

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

Scenario: 1-6jp_ujpp__g__h0ss_h0nss

Scale factor: 2 GB

The 1000-query pack was generated on: 2023-11-10

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, 179] 16 / 24
SELECT_n_of_non_aggr_func__RTRIM	
0	631 (63%)
1	243 (24%)
2	75 (7.5%)
3	22 (2.2%)
4	16 (1.6%)
5	7 (0.7%)
6	3 (0.3%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	610 (61%)
1	266 (27%)
2	81 (8.1%)

Query parameter or value	Descriptive statistics
3	17 (1.7%)
4	13 (1.3%)
5	8 (0.8%)
6	5 (0.5%)
SELECT_n_of_non_aggr_func___DOW	
0	835 (84%)
1	125 (13%)
2	24 (2.4%)
3	12 (1.2%)
4	1 (0.1%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func___MONTH	
0	813 (81%)
1	128 (13%)
2	31 (3.1%)
3	9 (0.9%)
4	8 (0.8%)
5	6 (0.6%)
6	3 (0.3%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	[0.00, 0.00, 11.00] 0.63 / 1.11
SELECT_n_of_non_aggr_func___YEAR	
0	829 (83%)
1	124 (12%)
2	27 (2.7%)
3	10 (1.0%)
4	5 (0.5%)
5	2 (0.2%)
6	3 (0.3%)
SELECT_n_of_non_aggr_func___LTRIM	
0	599 (60%)
1	260 (26%)
2	91 (9.1%)
3	24 (2.4%)
4	15 (1.5%)
5	8 (0.8%)
6	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	849 (85%)
1	118 (12%)
2	19 (1.9%)
3	9 (0.9%)
4	2 (0.2%)
5	2 (0.2%)
7	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non__aggr__func__LOWER	
0	612 (61%)
1	258 (26%)
2	75 (7.5%)
3	24 (2.4%)
4	14 (1.4%)
5	8 (0.8%)
6	4 (0.4%)
7	2 (0.2%)
8	3 (0.3%)
SELECT__n_of__non__aggr__func__DAY	
0	907 (91%)
1	70 (7.0%)
2	14 (1.4%)
3	5 (0.5%)
4	3 (0.3%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__ABS	
0	919 (92%)
1	68 (6.8%)
2	7 (0.7%)
3	4 (0.4%)
4	2 (0.2%)
SELECT__n_of__non__aggr__func__FLOOR	
0	919 (92%)
1	68 (6.8%)
2	11 (1.1%)
3	1 (0.1%)
4	1 (0.1%)
SELECT__n_of__non__aggr__func__SQRT	
0	818 (82%)
1	129 (13%)
10	1 (0.1%)
2	23 (2.3%)
3	14 (1.4%)
4	8 (0.8%)
5	4 (0.4%)
6	2 (0.2%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__ROUND	
0	929 (93%)
1	63 (6.3%)
2	6 (0.6%)
3	1 (0.1%)
4	1 (0.1%)
SELECT__n_of__non__aggr__func__TRUNC	

Query parameter or value	Descriptive statistics
0	921 (92%)
1	67 (6.7%)
2	8 (0.8%)
3	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 71.0] 4.9 / 7.5
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 11.00] 0.85 / 1.92
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 12.00] 0.85 / 1.99
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.81 / 1.86
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 13.00] 0.85 / 1.91
SELECT_n_of_aggr_func__AVG	
0	933 (93%)
1	41 (4.1%)
2	23 (2.3%)
3	3 (0.3%)
SELECT_n_of_aggr_func__SUM	
0	923 (92%)
1	42 (4.2%)
2	26 (2.6%)
3	5 (0.5%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 14.0] 3.6 / 3.1
FROM_n_of_join_paths	
1	166 (17%)
2	169 (17%)
3	163 (16%)
4	162 (16%)
5	165 (17%)
6	175 (18%)
FROM_n_of_super_joins__FULL	
0	535 (54%)
1	314 (31%)
2	121 (12%)
3	28 (2.8%)
4	2 (0.2%)
FROM_n_of_super_joins__RIGHT	
0	540 (54%)
1	311 (31%)
2	121 (12%)
3	27 (2.7%)
4	1 (0.1%)
FROM_n_of_super_joins__LEFT	
0	554 (55%)
1	303 (30%)
2	115 (12%)
3	27 (2.7%)

