

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

Scenario: 1-6jp_rjpp__g__h0-3ss_h0-3nss

Scale factor: 1000 GB

The 1000-query pack was generated on: 2023-11-13

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, 9, 184] 15 / 24
SELECT_n_of_non_aggr_func__RTRIM	
0	623 (62%)
1	252 (25%)
2	78 (7.8%)
3	25 (2.5%)
4	12 (1.2%)
5	5 (0.5%)
6	4 (0.4%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	822 (82%)
1	123 (12%)
11	1 (0.1%)
2	27 (2.7%)

Query parameter or value	Descriptive statistics
3	16 (1.6%)
4	4 (0.4%)
5	6 (0.6%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	631 (63%)
1	245 (25%)
10	1 (0.1%)
2	83 (8.3%)
3	17 (1.7%)
4	15 (1.5%)
5	5 (0.5%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	875 (88%)
1	99 (9.9%)
2	17 (1.7%)
3	7 (0.7%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___MONTH	
0	817 (82%)
1	123 (12%)
2	36 (3.6%)
3	10 (1.0%)
4	5 (0.5%)
5	3 (0.3%)
6	2 (0.2%)
7	3 (0.3%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	844 (84%)
1	107 (11%)
2	30 (3.0%)
3	10 (1.0%)
4	7 (0.7%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	933 (93%)
1	59 (5.9%)
2	5 (0.5%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___LOWER	
0	637 (64%)
1	243 (24%)
12	1 (0.1%)

Query parameter or value	Descriptive statistics
2	78 (7.8%)
3	21 (2.1%)
4	11 (1.1%)
5	5 (0.5%)
6	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	623 (62%)
1	253 (25%)
12	1 (0.1%)
2	78 (7.8%)
3	25 (2.5%)
4	10 (1.0%)
5	8 (0.8%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	[0.00, 0.00, 10.00] 0.60 / 1.12
SELECT_n_of_non_aggr_func__ABS	
0	913 (91%)
1	72 (7.2%)
2	13 (1.3%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__DOW	
0	877 (88%)
1	87 (8.7%)
2	24 (2.4%)
3	6 (0.6%)
4	3 (0.3%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	918 (92%)
1	68 (6.8%)
2	10 (1.0%)
3	2 (0.2%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__LOG	
0	868 (87%)
1	101 (10%)
2	12 (1.2%)
3	9 (0.9%)
4	8 (0.8%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__ROUND	
0	909 (91%)
1	74 (7.4%)
2	11 (1.1%)

Query parameter or value	Descriptive statistics
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 69.0] 4.7 / 7.5
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.85 / 1.94
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.80 / 1.79
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 15.00] 0.95 / 2.14
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 12.00] 0.90 / 2.01
SELECT_n_of_aggr_func__SUM	
0	926 (93%)
1	53 (5.3%)
2	13 (1.3%)
3	5 (0.5%)
4	1 (0.1%)
5	2 (0.2%)
SELECT_n_of_aggr_func__AVG	
0	932 (93%)
1	47 (4.7%)
2	18 (1.8%)
3	1 (0.1%)
4	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 15.0] 3.7 / 3.1
FROM_n_of_join_paths	
1	153 (15%)
2	178 (18%)
3	163 (16%)
4	165 (17%)
5	160 (16%)
6	181 (18%)
FROM_n_of_super_joins__FULL	
0	528 (53%)
1	341 (34%)
2	98 (9.8%)
3	32 (3.2%)
4	1 (0.1%)
FROM_n_of_super_joins__RIGHT	
0	540 (54%)
1	321 (32%)
2	110 (11%)
3	26 (2.6%)
4	3 (0.3%)
FROM_n_of_super_joins__LEFT	
0	543 (54%)
1	299 (30%)
2	131 (13%)
3	26 (2.6%)
4	1 (0.1%)

