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

Scenario: 1-3jp_ujjp__g__h0ss_h0nss

Scale factor: 5 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, 85] 10 / 12
SELECT_n_of_non_aggr_func___UPPER	
0	702 (70%)
1	222 (22%)
2	53 (5.3%)
3	13 (1.3%)
4	9 (0.9%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	718 (72%)
1	219 (22%)
2	47 (4.7%)
3	11 (1.1%)
4	4 (0.4%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func___LTRIM	
0	689 (69%)
1	236 (24%)
2	50 (5.0%)
3	16 (1.6%)
4	3 (0.3%)
5	2 (0.2%)
6	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	685 (69%)
1	254 (25%)
10	1 (0.1%)
2	47 (4.7%)
3	8 (0.8%)
4	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	946 (95%)
1	47 (4.7%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func___TRUNC	
0	948 (95%)
1	49 (4.9%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___RTRIM	
0	725 (73%)
1	221 (22%)
2	38 (3.8%)
3	13 (1.3%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func___SQRT	
0	866 (87%)
1	104 (10%)
2	25 (2.5%)
3	4 (0.4%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	903 (90%)
1	85 (8.5%)
2	10 (1.0%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___MONTH	
0	840 (84%)
1	127 (13%)
2	24 (2.4%)

Query parameter or value	Descriptive statistics
3	6 (0.6%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	877 (88%)
1	98 (9.8%)
2	17 (1.7%)
3	5 (0.5%)
4	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	946 (95%)
1	46 (4.6%)
2	8 (0.8%)
SELECT_n_of_non_aggr_func___DOW	
0	899 (90%)
1	79 (7.9%)
2	19 (1.9%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___LOG	
0	916 (92%)
1	68 (6.8%)
2	12 (1.2%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	946 (95%)
1	49 (4.9%)
2	5 (0.5%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 38.00] 3.08 / 4.13
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 9.00] 0.70 / 1.48
SELECT_n_of_aggr_func___AVG	
0	939 (94%)
1	42 (4.2%)
2	17 (1.7%)
3	2 (0.2%)
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 10.00] 0.65 / 1.49
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 9.00] 0.65 / 1.41
SELECT_n_of_aggr_func___SUM	
0	949 (95%)
1	39 (3.9%)
2	10 (1.0%)
3	2 (0.2%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.60 / 1.43
SELECT_n_of_all_aggr_func	[0.00, 2.00, 10.00] 2.74 / 2.26
FROM_n_of_join_paths	

Query parameter or value	Descriptive statistics
1	327 (33%)
2	339 (34%)
3	334 (33%)
FROM_n_of_super_joins___RIGHT	
0	767 (77%)
1	211 (21%)
2	22 (2.2%)
FROM_n_of_super_joins___LEFT	
0	767 (77%)
1	212 (21%)
2	21 (2.1%)
FROM_n_of_super_joins___FULL	
0	767 (77%)
1	206 (21%)
2	27 (2.7%)
FROM_n_of_joins___RIGHT	
0	227 (23%)
1	267 (27%)
2	245 (25%)
3	132 (13%)
4	89 (8.9%)
5	28 (2.8%)
6	9 (0.9%)
7	3 (0.3%)
FROM_n_of_joins___INNER	
0	235 (24%)
1	274 (27%)
2	221 (22%)
3	135 (14%)
4	71 (7.1%)
5	44 (4.4%)
6	14 (1.4%)
7	5 (0.5%)
8	1 (0.1%)
FROM_n_of_processed_rows	[5, 34,050,148, 114,749,827] 32,218,879 / 29,613,611
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.77 / 2.27
WHERE_n_of_attribs_of_type___integer	
0	454 (45%)
1	312 (31%)
2	153 (15%)
3	65 (6.5%)
4	14 (1.4%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type___numeric	
0	681 (68%)
1	220 (22%)

Query parameter or value	Descriptive statistics
2	81 (8.1%)
3	14 (1.4%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type___character_varying	
0	447 (45%)
1	320 (32%)
2	137 (14%)
3	78 (7.8%)
4	15 (1.5%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type___character	
0	726 (73%)
1	208 (21%)
2	53 (5.3%)
3	11 (1.1%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type___date	
0	840 (84%)
1	135 (14%)
2	24 (2.4%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	615 (62%)
1	277 (28%)
2	90 (9.0%)
3	16 (1.6%)
4	2 (0.2%)
WHERE_n_of_connect_OR	
0	467 (47%)
1	184 (18%)
2	120 (12%)
3	89 (8.9%)
4	68 (6.8%)
5	43 (4.3%)
6	22 (2.2%)
7	6 (0.6%)
8	1 (0.1%)
WHERE_n_of_operators___between	
0	632 (63%)
1	256 (26%)
2	86 (8.6%)
3	18 (1.8%)
4	6 (0.6%)
5	2 (0.2%)
WHERE_n_of_operators___greater_or_less	
0	414 (41%)

Query parameter or value	Descriptive statistics
1	291 (29%)
2	169 (17%)
3	91 (9.1%)
4	26 (2.6%)
5	6 (0.6%)
6	3 (0.3%)
WHERE_n_of_operators___in	
0	655 (66%)
1	234 (23%)
2	92 (9.2%)
3	19 (1.9%)
WHERE_n_of_operators___like	
0	837 (84%)
1	148 (15%)
2	13 (1.3%)
3	1 (0.1%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	933 (93%)
1	65 (6.5%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___ROUND	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func___YEAR	
0	972 (97%)
1	27 (2.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___SQRT	
0	898 (90%)
1	97 (9.7%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___MONTH	
0	980 (98%)
1	20 (2.0%)
WHERE_n_of_non_aggr_func___DAY	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func___ABS	
0	976 (98%)
1	23 (2.3%)
2	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func___DOW	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func___FLOOR	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_all_non_aggr_func	
0	701 (70%)
1	228 (23%)
2	51 (5.1%)
3	18 (1.8%)
4	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.38 / 2.09
HAVING_n_of_main_predicates	
0	434 (43%)
1	191 (19%)
2	171 (17%)
3	204 (20%)
ORDER_BY_n_of_columns	
0	326 (33%)
1	359 (36%)
2	234 (23%)
3	60 (6.0%)
4	14 (1.4%)
5	3 (0.3%)
6	2 (0.2%)
7	1 (0.1%)
8	1 (0.1%)
limit	[1, 494, 1,000] 495 / 292
offset	[0, 40, 998] 262 / 325