

Distribution of Parameters for Generated SQL Queries on the TPC-H Database

Scenario: 1-6jp_ujpp__g__h0-3ss_h0nss

Scale factor: 50 GB

The 1000-query pack was generated on: 2023-11-11

First column ('Query parameter or value') displays either:

- a parameter name, or

- a value of the current parameter (only for low-cardinality parameters).

Second column ('Descriptive statistics') displays:

- for low cardinality parameters: frequency and proportion from total (of 1000 queries)

- for higher cardinality parameters: [min, median, max] mean / SD

Query parameter or value	Descriptive statistics
SELECT_n_of_columns	[1, 8, 149] 14 / 20
SELECT_n_of_non_aggr_func__LTRIM	
0	635 (64%)
1	252 (25%)
2	80 (8.0%)
3	20 (2.0%)
4	6 (0.6%)
5	4 (0.4%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	849 (85%)
1	118 (12%)
2	21 (2.1%)
3	7 (0.7%)

Query parameter or value	Descriptive statistics
4	4 (0.4%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__RTRIM	
0	591 (59%)
1	278 (28%)
2	81 (8.1%)
3	30 (3.0%)
4	10 (1.0%)
5	7 (0.7%)
6	2 (0.2%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__UPPER	
0	619 (62%)
1	250 (25%)
2	82 (8.2%)
3	26 (2.6%)
4	12 (1.2%)
5	7 (0.7%)
6	1 (0.1%)
7	2 (0.2%)
9	1 (0.1%)
SELECT__n_of__non__aggr__func__SUBSTR	
0	623 (62%)
1	273 (27%)
2	61 (6.1%)
3	16 (1.6%)
4	12 (1.2%)
5	5 (0.5%)
6	6 (0.6%)
7	3 (0.3%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__LOWER	[0.00, 0.00, 11.00] 0.67 / 1.11
SELECT__n_of__non__aggr__func__ABS	
0	927 (93%)
1	64 (6.4%)
2	7 (0.7%)
3	2 (0.2%)
SELECT__n_of__non__aggr__func__SQRT	
0	801 (80%)
1	152 (15%)
2	31 (3.1%)
3	7 (0.7%)
4	7 (0.7%)
5	1 (0.1%)
9	1 (0.1%)
SELECT__n_of__non__aggr__func__DAY	

Query parameter or value	Descriptive statistics
0	892 (89%)
1	82 (8.2%)
2	17 (1.7%)
3	8 (0.8%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	813 (81%)
1	132 (13%)
2	28 (2.8%)
3	14 (1.4%)
4	4 (0.4%)
5	6 (0.6%)
6	1 (0.1%)
7	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	846 (85%)
1	106 (11%)
2	29 (2.9%)
3	10 (1.0%)
4	2 (0.2%)
5	6 (0.6%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	880 (88%)
1	79 (7.9%)
2	23 (2.3%)
3	9 (0.9%)
4	3 (0.3%)
5	3 (0.3%)
6	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	924 (92%)
1	63 (6.3%)
2	10 (1.0%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__ROUND	
0	922 (92%)
1	67 (6.7%)
2	8 (0.8%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	927 (93%)
1	65 (6.5%)

Query parameter or value	Descriptive statistics
2	7 (0.7%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 73.0] 4.7 / 6.9
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 14.00] 0.79 / 1.87
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 15.00] 0.92 / 2.07
SELECT_n_of_aggr_func__AVG	
0	930 (93%)
1	40 (4.0%)
2	19 (1.9%)
3	6 (0.6%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 16.00] 0.83 / 1.86
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 12.00] 0.92 / 2.10
SELECT_n_of_aggr_func__SUM	
0	932 (93%)
1	43 (4.3%)
2	17 (1.7%)
3	7 (0.7%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 16.0] 3.7 / 3.1
FROM_n_of_join_paths	
1	164 (16%)
2	176 (18%)
3	171 (17%)
4	164 (16%)
5	161 (16%)
6	164 (16%)
FROM_n_of_super_joins__RIGHT	
0	541 (54%)
1	306 (31%)
2	123 (12%)
3	24 (2.4%)
4	5 (0.5%)
5	1 (0.1%)
FROM_n_of_super_joins__FULL	
0	556 (56%)
1	313 (31%)
2	116 (12%)
3	13 (1.3%)
4	2 (0.2%)
FROM_n_of_super_joins__LEFT	
0	563 (56%)
1	304 (30%)
2	110 (11%)
3	20 (2.0%)

