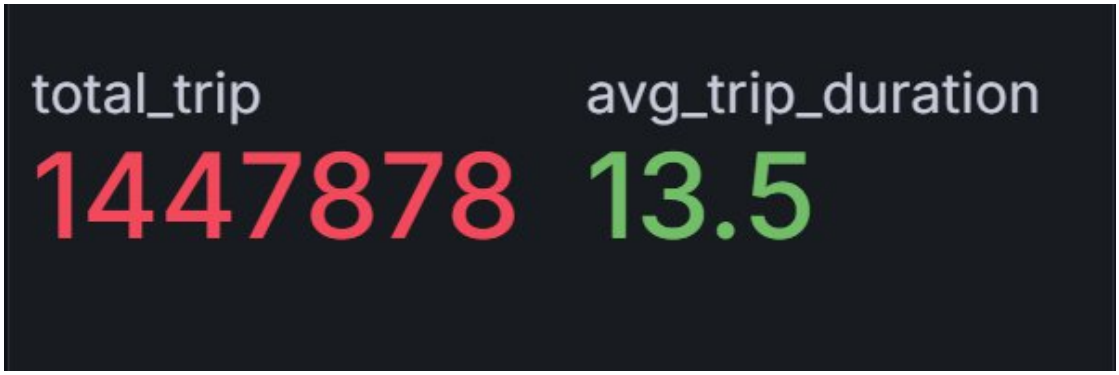
Общее количество поездок, среднее время поездки в мин:

```
SELECT
        COUNT(*) AS total_trip,
        ROUND(AVG(trip_duration_minute),2)
              AS avg_trip_duration
FROM dm.taxi_trip_and_weather_nyc;
```

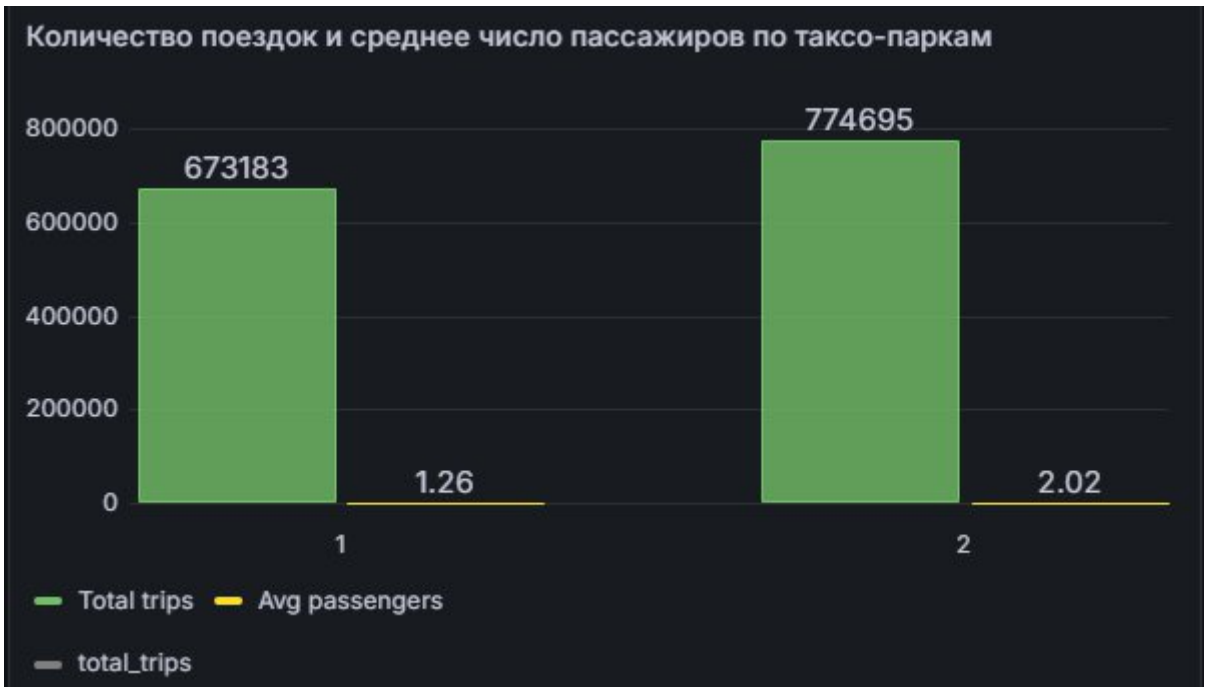
total_trip	avg_trip_duration
1447878	13.5



