

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

Scenario: 1-6jp_ujpp__g__h0-3ss_h0-3nss

Scale factor: 1 GB

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

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, 162] 15 / 22
SELECT_n_of_non_aggr_func__FLOOR	
0	912 (91%)
1	74 (7.4%)
2	12 (1.2%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	
0	653 (65%)
1	241 (24%)
2	62 (6.2%)
3	23 (2.3%)
4	11 (1.1%)
5	4 (0.4%)
6	4 (0.4%)
8	2 (0.2%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non__aggr__func__RTRIM	
0	596 (60%)
1	283 (28%)
2	75 (7.5%)
3	25 (2.5%)
4	12 (1.2%)
5	6 (0.6%)
6	2 (0.2%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__SUBSTR	[0.00, 0.00, 9.00] 0.58 / 1.02
SELECT__n_of__non__aggr__func__UPPER	
0	609 (61%)
1	259 (26%)
2	83 (8.3%)
3	24 (2.4%)
4	11 (1.1%)
5	11 (1.1%)
6	2 (0.2%)
9	1 (0.1%)
SELECT__n_of__non__aggr__func__YEAR	
0	825 (83%)
1	121 (12%)
2	28 (2.8%)
3	13 (1.3%)
4	10 (1.0%)
5	3 (0.3%)
SELECT__n_of__non__aggr__func__LOWER	
0	657 (66%)
1	233 (23%)
2	70 (7.0%)
3	19 (1.9%)
4	9 (0.9%)
5	9 (0.9%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__SQRT	[0.00, 0.00, 9.00] 0.33 / 0.89
SELECT__n_of__non__aggr__func__DAY	
0	896 (90%)
1	77 (7.7%)
2	17 (1.7%)
3	8 (0.8%)
4	2 (0.2%)
SELECT__n_of__non__aggr__func__LOG	
0	868 (87%)
1	96 (9.6%)

Query parameter or value	Descriptive statistics
2	21 (2.1%)
3	6 (0.6%)
4	6 (0.6%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 9.00] 0.32 / 0.89
SELECT_n_of_non_aggr_func__ROUND	
0	922 (92%)
1	65 (6.5%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__DOW	
0	843 (84%)
1	112 (11%)
2	27 (2.7%)
3	10 (1.0%)
4	4 (0.4%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	919 (92%)
1	70 (7.0%)
2	5 (0.5%)
3	6 (0.6%)
SELECT_n_of_non_aggr_func__ABS	
0	935 (94%)
1	56 (5.6%)
2	7 (0.7%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 68.0] 4.7 / 7.3
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.80 / 1.76
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 12.00] 0.97 / 2.05
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.75 / 1.78
SELECT_n_of_aggr_func__AVG	
0	924 (92%)
1	42 (4.2%)
2	20 (2.0%)
3	10 (1.0%)
4	4 (0.4%)
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 12.00] 0.72 / 1.76
SELECT_n_of_aggr_func__SUM	
0	931 (93%)
1	42 (4.2%)
2	18 (1.8%)
3	9 (0.9%)

Query parameter or value	Descriptive statistics
SELECT_n_of_all_aggr_func	[0.00, 3.00, 13.00] 3.48 / 3.04
FROM_n_of_join_paths	
1	191 (19%)
2	177 (18%)
3	155 (16%)
4	162 (16%)
5	162 (16%)
6	153 (15%)
FROM_n_of_super_joins___FULL	
0	583 (58%)
1	284 (28%)
2	113 (11%)
3	14 (1.4%)
4	5 (0.5%)
5	1 (0.1%)
FROM_n_of_super_joins___RIGHT	
0	544 (54%)
1	325 (33%)
2	117 (12%)
3	12 (1.2%)
4	2 (0.2%)
FROM_n_of_super_joins___LEFT	
0	575 (58%)
1	292 (29%)
2	108 (11%)
3	21 (2.1%)
4	4 (0.4%)
FROM_n_of_joins___INNER	[0.00, 3.00, 13.00] 3.20 / 2.49
FROM_n_of_joins___RIGHT	[0.00, 3.00, 12.00] 3.08 / 2.41
FROM_n_of_processed_rows	[5, 9,150,591, 43,373,661] 11,437,676 / 9,565,124
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.69 / 2.85
WHERE_n_of_attribs_of_type___character	
0	627 (63%)
1	279 (28%)
2	79 (7.9%)
3	13 (1.3%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type___character_varying	
0	383 (38%)
1	300 (30%)
2	152 (15%)
3	103 (10%)
4	43 (4.3%)
5	13 (1.3%)
6	4 (0.4%)
7	2 (0.2%)