Query parameter or value	Descriptive statistics
4	1 (0.1%)
FROM_n_of_joins__INNER	[0.00, 3.00, 15.00] 3.28 / 2.49
FROM_n_of_joins__RIGHT	[0.00, 3.00, 12.00] 3.24 / 2.46
FROM_n_of_processed_rows	[5, 18,555,550, 85,114,719] 23,159,188 / 19,452,406
WHERE_n_of_predicates	[0.0, 3.0, 13.0] 3.9 / 3.0
WHERE_n_of_attribs_of_type__character_varying	
0	354 (35%)
1	295 (30%)
2	176 (18%)
3	95 (9.5%)
4	49 (4.9%)
5	23 (2.3%)
6	7 (0.7%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	783 (78%)
1	182 (18%)
2	30 (3.0%)
3	4 (0.4%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	369 (37%)
1	299 (30%)
2	174 (17%)
3	101 (10%)
4	31 (3.1%)
5	17 (1.7%)
6	5 (0.5%)
7	4 (0.4%)
WHERE_n_of_attribs_of_type__numeric	
0	560 (56%)
1	282 (28%)
2	113 (11%)
3	34 (3.4%)
4	8 (0.8%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type__character	
0	669 (67%)
1	249 (25%)
2	62 (6.2%)
3	17 (1.7%)
4	3 (0.3%)
WHERE_n_of_pkey_attribs	
0	543 (54%)
1	299 (30%)
2	115 (12%)

Query parameter or value	Descriptive statistics
3	32 (3.2%)
4	8 (0.8%)
5	3 (0.3%)
WHERE_n_of_connect_OR	[0.00, 1.00, 12.00] 2.20 / 2.50
WHERE_n_of_operators__between	
0	512 (51%)
1	320 (32%)
2	114 (11%)
3	35 (3.5%)
4	18 (1.8%)
5	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	318 (32%)
1	282 (28%)
2	181 (18%)
3	100 (10%)
4	70 (7.0%)
5	33 (3.3%)
6	13 (1.3%)
7	1 (0.1%)
8	2 (0.2%)
WHERE_n_of_operators__in	
0	539 (54%)
1	289 (29%)
2	124 (12%)
3	38 (3.8%)
4	8 (0.8%)
5	2 (0.2%)
WHERE_n_of_operators__like	
0	786 (79%)
1	187 (19%)
2	25 (2.5%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func__SQRT	
0	861 (86%)
1	123 (12%)
2	14 (1.4%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func__DOW	
0	953 (95%)
1	46 (4.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	960 (96%)
1	38 (3.8%)
2	2 (0.2%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__YEAR	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func__LOG	
0	936 (94%)
1	58 (5.8%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func__FLOOR	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	958 (96%)
1	42 (4.2%)
WHERE_n_of_non_aggr_func__TRUNC	
0	959 (96%)
1	39 (3.9%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__MONTH	
0	967 (97%)
1	32 (3.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_all_non_aggr_func	
0	613 (61%)
1	277 (28%)
2	83 (8.3%)
3	16 (1.6%)
4	9 (0.9%)
5	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 3.04 / 2.95
HAVING_n_of_main_predicates	
0	468 (47%)
1	197 (20%)
2	150 (15%)
3	185 (19%)
ORDER_BY_n_of_columns	
0	271 (27%)
1	345 (35%)
2	245 (25%)
3	104 (10%)
4	23 (2.3%)
5	8 (0.8%)
6	2 (0.2%)

Query parameter or value	Descriptive statistics
δ	1 (0.1%)
ρ	1 (0.1%)
limit	[2, 500, 998] 505 / 284
offset	[0, 0, 995] 251 / 324