Query parameter or value	Descriptive statistics
FROM_n_of_joins__RIGHT	[0.00, 3.00, 12.00] 3.33 / 2.44
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.27 / 2.49
FROM_n_of_processed_rows	[5, 9,254,989,824, 45,749,938,314] 11,705,675,179 / 9,639,055,054
WHERE_n_of_predicates	[0.00, 3.00, 14.00] 3.81 / 2.83
WHERE_n_of_attribs_of_type__integer	
0	369 (37%)
1	313 (31%)
2	185 (19%)
3	76 (7.6%)
4	40 (4.0%)
5	14 (1.4%)
6	1 (0.1%)
7	2 (0.2%)
WHERE_n_of_attribs_of_type__character	
0	646 (65%)
1	235 (24%)
2	82 (8.2%)
3	32 (3.2%)
4	3 (0.3%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__character_varying	
0	349 (35%)
1	305 (31%)
2	190 (19%)
3	85 (8.5%)
4	54 (5.4%)
5	10 (1.0%)
6	6 (0.6%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	775 (78%)
1	174 (17%)
2	45 (4.5%)
3	6 (0.6%)
WHERE_n_of_attribs_of_type__numeric	
0	588 (59%)
1	272 (27%)
2	98 (9.8%)
3	38 (3.8%)
4	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	549 (55%)
1	321 (32%)
2	98 (9.8%)

Query parameter or value	Descriptive statistics
3	28 (2.8%)
4	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 12.00] 2.16 / 2.38
WHERE_n_of_operators___greater_or_less	
0	326 (33%)
1	284 (28%)
2	189 (19%)
3	113 (11%)
4	52 (5.2%)
5	23 (2.3%)
6	13 (1.3%)
WHERE_n_of_operators___between	
0	520 (52%)
1	295 (30%)
2	129 (13%)
3	45 (4.5%)
4	10 (1.0%)
5	1 (0.1%)
WHERE_n_of_operators___in	
0	527 (53%)
1	287 (29%)
2	137 (14%)
3	39 (3.9%)
4	9 (0.9%)
5	1 (0.1%)
WHERE_n_of_operators___like	
0	774 (77%)
1	192 (19%)
2	30 (3.0%)
3	4 (0.4%)
WHERE_n_of_non_aggr_func___SQRT	
0	865 (87%)
1	124 (12%)
2	10 (1.0%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	946 (95%)
1	51 (5.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___ABS	
0	963 (96%)
1	37 (3.7%)
WHERE_n_of_non_aggr_func___LOG	
0	927 (93%)
1	69 (6.9%)

Query parameter or value	Descriptive statistics
2	4 (0.4%)
WHERE_n_of_non_aggr_func__FLOOR	
0	952 (95%)
1	47 (4.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func__TRUNC	
0	972 (97%)
1	27 (2.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	953 (95%)
1	44 (4.4%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__ROUND	
0	959 (96%)
1	41 (4.1%)
WHERE_n_of_non_aggr_func__DAY	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	607 (61%)
1	263 (26%)
2	102 (10%)
3	22 (2.2%)
4	6 (0.6%)
GROUP_BY_n_of_columns	[0.00, 2.00, 11.00] 3.13 / 2.85
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.45 / 2.59
HAVING_n_of_main_predicates__non_scalar_subquery	
0	444 (44%)
1	173 (17%)
2	191 (19%)
3	192 (19%)
HAVING_n_of_main_predicates__scalar_subquery	
0	425 (43%)
1	182 (18%)
2	203 (20%)
3	190 (19%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	444 (44%)
1	173 (17%)
2	191 (19%)
3	192 (19%)

Query parameter or value	Descriptive statistics
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 20.0] 2.6 / 3.6
HAVING_n_of_processed_rows_by_subquery_order	[0, 14, 594, 990, 149, 548, 249, 302, 167] 42,841,466,001 / 73,536,548,026
ORDER_BY_n_of_columns	
0	319 (32%)
1	330 (33%)
2	217 (22%)
3	100 (10%)
4	18 (1.8%)
5	10 (1.0%)
6	4 (0.4%)
7	1 (0.1%)
8	1 (0.1%)
limit	[1, 492, 999] 495 / 290
offset	[0, 31, 994] 266 / 330