

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

Scenario: 1-3jp_rjpp__g__h0-3ss_h0nss

Scale factor: 50 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, 6, 93] 10 / 13
SELECT_n_of_non_aggr_func__LOWER	
0	711 (71%)
1	219 (22%)
2	48 (4.8%)
3	15 (1.5%)
4	3 (0.3%)
5	2 (0.2%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	
0	832 (83%)
1	126 (13%)
2	24 (2.4%)
3	8 (0.8%)
4	5 (0.5%)

Query parameter or value	Descriptive statistics
5	3 (0.3%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	696 (70%)
1	226 (23%)
2	58 (5.8%)
3	15 (1.5%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	945 (95%)
1	47 (4.7%)
2	5 (0.5%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__YEAR	
0	848 (85%)
1	108 (11%)
2	29 (2.9%)
3	10 (1.0%)
4	5 (0.5%)
SELECT_n_of_non_aggr_func__ABS	
0	950 (95%)
1	43 (4.3%)
2	5 (0.5%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	878 (88%)
1	98 (9.8%)
2	16 (1.6%)
3	6 (0.6%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__LOG	
0	892 (89%)
1	92 (9.2%)
2	14 (1.4%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	
0	677 (68%)
1	252 (25%)
2	55 (5.5%)
3	9 (0.9%)
4	3 (0.3%)
5	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	

Query parameter or value	Descriptive statistics
0	936 (94%)
1	57 (5.7%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	841 (84%)
1	128 (13%)
2	20 (2.0%)
3	7 (0.7%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	719 (72%)
1	217 (22%)
2	51 (5.1%)
3	6 (0.6%)
4	5 (0.5%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	
0	713 (71%)
1	220 (22%)
2	48 (4.8%)
3	9 (0.9%)
4	8 (0.8%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	898 (90%)
1	81 (8.1%)
2	18 (1.8%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___FLOOR	
0	945 (95%)
1	49 (4.9%)
2	6 (0.6%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 38.0] 3.3 / 4.7
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.66 / 1.46
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 9.00] 0.62 / 1.45
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 9.00] 0.68 / 1.50
SELECT_n_of_aggr_func___AVG	
0	942 (94%)
1	47 (4.7%)
2	9 (0.9%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 10.00] 0.61 / 1.42

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func__SUM	
0	939 (94%)
1	44 (4.4%)
2	11 (1.1%)
3	6 (0.6%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.73 / 2.34
FROM_n_of_join_paths	
1	333 (33%)
2	346 (35%)
3	321 (32%)
FROM_n_of_super_joins__LEFT	
0	790 (79%)
1	185 (19%)
2	25 (2.5%)
FROM_n_of_super_joins__FULL	
0	773 (77%)
1	206 (21%)
2	21 (2.1%)
FROM_n_of_super_joins__RIGHT	
0	763 (76%)
1	216 (22%)
2	21 (2.1%)
FROM_n_of_joins__INNER	
0	238 (24%)
1	259 (26%)
2	219 (22%)
3	151 (15%)
4	83 (8.3%)
5	36 (3.6%)
6	11 (1.1%)
7	2 (0.2%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	235 (24%)
1	277 (28%)
2	222 (22%)
3	133 (13%)
4	77 (7.7%)
5	34 (3.4%)
6	12 (1.2%)
7	8 (0.8%)
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 340,491,970, 1,147,549,221] 326,551,560 / 301,017,159
WHERE_n_of_predicates	[0.00, 3.00, 9.00] 2.78 / 2.14
WHERE_n_of_attrbs_of_type__character_varying	

Query parameter or value	Descriptive statistics
0	448 (45%)
1	318 (32%)
2	157 (16%)
3	59 (5.9%)
4	16 (1.6%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	653 (65%)
1	247 (25%)
2	78 (7.8%)
3	20 (2.0%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	
0	437 (44%)
1	311 (31%)
2	179 (18%)
3	57 (5.7%)
4	12 (1.2%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type__character	
0	734 (73%)
1	209 (21%)
2	49 (4.9%)
3	8 (0.8%)
WHERE_n_of_attribs_of_type__date	
0	837 (84%)
1	140 (14%)
2	20 (2.0%)
3	3 (0.3%)
WHERE_n_of_pkey_attribs	
0	586 (59%)
1	324 (32%)
2	74 (7.4%)
3	14 (1.4%)
4	2 (0.2%)
WHERE_n_of_connect_OR	
0	452 (45%)
1	185 (19%)
2	141 (14%)
3	91 (9.1%)
4	69 (6.9%)
5	38 (3.8%)
6	16 (1.6%)
7	4 (0.4%)
8	4 (0.4%)
WHERE_n_of_operators__greater_or_less	

Query parameter or value	Descriptive statistics
0	400 (40%)
1	315 (32%)
2	174 (17%)
3	66 (6.6%)
4	29 (2.9%)
5	9 (0.9%)
6	7 (0.7%)
WHERE_n_of_operators__like	
0	848 (85%)
1	133 (13%)
2	18 (1.8%)
3	1 (0.1%)
WHERE_n_of_operators__between	
0	619 (62%)
1	278 (28%)
2	81 (8.1%)
3	14 (1.4%)
4	7 (0.7%)
5	1 (0.1%)
WHERE_n_of_operators__in	
0	621 (62%)
1	273 (27%)
2	88 (8.8%)
3	17 (1.7%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	898 (90%)
1	92 (9.2%)
2	10 (1.0%)
WHERE_n_of_non_aggr_func__TRUNC	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func__ROUND	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func__DOW	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func__FLOOR	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func__ABS	
0	964 (96%)
1	35 (3.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__LOG	

Query parameter or value	Descriptive statistics
0	948 (95%)
1	49 (4.9%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__YEAR	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func__DAY	
0	974 (97%)
1	25 (2.5%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_all_non_aggr_func	
0	704 (70%)
1	216 (22%)
2	54 (5.4%)
3	23 (2.3%)
4	2 (0.2%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.25 / 2.05
HAVING_n_of_main_predicates	
0	321 (32%)
1	108 (11%)
2	135 (14%)
3	188 (19%)
4	128 (13%)
5	79 (7.9%)
6	41 (4.1%)
HAVING_n_of_main_predicates__scalar_subquery	
0	470 (47%)
1	194 (19%)
2	175 (18%)
3	161 (16%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 21.0] 2.2 / 3.5
HAVING_n_of_processed_rows_by_subqueries	[0, 500,035, 17,213,238,315] 740,869,187 / 1,721,949,171
ORDER_BY_n_of_columns	
0	300 (30%)
1	383 (38%)
2	228 (23%)
3	69 (6.9%)
4	13 (1.3%)
5	5 (0.5%)
6	2 (0.2%)
limit	[3, 488, 999] 498 / 290

Query parameter or value	Descriptive statistics
offset	[0, 0, 998] 250 / 329