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

Scenario: 1-3jp_ujjp__g__h0ss_h0nss

Scale factor: 50 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, 112] 10 / 14
SELECT_n_of_non_aggr_func___DOW	
0	892 (89%)
1	83 (8.3%)
2	20 (2.0%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func___LTRIM	
0	686 (69%)
1	226 (23%)
2	65 (6.5%)
3	17 (1.7%)
4	5 (0.5%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	701 (70%)

Query parameter or value	Descriptive statistics
1	228 (23%)
2	48 (4.8%)
3	15 (1.5%)
4	6 (0.6%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	689 (69%)
1	234 (23%)
2	53 (5.3%)
3	16 (1.6%)
4	7 (0.7%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	868 (87%)
1	101 (10%)
2	22 (2.2%)
3	3 (0.3%)
4	3 (0.3%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func___FLOOR	
0	944 (94%)
1	53 (5.3%)
2	2 (0.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___LOWER	
0	672 (67%)
1	250 (25%)
2	65 (6.5%)
3	2 (0.2%)
4	9 (0.9%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	689 (69%)
1	229 (23%)
2	58 (5.8%)
3	17 (1.7%)
4	5 (0.5%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	916 (92%)
1	67 (6.7%)
2	15 (1.5%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___YEAR	

Query parameter or value	Descriptive statistics
0	869 (87%)
1	96 (9.6%)
2	23 (2.3%)
3	11 (1.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	843 (84%)
1	123 (12%)
2	16 (1.6%)
3	12 (1.2%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___ROUND	
0	936 (94%)
1	56 (5.6%)
2	8 (0.8%)
SELECT_n_of_non_aggr_func___LOG	
0	904 (90%)
1	83 (8.3%)
2	11 (1.1%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	949 (95%)
1	47 (4.7%)
2	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	957 (96%)
1	40 (4.0%)
2	1 (0.1%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 36.0] 3.3 / 4.4
SELECT_n_of_aggr_func___COUNT	
0	766 (77%)
1	83 (8.3%)
2	61 (6.1%)
3	38 (3.8%)
4	20 (2.0%)
5	19 (1.9%)
6	10 (1.0%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 9.00] 0.61 / 1.39
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 9.00] 0.65 / 1.44

Query parameter or value	Descriptive statistics
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.60 / 1.39
SELECT_n_of_aggr_func___AVG	
0	945 (95%)
1	46 (4.6%)
2	3 (0.3%)
3	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_aggr_func___SUM	
0	941 (94%)
1	41 (4.1%)
2	13 (1.3%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.59 / 2.21
FROM_n_of_join_paths	
1	345 (35%)
2	351 (35%)
3	304 (30%)
FROM_n_of_super_joins___LEFT	
0	779 (78%)
1	199 (20%)
2	22 (2.2%)
FROM_n_of_super_joins___FULL	
0	762 (76%)
1	219 (22%)
2	19 (1.9%)
FROM_n_of_super_joins___RIGHT	
0	783 (78%)
1	198 (20%)
2	19 (1.9%)
FROM_n_of_joins___INNER	
0	252 (25%)
1	274 (27%)
2	203 (20%)
3	144 (14%)
4	79 (7.9%)
5	28 (2.8%)
6	15 (1.5%)
7	5 (0.5%)
FROM_n_of_joins___RIGHT	
0	239 (24%)
1	287 (29%)
2	208 (21%)
3	140 (14%)
4	71 (7.1%)

Query parameter or value	Descriptive statistics
5	34 (3.4%)
6	13 (1.3%)
7	7 (0.7%)
8	1 (0.1%)
FROM_n_of_processed_rows	[5, 340,491,940, 1,147,549,276] 309,422,312 / 297,729,612
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.73 / 2.17
WHERE_n_of_attribs_of_type___character_varying	
0	455 (46%)
1	306 (31%)
2	160 (16%)
3	66 (6.6%)
4	12 (1.2%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___integer	
0	463 (46%)
1	305 (31%)
2	160 (16%)
3	51 (5.1%)
4	18 (1.8%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type___character	
0	727 (73%)
1	210 (21%)
2	51 (5.1%)
3	11 (1.1%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	659 (66%)
1	253 (25%)
2	72 (7.2%)
3	11 (1.1%)
4	5 (0.5%)
WHERE_n_of_attribs_of_type___date	
0	845 (85%)
1	124 (12%)
2	26 (2.6%)
3	5 (0.5%)
WHERE_n_of_pkey_attribs	
0	618 (62%)
1	286 (29%)
2	83 (8.3%)
3	12 (1.2%)
4	1 (0.1%)
WHERE_n_of_connect_OR	
0	461 (46%)

Query parameter or value	Descriptive statistics
1	192 (19%)
2	131 (13%)
3	97 (9.7%)
4	58 (5.8%)
5	35 (3.5%)
6	18 (1.8%)
7	6 (0.6%)
8	2 (0.2%)
WHERE_n_of_operators___greater_or_less	
0	414 (41%)
1	298 (30%)
2	179 (18%)
3	67 (6.7%)
4	30 (3.0%)
5	11 (1.1%)
6	1 (0.1%)
WHERE_n_of_operators___in	
0	611 (61%)
1	294 (29%)
2	74 (7.4%)
3	17 (1.7%)
4	4 (0.4%)
WHERE_n_of_operators___between	
0	631 (63%)
1	266 (27%)
2	75 (7.5%)
3	26 (2.6%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_operators___like	
0	840 (84%)
1	148 (15%)
2	11 (1.1%)
4	1 (0.1%)
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	971 (97%)
1	28 (2.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___LOG	
0	954 (95%)
1	43 (4.3%)
2	3 (0.3%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func___MONTH	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func___YEAR	
0	968 (97%)
1	30 (3.0%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	930 (93%)
1	67 (6.7%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___DOW	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func___ABS	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	979 (98%)
1	21 (2.1%)
WHERE_n_of_all_non_aggr_func	
0	737 (74%)
1	200 (20%)
2	49 (4.9%)
3	13 (1.3%)
4	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.26 / 2.05
HAVING_n_of_main_predicates	
0	426 (43%)
1	187 (19%)
2	181 (18%)
3	206 (21%)
ORDER_BY_n_of_columns	
0	310 (31%)
1	369 (37%)
2	232 (23%)
3	57 (5.7%)
4	20 (2.0%)
5	9 (0.9%)
6	1 (0.1%)
7	2 (0.2%)

Query parameter or value	Descriptive statistics
limit	[2, 512, 999] 502 / 290
offset	[0, 0, 1,000] 241 / 314