

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

Scenario: 1-6jp_rjpp__g__h0-3ss_h0nss

Scale factor: 100 GB

The 1000-query pack was generated on: 2023-11-07

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, 146] 15 / 21
SELECT_n_of_non_aggr_func__LOWER	
0	612 (61%)
1	255 (26%)
2	81 (8.1%)
3	27 (2.7%)
4	12 (1.2%)
5	8 (0.8%)
6	3 (0.3%)
7	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	[0.00, 0.00, 13.00] 0.61 / 1.17
SELECT_n_of_non_aggr_func__ROUND	
0	912 (91%)
1	68 (6.8%)

Query parameter or value	Descriptive statistics
2	11 (1.1%)
3	7 (0.7%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	613 (61%)
1	249 (25%)
2	90 (9.0%)
3	25 (2.5%)
4	7 (0.7%)
5	12 (1.2%)
6	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 11.00] 0.62 / 1.18
SELECT_n_of_non_aggr_func__DOW	
0	842 (84%)
1	115 (12%)
2	30 (3.0%)
3	7 (0.7%)
4	2 (0.2%)
5	1 (0.1%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	786 (79%)
1	145 (15%)
2	41 (4.1%)
3	13 (1.3%)
4	7 (0.7%)
5	3 (0.3%)
6	2 (0.2%)
7	1 (0.1%)
9	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	810 (81%)
1	140 (14%)
2	27 (2.7%)
3	12 (1.2%)
4	8 (0.8%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__UPPER	
0	647 (65%)
1	225 (23%)
2	82 (8.2%)
3	18 (1.8%)
4	9 (0.9%)
5	11 (1.1%)

Query parameter or value	Descriptive statistics
6	6 (0.6%)
7	1 (0.1%)
8	1 (0.1%)
SELECT__n_of__non_aggr_func__DAY	
0	891 (89%)
1	84 (8.4%)
2	12 (1.2%)
3	8 (0.8%)
4	3 (0.3%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non_aggr_func__ABS	
0	921 (92%)
1	71 (7.1%)
2	6 (0.6%)
3	1 (0.1%)
4	1 (0.1%)
SELECT__n_of__non_aggr_func__SQRT	
0	789 (79%)
1	146 (15%)
2	32 (3.2%)
3	17 (1.7%)
4	8 (0.8%)
5	4 (0.4%)
6	1 (0.1%)
7	2 (0.2%)
8	1 (0.1%)
SELECT__n_of__non_aggr_func__FLOOR	
0	909 (91%)
1	71 (7.1%)
2	17 (1.7%)
3	2 (0.2%)
4	1 (0.1%)
SELECT__n_of__non_aggr_func__LOG	
0	855 (86%)
1	105 (11%)
2	22 (2.2%)
3	12 (1.2%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non_aggr_func__TRUNC	
0	897 (90%)
1	74 (7.4%)
2	23 (2.3%)
3	5 (0.5%)

Query parameter or value	Descriptive statistics
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 66.0] 5.1 / 8.1
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 12.00] 0.74 / 1.80
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.90 / 1.97
SELECT_n_of_aggr_func__AVG	
0	942 (94%)
1	37 (3.7%)
2	13 (1.3%)
3	5 (0.5%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 14.00] 0.84 / 1.98
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.82 / 1.91
SELECT_n_of_aggr_func__SUM	
0	942 (94%)
1	41 (4.1%)
2	12 (1.2%)
3	5 (0.5%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 15.0] 3.5 / 3.1
FROM_n_of_join_paths	
1	181 (18%)
2	167 (17%)
3	147 (15%)
4	163 (16%)
5	179 (18%)
6	163 (16%)
FROM_n_of_super_joins__RIGHT	
0	549 (55%)
1	315 (32%)
2	116 (12%)
3	19 (1.9%)
4	1 (0.1%)
FROM_n_of_super_joins__LEFT	
0	543 (54%)
1	315 (32%)
2	115 (12%)
3	23 (2.3%)
4	4 (0.4%)
FROM_n_of_super_joins__FULL	
0	544 (54%)
1	315 (32%)
2	108 (11%)
3	31 (3.1%)
4	2 (0.2%)
FROM_n_of_joins__INNER	[0.00, 3.00, 12.00] 3.17 / 2.46
FROM_n_of_joins__RIGHT	[0.00, 3.00, 14.00] 3.22 / 2.45

