

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

Scenario: 1-3jp_ujpp__g__h0-3ss_h0-3nss

Scale factor: 200 GB

The 1000-query pack was generated on: 2023-11-10

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, 94] 10 / 13
SELECT_n_of_non_aggr_func__LOG	
0	881 (88%)
1	101 (10%)
2	13 (1.3%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__LOWER	
0	673 (67%)
1	242 (24%)
2	73 (7.3%)
3	7 (0.7%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	

Query parameter or value	Descriptive statistics
0	712 (71%)
1	225 (23%)
2	43 (4.3%)
3	10 (1.0%)
4	5 (0.5%)
5	2 (0.2%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	947 (95%)
1	50 (5.0%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func___UPPER	
0	704 (70%)
1	219 (22%)
2	65 (6.5%)
3	11 (1.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	700 (70%)
1	227 (23%)
2	56 (5.6%)
3	15 (1.5%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	696 (70%)
1	227 (23%)
2	62 (6.2%)
3	10 (1.0%)
4	5 (0.5%)
SELECT_n_of_non_aggr_func___SQRT	
0	863 (86%)
1	113 (11%)
2	13 (1.3%)
3	8 (0.8%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func___DOW	
0	894 (89%)
1	86 (8.6%)
2	15 (1.5%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func___FLOOR	
0	955 (96%)
1	39 (3.9%)
2	4 (0.4%)

Query parameter or value	Descriptive statistics
3	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	881 (88%)
1	87 (8.7%)
2	24 (2.4%)
3	4 (0.4%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	858 (86%)
1	107 (11%)
2	21 (2.1%)
3	10 (1.0%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__ROUND	
0	948 (95%)
1	49 (4.9%)
2	3 (0.3%)
SELECT_n_of_non_aggr_func__DAY	
0	925 (93%)
1	57 (5.7%)
2	17 (1.7%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	956 (96%)
1	41 (4.1%)
2	3 (0.3%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 38.0] 3.1 / 4.3
SELECT_n_of_aggr_func__MIN	
0	742 (74%)
1	102 (10%)
2	62 (6.2%)
3	39 (3.9%)
4	25 (2.5%)
5	13 (1.3%)
6	11 (1.1%)
7	5 (0.5%)
8	1 (0.1%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.58 / 1.33
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 11.00] 0.65 / 1.46
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.55 / 1.32
SELECT_n_of_aggr_func__AVG	
0	955 (96%)
1	31 (3.1%)

Query parameter or value	Descriptive statistics
2	12 (1.2%)
3	2 (0.2%)
SELECT_n_of_aggr_func__SUM	
0	950 (95%)
1	34 (3.4%)
2	12 (1.2%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.52 / 2.16
FROM_n_of_join_paths	
1	356 (36%)
2	351 (35%)
3	293 (29%)
FROM_n_of_super_joins__LEFT	
0	786 (79%)
1	199 (20%)
2	15 (1.5%)
FROM_n_of_super_joins__FULL	
0	777 (78%)
1	202 (20%)
2	21 (2.1%)
FROM_n_of_super_joins__RIGHT	
0	796 (80%)
1	180 (18%)
2	24 (2.4%)
FROM_n_of_joins__INNER	[0.00, 2.00, 9.00] 1.76 / 1.53
FROM_n_of_joins__RIGHT	
0	266 (27%)
1	263 (26%)
2	198 (20%)
3	144 (14%)
4	72 (7.2%)
5	40 (4.0%)
6	12 (1.2%)
7	2 (0.2%)
8	3 (0.3%)
FROM_n_of_processed_rows	[5, 1,364,018,464, 4,590,055,377] 1,293,096,870 / 1,172,433,837
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.74 / 2.19
WHERE_n_of_attris_of_type__character_varying	
0	427 (43%)
1	344 (34%)
2	150 (15%)
3	61 (6.1%)
4	14 (1.4%)
5	3 (0.3%)

Query parameter or value	Descriptive statistics
6	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	729 (73%)
1	213 (21%)
2	50 (5.0%)
3	7 (0.7%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	463 (46%)
1	321 (32%)
2	135 (14%)
3	58 (5.8%)
4	21 (2.1%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	663 (66%)
1	254 (25%)
2	66 (6.6%)
3	13 (1.3%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type__date	
0	835 (84%)
1	134 (13%)
2	30 (3.0%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	615 (62%)
1	290 (29%)
2	77 (7.7%)
3	15 (1.5%)
4	3 (0.3%)
WHERE_n_of_connect_OR	
0	462 (46%)
1	188 (19%)
2	122 (12%)
3	104 (10%)
4	62 (6.2%)
5	34 (3.4%)
6	21 (2.1%)
7	4 (0.4%)
8	3 (0.3%)
WHERE_n_of_operators__between	
0	627 (63%)
1	264 (26%)
2	87 (8.7%)
3	18 (1.8%)

Query parameter or value	Descriptive statistics
4	4 (0.4%)
WHERE_n_of_operators__greater_or_less	
0	427 (43%)
1	296 (30%)
2	159 (16%)
3	78 (7.8%)
4	27 (2.7%)
5	9 (0.9%)
6	4 (0.4%)
WHERE_n_of_operators__in	
0	621 (62%)
1	278 (28%)
2	76 (7.6%)
3	21 (2.1%)
4	4 (0.4%)
WHERE_n_of_operators__like	
0	845 (85%)
1	143 (14%)
2	12 (1.2%)
WHERE_n_of_non_aggr_func__ABS	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func__DAY	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func__LOG	
0	947 (95%)
1	50 (5.0%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__TRUNC	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	917 (92%)
1	79 (7.9%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func__MONTH	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func__YEAR	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func__FLOOR	
0	977 (98%)
1	23 (2.3%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__DOW	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func__ROUND	
0	978 (98%)
1	22 (2.2%)
WHERE_n_of_all_non_aggr_func	
0	721 (72%)
1	214 (21%)
2	51 (5.1%)
3	14 (1.4%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.34 / 2.10
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.39 / 2.58
HAVING_n_of_main_predicates__non_scalar_subquery	
0	430 (43%)
1	167 (17%)
2	208 (21%)
3	195 (20%)
HAVING_n_of_main_predicates__scalar_subquery	
0	422 (42%)
1	203 (20%)
2	212 (21%)
3	163 (16%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	430 (43%)
1	167 (17%)
2	208 (21%)
3	195 (20%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 20.0] 2.4 / 3.7
HAVING_n_of_processed_rows_by_subquery	[0, 972,000,150, 60,102,774,753] 4,569,395,835 / 7,701,353,509
ORDER_BY_n_of_columns	
0	298 (30%)
1	374 (37%)
2	247 (25%)
3	57 (5.7%)
4	13 (1.3%)
5	8 (0.8%)
6	3 (0.3%)
limit	[2, 504, 1,000] 503 / 284
offset	[0, 0, 999] 226 / 308