Query parameter or value	Descriptive statistics
WHERE_n_of_attribs_of_type__date	
0	798 (80%)
1	159 (16%)
2	40 (4.0%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	379 (38%)
1	301 (30%)
2	174 (17%)
3	89 (8.9%)
4	43 (4.3%)
5	12 (1.2%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	584 (58%)
1	273 (27%)
2	102 (10%)
3	33 (3.3%)
4	6 (0.6%)
5	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	538 (54%)
1	311 (31%)
2	110 (11%)
3	32 (3.2%)
4	8 (0.8%)
5	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.08 / 2.28
WHERE_n_of_operators__between	
0	554 (55%)
1	277 (28%)
2	113 (11%)
3	43 (4.3%)
4	10 (1.0%)
5	3 (0.3%)
WHERE_n_of_operators__greater_or_less	
0	341 (34%)
1	276 (28%)
2	194 (19%)
3	113 (11%)
4	48 (4.8%)
5	16 (1.6%)
6	8 (0.8%)
7	1 (0.1%)
8	3 (0.3%)
WHERE_n_of_operators__in	

Query parameter or value	Descriptive statistics
0	542 (54%)
1	280 (28%)
2	113 (11%)
3	49 (4.9%)
4	12 (1.2%)
5	4 (0.4%)
WHERE_n_of_operators_like	
0	806 (81%)
1	173 (17%)
2	18 (1.8%)
3	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func_YEAR	
0	961 (96%)
1	39 (3.9%)
WHERE_n_of_non_aggr_func_ABS	
0	960 (96%)
1	40 (4.0%)
WHERE_n_of_non_aggr_func_FLOOR	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func_SQRT	
0	879 (88%)
1	108 (11%)
2	11 (1.1%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func_DAY	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func_LOG	
0	928 (93%)
1	70 (7.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_ROUND	
0	953 (95%)
1	47 (4.7%)
WHERE_n_of_non_aggr_func_TRUNC	
0	959 (96%)
1	41 (4.1%)
WHERE_n_of_non_aggr_func_DOW	
0	962 (96%)
1	36 (3.6%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_MONTH	
0	966 (97%)
1	34 (3.4%)

Query parameter or value	Descriptive statistics
WHERE_n_of_all_non_aggr_func	
0	636 (64%)
1	252 (25%)
2	85 (8.5%)
3	21 (2.1%)
4	6 (0.6%)
GROUP_BY_n_of_columns	[0.00, 2.00, 13.00] 3.06 / 2.89
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.46 / 2.64
HAVING_n_of_main_predicates__non_scalar_subquery	
0	429 (43%)
1	157 (16%)
2	212 (21%)
3	202 (20%)
HAVING_n_of_main_predicates__scalar_subquery	
0	434 (43%)
1	184 (18%)
2	183 (18%)
3	199 (20%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	429 (43%)
1	157 (16%)
2	212 (21%)
3	202 (20%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 23.0] 2.6 / 3.8
HAVING_n_of_processed_rows_by_subqueries	[0, 13,611,182, 893,834,991] 41,526,597 / 75,017,528
ORDER_BY_n_of_columns	
0	316 (32%)
1	312 (31%)
2	236 (24%)
3	91 (9.1%)
4	25 (2.5%)
5	11 (1.1%)
6	7 (0.7%)
7	2 (0.2%)
limit	[1, 478, 999] 487 / 287
offset	[0, 0, 999] 240 / 316