Query parameter or value	Descriptive statistics
FROM_n_of_processed_rows	[5, 881,537,927, 4,590,227,517] 1,122,236,362 / 943,303,095
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.76 / 2.95
WHERE_n_of_attribs_of_type___character_varying	
0	357 (36%)
1	275 (28%)
2	191 (19%)
3	114 (11%)
4	42 (4.2%)
5	16 (1.6%)
6	5 (0.5%)
WHERE_n_of_attribs_of_type___integer	
0	390 (39%)
1	278 (28%)
2	171 (17%)
3	97 (9.7%)
4	48 (4.8%)
5	9 (0.9%)
6	4 (0.4%)
7	3 (0.3%)
WHERE_n_of_attribs_of_type___character	
0	676 (68%)
1	239 (24%)
2	74 (7.4%)
3	9 (0.9%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	586 (59%)
1	264 (26%)
2	107 (11%)
3	26 (2.6%)
4	13 (1.3%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type___date	
0	799 (80%)
1	168 (17%)
2	29 (2.9%)
3	3 (0.3%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	547 (55%)
1	296 (30%)
2	113 (11%)
3	35 (3.5%)
4	8 (0.8%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 12.00] 2.20 / 2.44
WHERE_n_of_operators__between	
0	523 (52%)
1	291 (29%)
2	138 (14%)
3	37 (3.7%)
4	9 (0.9%)
5	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	345 (35%)
1	254 (25%)
2	186 (19%)
3	110 (11%)
4	68 (6.8%)
5	26 (2.6%)
6	11 (1.1%)
WHERE_n_of_operators__like	
0	778 (78%)
1	187 (19%)
2	32 (3.2%)
3	3 (0.3%)
WHERE_n_of_operators__in	
0	546 (55%)
1	294 (29%)
2	110 (11%)
3	33 (3.3%)
4	14 (1.4%)
5	3 (0.3%)
WHERE_n_of_non_aggr_func__ABS	
0	965 (97%)
1	33 (3.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__LOG	
0	929 (93%)
1	66 (6.6%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func__TRUNC	
0	953 (95%)
1	46 (4.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	861 (86%)
1	125 (13%)
2	13 (1.3%)

Query parameter or value	Descriptive statistics
3	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func__DAY	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__FLOOR	
0	959 (96%)
1	39 (3.9%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__MONTH	
0	959 (96%)
1	41 (4.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	967 (97%)
1	31 (3.1%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	617 (62%)
1	259 (26%)
2	92 (9.2%)
3	20 (2.0%)
4	10 (1.0%)
5	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 2.96 / 2.78
HAVING_n_of_main_predicates	
0	325 (33%)
1	87 (8.7%)
2	153 (15%)
3	169 (17%)
4	127 (13%)
5	93 (9.3%)
6	46 (4.6%)
HAVING_n_of_main_predicates__scalar_subquery	
0	464 (46%)
1	162 (16%)
2	194 (19%)
3	180 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 19.0] 2.4 / 3.6

Query parameter or value	Descriptive statistics
HAVING_n_of_processed_rows_by_subqueries	[0, 16,000,033, 48,890,653,140] 2,791,695,438 / 6,222,750,132
ORDER_BY_n_of_columns	
0	298 (30%)
1	334 (33%)
2	234 (23%)
3	101 (10%)
4	21 (2.1%)
5	6 (0.6%)
6	5 (0.5%)
9	1 (0.1%)
limit	[1, 484, 1,000] 492 / 288
offset	[0, 21, 1,000] 257 / 325