Query parameter or value	Descriptive statistics
4	2 (0.2%)
5	1 (0.1%)
FROM_n_of_joins__INNER	[0.00, 3.00, 12.00] 3.31 / 2.43
FROM_n_of_joins__RIGHT	[0.00, 3.00, 16.00] 3.19 / 2.36
FROM_n_of_processed_rows	[5, 465,031,228, 2,272,598,440] 581,592,885 / 481,958,623
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.76 / 2.83
WHERE_n_of_attribs_of_type__character	
0	648 (65%)
1	257 (26%)
2	74 (7.4%)
3	19 (1.9%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	342 (34%)
1	314 (31%)
2	190 (19%)
3	93 (9.3%)
4	47 (4.7%)
5	6 (0.6%)
6	7 (0.7%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	369 (37%)
1	300 (30%)
2	182 (18%)
3	86 (8.6%)
4	41 (4.1%)
5	14 (1.4%)
6	4 (0.4%)
7	4 (0.4%)
WHERE_n_of_attribs_of_type__numeric	
0	583 (58%)
1	281 (28%)
2	91 (9.1%)
3	36 (3.6%)
4	5 (0.5%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type__date	
0	809 (81%)
1	158 (16%)
2	28 (2.8%)
3	4 (0.4%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	

Query parameter or value	Descriptive statistics
0	548 (55%)
1	295 (30%)
2	109 (11%)
3	36 (3.6%)
4	11 (1.1%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.11 / 2.34
WHERE_n_of_operators___greater_or_less	
0	344 (34%)
1	257 (26%)
2	187 (19%)
3	113 (11%)
4	56 (5.6%)
5	26 (2.6%)
6	11 (1.1%)
7	6 (0.6%)
WHERE_n_of_operators___between	
0	522 (52%)
1	294 (29%)
2	132 (13%)
3	45 (4.5%)
4	5 (0.5%)
5	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_operators___in	
0	542 (54%)
1	296 (30%)
2	112 (11%)
3	44 (4.4%)
4	4 (0.4%)
5	2 (0.2%)
WHERE_n_of_operators___like	
0	772 (77%)
1	193 (19%)
2	32 (3.2%)
3	3 (0.3%)
WHERE_n_of_non_aggr_func___ROUND	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func___LOG	
0	914 (91%)
1	85 (8.5%)

Query parameter or value	Descriptive statistics
2	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	862 (86%)
1	132 (13%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func__TRUNC	
0	955 (96%)
1	44 (4.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	974 (97%)
1	23 (2.3%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__ABS	
0	956 (96%)
1	42 (4.2%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__YEAR	
0	965 (97%)
1	33 (3.3%)
2	1 (0.1%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	969 (97%)
1	30 (3.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__FLOOR	
0	960 (96%)
1	39 (3.9%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	630 (63%)
1	242 (24%)
2	96 (9.6%)
3	26 (2.6%)
4	3 (0.3%)
5	3 (0.3%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.14 / 2.96
HAVING_n_of_main_predicates	
0	287 (29%)
1	85 (8.5%)
2	159 (16%)
3	196 (20%)
4	154 (15%)
5	72 (7.2%)
6	47 (4.7%)

Query parameter or value	Descriptive statistics
HAVING_n_of_main_predicates__scalar_subquery	
0	427 (43%)
1	209 (21%)
2	180 (18%)
3	184 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 21.0] 2.3 / 3.5
HAVING_n_of_processed_rows_by_subqueries	[0, 80,980,747, 32,454,153,216] 1,463,824,540 / 3,054,775,157
ORDER_BY_n_of_columns	
0	272 (27%)
1	367 (37%)
2	235 (24%)
3	98 (9.8%)
4	15 (1.5%)
5	9 (0.9%)
6	3 (0.3%)
7	1 (0.1%)
limit	[1, 512, 999] 509 / 286
offset	[0, 0, 1,000] 254 / 327