифровка кодов качества приведена после этого списка во второй таблице.

Элементам массива вида temperature\_###cm присвоены признаки качества temperature\_###cm\_qual, которые могут принимать следующие значения:

Код качества	Описание
0	значение достоверно
1	значение ошибочно, забраковано методом гистограмм
2	значение ошибочно, забраковано методом проверки на сигмы
3	значение ошибочно, забраковано проверкой на связанность соседних по времени значений
4	значение ошибочно, забраковано проверкой соседних значений на ошибку обратного знака
5	значение сомнительно согласно методу гистограмм (т.к. данный метод контроля применить не удалось)
6	значение сомнительно согласно методу проверки на сигмы
7	значение сомнительно согласно проверке на связанность соседних по времени значений
8	значение ошибочно: забраковано более, чем одним методом одновременно
9	отсутствие данных

Элементам массива вида temperature\_ks\_##cm присвоены признаки качества temperature\_ks\_##cm\_qual, которые могут принимать следующие значения:

Код качества	Описание
0	значение достоверно
1–8	значение ошибочно и было забраковано на разных этапах создания массива
9	отсутствие данных
10	значение ошибочно, забраковано методом предельных значений



## **meteo\_1month.csv — значения метеопараметров за месяц**

Здесь можно найти данные о скорректированных месячных суммах осадков и суммарной за месяц продолжительности солнечного сияния.

- `station_id` (INT) — синоптический индекс метеостанции. Соответствует полю `station_id` в таблице `meteo_coord.csv`. Примеры: 24944, 31004.
- `year` (INT) — год. Пример: 1998.
- `month` (INT) — месяц. Пример: 3.
- `data_qual` (INT) — признак полноты. Отражает наличие пропусков отдельных сроков наблюдений в массиве исходных срочных данных:
  - 0 — пропуски в срочных наблюдениях отсутствуют;
  - 1 — число пропусков 1–10 %;
  - 2 — число пропусков более 10%;
  - 9 — данные отсутствуют полностью.
- `precipitation_observed` (FLOAT) — месячная сумма осадков по данным наблюдений с поправкой на смачивание, мм. Константа отсутствия — 9999.9.
- `precipitation_corrected` (FLOAT) — месячная сумма осадков с устранением систематических погрешностей осадкомерных приборов, мм. Константа отсутствия — 9999.9.
- `precipitation_corrected_liquid` (FLOAT) — месячная сумма жидких осадков с устранением систематических погрешностей осадкомерных приборов, мм. Константа отсутствия — 9999.9.
- `precipitation_corrected_mixed` (FLOAT) — месячная сумма смешанных осадков с устранением систематических погрешностей осадкомерных приборов, мм. Константа отсутствия — 9999.9.
- `precipitation_corrected_solid` (FLOAT) — месячная сумма твердых осадков с устранением систематических погрешностей осадкомерных приборов, мм. Константа отсутствия — 9999.9.
- `sunshine_hours` (INT) — продолжительность солнечного сияния, час. Если наблюдения не производились, то значение продолжительности отсутствует. Если продолжительность равна нулю, значит, солнце в течение месяца не появлялось или же продолжительность его сияния составила менее получаса.

## **meteo\_coord.csv — географические координаты метеостанций**

- `station_id` (INT) — синоптический индекс метеостанции. Соответствует полю `station_id` в таблицах `meteo_3hours.csv`, `meteo_1day.csv`, `meteo_1month.csv`. Примеры: 24538, 31102.
- `name` (TEXT) — название метеостанции. Примеры: 'БОДАЙБО', 'Средний Калар'.
- `lat` (FLOAT) — широта, градусы. Примеры: 60.82, 62.01370819.
- `lon` (FLOAT) — долгота, градусы. Примеры: 106.17, 112.58333333.
- `z` (INT) — высота метеоплощадки, м. Пример: 225.

## **ice\_saw.csv — мероприятия по ослаблению льда**

- `date` (DATE) — дата проведения мероприятия. Примеры: '2011-04-11', '2013-04-24'.
- `place` (TEXT) — название места проведения мероприятия. Примерно соответствует названию гидропоста, но не всегда. Примеры: 'Олёкминск', 'Саха (Якутия)'.
- `name` (TEXT) — наименование мероприятия с кратким описанием. Пример: 'Ослабление прочности льда на р. Лена в районе г. Якутск, Республика Саха (Якутия)'.
- `lat` (FLOAT) — широта, градусы. Примеры: 60.383333, 61.48333333333333.
- `lon` (FLOAT) — долгота, градусы. Примеры: 120.43333333333334, 114.565.

## reference\_water\_codes.csv — коды режимных групп

Справочная таблица с расшифровками номеров режимных групп от Росгидромета и группировками эксперта Светланы Агафоновой. Режимная группа объединяет несколько КСВО (кодов состояния водного объекта), то есть в целом характеризует состояние реки. Номера режимных групп можно встретить в поле `water_code` таблиц `train.csv` и `test.csv`.

- `water_code` (TEXT) — номер режимной группы.
- `descr` (TEXT) — описание режимной группы.
- `group` (TEXT) — Светланой сгруппированы явления, описывающие вскрытие реки весной. Как я понимаю, хронология такая: 'подготовка к вскрытию' — 'подвижка льда' — 'ледоход', 'вторичный ледоход' — 'затор выше поста', 'затор ниже поста', 'последствия ледохода'.
- `phase` (TEXT) — Светланой разбиты режимные группы в зависимости от сезона: 'замерзание (осень)', 'ледостав (зима)', 'вскрытие (весна)', 'не лед'.

Пара строк таблицы в качестве примера:

<code>water_code</code>	<code>descr</code>	<code>group</code>	<code>phase</code>
7	Средняя, густая снежура	None	замерзание (осень)
8	Редкий ледоход	ледоход	вскрытие (весна)

## reference\_horiz\_visib.csv — коды дальности горизонтальной видимости

Просто таблица для быстрой расшифровки кодов в поле `horizontal_visibility` таблицы `meteo_3hours.csv`.

- `code` (INT) — код.
- `km` (TEXT) — дальность видимости, км.