

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

Scenario: 1-3jp_rjpp__g__h0-3ss_h0nss

Scale factor: 1000 GB

The 1000-query pack was generated on: 2023-11-07

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, 110] 10 / 13
SELECT_n_of_non_aggr_func__UPPER	
0	681 (68%)
1	249 (25%)
2	52 (5.2%)
3	11 (1.1%)
4	2 (0.2%)
5	2 (0.2%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	718 (72%)
1	203 (20%)
2	57 (5.7%)
3	16 (1.6%)

Query parameter or value	Descriptive statistics
4	5 (0.5%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__MONTH	
0	846 (85%)
1	112 (11%)
2	23 (2.3%)
3	7 (0.7%)
4	7 (0.7%)
5	4 (0.4%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__LOG	
0	891 (89%)
1	99 (9.9%)
2	8 (0.8%)
4	2 (0.2%)
SELECT__n_of__non__aggr__func__RTRIM	
0	708 (71%)
1	234 (23%)
2	43 (4.3%)
3	8 (0.8%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
SELECT__n_of__non__aggr__func__SQRT	
0	843 (84%)
1	129 (13%)
2	16 (1.6%)
3	7 (0.7%)
4	2 (0.2%)
5	3 (0.3%)
SELECT__n_of__non__aggr__func__ROUND	
0	939 (94%)
1	53 (5.3%)
2	4 (0.4%)
3	3 (0.3%)
4	1 (0.1%)
SELECT__n_of__non__aggr__func__ABS	
0	944 (94%)
1	52 (5.2%)
2	3 (0.3%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func__DOW	
0	900 (90%)
1	75 (7.5%)
2	19 (1.9%)

Query parameter or value	Descriptive statistics
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	931 (93%)
1	64 (6.4%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func___LOWER	
0	676 (68%)
1	240 (24%)
2	59 (5.9%)
3	19 (1.9%)
4	4 (0.4%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func___YEAR	
0	872 (87%)
1	97 (9.7%)
2	17 (1.7%)
3	11 (1.1%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	944 (94%)
1	52 (5.2%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	694 (69%)
1	234 (23%)
2	51 (5.1%)
3	17 (1.7%)
4	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	905 (91%)
1	82 (8.2%)
2	11 (1.1%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 44.00] 3.29 / 4.58
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 11.00] 0.72 / 1.54
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 10.00] 0.65 / 1.48
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 10.00] 0.56 / 1.35
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.67 / 1.55
SELECT_n_of_aggr_func___AVG	
0	939 (94%)
1	49 (4.9%)
2	7 (0.7%)

Query parameter or value	Descriptive statistics
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	929 (93%)
1	48 (4.8%)
2	19 (1.9%)
3	4 (0.4%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.77 / 2.40
FROM_n_of_join_paths	
1	343 (34%)
2	309 (31%)
3	348 (35%)
FROM_n_of_super_joins__FULL	
0	758 (76%)
1	227 (23%)
2	15 (1.5%)
FROM_n_of_super_joins__RIGHT	
0	779 (78%)
1	200 (20%)
2	21 (2.1%)
FROM_n_of_super_joins__LEFT	
0	772 (77%)
1	214 (21%)
2	14 (1.4%)
FROM_n_of_joins__INNER	
0	261 (26%)
1	249 (25%)
2	182 (18%)
3	154 (15%)
4	89 (8.9%)
5	45 (4.5%)
6	11 (1.1%)
7	8 (0.8%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	262 (26%)
1	249 (25%)
2	217 (22%)
3	129 (13%)
4	75 (7.5%)
5	44 (4.4%)
6	18 (1.8%)
7	6 (0.6%)
FROM_n_of_processed_rows	[5, 6,809,989,754, 22,949,969,182] 6,399,901,942 / 5,944,563,015
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.74 / 2.17

Query parameter or value	Descriptive statistics
WHERE_n_of_attribs_of_type___character_varying	
0	439 (44%)
1	300 (30%)
2	177 (18%)
3	63 (6.3%)
4	16 (1.6%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type___numeric	
0	663 (66%)
1	265 (27%)
2	60 (6.0%)
3	10 (1.0%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type___character	
0	732 (73%)
1	217 (22%)
2	46 (4.6%)
3	5 (0.5%)
WHERE_n_of_attribs_of_type___integer	
0	448 (45%)
1	326 (33%)
2	159 (16%)
3	43 (4.3%)
4	21 (2.1%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type___date	
0	837 (84%)
1	137 (14%)
2	24 (2.4%)
3	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	610 (61%)
1	291 (29%)
2	78 (7.8%)
3	18 (1.8%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_connect_OR	
0	456 (46%)
1	201 (20%)
2	130 (13%)
3	89 (8.9%)
4	61 (6.1%)
5	36 (3.6%)
6	18 (1.8%)

Query parameter or value	Descriptive statistics
7	8 (0.8%)
8	1 (0.1%)
WHERE_n_of_operators___between	
0	619 (62%)
1	285 (29%)
2	72 (7.2%)
3	20 (2.0%)
4	4 (0.4%)
WHERE_n_of_operators___greater_or_less	
0	436 (44%)
1	288 (29%)
2	162 (16%)
3	69 (6.9%)
4	33 (3.3%)
5	12 (1.2%)
WHERE_n_of_operators___in	
0	629 (63%)
1	256 (26%)
2	84 (8.4%)
3	30 (3.0%)
4	1 (0.1%)
WHERE_n_of_operators___like	
0	816 (82%)
1	158 (16%)
2	26 (2.6%)
WHERE_n_of_non_aggr_func___SQRT	
0	921 (92%)
1	75 (7.5%)
2	3 (0.3%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	960 (96%)
1	38 (3.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___TRUNC	
0	970 (97%)
1	28 (2.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___DAY	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func___LOG	
0	950 (95%)
1	47 (4.7%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___DOW	

Query parameter or value	Descriptive statistics
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_non_aggr_func__MONTH	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func__YEAR	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func__ABS	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func__ROUND	
0	969 (97%)
1	28 (2.8%)
2	3 (0.3%)
WHERE_n_of_all_non_aggr_func	
0	715 (72%)
1	216 (22%)
2	56 (5.6%)
3	11 (1.1%)
4	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.36 / 2.17
HAVING_n_of_main_predicates	
0	296 (30%)
1	105 (11%)
2	147 (15%)
3	200 (20%)
4	123 (12%)
5	91 (9.1%)
6	38 (3.8%)
HAVING_n_of_main_predicates__scalar_subquery	
0	451 (45%)
1	178 (18%)
2	206 (21%)
3	165 (17%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 18.0] 2.4 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 135,000,013, 379,099,475,584] 15,253,191,057 / 35,139,626,269
ORDER_BY_n_of_columns	
0	292 (29%)
1	406 (41%)
2	214 (21%)
3	62 (6.2%)
4	15 (1.5%)
5	8 (0.8%)
6	3 (0.3%)

Query parameter or value	Descriptive statistics
limit	[3, 517, 999] 508 / 289
offset	[0, 0, 1,000] 236 / 316