Количество поездок и среднее число пассажиров по таксопаркам:

```
SELECT
        vendor_id,
        COUNT(*) AS total_trips,
        ROUND(AVG(passenger_count),2) AS avg_pass_count
FROM dm.taxi_trip_and_weather_nyc
GROUP BY vendor_id
```

Vendor	Total trips	Avg passengers
1	673183	1.26
2	774695	2.02

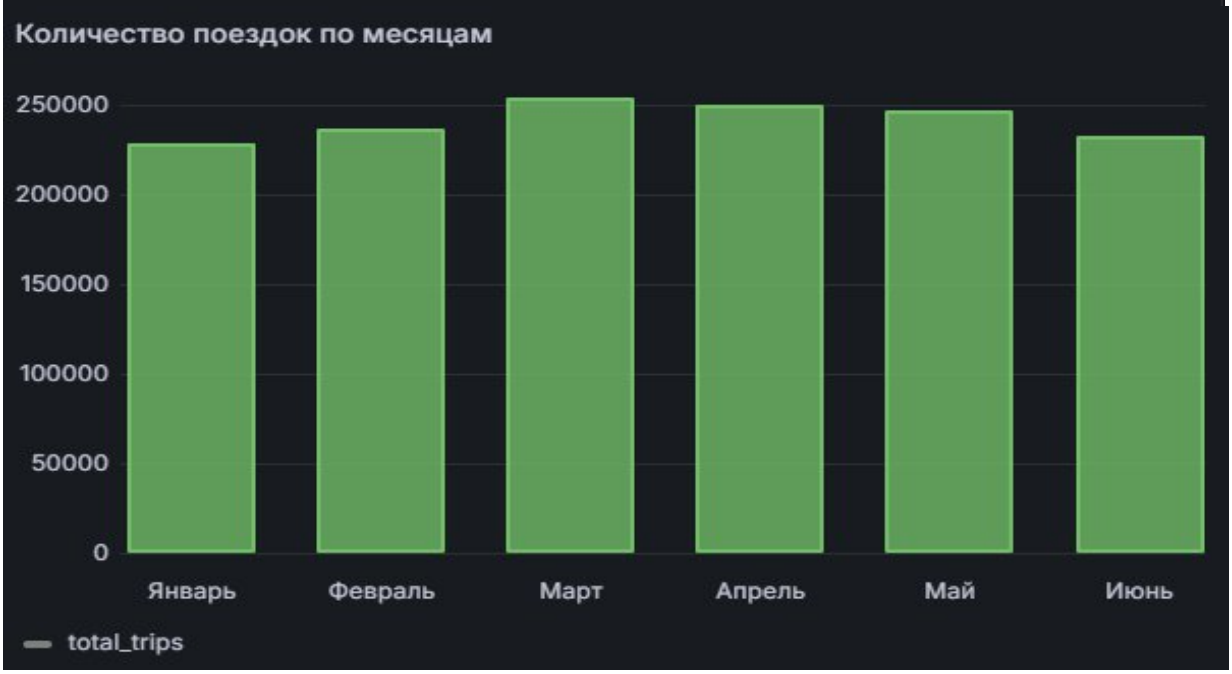


Количество поездок по месяцам и среднее число пассажиров:

```
SELECT
    CASE
        WHEN month_dim = 1 THEN 'Январь'
        WHEN month_dim = 2 THEN 'Февраль'
        WHEN month_dim = 3 THEN 'Март'
        WHEN month_dim = 4 THEN 'Апрель'
        WHEN month_dim = 5 THEN 'Май'
        WHEN month_dim = 6 THEN 'Июнь'
    END AS month_name,
    COUNT(*) AS total_trip_month,
    ROUND(AVG(passenger_count),1) AS avg_passenger_trip,
    ROUND(AVG(trip_duration_minute),2)
    AS avg_trip_duration
FROM dm.taxi_trip_and_weather_nyc
```

```
GROUP BY month_dim
ORDER BY month_dim;
```

