

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

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

Scale factor: 2 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, 89] 10 / 13
SELECT_n_of_non_aggr_func__ABS	
0	934 (93%)
1	61 (6.1%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	704 (70%)
1	221 (22%)
2	61 (6.1%)
3	10 (1.0%)
4	3 (0.3%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	697 (70%)
1	235 (24%)

Query parameter or value	Descriptive statistics
2	51 (5.1%)
3	8 (0.8%)
4	6 (0.6%)
5	1 (0.1%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__DAY	
0	905 (91%)
1	77 (7.7%)
2	16 (1.6%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	903 (90%)
1	79 (7.9%)
2	14 (1.4%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	725 (73%)
1	201 (20%)
2	55 (5.5%)
3	15 (1.5%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	847 (85%)
1	118 (12%)
2	29 (2.9%)
3	4 (0.4%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__ROUND	
0	943 (94%)
1	51 (5.1%)
2	6 (0.6%)
SELECT_n_of_non_aggr_func__RTRIM	
0	692 (69%)
1	230 (23%)
2	56 (5.6%)
3	14 (1.4%)
4	5 (0.5%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	842 (84%)
1	125 (13%)
2	24 (2.4%)

Query parameter or value	Descriptive statistics
3	6 (0.6%)
4	1 (0.1%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__LOWER	
0	691 (69%)
1	236 (24%)
2	59 (5.9%)
3	10 (1.0%)
4	3 (0.3%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__TRUNC	
0	949 (95%)
1	49 (4.9%)
2	2 (0.2%)
SELECT__n_of__non__aggr__func__YEAR	
0	866 (87%)
1	108 (11%)
2	21 (2.1%)
3	5 (0.5%)
SELECT__n_of__non__aggr__func__LOG	
0	901 (90%)
1	85 (8.5%)
2	8 (0.8%)
3	5 (0.5%)
4	1 (0.1%)
SELECT__n_of__non__aggr__func__FLOOR	
0	953 (95%)
1	44 (4.4%)
2	3 (0.3%)
SELECT__n_of__all__non__aggr__func	[0.0, 2.0, 35.0] 3.1 / 4.1
SELECT__n_of__aggr__func__AVG	
0	950 (95%)
1	39 (3.9%)
2	10 (1.0%)
3	1 (0.1%)
SELECT__n_of__aggr__func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.55 / 1.32
SELECT__n_of__aggr__func__COUNT	
0	770 (77%)
1	83 (8.3%)
2	52 (5.2%)
3	34 (3.4%)
4	23 (2.3%)
5	17 (1.7%)
6	11 (1.1%)
7	6 (0.6%)

Query parameter or value	Descriptive statistics
8	4 (0.4%)
SELECT_n_of_aggr_func__SUM	
0	950 (95%)
1	36 (3.6%)
2	9 (0.9%)
3	5 (0.5%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 11.00] 0.61 / 1.44
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 9.00] 0.72 / 1.59
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.63 / 2.32
FROM_n_of_join_paths	
1	330 (33%)
2	325 (33%)
3	345 (35%)
FROM_n_of_super_joins__FULL	
0	764 (76%)
1	216 (22%)
2	20 (2.0%)
FROM_n_of_super_joins__RIGHT	
0	756 (76%)
1	217 (22%)
2	27 (2.7%)
FROM_n_of_super_joins__LEFT	
0	764 (76%)
1	212 (21%)
2	24 (2.4%)
FROM_n_of_joins__INNER	[0.00, 2.00, 9.00] 1.76 / 1.57
FROM_n_of_joins__RIGHT	
0	244 (24%)
1	267 (27%)
2	211 (21%)
3	133 (13%)
4	81 (8.1%)
5	47 (4.7%)
6	15 (1.5%)
7	2 (0.2%)
FROM_n_of_processed_rows	[5, 13,615,796, 45,887,236] 13,035,186 / 12,058,970
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.83 / 2.22
WHERE_n_of_attribs_of_type__character	
0	732 (73%)
1	217 (22%)
2	42 (4.2%)
3	8 (0.8%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	435 (44%)
1	303 (30%)

Query parameter or value	Descriptive statistics
2	174 (17%)
3	63 (6.3%)
4	20 (2.0%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type__integer	
0	433 (43%)
1	335 (34%)
2	152 (15%)
3	63 (6.3%)
4	14 (1.4%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type__numeric	
0	651 (65%)
1	264 (26%)
2	65 (6.5%)
3	19 (1.9%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	828 (83%)
1	147 (15%)
2	24 (2.4%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	610 (61%)
1	306 (31%)
2	69 (6.9%)
3	14 (1.4%)
5	1 (0.1%)
WHERE_n_of_connect_OR	
0	455 (46%)
1	189 (19%)
2	137 (14%)
3	92 (9.2%)
4	62 (6.2%)
5	41 (4.1%)
6	14 (1.4%)
7	9 (0.9%)
8	1 (0.1%)
WHERE_n_of_operators__between	
0	615 (62%)
1	270 (27%)
2	82 (8.2%)
3	28 (2.8%)
4	5 (0.5%)
WHERE_n_of_operators__greater_or_less	
0	389 (39%)

Query parameter or value	Descriptive statistics
1	311 (31%)
2	180 (18%)
3	85 (8.5%)
4	22 (2.2%)
5	7 (0.7%)
6	6 (0.6%)
WHERE_n_of_operators__in	
0	628 (63%)
1	268 (27%)
2	78 (7.8%)
3	19 (1.9%)
4	5 (0.5%)
5	2 (0.2%)
WHERE_n_of_operators__like	
0	853 (85%)
1	126 (13%)
2	21 (2.1%)
WHERE_n_of_non_aggr_func__ABS	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func__SQRT	
0	900 (90%)
1	95 (9.5%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func__DOW	
0	973 (97%)
1	26 (2.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__FLOOR	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func__YEAR	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func__LOG	
0	950 (95%)
1	50 (5.0%)
WHERE_n_of_non_aggr_func__TRUNC	
0	972 (97%)
1	27 (2.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	

Query parameter or value	Descriptive statistics
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	972 (97%)
1	27 (2.7%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	695 (70%)
1	217 (22%)
2	78 (7.8%)
3	9 (0.9%)
4	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 11.00] 2.32 / 2.13
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.32 / 2.54
HAVING_n_of_main_predicates__non_scalar_subquery	
0	438 (44%)
1	203 (20%)
2	180 (18%)
3	179 (18%)
HAVING_n_of_main_predicates__scalar_subquery	
0	433 (43%)
1	188 (19%)
2	196 (20%)
3	183 (18%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	438 (44%)
1	203 (20%)
2	180 (18%)
3	179 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 29.0] 2.4 / 3.8
HAVING_n_of_processed_rows_by_subqueries	[0, 8,450,098, 792,148,960] 45,823,714 / 84,550,828
ORDER_BY_n_of_columns	
0	300 (30%)
1	383 (38%)
2	221 (22%)
3	76 (7.6%)
4	11 (1.1%)
5	6 (0.6%)
6	2 (0.2%)
7	1 (0.1%)
limit	[1, 498, 1,000] 494 / 293
offset	[0, 0, 1,000] 244 / 315