

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

Scenario: 1-3jp_ujpp__g__h0-3ss_h0nss

Scale factor: 5 GB

The 1000-query pack was generated on: 2023-11-09

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, 7, 86] 11 / 14
SELECT_n_of_non_aggr_func__LOG	
0	901 (90%)
1	81 (8.1%)
2	14 (1.4%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	691 (69%)
1	240 (24%)
2	48 (4.8%)
3	16 (1.6%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	

Query parameter or value	Descriptive statistics
0	719 (72%)
1	210 (21%)
2	52 (5.2%)
3	11 (1.1%)
4	4 (0.4%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	892 (89%)
1	88 (8.8%)
2	15 (1.5%)
3	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___LTRIM	
0	697 (70%)
1	221 (22%)
2	57 (5.7%)
3	13 (1.3%)
4	6 (0.6%)
5	1 (0.1%)
6	3 (0.3%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	694 (69%)
1	241 (24%)
2	47 (4.7%)
3	14 (1.4%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___SQRT	
0	855 (86%)
1	114 (11%)
2	16 (1.6%)
3	11 (1.1%)
4	4 (0.4%)
SELECT_n_of_non_aggr_func___UPPER	
0	689 (69%)
1	234 (23%)
2	58 (5.8%)
3	17 (1.7%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	851 (85%)
1	107 (11%)

Query parameter or value	Descriptive statistics
2	27 (2.7%)
3	12 (1.2%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func__YEAR	
0	859 (86%)
1	110 (11%)
2	28 (2.8%)
3	1 (0.1%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__DAY	
0	919 (92%)
1	70 (7.0%)
2	11 (1.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	927 (93%)
1	68 (6.8%)
2	3 (0.3%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	946 (95%)
1	46 (4.6%)
2	7 (0.7%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	945 (95%)
1	49 (4.9%)
2	4 (0.4%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	941 (94%)
1	53 (5.3%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 39.0] 3.2 / 4.3
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 11.00] 0.61 / 1.38
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 10.00] 0.63 / 1.44
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.69 / 1.53
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 9.00] 0.65 / 1.44
SELECT_n_of_aggr_func__SUM	
0	946 (95%)
1	41 (4.1%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_aggr_func__AVG	

Query parameter or value	Descriptive statistics
0	949 (95%)
1	36 (3.6%)
2	15 (1.5%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.72 / 2.28
FROM_n_of_join_paths	
1	309 (31%)
2	365 (37%)
3	326 (33%)
FROM_n_of_super_joins__LEFT	
0	769 (77%)
1	210 (21%)
2	21 (2.1%)
FROM_n_of_super_joins__RIGHT	
0	777 (78%)
1	206 (21%)
2	17 (1.7%)
FROM_n_of_super_joins__FULL	
0	747 (75%)
1	238 (24%)
2	15 (1.5%)
FROM_n_of_joins__INNER	
0	219 (22%)
1	272 (27%)
2	218 (22%)
3	154 (15%)
4	79 (7.9%)
5	40 (4.0%)
6	13 (1.3%)
7	5 (0.5%)
FROM_n_of_joins__RIGHT	
0	225 (23%)
1	253 (25%)
2	231 (23%)
3	153 (15%)
4	77 (7.7%)
5	41 (4.1%)
6	10 (1.0%)
7	8 (0.8%)
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 34,100,193, 114,749,827] 33,975,382 / 29,134,661
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.81 / 2.20
WHERE_n_of_attribs_of_type__character_varying	
0	446 (45%)
1	312 (31%)
2	153 (15%)
3	65 (6.5%)

Query parameter or value	Descriptive statistics
4	23 (2.3%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	483 (48%)
1	278 (28%)
2	164 (16%)
3	58 (5.8%)
4	13 (1.3%)
5	2 (0.2%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	628 (63%)
1	265 (27%)
2	80 (8.0%)
3	27 (2.7%)
WHERE_n_of_attribs_of_type__character	
0	732 (73%)
1	217 (22%)
2	44 (4.4%)
3	6 (0.6%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	818 (82%)
1	152 (15%)
2	26 (2.6%)
3	4 (0.4%)
WHERE_n_of_pkey_attribs	
0	645 (65%)
1	256 (26%)
2	85 (8.5%)
3	11 (1.1%)
4	3 (0.3%)
WHERE_n_of_connect_OR	
0	462 (46%)
1	157 (16%)
2	147 (15%)
3	96 (9.6%)
4	75 (7.5%)
5	45 (4.5%)
6	12 (1.2%)
7	4 (0.4%)
8	2 (0.2%)
WHERE_n_of_operators__greater_or_less	
0	436 (44%)
1	274 (27%)
2	172 (17%)

Query parameter or value	Descriptive statistics
3	81 (8.1%)
4	26 (2.6%)
5	7 (0.7%)
6	3 (0.3%)
7	1 (0.1%)
WHERE_n_of_operators__like	
0	817 (82%)
1	164 (16%)
2	15 (1.5%)
3	4 (0.4%)
WHERE_n_of_operators__between	
0	615 (62%)
1	280 (28%)
2	81 (8.1%)
3	21 (2.1%)
4	3 (0.3%)
WHERE_n_of_operators__in	
0	608 (61%)
1	286 (29%)
2	77 (7.7%)
3	25 (2.5%)
4	4 (0.4%)
WHERE_n_of_non_aggr_func__LOG	
0	953 (95%)
1	46 (4.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	908 (91%)
1	86 (8.6%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func__DOW	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	977 (98%)
1	22 (2.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	956 (96%)
1	43 (4.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__DAY	
0	966 (97%)
1	31 (3.1%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__FLOOR	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	971 (97%)
1	28 (2.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	976 (98%)
1	23 (2.3%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	685 (69%)
1	230 (23%)
2	67 (6.7%)
3	16 (1.6%)
4	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 10.00] 2.37 / 2.14
HAVING_n_of_main_predicates	
0	298 (30%)
1	96 (9.6%)
2	157 (16%)
3	184 (18%)
4	144 (14%)
5	78 (7.8%)
6	43 (4.3%)
HAVING_n_of_main_predicates__scalar_subquery	
0	430 (43%)
1	213 (21%)
2	171 (17%)
3	186 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 19.0] 2.3 / 3.3
HAVING_n_of_processed_rows_by_subquery	[0, 2,250,038, 1,317,599,652] 75,480,663 / 161,064,820
ORDER_BY_n_of_columns	
0	286 (29%)
1	378 (38%)
2	246 (25%)
3	57 (5.7%)
4	19 (1.9%)
5	6 (0.6%)
6	6 (0.6%)

Query parameter or value	Descriptive statistics
γ	2 (0.2%)
limit	[1, 469, 996] 489 / 291
offset	[0, 8, 1,000] 239 / 313