month_name	total_trip_month	avg_passenger...	avg_trip_duration
Январь	228059	1.70	12.8
Февраль	236609	1.70	12.8
Март	254321	1.70	13.2
Апрель	249725	1.70	13.7
Май	246618	1.70	14.3
Июнь	232546	1.70	14.4



Количество поездок, среднее число пассажиров и среднее время поездки по дням недели:

```
SELECT
        dow_dim,
        CASE
                WHEN dow_dim = 0 THEN 'Воскресенье'
                WHEN dow_dim = 1 THEN 'Понедельник'
                WHEN dow_dim = 2 THEN 'Вторник'
                WHEN dow_dim = 3 THEN 'Среда'
                WHEN dow_dim = 4 THEN 'Четверг'
                WHEN dow_dim = 5 THEN 'Пятница'
                WHEN dow_dim = 6 THEN 'Суббота'
        END AS dow_day,
        COUNT(*) AS total_trips_dow,
        ROUND(AVG(passenger_count), 1)
        AS   avg_passenger_trip_dow,
        ROUND(AVG(trip_duration_minute),2)
                AS avg_trip_duration
FROM dm.taxi_trip_and_weather_nyc
GROUP BY dow_dim, dow_day
```

```
ORDER BY dow_dim;
```

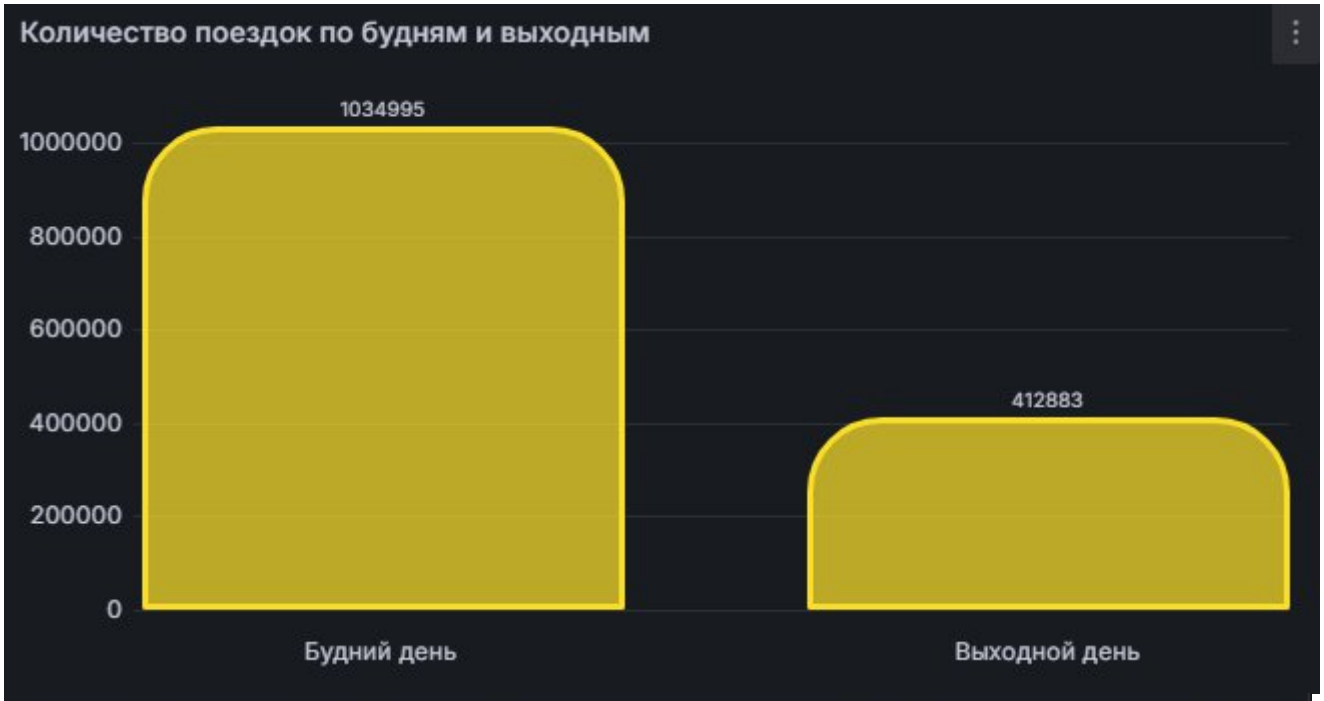
dow_dim	dow_day	total_trips_dow	avg_passenger_tri...	avg_trip_duration
0	Воскресенье	193682	1.70	12.3
1	Понедельник	186059	1.60	13.1
2	Вторник	201336	1.60	13.8
3	Среда	208768	1.60	14.3
4	Четверг	216955	1.60	14.6
5	Пятница	221877	1.70	14.0
6	Суббота	219201	1.70	12.6

