

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

Scenario: 1-3jp_rjpp__g__h0-3ss_h0-3nss

Scale factor: 500 GB

The 1000-query pack was generated on: 2023-11-06

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, 117] 10 / 13
SELECT_n_of_non_aggr_func__FLOOR	
0	953 (95%)
1	44 (4.4%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func__ROUND	
0	957 (96%)
1	42 (4.2%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	685 (69%)
1	251 (25%)
2	47 (4.7%)
3	12 (1.2%)
4	5 (0.5%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func___SQRT	
0	877 (88%)
1	108 (11%)
2	8 (0.8%)
3	3 (0.3%)
4	4 (0.4%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	695 (70%)
1	233 (23%)
2	56 (5.6%)
3	11 (1.1%)
4	3 (0.3%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	682 (68%)
1	255 (26%)
2	49 (4.9%)
3	8 (0.8%)
4	2 (0.2%)
5	4 (0.4%)
SELECT_n_of_non_aggr_func___LOWER	
0	722 (72%)
1	217 (22%)
2	54 (5.4%)
3	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	836 (84%)
1	125 (13%)
2	31 (3.1%)
3	5 (0.5%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func___YEAR	
0	877 (88%)
1	102 (10%)
2	15 (1.5%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___ABS	
0	949 (95%)
1	50 (5.0%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	920 (92%)
1	69 (6.9%)

Query parameter or value	Descriptive statistics
2	10 (1.0%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	932 (93%)
1	59 (5.9%)
2	7 (0.7%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	941 (94%)
1	57 (5.7%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func__UPPER	
0	697 (70%)
1	234 (23%)
2	56 (5.6%)
3	12 (1.2%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	880 (88%)
1	104 (10%)
2	7 (0.7%)
3	7 (0.7%)
4	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 32.00] 2.98 / 3.58
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 11.00] 0.65 / 1.49
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.71 / 1.56
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.61 / 1.46
SELECT_n_of_aggr_func__SUM	
0	929 (93%)
1	53 (5.3%)
2	13 (1.3%)
3	5 (0.5%)
SELECT_n_of_aggr_func__AVG	
0	945 (95%)
1	42 (4.2%)
2	10 (1.0%)
3	3 (0.3%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 15.00] 0.65 / 1.51
SELECT_n_of_all_aggr_func	[0.00, 2.00, 15.00] 2.79 / 2.37
FROM_n_of_join_paths	
1	348 (35%)
2	326 (33%)
3	326 (33%)
FROM_n_of_super_joins__RIGHT	
0	783 (78%)
1	188 (19%)

Query parameter or value	Descriptive statistics
2	29 (2.9%)
FROM_n_of_super_joins__LEFT	
0	763 (76%)
1	218 (22%)
2	19 (1.9%)
FROM_n_of_super_joins__FULL	
0	793 (79%)
1	190 (19%)
2	17 (1.7%)
FROM_n_of_joins__INNER	
0	253 (25%)
1	253 (25%)
2	229 (23%)
3	136 (14%)
4	79 (7.9%)
5	25 (2.5%)
6	18 (1.8%)
7	6 (0.6%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	241 (24%)
1	258 (26%)
2	227 (23%)
3	137 (14%)
4	84 (8.4%)
5	31 (3.1%)
6	20 (2.0%)
8	1 (0.1%)
9	1 (0.1%)
FROM_n_of_processed_rows	[5, 3,405,028,267, 11,475,084,811] 3,152,362,004 / 2,906,878,311
WHERE_n_of_predicates	[0.00, 2.00, 9.00] 2.78 / 2.19
WHERE_n_of_attribs_of_type__character_varying	
0	430 (43%)
1	333 (33%)
2	158 (16%)
3	52 (5.2%)
4	24 (2.4%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type__date	
0	833 (83%)
1	141 (14%)
2	24 (2.4%)
3	1 (0.1%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	

Query parameter or value	Descriptive statistics
0	472 (47%)
1	306 (31%)
2	162 (16%)
3	41 (4.1%)
4	16 (1.6%)
5	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	658 (66%)
1	252 (25%)
2	77 (7.7%)
3	11 (1.1%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__character	
0	708 (71%)
1	213 (21%)
2	64 (6.4%)
3	12 (1.2%)
4	3 (0.3%)
WHERE_n_of_pkey_attribs	
0	642 (64%)
1	260 (26%)
2	82 (8.2%)
3	14 (1.4%)
4	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_connect_OR	
0	466 (47%)
1	190 (19%)
2	120 (12%)
3	102 (10%)
4	58 (5.8%)
5	44 (4.4%)
6	16 (1.6%)
7	4 (0.4%)
WHERE_n_of_operators__greater_or_less	
0	423 (42%)
1	288 (29%)
2	171 (17%)
3	80 (8.0%)
4	28 (2.8%)
5	10 (1.0%)
WHERE_n_of_operators__in	
0	635 (64%)
1	251 (25%)

Query parameter or value	Descriptive statistics
2	84 (8.4%)
3	23 (2.3%)
4	7 (0.7%)
WHERE_n_of_operators__like	
0	839 (84%)
1	147 (15%)
2	13 (1.3%)
3	1 (0.1%)
WHERE_n_of_operators__between	
0	600 (60%)
1	284 (28%)
2	84 (8.4%)
3	28 (2.8%)
4	3 (0.3%)
5	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	977 (98%)
1	22 (2.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__LOG	
0	950 (95%)
1	49 (4.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func__SQRT	
0	923 (92%)
1	72 (7.2%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func__ROUND	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func__YEAR	
0	969 (97%)
1	30 (3.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func__DAY	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func__FLOOR	
0	980 (98%)
1	20 (2.0%)

