

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

Scenario: 1-6jp_rjpp_g_h0-3ss_h0-3nss

Scale factor: 10 GB

The 1000-query pack was generated on: 2023-11-12

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, 185] 15 / 22
SELECT_n_of_non_aggr_func__SQRT	[0.00, 0.00, 9.00] 0.31 / 0.84
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 10.00] 0.60 / 1.07
SELECT_n_of_non_aggr_func__UPPER	[0.00, 0.00, 10.00] 0.58 / 1.04
SELECT_n_of_non_aggr_func__LTRIM	
0	616 (62%)
1	278 (28%)
2	64 (6.4%)
3	19 (1.9%)
4	11 (1.1%)
5	7 (0.7%)
6	1 (0.1%)
7	2 (0.2%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	

Query parameter or value	Descriptive statistics
0	812 (81%)
1	125 (13%)
2	38 (3.8%)
3	11 (1.1%)
4	4 (0.4%)
5	3 (0.3%)
6	4 (0.4%)
7	2 (0.2%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	612 (61%)
1	279 (28%)
2	62 (6.2%)
3	24 (2.4%)
4	12 (1.2%)
5	6 (0.6%)
6	2 (0.2%)
7	1 (0.1%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	836 (84%)
1	116 (12%)
2	24 (2.4%)
3	16 (1.6%)
4	1 (0.1%)
5	5 (0.5%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	888 (89%)
1	84 (8.4%)
2	23 (2.3%)
3	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	861 (86%)
1	108 (11%)
2	18 (1.8%)
3	5 (0.5%)
4	2 (0.2%)
5	1 (0.1%)
6	2 (0.2%)
7	3 (0.3%)
SELECT_n_of_non_aggr_func__LOG	
0	862 (86%)
1	110 (11%)

Query parameter or value	Descriptive statistics
2	20 (2.0%)
3	5 (0.5%)
4	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	612 (61%)
1	266 (27%)
10	1 (0.1%)
2	84 (8.4%)
3	24 (2.4%)
4	4 (0.4%)
5	5 (0.5%)
6	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	935 (94%)
1	57 (5.7%)
2	7 (0.7%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	918 (92%)
1	72 (7.2%)
2	5 (0.5%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__ABS	
0	936 (94%)
1	52 (5.2%)
2	10 (1.0%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	910 (91%)
1	80 (8.0%)
2	8 (0.8%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 71.0] 4.7 / 7.3
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 12.00] 0.81 / 1.85
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 1.03 / 2.17
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.84 / 2.01
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 13.00] 0.86 / 1.99
SELECT_n_of_aggr_func__AVG	
0	931 (93%)
1	42 (4.2%)
2	18 (1.8%)
3	6 (0.6%)

Query parameter or value	Descriptive statistics
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	923 (92%)
1	51 (5.1%)
2	16 (1.6%)
3	10 (1.0%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 15.0] 3.8 / 3.2
FROM_n_of_join_paths	
1	170 (17%)
2	141 (14%)
3	183 (18%)
4	161 (16%)
5	188 (19%)
6	157 (16%)
FROM_n_of_super_joins__RIGHT	
0	552 (55%)
1	306 (31%)
2	106 (11%)
3	32 (3.2%)
4	4 (0.4%)
FROM_n_of_super_joins__FULL	
0	544 (54%)
1	314 (31%)
2	114 (11%)
3	25 (2.5%)
4	3 (0.3%)
FROM_n_of_super_joins__LEFT	
0	566 (57%)
1	290 (29%)
2	107 (11%)
3	34 (3.4%)
4	3 (0.3%)
FROM_n_of_joins__INNER	[0.00, 3.00, 15.00] 3.36 / 2.52
FROM_n_of_joins__RIGHT	[0.00, 3.00, 12.00] 3.35 / 2.49
FROM_n_of_processed_rows	[5, 94,382,624, 448,881,778] 122,118,318 / 98,466,147
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.83 / 3.01
WHERE_n_of_attribs_of_type__date	
0	783 (78%)
1	172 (17%)
2	37 (3.7%)
3	5 (0.5%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	378 (38%)

Query parameter or value	Descriptive statistics
1	289 (29%)
2	176 (18%)
3	99 (9.9%)
4	41 (4.1%)
5	13 (1.3%)
6	3 (0.3%)
8	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	380 (38%)
1	274 (27%)
2	176 (18%)
3	114 (11%)
4	36 (3.6%)
5	12 (1.2%)
6	6 (0.6%)
7	2 (0.2%)
WHERE_n_of_attribs_of_type__character	
0	634 (63%)
1	258 (26%)
2	80 (8.0%)
3	24 (2.4%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__numeric	
0	583 (58%)
1	257 (26%)
2	113 (11%)
3	37 (3.7%)
4	8 (0.8%)
5	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	529 (53%)
1	311 (31%)
2	114 (11%)
3	36 (3.6%)
4	5 (0.5%)
5	4 (0.4%)
6	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.21 / 2.50
WHERE_n_of_operators__in	
0	504 (50%)
1	301 (30%)
2	129 (13%)
3	49 (4.9%)
4	16 (1.6%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators__between	
0	547 (55%)
1	286 (29%)
2	111 (11%)
3	42 (4.2%)
4	13 (1.3%)
5	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	334 (33%)
1	272 (27%)
2	187 (19%)
3	114 (11%)
4	55 (5.5%)
5	31 (3.1%)
6	5 (0.5%)
7	2 (0.2%)
WHERE_n_of_operators__like	
0	783 (78%)
1	173 (17%)
2	43 (4.3%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	971 (97%)
1	27 (2.7%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__FLOOR	
0	950 (95%)
1	48 (4.8%)
2	1 (0.1%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	878 (88%)
1	110 (11%)
2	12 (1.2%)
WHERE_n_of_non_aggr_func__YEAR	
0	956 (96%)
1	42 (4.2%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__DOW	
0	961 (96%)
1	37 (3.7%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__ABS	
0	964 (96%)
1	36 (3.6%)
WHERE_n_of_non_aggr_func__LOG	

Query parameter or value	Descriptive statistics
0	921 (92%)
1	72 (7.2%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func__MONTH	
0	953 (95%)
1	43 (4.3%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func__TRUNC	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_non_aggr_func__ROUND	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	628 (63%)
1	227 (23%)
2	107 (11%)
3	25 (2.5%)
4	11 (1.1%)
5	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.16 / 2.88
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.33 / 2.50
HAVING_n_of_main_predicates__non_scalar_subquery	
0	429 (43%)
1	206 (21%)
2	192 (19%)
3	173 (17%)
HAVING_n_of_main_predicates__scalar_subquery	
0	439 (44%)
1	193 (19%)
2	187 (19%)
3	181 (18%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	429 (43%)
1	206 (21%)
2	192 (19%)
3	173 (17%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 21.0] 2.3 / 3.4
HAVING_n_of_processed_rows_by_subqueries	[0, 149,691,884, 6,048,374,178] 425,786,693 / 708,253,080
ORDER_BY_n_of_columns	
0	316 (32%)
1	303 (30%)
2	226 (23%)
3	118 (12%)

Query parameter or value	Descriptive statistics
4	29 (2.9%)
5	6 (0.6%)
7	1 (0.1%)
8	1 (0.1%)
limit	[2, 490, 999] 488 / 288
offset	[0, 5, 998] 262 / 330