Количество поездок и количество пассажиров по будням и выходным:

```
SELECT
    CASE
        WHEN dow_dim = 0 OR dow_dim = 6
            THEN 'Выходной день'
    ELSE 'Будний день'
    END AS dow_day,
    COUNT(*) AS total_trips,
    ROUND(AVG(passenger_count), 1) AS avg_passenger,
    ROUND(AVG(trip_duration_minute),2)
        AS avg_trip_duration
```

```
FROM dm.taxi_trip_and_weather_nyc
GROUP BY 1
```

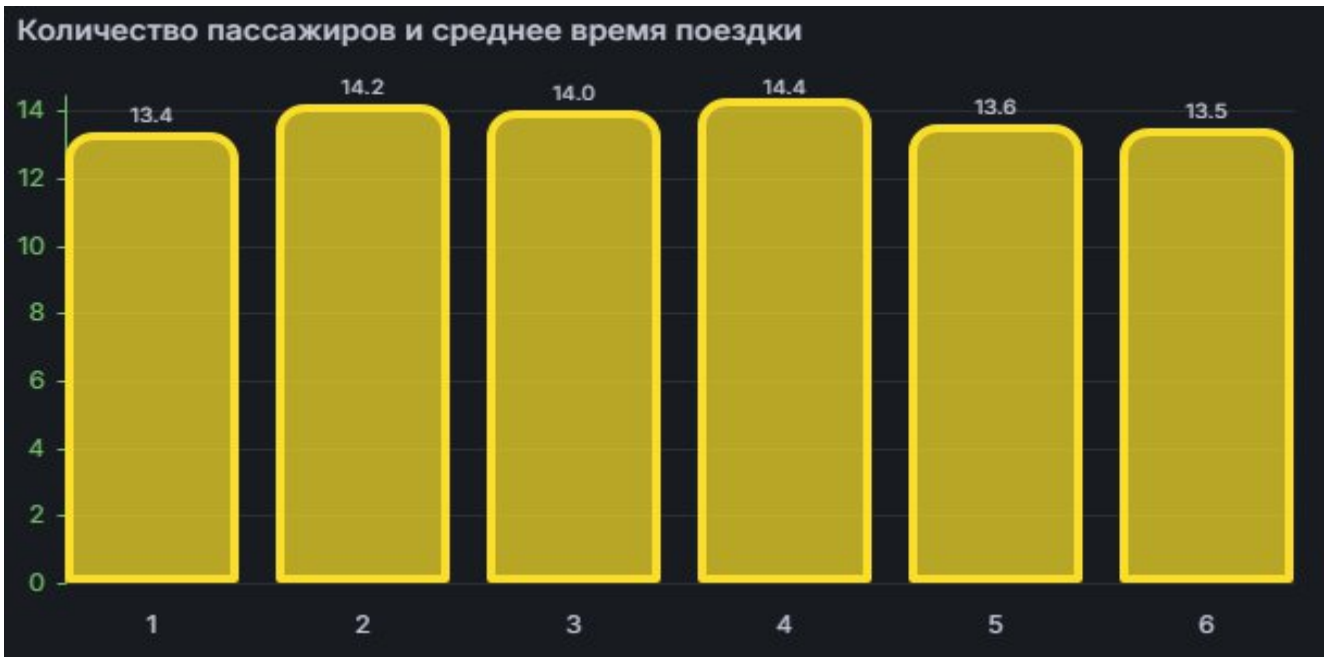
dow_day	total_trips	avg_passenger	avg_trip_duration
Будний день	1034995	1.60	14.0
Выходной день	412883	1.70	12.4



