

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

Scenario: 1-3jp_rjpp__g__h0-3ss_h0nss

Scale factor: 5 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, 81] 10 / 12
SELECT_n_of_non_aggr_func__LOG	
0	901 (90%)
1	85 (8.5%)
2	12 (1.2%)
3	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	685 (69%)
1	232 (23%)
2	62 (6.2%)
3	16 (1.6%)
4	2 (0.2%)
5	2 (0.2%)
6	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non__aggr__func__RTRIM	
0	684 (68%)
1	245 (25%)
2	50 (5.0%)
3	17 (1.7%)
4	2 (0.2%)
5	1 (0.1%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__LOWER	
0	696 (70%)
1	235 (24%)
2	47 (4.7%)
3	14 (1.4%)
4	4 (0.4%)
5	1 (0.1%)
6	3 (0.3%)
SELECT__n_of__non__aggr__func__LTRIM	
0	681 (68%)
1	239 (24%)
2	63 (6.3%)
3	12 (1.2%)
4	5 (0.5%)
SELECT__n_of__non__aggr__func__UPPER	
0	672 (67%)
1	251 (25%)
2	56 (5.6%)
3	14 (1.4%)
4	4 (0.4%)
5	2 (0.2%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__DAY	
0	896 (90%)
1	82 (8.2%)
2	17 (1.7%)
3	5 (0.5%)
SELECT__n_of__non__aggr__func__MONTH	
0	842 (84%)
1	122 (12%)
2	25 (2.5%)
3	8 (0.8%)
4	2 (0.2%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__SQRT	
0	848 (85%)
1	123 (12%)
2	19 (1.9%)

Query parameter or value	Descriptive statistics
3	5 (0.5%)
4	5 (0.5%)
SELECT_n_of_non_aggr_func__TRUNC	
0	944 (94%)
1	51 (5.1%)
2	4 (0.4%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	870 (87%)
1	102 (10%)
2	22 (2.2%)
3	4 (0.4%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	889 (89%)
1	89 (8.9%)
2	15 (1.5%)
3	5 (0.5%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__ABS	
0	945 (95%)
1	49 (4.9%)
2	4 (0.4%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	947 (95%)
1	48 (4.8%)
2	3 (0.3%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	942 (94%)
1	51 (5.1%)
2	7 (0.7%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 40.0] 3.3 / 4.3
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.69 / 1.59
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 11.00] 0.68 / 1.52
SELECT_n_of_aggr_func__SUM	
0	951 (95%)
1	36 (3.6%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 10.00] 0.72 / 1.55
SELECT_n_of_aggr_func__COUNT	
0	777 (78%)

Query parameter or value	Descriptive statistics
1	90 (9.0%)
2	46 (4.6%)
3	32 (3.2%)
4	21 (2.1%)
5	10 (1.0%)
6	18 (1.8%)
7	5 (0.5%)
8	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	929 (93%)
1	47 (4.7%)
2	23 (2.3%)
3	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.80 / 2.34
FROM_n_of_join_paths	
1	352 (35%)
2	311 (31%)
3	337 (34%)
FROM_n_of_super_joins__RIGHT	
0	780 (78%)
1	199 (20%)
2	21 (2.1%)
FROM_n_of_super_joins__LEFT	
0	786 (79%)
1	191 (19%)
2	23 (2.3%)
FROM_n_of_super_joins__FULL	
0	756 (76%)
1	228 (23%)
2	16 (1.6%)
FROM_n_of_joins__INNER	
0	219 (22%)
1	276 (28%)
2	201 (20%)
3	160 (16%)
4	88 (8.8%)
5	35 (3.5%)
6	18 (1.8%)
7	3 (0.3%)
FROM_n_of_joins__RIGHT	[0.00, 1.00, 9.00] 1.75 / 1.57
FROM_n_of_processed_rows	[5, 34,050,118, 114,749,827] 31,807,732 / 29,685,286
WHERE_n_of_predicates	[0.00, 2.00, 9.00] 2.77 / 2.15
WHERE_n_of_attribs_of_type__character	
0	736 (74%)
1	213 (21%)
2	44 (4.4%)

Query parameter or value	Descriptive statistics
3	7 (0.7%)
WHERE_n_of_attribs_of_type__character_varying	
0	439 (44%)
1	304 (30%)
2	169 (17%)
3	63 (6.3%)
4	21 (2.1%)
5	3 (0.3%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	854 (85%)
1	126 (13%)
2	19 (1.9%)
3	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	670 (67%)
1	241 (24%)
2	66 (6.6%)
3	20 (2.0%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__integer	
0	440 (44%)
1	316 (32%)
2	166 (17%)
3	64 (6.4%)
4	8 (0.8%)
5	5 (0.5%)
6	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	621 (62%)
1	285 (29%)
2	80 (8.0%)
3	11 (1.1%)
4	3 (0.3%)
WHERE_n_of_connect_OR	
0	449 (45%)
1	175 (18%)
2	149 (15%)
3	107 (11%)
4	63 (6.3%)
5	34 (3.4%)
6	19 (1.9%)
7	3 (0.3%)
8	1 (0.1%)
WHERE_n_of_operators__between	
0	635 (64%)

Query parameter or value	Descriptive statistics
1	255 (26%)
2	87 (8.7%)
3	20 (2.0%)
4	3 (0.3%)
WHERE_n_of_operators__greater_or_less	
0	433 (43%)
1	288 (29%)
2	180 (18%)
3	73 (7.3%)
4	23 (2.3%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_operators__in	
0	616 (62%)
1	261 (26%)
2	99 (9.9%)
3	22 (2.2%)
4	2 (0.2%)
WHERE_n_of_operators__like	
0	827 (83%)
1	159 (16%)
2	12 (1.2%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func__DOW	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func__SQRT	
0	901 (90%)
1	94 (9.4%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func__YEAR	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func__MONTH	
0	979 (98%)
1	21 (2.1%)
WHERE_n_of_non_aggr_func__ABS	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func__DAY	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	977 (98%)
1	23 (2.3%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__TRUNC	
0	964 (96%)
1	36 (3.6%)
WHERE_n_of_non_aggr_func__LOG	
0	944 (94%)
1	54 (5.4%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__FLOOR	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	706 (71%)
1	221 (22%)
2	60 (6.0%)
3	10 (1.0%)
4	3 (0.3%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.30 / 2.11
HAVING_n_of_main_predicates	
0	317 (32%)
1	92 (9.2%)
2	149 (15%)
3	154 (15%)
4	134 (13%)
5	103 (10%)
6	51 (5.1%)
HAVING_n_of_main_predicates__scalar_subquery	
0	443 (44%)
1	177 (18%)
2	204 (20%)
3	176 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 21.0] 2.3 / 3.5
HAVING_n_of_processed_rows_by_subquery	[0, 1,000,000, 1,156,801,312] 72,217,642 / 159,296,910
ORDER_BY_n_of_columns	
0	300 (30%)
1	386 (39%)
2	239 (24%)
3	50 (5.0%)
4	16 (1.6%)
5	5 (0.5%)
6	4 (0.4%)
limit	[1, 492, 1,000] 490 / 288
offset	[0, 16, 999] 258 / 324