

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

Scenario: 1-3jp_rjpp__g__h0-3ss_h0nss

Scale factor: 1 GB

The 1000-query pack was generated on: 2023-11-06

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, 99] 10 / 12
SELECT_n_of_non_aggr_func__DOW	
0	889 (89%)
1	85 (8.5%)
2	22 (2.2%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func__LOWER	
0	695 (70%)
1	234 (23%)
2	57 (5.7%)
3	9 (0.9%)
4	5 (0.5%)
SELECT_n_of_non_aggr_func__MONTH	
0	851 (85%)
1	109 (11%)

Query parameter or value	Descriptive statistics
2	25 (2.5%)
3	6 (0.6%)
4	7 (0.7%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	727 (73%)
1	219 (22%)
2	39 (3.9%)
3	8 (0.8%)
4	3 (0.3%)
5	4 (0.4%)
SELECT_n_of_non_aggr_func__UPPER	
0	666 (67%)
1	264 (26%)
2	58 (5.8%)
3	9 (0.9%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func__YEAR	
0	871 (87%)
1	100 (10%)
2	16 (1.6%)
3	11 (1.1%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	
0	715 (72%)
1	222 (22%)
2	48 (4.8%)
3	13 (1.3%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	907 (91%)
1	81 (8.1%)
2	10 (1.0%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	712 (71%)
1	232 (23%)
2	44 (4.4%)
3	11 (1.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	946 (95%)
1	47 (4.7%)
2	3 (0.3%)
3	4 (0.4%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func__DAY	
0	932 (93%)
1	54 (5.4%)
2	13 (1.3%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	852 (85%)
1	119 (12%)
2	20 (2.0%)
3	4 (0.4%)
4	5 (0.5%)
SELECT_n_of_non_aggr_func__TRUNC	
0	944 (94%)
1	51 (5.1%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__ABS	
0	942 (94%)
1	54 (5.4%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func__FLOOR	
0	944 (94%)
1	50 (5.0%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 35.00] 3.05 / 3.74
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 10.00] 0.57 / 1.37
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 11.00] 0.71 / 1.64
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.63 / 1.43
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.67 / 1.49
SELECT_n_of_aggr_func__AVG	
0	943 (94%)
1	44 (4.4%)
2	12 (1.2%)
3	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	953 (95%)
1	34 (3.4%)
2	9 (0.9%)
3	4 (0.4%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.71 / 2.39
FROM_n_of_join_paths	
1	345 (35%)
2	322 (32%)
3	333 (33%)
FROM_n_of_super_joins__FULL	
0	777 (78%)

Query parameter or value	Descriptive statistics
1	204 (20%)
2	19 (1.9%)
FROM_n_of_super_joins__LEFT	
0	791 (79%)
1	188 (19%)
2	21 (2.1%)
FROM_n_of_super_joins__RIGHT	
0	761 (76%)
1	222 (22%)
2	17 (1.7%)
FROM_n_of_joins__INNER	
0	260 (26%)
1	256 (26%)
2	205 (21%)
3	134 (13%)
4	78 (7.8%)
5	47 (4.7%)
6	14 (1.4%)
7	6 (0.6%)
FROM_n_of_joins__RIGHT	
0	268 (27%)
1	241 (24%)
2	213 (21%)
3	140 (14%)
4	80 (8.0%)
5	33 (3.3%)
6	21 (2.1%)
7	4 (0.4%)
FROM_n_of_processed_rows	[5, 6,810,591, 22,951,828] 6,306,893 / 6,079,348
WHERE_n_of_predicates	[0.00, 3.00, 10.00] 2.80 / 2.16
WHERE_n_of_attrbs_of_type__character_varying	
0	426 (43%)
1	309 (31%)
2	173 (17%)
3	73 (7.3%)
4	16 (1.6%)
5	2 (0.2%)
8	1 (0.1%)
WHERE_n_of_attrbs_of_type__integer	
0	442 (44%)
1	329 (33%)
2	155 (16%)
3	64 (6.4%)
4	8 (0.8%)
5	2 (0.2%)
WHERE_n_of_attrbs_of_type__numeric	

Query parameter or value	Descriptive statistics
0	637 (64%)
1	272 (27%)
2	70 (7.0%)
3	20 (2.0%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	853 (85%)
1	127 (13%)
2	19 (1.9%)
3	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	739 (74%)
1	201 (20%)
2	53 (5.3%)
3	7 (0.7%)
WHERE_n_of_pkey_attribs	
0	613 (61%)
1	297 (30%)
2	79 (7.9%)
3	10 (1.0%)
4	1 (0.1%)
WHERE_n_of_connect_OR	
0	447 (45%)
1	183 (18%)
2	131 (13%)
3	123 (12%)
4	56 (5.6%)
5	36 (3.6%)
6	19 (1.9%)
7	5 (0.5%)
WHERE_n_of_operators__between	
0	620 (62%)
1	270 (27%)
2	90 (9.0%)
3	17 (1.7%)
4	3 (0.3%)
WHERE_n_of_operators__in	
0	613 (61%)
1	269 (27%)
2	95 (9.5%)
3	19 (1.9%)
4	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	423 (42%)
1	298 (30%)

Query parameter or value	Descriptive statistics
2	168 (17%)
3	77 (7.7%)
4	28 (2.8%)
5	4 (0.4%)
6	2 (0.2%)
WHERE_n_of_operators__like	
0	822 (82%)
1	153 (15%)
2	25 (2.5%)
WHERE_n_of_non_aggr_func__LOG	
0	941 (94%)
1	54 (5.4%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func__MONTH	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func__DAY	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func__SQRT	
0	904 (90%)
1	88 (8.8%)
2	8 (0.8%)
WHERE_n_of_non_aggr_func__ABS	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func__FLOOR	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_non_aggr_func__TRUNC	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func__YEAR	
0	981 (98%)
1	19 (1.9%)
WHERE_n_of_non_aggr_func__ROUND	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func__DOW	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_all_non_aggr_func	
0	724 (72%)
1	206 (21%)
2	54 (5.4%)
3	14 (1.4%)

Query parameter or value	Descriptive statistics
4	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.26 / 2.09
HAVING_n_of_main_predicates	
0	292 (29%)
1	98 (9.8%)
2	153 (15%)
3	173 (17%)
4	151 (15%)
5	84 (8.4%)
6	49 (4.9%)
HAVING_n_of_main_predicates__scalar_subquery	
0	442 (44%)
1	196 (20%)
2	168 (17%)
3	194 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 21.0] 2.4 / 3.5
HAVING_n_of_processed_rows_by_subqueries	[0, 200,000, 296,423,699] 15,416,534 / 35,608,638
ORDER_BY_n_of_columns	
0	275 (28%)
1	369 (37%)
2	265 (27%)
3	69 (6.9%)
4	15 (1.5%)
5	3 (0.3%)
6	4 (0.4%)
limit	[3, 526, 1,000] 508 / 277
offset	[0, 0, 995] 243 / 322