Количество поездок и среднее время поездки по количеству пассажиров:

```
SELECT
    passenger_count,
    count(*) AS count_trips,
    ROUND(avg(trip_duration_minute),2)
    AS avg_trip_duration
FROM dm.taxi_trip_and_weather_nyc
GROUP BY passenger_count
ORDER BY count_trips DESC;
```

count_passenger	count_trips	avg_trip_duration
1	1025386	13.4
2	209098	14.2
5	77580	13.6
3	59552	14.0
6	48035	13.5
4	28227	14.4



Средняя температура по месяцам и количество поездок:

```
SELECT
    CASE
        WHEN month_dim = 1 THEN 'Январь'
        WHEN month_dim = 2 THEN 'Февраль'
        WHEN month_dim = 3 THEN 'Март'
        WHEN month_dim = 4 THEN 'Апрель'
        WHEN month_dim = 5 THEN 'Май'
        WHEN month_dim = 6 THEN 'Июнь'
    END AS month_name,
    ROUND(AVG(avg_temp),1) AS avg_temp_f,
```

```
    ROUND((AVG(avg_temp) - 32) / 1.8, 1) AS avg_temp_c,
    COUNT(*) AS total_trips
FROM dm.taxi_trip_and_weather_nyc
GROUP BY month_dim
ORDER BY month_dim;
```

Средняя темпиратура по месяцам и количество поездок

month_name	avg_temp_f	avg_temp_c	total_trips
Январь	34.9	1.60	228059
Февраль	37.3	2.90	236609
Март	48.7	9.30	254321
Апрель	53.2	11.8	249725
Май	62.2	16.8	246618
Июнь	72.2	22.3	232546

Количество поездок по месяцам

Среднее количество осадков, выповшего снега и снега на дорогах по месяцам:

```
SELECT
    CASE
        WHEN month_dim = 1 THEN 'Январь'
        WHEN month_dim = 2 THEN 'Февраль'
        WHEN month_dim = 3 THEN 'Март'
        WHEN month_dim = 4 THEN 'Апрель'
        WHEN month_dim = 5 THEN 'Май'
        WHEN month_dim = 6 THEN 'Июнь'
    END AS month_name,
```

