

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

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

Scale factor: 10 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, 108] 10 / 13
SELECT_n_of_non_aggr_func__DOW	
0	886 (89%)
1	96 (9.6%)
2	14 (1.4%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	888 (89%)
1	99 (9.9%)
2	10 (1.0%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__LOWER	
0	703 (70%)
1	224 (22%)

Query parameter or value	Descriptive statistics
2	57 (5.7%)
3	13 (1.3%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func___LTRIM	
0	718 (72%)
1	218 (22%)
2	42 (4.2%)
3	12 (1.2%)
4	5 (0.5%)
5	5 (0.5%)
SELECT_n_of_non_aggr_func___ROUND	
0	953 (95%)
1	40 (4.0%)
2	5 (0.5%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___RTRIM	
0	702 (70%)
1	222 (22%)
2	61 (6.1%)
3	10 (1.0%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___SQRT	
0	832 (83%)
1	137 (14%)
2	24 (2.4%)
3	4 (0.4%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	667 (67%)
1	246 (25%)
2	73 (7.3%)
3	9 (0.9%)
4	3 (0.3%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___TRUNC	
0	949 (95%)
1	43 (4.3%)
2	7 (0.7%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	685 (69%)
1	243 (24%)
2	56 (5.6%)
3	9 (0.9%)
4	5 (0.5%)

Query parameter or value	Descriptive statistics
5	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	861 (86%)
1	110 (11%)
2	20 (2.0%)
3	7 (0.7%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	934 (93%)
1	60 (6.0%)
2	4 (0.4%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	
0	844 (84%)
1	120 (12%)
2	24 (2.4%)
3	8 (0.8%)
4	1 (0.1%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__DAY	
0	915 (92%)
1	72 (7.2%)
2	10 (1.0%)
3	1 (0.1%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	953 (95%)
1	40 (4.0%)
2	7 (0.7%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 31.0] 3.2 / 4.1
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.63 / 1.44
SELECT_n_of_aggr_func__AVG	
0	942 (94%)
1	44 (4.4%)
2	13 (1.3%)
3	1 (0.1%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.59 / 1.34
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.57 / 1.35
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 13.00] 0.66 / 1.56
SELECT_n_of_aggr_func__SUM	
0	956 (96%)
1	24 (2.4%)
2	18 (1.8%)
4	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 13.00] 2.59 / 2.26

Query parameter or value	Descriptive statistics
FROM_n_of_join_paths	
1	353 (35%)
2	327 (33%)
3	320 (32%)
FROM_n_of_super_joins___FULL	
0	779 (78%)
1	200 (20%)
2	21 (2.1%)
FROM_n_of_super_joins___LEFT	
0	777 (78%)
1	212 (21%)
2	11 (1.1%)
FROM_n_of_super_joins___RIGHT	
0	776 (78%)
1	203 (20%)
2	21 (2.1%)
FROM_n_of_joins___INNER	
0	239 (24%)
1	266 (27%)
2	216 (22%)
3	142 (14%)
4	81 (8.1%)
5	37 (3.7%)
6	13 (1.3%)
7	5 (0.5%)
9	1 (0.1%)
FROM_n_of_joins___RIGHT	
0	233 (23%)
1	261 (26%)
2	222 (22%)
3	159 (16%)
4	69 (6.9%)
5	36 (3.6%)
6	16 (1.6%)
7	4 (0.4%)
FROM_n_of_processed_rows	[5, 68,079,156, 229,441,504] 64,312,964 / 59,132,298
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.77 / 2.16
WHERE_n_of_attribs_of_type___character_varying	
0	427 (43%)
1	325 (33%)
2	175 (18%)
3	53 (5.3%)
4	17 (1.7%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type___date	
0	849 (85%)

Query parameter or value	Descriptive statistics
1	128 (13%)
2	22 (2.2%)
3	1 (0.1%)
WHERE_n_of_attribs_of_type__integer	
0	454 (45%)
1	328 (33%)
2	141 (14%)
3	59 (5.9%)
4	13 (1.3%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type__numeric	
0	659 (66%)
1	239 (24%)
2	88 (8.8%)
3	12 (1.2%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__character	
0	725 (73%)
1	210 (21%)
2	55 (5.5%)
3	9 (0.9%)
5	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	623 (62%)
1	280 (28%)
2	80 (8.0%)
3	14 (1.4%)
4	3 (0.3%)
WHERE_n_of_connect_OR	
0	455 (46%)
1	178 (18%)
2	146 (15%)
3	95 (9.5%)
4	54 (5.4%)
5	54 (5.4%)
6	11 (1.1%)
7	6 (0.6%)
9	1 (0.1%)
WHERE_n_of_operators__between	
0	611 (61%)
1	294 (29%)
2	76 (7.6%)
3	14 (1.4%)
4	5 (0.5%)
WHERE_n_of_operators__greater_or_less	
0	432 (43%)

Query parameter or value	Descriptive statistics
1	280 (28%)
2	174 (17%)
3	72 (7.2%)
4	28 (2.8%)
5	10 (1.0%)
6	4 (0.4%)
WHERE_n_of_operators_like	
0	820 (82%)
1	160 (16%)
2	18 (1.8%)
3	2 (0.2%)
WHERE_n_of_operators_in	
0	626 (63%)
1	258 (26%)
2	94 (9.4%)
3	19 (1.9%)
4	3 (0.3%)
WHERE_n_of_non_aggr_func_DOW	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func_ABS	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_SQRT	
0	921 (92%)
1	73 (7.3%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func_MONTH	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func_ROUND	
0	964 (96%)
1	36 (3.6%)
WHERE_n_of_non_aggr_func_LOG	
0	942 (94%)
1	55 (5.5%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func_YEAR	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_non_aggr_func_DAY	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func_TRUNC	
0	968 (97%)

Query parameter or value	Descriptive statistics
1	32 (3.2%)
WHERE_n_of_non_aggr_func__FLOOR	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_all_non_aggr_func	
0	703 (70%)
1	221 (22%)
2	64 (6.4%)
3	11 (1.1%)
4	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.37 / 2.14
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.32 / 2.56
HAVING_n_of_main_predicates__non_scalar_subquery	
0	463 (46%)
1	190 (19%)
2	184 (18%)
3	163 (16%)
HAVING_n_of_main_predicates__scalar_subquery	
0	417 (42%)
1	210 (21%)
2	200 (20%)
3	173 (17%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	463 (46%)
1	190 (19%)
2	184 (18%)
3	163 (16%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 20.0] 2.4 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 48,606,494, 4,786,775,406] 228,504,701 / 415,703,836
ORDER_BY_n_of_columns	
0	303 (30%)
1	384 (38%)
2	230 (23%)
3	62 (6.2%)
4	12 (1.2%)
5	4 (0.4%)
6	3 (0.3%)
7	1 (0.1%)
8	1 (0.1%)
limit	[1, 503, 998] 493 / 292
offset	[0, 63, 1,000] 267 / 331