

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

Scenario: 1-3jp_rjpp__g__h0-3ss_h0nss

Scale factor: 0.01 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, 97] 10 / 12
SELECT_n_of_non_aggr_func__SUBSTR	
0	705 (71%)
1	232 (23%)
2	47 (4.7%)
3	8 (0.8%)
4	4 (0.4%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	945 (95%)
1	47 (4.7%)
2	8 (0.8%)
SELECT_n_of_non_aggr_func__ROUND	

Query parameter or value	Descriptive statistics
0	949 (95%)
1	45 (4.5%)
2	4 (0.4%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	683 (68%)
1	239 (24%)
2	61 (6.1%)
3	13 (1.3%)
4	1 (0.1%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	692 (69%)
1	237 (24%)
2	51 (5.1%)
3	13 (1.3%)
4	6 (0.6%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	704 (70%)
1	220 (22%)
2	57 (5.7%)
3	12 (1.2%)
4	5 (0.5%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	686 (69%)
1	261 (26%)
2	40 (4.0%)
3	10 (1.0%)
4	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	906 (91%)
1	75 (7.5%)
2	12 (1.2%)
3	4 (0.4%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func___YEAR	
0	878 (88%)
1	92 (9.2%)
2	22 (2.2%)
3	6 (0.6%)

Query parameter or value	Descriptive statistics
4	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	
0	862 (86%)
1	110 (11%)
2	19 (1.9%)
3	4 (0.4%)
4	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__SQRT	
0	869 (87%)
1	98 (9.8%)
2	24 (2.4%)
3	7 (0.7%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	948 (95%)
1	49 (4.9%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func__DAY	
0	917 (92%)
1	75 (7.5%)
2	8 (0.8%)
SELECT_n_of_non_aggr_func__LOG	
0	904 (90%)
1	79 (7.9%)
2	14 (1.4%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	949 (95%)
1	44 (4.4%)
2	5 (0.5%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 38.00] 3.09 / 4.06
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.66 / 1.47
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.64 / 1.41
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 10.00] 0.62 / 1.37
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.65 / 1.50
SELECT_n_of_aggr_func__SUM	
0	946 (95%)
1	37 (3.7%)
2	16 (1.6%)
3	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	948 (95%)

Query parameter or value	Descriptive statistics
1	41 (4.1%)
2	9 (0.9%)
3	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.70 / 2.30
FROM_n_of_join_paths	
1	336 (34%)
2	301 (30%)
3	363 (36%)
FROM_n_of_super_joins__RIGHT	
0	783 (78%)
1	201 (20%)
2	16 (1.6%)
FROM_n_of_super_joins__FULL	
0	774 (77%)
1	203 (20%)
2	23 (2.3%)
FROM_n_of_super_joins__LEFT	
0	762 (76%)
1	212 (21%)
2	26 (2.6%)
FROM_n_of_joins__INNER	
0	234 (23%)
1	251 (25%)
2	228 (23%)
3	137 (14%)
4	79 (7.9%)
5	50 (5.0%)
6	15 (1.5%)
7	5 (0.5%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	230 (23%)
1	269 (27%)
2	207 (21%)
3	148 (15%)
4	92 (9.2%)
5	37 (3.7%)
6	11 (1.1%)
7	2 (0.2%)
8	4 (0.4%)
FROM_n_of_processed_rows	[5, 68,308, 230,110] 63,963 / 58,606
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.72 / 2.25
WHERE_n_of_attribs_of_type__integer	
0	462 (46%)
1	306 (31%)
2	157 (16%)

Query parameter or value	Descriptive statistics
3	54 (5.4%)
4	20 (2.0%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	458 (46%)
1	295 (30%)
2	156 (16%)
3	60 (6.0%)
4	26 (2.6%)
5	4 (0.4%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	736 (74%)
1	216 (22%)
2	39 (3.9%)
3	8 (0.8%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	674 (67%)
1	237 (24%)
2	72 (7.2%)
3	13 (1.3%)
4	3 (0.3%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	854 (85%)
1	123 (12%)
2	17 (1.7%)
3	6 (0.6%)
WHERE_n_of_pkey_attribs	
0	621 (62%)
1	280 (28%)
2	82 (8.2%)
3	16 (1.6%)
4	1 (0.1%)
WHERE_n_of_connect_OR	[0.00, 1.00, 9.00] 1.35 / 1.73
WHERE_n_of_operators__between	
0	616 (62%)
1	262 (26%)
2	91 (9.1%)
3	25 (2.5%)
4	5 (0.5%)
5	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	431 (43%)
1	302 (30%)

Query parameter or value	Descriptive statistics
2	160 (16%)
3	74 (7.4%)
4	24 (2.4%)
5	9 (0.9%)
WHERE_n_of_operators_in	
0	637 (64%)
1	253 (25%)
2	81 (8.1%)
3	22 (2.2%)
4	7 (0.7%)
WHERE_n_of_operators_like	
0	842 (84%)
1	137 (14%)
2	20 (2.0%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_LOG	
0	937 (94%)
1	63 (6.3%)
WHERE_n_of_non_aggr_func_SQRT	
0	920 (92%)
1	76 (7.6%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func_DAY	
0	978 (98%)
1	20 (2.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_FLOOR	
0	975 (98%)
1	23 (2.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_TRUNC	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func_ABS	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func_YEAR	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func_DOW	
0	977 (98%)
1	23 (2.3%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__MONTH	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	733 (73%)
1	204 (20%)
2	50 (5.0%)
3	12 (1.2%)
4	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.39 / 2.10
HAVING_n_of_main_predicates	
0	291 (29%)
1	103 (10%)
2	146 (15%)
3	180 (18%)
4	147 (15%)
5	80 (8.0%)
6	53 (5.3%)
HAVING_n_of_main_predicates__scalar_subquery	
0	422 (42%)
1	215 (22%)
2	174 (17%)
3	189 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 23.0] 2.4 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 4,673, 2,823,300] 158,573 / 339,907
ORDER_BY_n_of_columns	
0	312 (31%)
1	372 (37%)
2	228 (23%)
3	69 (6.9%)
4	7 (0.7%)
5	7 (0.7%)
6	5 (0.5%)
limit	[1, 487, 1,000] 488 / 286
offset	[0, 0, 999] 237 / 317