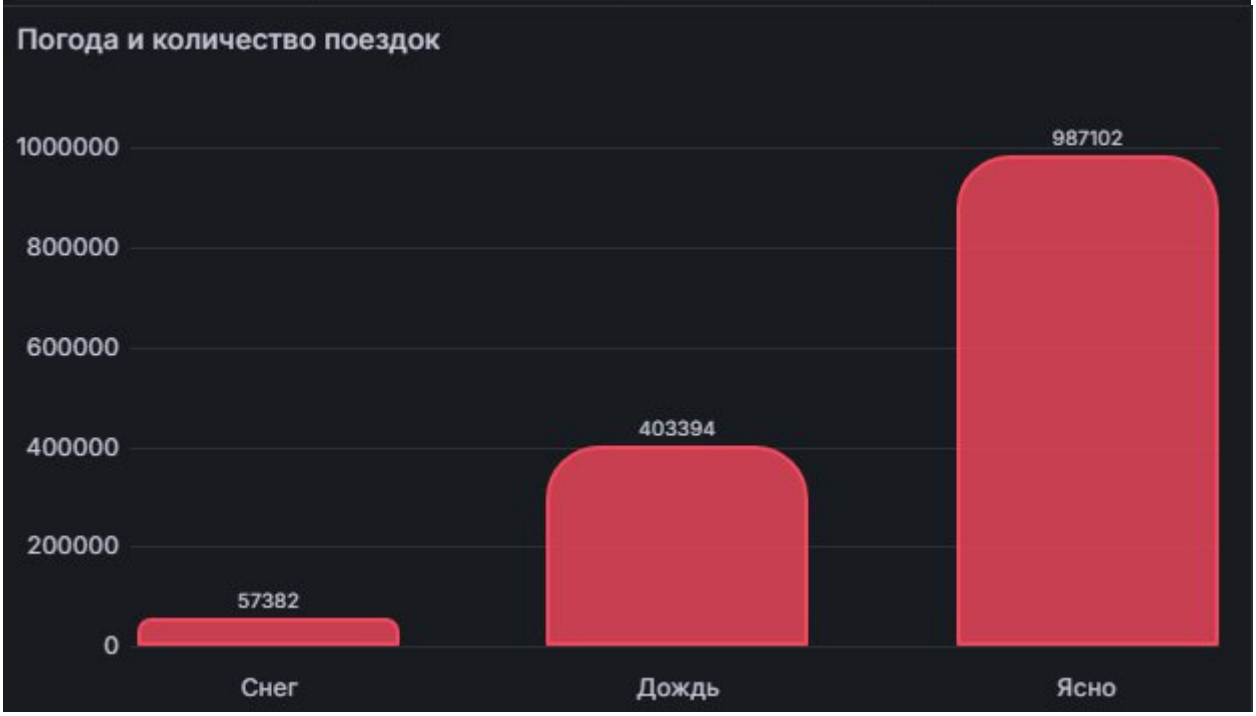
```
    ROUND(AVG(precipitation), 2) AS avg_precipitation,
    ROUND(AVG(snow_fall), 2) AS avg_snow_fall,
    ROUND(AVG(snow_depth), 2) AS avg_snow_depth
FROM dm.taxi_trip_and_weather_nyc
GROUP BY month_dim
ORDER BY month_dim;
```

month_name	avg_precipitation	avg_snow_fall	avg_snow_depth
Январь	0.0900	0.220	2.49
Февраль	0.150	0.140	0.100
Март	0.0400	0.0300	0
Апрель	0.0500	0	0
Май	0.110	0	0
Июнь	0.0900	0	0

Погода и количество поездок:

```
SELECT
    CASE
        WHEN precipitation = 0 AND snow_fall = 0
        THEN 'Ясно'
        WHEN precipitation > 0 AND snow_fall = 0
        THEN 'Дождь'
        WHEN precipitation > 0 AND snow_fall > 0
        THEN 'Снег'
    END AS weather_type,
    COUNT(*) AS count_trips,
    ROUND(AVG(trip_duration_minute), 2) AS avg_duration,
    ROUND(AVG(passenger_count), 2) AS avg_passengers
FROM dm.taxi_trip_and_weather_nyc
GROUP BY 1
ORDER BY count_trips;
```

weather_type	count_trips	avg_duration	avg_passengers
Снег	57382	12.3	1.66
Дождь	403394	13.5	1.66
Ясно	987102	13.6	1.67



День с максимальной и минимальной температурой и показателями:

