

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

Scenario: 1-3jp_rjpp__g__h0ss_h0nss

Scale factor: 500 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, 102] 10 / 12
SELECT_n_of_non_aggr_func__RTRIM	
0	675 (68%)
1	252 (25%)
2	60 (6.0%)
3	12 (1.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	847 (85%)
1	129 (13%)
2	15 (1.5%)
3	7 (0.7%)
4	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	

Query parameter or value	Descriptive statistics
0	673 (67%)
1	245 (25%)
2	61 (6.1%)
3	14 (1.4%)
4	3 (0.3%)
5	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	724 (72%)
1	215 (22%)
2	36 (3.6%)
3	20 (2.0%)
4	1 (0.1%)
5	2 (0.2%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___ABS	
0	929 (93%)
1	65 (6.5%)
2	4 (0.4%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___LOG	
0	928 (93%)
1	65 (6.5%)
2	3 (0.3%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	677 (68%)
1	253 (25%)
2	55 (5.5%)
3	11 (1.1%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	902 (90%)
1	74 (7.4%)
2	18 (1.8%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___MONTH	
0	841 (84%)
1	118 (12%)
2	27 (2.7%)
3	10 (1.0%)
4	3 (0.3%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func__YEAR	
0	870 (87%)
1	96 (9.6%)
2	27 (2.7%)
3	5 (0.5%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	942 (94%)
1	53 (5.3%)
2	4 (0.4%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	690 (69%)
1	244 (24%)
2	51 (5.1%)
3	11 (1.1%)
4	4 (0.4%)
SELECT_n_of_non_aggr_func__DAY	
0	911 (91%)
1	75 (7.5%)
2	13 (1.3%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	935 (94%)
1	60 (6.0%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__ROUND	
0	942 (94%)
1	52 (5.2%)
2	6 (0.6%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 41.0] 3.2 / 4.2
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.67 / 1.58
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 9.00] 0.72 / 1.51
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.67 / 1.42
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.57 / 1.33
SELECT_n_of_aggr_func__AVG	
0	948 (95%)
1	36 (3.6%)
2	13 (1.3%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	943 (94%)
1	43 (4.3%)
2	11 (1.1%)
3	3 (0.3%)

Query parameter or value	Descriptive statistics
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.77 / 2.26
FROM_n_of_join_paths	
1	344 (34%)
2	363 (36%)
3	293 (29%)
FROM_n_of_super_joins__FULL	
0	782 (78%)
1	203 (20%)
2	15 (1.5%)
FROM_n_of_super_joins__LEFT	
0	773 (77%)
1	204 (20%)
2	23 (2.3%)
FROM_n_of_super_joins__RIGHT	
0	788 (79%)
1	196 (20%)
2	16 (1.6%)
FROM_n_of_joins__INNER	
0	268 (27%)
1	235 (24%)
2	217 (22%)
3	156 (16%)
4	79 (7.9%)
5	32 (3.2%)
6	7 (0.7%)
7	6 (0.6%)
FROM_n_of_joins__RIGHT	
0	245 (25%)
1	274 (27%)
2	214 (21%)
3	134 (13%)
4	67 (6.7%)
5	43 (4.3%)
6	13 (1.3%)
7	9 (0.9%)
8	1 (0.1%)
FROM_n_of_processed_rows	[5, 3,405,028,242, 11,475,084,786] 3,136,126,664 / 2,929,247,041
WHERE_n_of_predicates	[0.00, 3.00, 10.00] 2.81 / 2.19
WHERE_n_of_attribs_of_type__character_varying	
0	433 (43%)
1	314 (31%)
2	175 (18%)
3	55 (5.5%)
4	16 (1.6%)
5	6 (0.6%)

Query parameter or value	Descriptive statistics
6	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	835 (84%)
1	136 (14%)
2	27 (2.7%)
3	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	
0	463 (46%)
1	270 (27%)
2	187 (19%)
3	62 (6.2%)
4	11 (1.1%)
5	7 (0.7%)
WHERE_n_of_attribs_of_type__character	
0	727 (73%)
1	220 (22%)
2	44 (4.4%)
3	8 (0.8%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	670 (67%)
1	238 (24%)
2	77 (7.7%)
3	14 (1.4%)
4	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	617 (62%)
1	271 (27%)
2	94 (9.4%)
3	15 (1.5%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_connect_OR	
0	453 (45%)
1	194 (19%)
2	151 (15%)
3	85 (8.5%)
4	52 (5.2%)
5	37 (3.7%)
6	18 (1.8%)
7	9 (0.9%)
9	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	404 (40%)
1	296 (30%)
2	168 (17%)

Query parameter or value	Descriptive statistics
3	85 (8.5%)
4	35 (3.5%)
5	11 (1.1%)
6	1 (0.1%)
WHERE_n_of_operators_in	
0	625 (63%)
1	276 (28%)
2	71 (7.1%)
3	23 (2.3%)
4	4 (0.4%)
5	1 (0.1%)
WHERE_n_of_operators_like	
0	852 (85%)
1	131 (13%)
2	16 (1.6%)
3	1 (0.1%)
WHERE_n_of_operators_between	
0	623 (62%)
1	269 (27%)
2	83 (8.3%)
3	22 (2.2%)
4	3 (0.3%)
WHERE_n_of_non_aggr_func_YEAR	
0	969 (97%)
1	29 (2.9%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_DAY	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func_LOG	
0	950 (95%)
1	50 (5.0%)
WHERE_n_of_non_aggr_func_SQRT	
0	911 (91%)
1	83 (8.3%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func_FLOOR	
0	965 (97%)
1	35 (3.5%)
WHERE_n_of_non_aggr_func_ABS	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func_ROUND	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func_TRUNC	

Query parameter or value	Descriptive statistics
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func__MONTH	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func__DOW	
0	975 (98%)
1	23 (2.3%)
2	2 (0.2%)
WHERE_n_of_all_non_aggr_func	
0	714 (71%)
1	208 (21%)
2	66 (6.6%)
3	11 (1.1%)
4	1 (0.1%)
GROUP_BY_n_of_columns	
0	222 (22%)
1	165 (17%)
2	166 (17%)
3	154 (15%)
4	105 (11%)
5	89 (8.9%)
6	58 (5.8%)
7	32 (3.2%)
8	9 (0.9%)
HAVING_n_of_main_predicates	
0	413 (41%)
1	185 (19%)
2	195 (20%)
3	207 (21%)
ORDER_BY_n_of_columns	
0	268 (27%)
1	393 (39%)
2	261 (26%)
3	63 (6.3%)
4	9 (0.9%)
5	5 (0.5%)
7	1 (0.1%)
limit	[1, 485, 999] 489 / 279
offset	[0, 10, 1,000] 252 / 330