

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

Scenario: 1-3jp_ujpp__g__h0-3ss_h0nss

Scale factor: 2 GB

The 1000-query pack was generated on: 2023-11-09

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, 113] 10 / 13
SELECT_n_of_non_aggr_func__LOWER	
0	711 (71%)
1	218 (22%)
2	47 (4.7%)
3	15 (1.5%)
4	5 (0.5%)
5	4 (0.4%)
SELECT_n_of_non_aggr_func__RTRIM	
0	669 (67%)
1	256 (26%)
2	50 (5.0%)
3	19 (1.9%)
4	5 (0.5%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non__aggr__func__SQRT	
0	837 (84%)
1	126 (13%)
2	29 (2.9%)
3	7 (0.7%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__LOG	
0	889 (89%)
1	92 (9.2%)
2	11 (1.1%)
3	8 (0.8%)
SELECT__n_of__non__aggr__func__LTRIM	
0	686 (69%)
1	243 (24%)
2	48 (4.8%)
3	13 (1.3%)
4	7 (0.7%)
5	1 (0.1%)
6	1 (0.1%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__ROUND	
0	944 (94%)
1	51 (5.1%)
2	5 (0.5%)
SELECT__n_of__non__aggr__func__SUBSTR	
0	675 (68%)
1	235 (24%)
2	65 (6.5%)
3	19 (1.9%)
4	4 (0.4%)
6	2 (0.2%)
SELECT__n_of__non__aggr__func__UPPER	
0	694 (69%)
1	241 (24%)
2	42 (4.2%)
3	12 (1.2%)
4	7 (0.7%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__ABS	
0	945 (95%)
1	49 (4.9%)
2	5 (0.5%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func__MONTH	

Query parameter or value	Descriptive statistics
0	846 (85%)
1	115 (12%)
2	23 (2.3%)
3	6 (0.6%)
4	5 (0.5%)
5	3 (0.3%)
6	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	926 (93%)
1	68 (6.8%)
2	6 (0.6%)
SELECT_n_of_non_aggr_func__DAY	
0	920 (92%)
1	62 (6.2%)
2	13 (1.3%)
3	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	861 (86%)
1	113 (11%)
2	18 (1.8%)
3	4 (0.4%)
4	2 (0.2%)
6	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	885 (89%)
1	90 (9.0%)
2	14 (1.4%)
3	7 (0.7%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	941 (94%)
1	53 (5.3%)
2	6 (0.6%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 50.0] 3.4 / 4.9
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.68 / 1.59
SELECT_n_of_aggr_func__MAX	
0	780 (78%)
1	84 (8.4%)
2	52 (5.2%)
3	33 (3.3%)
4	21 (2.1%)
5	16 (1.6%)

Query parameter or value	Descriptive statistics
6	4 (0.4%)
7	5 (0.5%)
8	5 (0.5%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.64 / 1.46
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.67 / 1.47
SELECT_n_of_aggr_func__SUM	
0	944 (94%)
1	35 (3.5%)
2	17 (1.7%)
3	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	946 (95%)
1	40 (4.0%)
2	12 (1.2%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 10.00] 2.69 / 2.32
FROM_n_of_join_paths	
1	317 (32%)
2	359 (36%)
3	324 (32%)
FROM_n_of_super_joins__RIGHT	
0	736 (74%)
1	241 (24%)
2	23 (2.3%)
FROM_n_of_super_joins__FULL	
0	779 (78%)
1	203 (20%)
2	18 (1.8%)
FROM_n_of_super_joins__LEFT	
0	780 (78%)
1	207 (21%)
2	13 (1.3%)
FROM_n_of_joins__INNER	[0.00, 1.00, 9.00] 1.71 / 1.49
FROM_n_of_joins__RIGHT	
0	235 (24%)
1	268 (27%)
2	205 (21%)
3	131 (13%)
4	98 (9.8%)
5	38 (3.8%)
6	19 (1.9%)
7	6 (0.6%)
FROM_n_of_processed_rows	[5, 13,615,796, 45,887,241] 12,804,519 / 11,697,003
WHERE_n_of_predicates	[0.00, 2.00, 9.00] 2.82 / 2.23

Query parameter or value	Descriptive statistics
WHERE_n_of_attribs_of_type__character	
0	736 (74%)
1	207 (21%)
2	45 (4.5%)
3	10 (1.0%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__character_varying	
0	436 (44%)
1	307 (31%)
2	173 (17%)
3	61 (6.1%)
4	21 (2.1%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	828 (83%)
1	163 (16%)
2	9 (0.9%)
WHERE_n_of_attribs_of_type__numeric	
0	659 (66%)
1	222 (22%)
2	92 (9.2%)
3	23 (2.3%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__integer	
0	450 (45%)
1	310 (31%)
2	166 (17%)
3	58 (5.8%)
4	14 (1.4%)
5	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	599 (60%)
1	311 (31%)
2	71 (7.1%)
3	16 (1.6%)
4	3 (0.3%)
WHERE_n_of_connect_OR	
0	450 (45%)
1	195 (20%)
2	127 (13%)
3	95 (9.5%)
4	62 (6.2%)
5	40 (4.0%)
6	22 (2.2%)
7	4 (0.4%)
8	5 (0.5%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators___greater_or_less	
0	402 (40%)
1	323 (32%)
2	153 (15%)
3	91 (9.1%)
4	28 (2.8%)
5	3 (0.3%)
WHERE_n_of_operators___in	
0	628 (63%)
1	256 (26%)
2	89 (8.9%)
3	21 (2.1%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_operators___like	
0	849 (85%)
1	133 (13%)
2	18 (1.8%)
WHERE_n_of_operators___between	
0	623 (62%)
1	255 (26%)
2	96 (9.6%)
3	23 (2.3%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_non_aggr_func___ABS	
0	979 (98%)
1	19 (1.9%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	901 (90%)
1	91 (9.1%)
2	8 (0.8%)
WHERE_n_of_non_aggr_func___FLOOR	
0	961 (96%)
1	38 (3.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	943 (94%)
1	56 (5.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___MONTH	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func___DOW	

Query parameter or value	Descriptive statistics
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func__TRUNC	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func__YEAR	
0	978 (98%)
1	22 (2.2%)
WHERE_n_of_non_aggr_func__ROUND	
0	970 (97%)
1	28 (2.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__DAY	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_all_non_aggr_func	
0	690 (69%)
1	240 (24%)
2	53 (5.3%)
3	15 (1.5%)
4	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.29 / 2.08
HAVING_n_of_main_predicates	
0	315 (32%)
1	78 (7.8%)
2	129 (13%)
3	193 (19%)
4	146 (15%)
5	92 (9.2%)
6	47 (4.7%)
HAVING_n_of_main_predicates__scalar_subquery	
0	452 (45%)
1	175 (18%)
2	185 (19%)
3	188 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 20.0] 2.4 / 3.5
HAVING_n_of_processed_rows_by_subqueries	[0, 190,285, 643,609,335] 28,848,818 / 63,429,131
ORDER_BY_n_of_columns	
0	307 (31%)
1	375 (38%)
2	251 (25%)
3	47 (4.7%)
4	11 (1.1%)
5	6 (0.6%)
6	2 (0.2%)
8	1 (0.1%)

Query parameter or value	Descriptive statistics
limit	[1, 475, 998] 485 / 293
offset	[0, 0, 999] 262 / 336