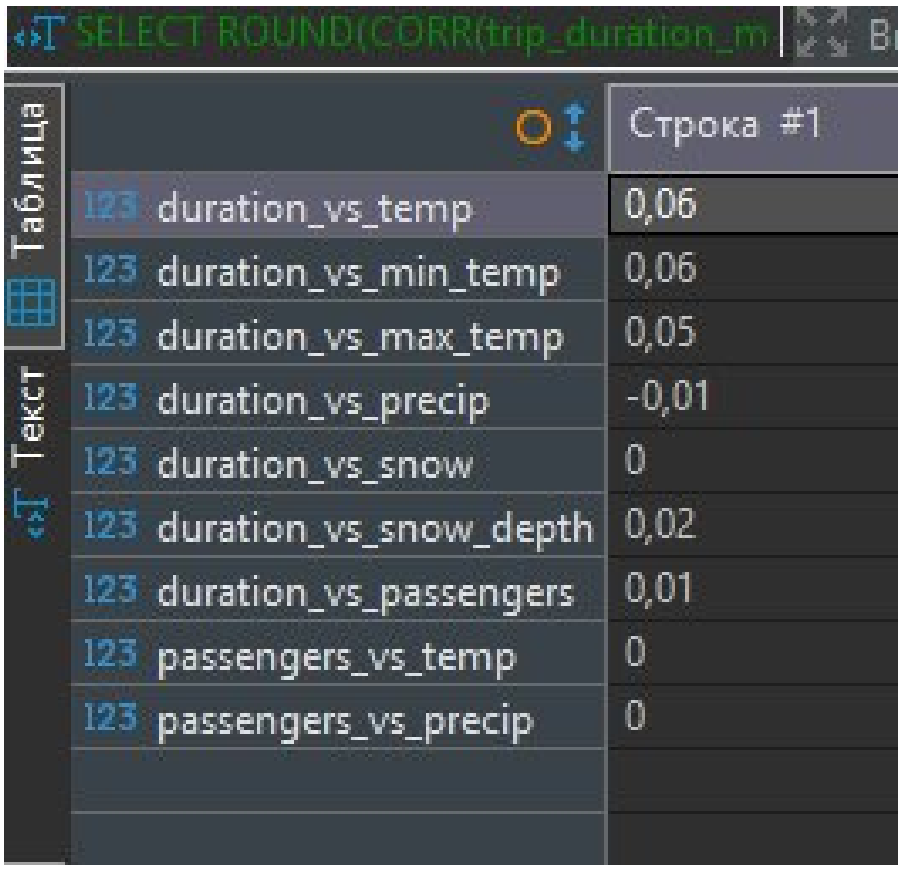
```
SELECT
        day,
        date_temp,
        temp_C,
        total_trips,
        avg_passenger,
        avg_duration
FROM(
        SELECT
                'Maximum_temp_day' AS day,
                date_key AS date_temp,
                ROUND((MAX(max_temp) - 32) / 1.8, 1) AS temp_C,
                COUNT(*) AS total_trips,
                ROUND(AVG(passenger_count),2) AS avg_passenger,
                ROUND(AVG(trip_duration_minute),2)
                AS avg_duration
        FROM dm.taxi_trip_and_weather_nyc
        WHERE max_temp = (SELECT MAX(max_temp)
                                FROM dm.taxi_trip_and_weather_nyc)
        GROUP BY date_key

UNION ALL

        SELECT
                'Minimum_temp_day' AS day,
                date_key AS date_temp,
                ROUND((MIN(min_temp) - 32) / 1.8, 1) AS temp_C,
                COUNT(*) AS total_trips,
                ROUND(AVG(passenger_count),2) AS avg_passenger,
```

```
        ROUND(AVG(trip_duration_minute),2)
        AS avg_duration
FROM dm.taxi_trip_and_weather_nyc
WHERE min_temp = (SELECT MIN(min_temp)
                  FROM dm.taxi_trip_and_weather_nyc)
GROUP BY date_key ) t1;
```

day	date_temp	temp_c	total_trips	avg_passenger	avg_duration
Maximum_temp_day	2016-05-28 03:00:00	33.3	6933	1.75	11.9
Minimum_temp_day	2016-02-14 03:00:00	-18.3	8397	1.71	11.6

Основные корреляции:

```
SELECT
    ROUND(CORR(trip_duration_minute,avg_temp)::decimal,2)
        AS duration_vs_temp,
    ROUND(CORR(trip_duration_minute, min_temp)::decimal,2)
        AS duration_vs_min_temp,
    ROUND(CORR(trip_duration_minute, max_temp)::decimal,2)
        AS duration_vs_max_temp,
    ROUND(CORR(trip_duration_minute, precipitation)::decimal,2)
    AS duration_vs_precip,
    ROUND(CORR(trip_duration_minute, snow_fall)::decimal,2)
        AS duration_vs_snow,
    ROUND(CORR(trip_duration_minute, snow_depth)::decimal,2)
        AS duration_vs_snow_depth,
    ROUND(CORR(trip_duration_minute,passenger_count)::decimal,2)
    AS duration_vs_passengers,
    ROUND(CORR(passenger_count,avg_temp)::decimal,2)
        AS passengers_vs_temp,
    ROUND(CORR(passenger_count, precipitation)::decimal,2)
        AS passengers_vs_precip
FROM dm.taxi_trip_and_weather_nyc;
```



