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

Scenario: 1-3jp_ujjp__g__h0ss_h0nss

Scale factor: 1000 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, 101] 10 / 12
SELECT_n_of_non_aggr_func___LOG	
0	878 (88%)
1	106 (11%)
2	13 (1.3%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func___LTRIM	
0	692 (69%)
1	231 (23%)
2	53 (5.3%)
3	17 (1.7%)
4	3 (0.3%)
5	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	

Query parameter or value	Descriptive statistics
0	659 (66%)
1	280 (28%)
2	42 (4.2%)
3	10 (1.0%)
4	5 (0.5%)
5	2 (0.2%)
6	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	838 (84%)
1	135 (14%)
2	18 (1.8%)
3	5 (0.5%)
4	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	683 (68%)
1	241 (24%)
2	60 (6.0%)
3	8 (0.8%)
4	4 (0.4%)
5	2 (0.2%)
6	1 (0.1%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	682 (68%)
1	243 (24%)
2	47 (4.7%)
3	18 (1.8%)
4	5 (0.5%)
5	2 (0.2%)
6	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	684 (68%)
1	238 (24%)
2	61 (6.1%)
3	14 (1.4%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	895 (90%)
1	85 (8.5%)
2	15 (1.5%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func___MONTH	

Query parameter or value	Descriptive statistics
0	854 (85%)
1	102 (10%)
2	29 (2.9%)
3	11 (1.1%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___YEAR	
0	875 (88%)
1	99 (9.9%)
2	19 (1.9%)
3	5 (0.5%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	948 (95%)
1	51 (5.1%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	944 (94%)
1	53 (5.3%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___DAY	
0	918 (92%)
1	72 (7.2%)
2	6 (0.6%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func___FLOOR	
0	941 (94%)
1	53 (5.3%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	952 (95%)
1	45 (4.5%)
2	3 (0.3%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 46.00] 3.31 / 4.38
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 9.00] 0.65 / 1.44
SELECT_n_of_aggr_func___MIN	
0	751 (75%)
1	90 (9.0%)
2	57 (5.7%)
3	36 (3.6%)
4	23 (2.3%)
5	19 (1.9%)
6	11 (1.1%)
7	11 (1.1%)
8	2 (0.2%)

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func___COUNT_DISTINCT	
0	739 (74%)
1	99 (9.9%)
11	1 (0.1%)
2	62 (6.2%)
3	34 (3.4%)
4	29 (2.9%)
5	17 (1.7%)
6	14 (1.4%)
7	5 (0.5%)
SELECT_n_of_aggr_func___MAX	
0	781 (78%)
1	77 (7.7%)
2	44 (4.4%)
3	46 (4.6%)
4	22 (2.2%)
5	17 (1.7%)
6	6 (0.6%)
7	2 (0.2%)
9	5 (0.5%)
SELECT_n_of_aggr_func___SUM	
0	945 (95%)
1	41 (4.1%)
2	11 (1.1%)
3	3 (0.3%)
SELECT_n_of_aggr_func___AVG	
0	944 (94%)
1	48 (4.8%)
2	6 (0.6%)
3	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.67 / 2.23
FROM_n_of_join_paths	
1	318 (32%)
2	344 (34%)
3	338 (34%)
FROM_n_of_super_joins___FULL	
0	763 (76%)
1	207 (21%)
2	30 (3.0%)
FROM_n_of_super_joins___LEFT	
0	776 (78%)
1	203 (20%)
2	21 (2.1%)
FROM_n_of_super_joins___RIGHT	
0	777 (78%)
1	206 (21%)

Query parameter or value	Descriptive statistics
2	17 (1.7%)
FROM_n_of_joins___INNER	
0	212 (21%)
1	261 (26%)
2	232 (23%)
3	154 (15%)
4	79 (7.9%)
5	35 (3.5%)
6	23 (2.3%)
7	3 (0.3%)
8	1 (0.1%)
FROM_n_of_joins___RIGHT	
0	230 (23%)
1	266 (27%)
2	222 (22%)
3	141 (14%)
4	75 (7.5%)
5	48 (4.8%)
6	15 (1.5%)
7	3 (0.3%)
FROM_n_of_processed_rows	[5, 6,819,989,759, 22,949,969,212] 6,744,541,358 / 5,938,482,380
WHERE_n_of_predicates	[0.00, 3.00, 10.00] 2.96 / 2.27
WHERE_n_of_attribs_of_type___character	
0	725 (73%)
1	209 (21%)
2	59 (5.9%)
3	6 (0.6%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___character_varying	
0	421 (42%)
1	311 (31%)
2	161 (16%)
3	74 (7.4%)
4	23 (2.3%)
5	9 (0.9%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type___integer	
0	423 (42%)
1	340 (34%)
2	155 (16%)
3	65 (6.5%)
4	13 (1.3%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type___numeric	
0	633 (63%)

Query parameter or value	Descriptive statistics
1	258 (26%)
2	88 (8.8%)
3	21 (2.1%)
WHERE_n_of_attribs_of_type___date	
0	833 (83%)
1	138 (14%)
2	25 (2.5%)
3	3 (0.3%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	603 (60%)
1	292 (29%)
2	88 (8.8%)
3	17 (1.7%)
WHERE_n_of_connect_OR	
0	422 (42%)
1	174 (17%)
2	148 (15%)
3	103 (10%)
4	74 (7.4%)
5	43 (4.3%)
6	22 (2.2%)
7	11 (1.1%)
8	3 (0.3%)
WHERE_n_of_operators___greater_or_less	
0	390 (39%)
1	292 (29%)
2	183 (18%)
3	95 (9.5%)
4	31 (3.1%)
5	5 (0.5%)
6	3 (0.3%)
7	1 (0.1%)
WHERE_n_of_operators___in	
0	608 (61%)
1	288 (29%)
2	69 (6.9%)
3	29 (2.9%)
4	6 (0.6%)
WHERE_n_of_operators___like	
0	832 (83%)
1	148 (15%)
2	17 (1.7%)
3	3 (0.3%)
WHERE_n_of_operators___between	
0	578 (58%)

Query parameter or value	Descriptive statistics
1	292 (29%)
2	101 (10%)
3	19 (1.9%)
4	6 (0.6%)
5	4 (0.4%)
WHERE_n_of_non_aggr_func___DOW	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func___SQRT	
0	888 (89%)
1	106 (11%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func___ABS	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func___DAY	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func___LOG	
0	934 (93%)
1	66 (6.6%)
WHERE_n_of_non_aggr_func___TRUNC	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func___MONTH	
0	983 (98%)
1	16 (1.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_non_aggr_func___ROUND	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	965 (97%)
1	32 (3.2%)
2	3 (0.3%)
WHERE_n_of_all_non_aggr_func	
0	661 (66%)
1	262 (26%)
2	64 (6.4%)
3	11 (1.1%)
4	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.51 / 2.17

Query parameter or value	Descriptive statistics
HAVING_n_of_main_predicates	
0	407 (41%)
1	209 (21%)
2	203 (20%)
3	181 (18%)
ORDER_BY_n_of_columns	
0	296 (30%)
1	381 (38%)
2	223 (22%)
3	75 (7.5%)
4	16 (1.6%)
5	6 (0.6%)
6	2 (0.2%)
7	1 (0.1%)
limit	[2, 482, 1,000] 495 / 289
offset	[0, 17, 999] 253 / 323