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

Scenario: 1-6jp_ujjp__g__h0-3ss_h0-3nss

Scale factor: 2 GB

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

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, 9, 183] 17 / 26
SELECT_n_of_non_aggr_func___LTRIM	[0.00, 0.00, 10.00] 0.63 / 1.10
SELECT_n_of_non_aggr_func___DAY	
0	871 (87%)
1	89 (8.9%)
2	25 (2.5%)
3	7 (0.7%)
4	3 (0.3%)
5	5 (0.5%)
SELECT_n_of_non_aggr_func___DOW	
0	853 (85%)
1	100 (10%)
2	25 (2.5%)
3	9 (0.9%)
4	8 (0.8%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
6	2 (0.2%)
7	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	[0.00, 0.00, 10.00] 0.67 / 1.11
SELECT_n_of_non_aggr_func___MONTH	[0.00, 0.00, 10.00] 0.37 / 1.01
SELECT_n_of_non_aggr_func___YEAR	
0	810 (81%)
1	133 (13%)
2	21 (2.1%)
3	20 (2.0%)
4	7 (0.7%)
5	4 (0.4%)
6	3 (0.3%)
7	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	903 (90%)
1	78 (7.8%)
2	17 (1.7%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	[0.00, 0.00, 9.00] 0.64 / 1.09
SELECT_n_of_non_aggr_func___ROUND	
0	912 (91%)
1	62 (6.2%)
2	19 (1.9%)
3	5 (0.5%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___RTRIM	[0.00, 0.00, 11.00] 0.61 / 1.15
SELECT_n_of_non_aggr_func___UPPER	[0.00, 0.00, 9.00] 0.62 / 1.13
SELECT_n_of_non_aggr_func___LOG	
0	848 (85%)
1	106 (11%)
2	28 (2.8%)
3	9 (0.9%)
4	5 (0.5%)
5	3 (0.3%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	791 (79%)
1	135 (14%)
2	29 (2.9%)
3	25 (2.5%)
4	11 (1.1%)
5	6 (0.6%)
8	2 (0.2%)

Query parameter or value	Descriptive statistics
9	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	908 (91%)
1	70 (7.0%)
2	18 (1.8%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	910 (91%)
1	73 (7.3%)
2	13 (1.3%)
3	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0, 3, 82] 5 / 9
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 11.00] 0.82 / 1.91
SELECT_n_of_aggr_func___AVG	
0	923 (92%)
1	45 (4.5%)
2	28 (2.8%)
3	2 (0.2%)
4	2 (0.2%)
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 15.00] 0.81 / 1.95
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 13.00] 0.85 / 1.93
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.82 / 1.81
SELECT_n_of_aggr_func___SUM	
0	929 (93%)
1	38 (3.8%)
2	30 (3.0%)
3	1 (0.1%)
4	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 3.00, 15.00] 3.52 / 3.10
FROM_n_of_join_paths	
1	156 (16%)
2	173 (17%)
3	171 (17%)
4	142 (14%)
5	174 (17%)
6	184 (18%)
FROM_n_of_super_joins___FULL	
0	542 (54%)
1	331 (33%)
2	102 (10%)
3	23 (2.3%)
4	1 (0.1%)
5	1 (0.1%)
FROM_n_of_super_joins___RIGHT	
0	527 (53%)

