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

Scenario: 1-1jp__g__h0-3ss_h0-3nss

Scale factor: 1 GB

The 1000-query pack was generated on: 2023-11-05

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	[1.0, 4.0, 38.0] 5.6 / 5.1
SELECT_n_of_non_aggr_func___SUBSTR	
0	760 (76%)
1	212 (21%)
2	25 (2.5%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	944 (94%)
1	50 (5.0%)
2	6 (0.6%)
SELECT_n_of_non_aggr_func___UPPER	
0	796 (80%)
1	186 (19%)
2	16 (1.6%)

Query parameter or value	Descriptive statistics
3	2 (0.2%)
SELECT_n_of_non_aggr_func___YEAR	
0	930 (93%)
1	54 (5.4%)
2	16 (1.6%)
SELECT_n_of_non_aggr_func___LOWER	
0	760 (76%)
1	213 (21%)
2	25 (2.5%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___RTRIM	
0	774 (77%)
1	198 (20%)
2	24 (2.4%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	904 (90%)
1	84 (8.4%)
2	9 (0.9%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___LTRIM	
0	802 (80%)
1	174 (17%)
2	21 (2.1%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___DAY	
0	950 (95%)
1	46 (4.6%)
2	3 (0.3%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	919 (92%)
1	67 (6.7%)
2	10 (1.0%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func___TRUNC	
0	967 (97%)
1	32 (3.2%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	936 (94%)
1	55 (5.5%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	

Query parameter or value	Descriptive statistics
0	967 (97%)
1	31 (3.1%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	968 (97%)
1	32 (3.2%)
SELECT_n_of_non_aggr_func___FLOOR	
0	969 (97%)
1	29 (2.9%)
2	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.00, 1.00, 18.00] 1.87 / 2.18
SELECT_n_of_aggr_func___COUNT	
0	780 (78%)
1	119 (12%)
2	50 (5.0%)
3	35 (3.5%)
4	12 (1.2%)
5	2 (0.2%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_aggr_func___SUM	
0	954 (95%)
1	41 (4.1%)
2	5 (0.5%)
SELECT_n_of_aggr_func___MIN	
0	773 (77%)
1	116 (12%)
2	49 (4.9%)
3	40 (4.0%)
4	9 (0.9%)
5	8 (0.8%)
6	4 (0.4%)
7	1 (0.1%)
SELECT_n_of_aggr_func___MAX	
0	762 (76%)
1	131 (13%)
2	51 (5.1%)
3	28 (2.8%)
4	23 (2.3%)
5	1 (0.1%)
6	4 (0.4%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	764 (76%)
1	107 (11%)
2	68 (6.8%)
3	37 (3.7%)

Query parameter or value	Descriptive statistics
4	15 (1.5%)
5	5 (0.5%)
6	3 (0.3%)
7	1 (0.1%)
SELECT_n_of_aggr_func___AVG	
0	962 (96%)
1	34 (3.4%)
2	4 (0.4%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 9.00] 1.83 / 1.48
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins___INNER	
0	520 (52%)
1	311 (31%)
2	128 (13%)
3	34 (3.4%)
4	7 (0.7%)
FROM_n_of_joins___RIGHT	
0	463 (46%)
1	346 (35%)
2	148 (15%)
3	37 (3.7%)
4	6 (0.6%)
FROM_n_of_processed_rows	[5, 1,000,000, 7,650,621] 2,751,090 / 3,069,012
WHERE_n_of_predicates	
0	251 (25%)
1	231 (23%)
2	214 (21%)
3	157 (16%)
4	100 (10%)
5	36 (3.6%)
6	11 (1.1%)
WHERE_n_of_attribs_of_type___integer	
0	561 (56%)
1	324 (32%)
2	98 (9.8%)
3	17 (1.7%)
WHERE_n_of_attribs_of_type___character_varying	
0	548 (55%)
1	328 (33%)
2	105 (11%)
3	19 (1.9%)
WHERE_n_of_attribs_of_type___character	
0	830 (83%)
1	147 (15%)
2	20 (2.0%)

Query parameter or value	Descriptive statistics
3	3 (0.3%)
WHERE_n_of_attribs_of_type___date	
0	896 (90%)
1	95 (9.5%)
2	8 (0.8%)
3	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	751 (75%)
1	202 (20%)
2	43 (4.3%)
3	4 (0.4%)
WHERE_n_of_pkey_attribs	
0	698 (70%)
1	264 (26%)
2	37 (3.7%)
3	1 (0.1%)
WHERE_n_of_connect_OR	
0	613 (61%)
1	210 (21%)
2	112 (11%)
3	51 (5.1%)
4	10 (1.0%)
5	4 (0.4%)
WHERE_n_of_operators___between	
0	707 (71%)
1	232 (23%)
2	53 (5.3%)
3	8 (0.8%)
WHERE_n_of_operators___greater_or_less	
0	549 (55%)
1	308 (31%)
2	105 (11%)
3	29 (2.9%)
4	9 (0.9%)
WHERE_n_of_operators___in	
0	722 (72%)
1	233 (23%)
2	38 (3.8%)
3	7 (0.7%)
WHERE_n_of_operators___like	
0	872 (87%)
1	123 (12%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___YEAR	
0	981 (98%)
1	18 (1.8%)

Query parameter or value	Descriptive statistics
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___LOG	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func___DOW	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func___SQRT	
0	943 (94%)
1	56 (5.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func___DAY	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func___ABS	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___MONTH	
0	988 (99%)
1	12 (1.2%)
WHERE_n_of_all_non_aggr_func	
0	797 (80%)
1	170 (17%)
2	31 (3.1%)
3	2 (0.2%)
GROUP_BY_n_of_columns	
0	250 (25%)
1	264 (26%)
2	221 (22%)
3	156 (16%)
4	70 (7.0%)
5	33 (3.3%)
6	5 (0.5%)
7	1 (0.1%)
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.36 / 2.50
HAVING_n_of_main_predicates___non_scalar_subquery	
0	429 (43%)

Query parameter or value	Descriptive statistics
1	190 (19%)
2	187 (19%)
3	194 (19%)
HAVING_n_of_main_predicates___scalar_subquery	
0	426 (43%)
1	201 (20%)
2	200 (20%)
3	173 (17%)
HAVING_n_of_subqueries___non_scalar_subquery	
0	429 (43%)
1	190 (19%)
2	187 (19%)
3	194 (19%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 24.0] 2.5 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 810,000, 177,075,366] 9,958,419 / 20,544,137
ORDER_BY_n_of_columns	
0	323 (32%)
1	443 (44%)
2	199 (20%)
3	24 (2.4%)
4	9 (0.9%)
5	2 (0.2%)
limit	[1, 510, 996] 514 / 285
offset	[0, 0, 1,000] 233 / 322