Корреляции по дням недели:

```
SELECT
    dow_dim,
    CASE
        WHEN dow_dim = 0 THEN 'Воскресенье'
        WHEN dow_dim = 1 THEN 'Понедельник'
        WHEN dow_dim = 2 THEN 'Вторник'
        WHEN dow_dim = 3 THEN 'Среда'
        WHEN dow_dim = 4 THEN 'Четверг'
        WHEN dow_dim = 5 THEN 'Пятница'
        WHEN dow_dim = 6 THEN 'Суббота'
    END AS dow_day,
    COUNT(*) AS total_trips_dow,
    ROUND(AVG(passenger_count), 1) AS avg_passenger_trip_dow,
    ROUND(AVG(trip_duration_minute), 2) AS avg_trip_duration,
    ROUND(CORR(trip_duration_minute,avg_temp)::decimal,2)
        AS corr_duration_temp,
    ROUND(CORR(trip_duration_minute, precipitation)::decimal,2)
    AS corr_duration_precip,
```

```
        ROUND(CORR(trip_duration_minute, snow_fall)::decimal,2)
            AS corr_duration_snow,
        ROUND(CORR(passenger_count, avg_temp)::decimal,2)
            AS corr_passengers_temp,
        ROUND(CORR(passenger_count, precipitation)::decimal,2)
            AS corr_passengers_precip
FROM dm.taxi_trip_and_weather_nyc
GROUP BY dow_dim, dow_day
ORDER BY dow_dim;
```

dow_dim	dow_day	total_trips_dow	avg_passenger_trip_dow	avg_trip_duration
0	Воскресенье	193 682	1,7	12,31
1	Понедельник	186 059	1,6	13,09
2	Вторник	201 336	1,6	13,84
3	Среда	208 768	1,6	14,25
4	Четверг	216 955	1,6	14,55
5	Пятница	221 877	1,7	14,04
6	Суббота	219 201	1,7	12,56

corr_duration_temp	corr_duration_precip	corr_duration_snow	corr_passengers_temp	corr_passengers_precip
0,04	-0,01	-0,02	0	0
0,04	-0,02	-0,04	0	0,01
0,04	0	[NULL]	0	0
0,07	-0,01	[NULL]	0	0
0,06	0,01	[NULL]	0	0
0,06	0	-0,03	0	0
0,04	0,01	0,01	0	0

Корреляции по будням и выходным:

```
SELECT
    CASE
        WHEN dow_dim = 0 OR dow_dim = 6 THEN 'Выходной день'
        ELSE 'Будний день'
    END AS dow_day,
    COUNT(*) AS total_trips,
    ROUND(AVG(trip_duration_minute), 2) AS avg_duration,
    ROUND(AVG(passenger_count), 2) AS avg_passengers,
    ROUND(CORR(trip_duration_minute, avg_temp)::decimal,2)
        AS duration_vs_temp,
    ROUND(CORR(trip_duration_minute, precipitation)::decimal,2)
    AS duration_vs_precip,
    ROUND(CORR(trip_duration_minute, snow_fall)::decimal,2)
        AS duration_vs_snow,
```

```
        ROUND(CORR(passenger_count, avg_temp)::decimal,2)
              AS passengers_vs_temp
FROM dm.taxi_trip_and_weather_nyc
GROUP BY dow_day
ORDER BY dow_day;
```

dow_day	total_trips	avg_duration	avg_passengers
Будний день	1 034 995	13,98	1,64
Выходной день	412 883	12,44	1,72

duration_vs_temp	duration_vs_precip	duration_vs_snow	passengers_vs_temp
0,06	-0,01	-0,02	0
0,04	0	0,01	0