Query parameter or value	Descriptive statistics
1	306 (31%)
2	131 (13%)
3	32 (3.2%)
4	4 (0.4%)
FROM_n_of_super_joins___LEFT	
0	544 (54%)
1	293 (29%)
2	131 (13%)
3	28 (2.8%)
4	4 (0.4%)
FROM_n_of_joins___RIGHT	[0.00, 3.00, 13.00] 3.35 / 2.47
FROM_n_of_joins___INNER	[0.00, 3.00, 13.00] 3.38 / 2.52
FROM_n_of_processed_rows	[5, 20,455,437, 89,194,406] 24,906,330 / 19,441,996
WHERE_n_of_predicates	[0.00, 4.00, 13.00] 3.99 / 2.95
WHERE_n_of_attribs_of_type___character	
0	639 (64%)
1	250 (25%)
2	84 (8.4%)
3	25 (2.5%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type___character_varying	
0	338 (34%)
1	300 (30%)
2	189 (19%)
3	100 (10%)
4	53 (5.3%)
5	14 (1.4%)
6	6 (0.6%)
WHERE_n_of_attribs_of_type___integer	
0	355 (36%)
1	297 (30%)
2	189 (19%)
3	91 (9.1%)
4	49 (4.9%)
5	14 (1.4%)
6	4 (0.4%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	564 (56%)
1	285 (29%)
2	102 (10%)
3	39 (3.9%)
4	6 (0.6%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type___date	
0	751 (75%)

Query parameter or value	Descriptive statistics
1	205 (21%)
2	39 (3.9%)
3	4 (0.4%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	530 (53%)
1	290 (29%)
2	128 (13%)
3	42 (4.2%)
4	8 (0.8%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 2.00, 11.00] 2.31 / 2.45
WHERE_n_of_operators___between	
0	528 (53%)
1	295 (30%)
2	121 (12%)
3	41 (4.1%)
4	15 (1.5%)
WHERE_n_of_operators___greater_or_less	
0	326 (33%)
1	262 (26%)
2	190 (19%)
3	123 (12%)
4	53 (5.3%)
5	27 (2.7%)
6	14 (1.4%)
7	2 (0.2%)
8	3 (0.3%)
WHERE_n_of_operators___like	
0	760 (76%)
1	192 (19%)
2	44 (4.4%)
3	3 (0.3%)
4	1 (0.1%)
WHERE_n_of_operators___in	
0	493 (49%)
1	299 (30%)
2	143 (14%)
3	43 (4.3%)
4	15 (1.5%)
5	5 (0.5%)
6	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	961 (96%)
1	35 (3.5%)

Query parameter or value	Descriptive statistics
2	4 (0.4%)
WHERE_n_of_non_aggr_func___ABS	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func___SQRT	
0	879 (88%)
1	111 (11%)
2	10 (1.0%)
WHERE_n_of_non_aggr_func___ROUND	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	958 (96%)
1	40 (4.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___MONTH	
0	948 (95%)
1	52 (5.2%)
WHERE_n_of_non_aggr_func___FLOOR	
0	965 (97%)
1	33 (3.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___LOG	
0	913 (91%)
1	80 (8.0%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func___DOW	
0	957 (96%)
1	41 (4.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___YEAR	
0	959 (96%)
1	40 (4.0%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	611 (61%)
1	255 (26%)
2	102 (10%)
3	24 (2.4%)
4	6 (0.6%)
5	1 (0.1%)
7	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 13.00] 2.97 / 2.84
HAVING_n_of_main_predicates	[0.00, 3.50, 9.00] 3.28 / 2.60
HAVING_n_of_main_predicates___non_scalar_subquery	

Query parameter or value	Descriptive statistics
0	458 (46%)
1	182 (18%)
2	181 (18%)
3	179 (18%)
HAVING_n_of_main_predicates___scalar_subquery	
0	452 (45%)
1	184 (18%)
2	184 (18%)
3	180 (18%)
HAVING_n_of_subqueries___non_scalar_subquery	
0	458 (46%)
1	182 (18%)
2	181 (18%)
3	179 (18%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 25.0] 2.5 / 3.9
HAVING_n_of_processed_rows_by_subqueries	[0, 27,559,075, 1,014,465,117] 88,358,260 / 143,956,817
ORDER_BY_n_of_columns	
0	306 (31%)
1	302 (30%)
2	247 (25%)
3	103 (10%)
4	24 (2.4%)
5	8 (0.8%)
6	6 (0.6%)
7	3 (0.3%)
8	1 (0.1%)
limit	[1, 474, 1,000] 484 / 290
offset	[0, 0, 1,000] 249 / 320