

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

Scenario: 1-6jp_rjpp__g__h0ss_h0nss

Scale factor: 500 GB

The 1000-query pack was generated on: 2023-11-12

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, 197] 17 / 27
SELECT_n_of_non_aggr_func__UPPER	[0.00, 0.00, 12.00] 0.62 / 1.13
SELECT_n_of_non_aggr_func__LOWER	[0.00, 0.00, 11.00] 0.61 / 1.09
SELECT_n_of_non_aggr_func__LTRIM	
0	589 (59%)
1	278 (28%)
2	84 (8.4%)
3	19 (1.9%)
4	16 (1.6%)
5	8 (0.8%)
6	3 (0.3%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 12.00] 0.64 / 1.12
SELECT_n_of_non_aggr_func__DAY	

Query parameter or value	Descriptive statistics
0	879 (88%)
1	90 (9.0%)
2	13 (1.3%)
3	8 (0.8%)
4	6 (0.6%)
5	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__LOG	
0	891 (89%)
1	86 (8.6%)
2	13 (1.3%)
3	4 (0.4%)
4	2 (0.2%)
5	2 (0.2%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 10.00] 0.36 / 1.05
SELECT_n_of_non_aggr_func__RTRIM	[0.00, 0.00, 12.00] 0.57 / 1.10
SELECT_n_of_non_aggr_func__YEAR	
0	824 (82%)
1	116 (12%)
2	26 (2.6%)
3	15 (1.5%)
4	10 (1.0%)
5	4 (0.4%)
6	1 (0.1%)
7	3 (0.3%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	926 (93%)
1	61 (6.1%)
2	9 (0.9%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	802 (80%)
1	143 (14%)
2	17 (1.7%)
3	23 (2.3%)
4	7 (0.7%)
5	5 (0.5%)
6	1 (0.1%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	930 (93%)
1	60 (6.0%)
2	5 (0.5%)

Query parameter or value	Descriptive statistics
3	3 (0.3%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	911 (91%)
1	69 (6.9%)
2	15 (1.5%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__DOW	
0	851 (85%)
1	103 (10%)
2	24 (2.4%)
3	11 (1.1%)
4	6 (0.6%)
5	4 (0.4%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	922 (92%)
1	68 (6.8%)
2	8 (0.8%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 75.0] 5.0 / 8.6
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 11.00] 0.82 / 1.80
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 11.00] 0.83 / 1.85
SELECT_n_of_aggr_func__AVG	
0	933 (93%)
1	39 (3.9%)
2	15 (1.5%)
3	9 (0.9%)
4	4 (0.4%)
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 13.00] 0.86 / 2.06
SELECT_n_of_aggr_func__SUM	
0	934 (93%)
1	39 (3.9%)
2	23 (2.3%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.78 / 1.84
SELECT_n_of_all_aggr_func	[0.0, 3.0, 13.0] 3.5 / 3.1
FROM_n_of_join_paths	
1	153 (15%)
2	170 (17%)
3	191 (19%)
4	151 (15%)
5	171 (17%)

Query parameter or value	Descriptive statistics
6	164 (16%)
FROM_n_of_super_joins___FULL	
0	539 (54%)
1	317 (32%)
2	113 (11%)
3	25 (2.5%)
4	6 (0.6%)
FROM_n_of_super_joins___LEFT	
0	541 (54%)
1	322 (32%)
2	113 (11%)
3	19 (1.9%)
4	5 (0.5%)
FROM_n_of_super_joins___RIGHT	
0	545 (55%)
1	327 (33%)
2	100 (10%)
3	25 (2.5%)
4	3 (0.3%)
FROM_n_of_joins___INNER	[0.00, 3.00, 12.00] 3.29 / 2.44
FROM_n_of_joins___RIGHT	[0.00, 3.00, 12.00] 3.35 / 2.50
FROM_n_of_processed_rows	[5, 4,650,028,272, 20,885,169,507] 5,957,412,026 / 4,845,105,682
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.90 / 2.97
WHERE_n_of_attribs_of_type___character_varying	
0	362 (36%)
1	292 (29%)
2	198 (20%)
3	85 (8.5%)
4	44 (4.4%)
5	16 (1.6%)
6	2 (0.2%)
8	1 (0.1%)
WHERE_n_of_attribs_of_type___integer	
0	333 (33%)
1	317 (32%)
2	197 (20%)
3	113 (11%)
4	23 (2.3%)
5	13 (1.3%)
6	4 (0.4%)
WHERE_n_of_attribs_of_type___numeric	
0	562 (56%)
1	284 (28%)
2	99 (9.9%)
3	43 (4.3%)

Query parameter or value	Descriptive statistics
4	11 (1.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	626 (63%)
1	259 (26%)
2	83 (8.3%)
3	24 (2.4%)
4	7 (0.7%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	786 (79%)
1	170 (17%)
2	38 (3.8%)
3	6 (0.6%)
WHERE_n_of_pkey_attribs	
0	487 (49%)
1	355 (36%)
2	121 (12%)
3	28 (2.8%)
4	7 (0.7%)
5	2 (0.2%)
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.28 / 2.48
WHERE_n_of_operators__between	
0	498 (50%)
1	330 (33%)
2	117 (12%)
3	44 (4.4%)
4	10 (1.0%)
5	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	333 (33%)
1	270 (27%)
2	179 (18%)
3	119 (12%)
4	51 (5.1%)
5	32 (3.2%)
6	15 (1.5%)
7	1 (0.1%)
WHERE_n_of_operators__in	
0	526 (53%)
1	290 (29%)
2	129 (13%)
3	48 (4.8%)
4	6 (0.6%)
5	1 (0.1%)
WHERE_n_of_operators__like	

Query parameter or value	Descriptive statistics
0	771 (77%)
1	192 (19%)
2	28 (2.8%)
3	8 (0.8%)
5	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	952 (95%)
1	46 (4.6%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__LOG	
0	925 (93%)
1	71 (7.1%)
2	3 (0.3%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	963 (96%)
1	37 (3.7%)
WHERE_n_of_non_aggr_func__FLOOR	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	967 (97%)
1	32 (3.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	893 (89%)
1	98 (9.8%)
2	9 (0.9%)
WHERE_n_of_non_aggr_func__DAY	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func__ROUND	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func__MONTH	
0	965 (97%)
1	33 (3.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__TRUNC	
0	961 (96%)
1	37 (3.7%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	643 (64%)
1	242 (24%)

Query parameter or value	Descriptive statistics
2	90 (9.0%)
3	21 (2.1%)
4	4 (0.4%)
GROUP_BY_n_of_columns	[0.00, 3.00, 12.00] 3.15 / 2.86
HAVING_n_of_main_predicates	
0	419 (42%)
1	193 (19%)
2	188 (19%)
3	200 (20%)
ORDER_BY_n_of_columns	
0	287 (29%)
1	336 (34%)
2	244 (24%)
3	97 (9.7%)
4	20 (2.0%)
5	10 (1.0%)
6	2 (0.2%)
7	2 (0.2%)
8	2 (0.2%)
limit	[1, 486, 1,000] 494 / 295
offset	[0, 22, 1,000] 255 / 323