

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

Scenario: 1-1jp__g__h0-3ss_h0-3nss

Scale factor: 5 GB

The 1000-query pack was generated on: 2023-11-05

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	[1.0, 4.0, 40.0] 5.7 / 5.5
SELECT_n_of_non_aggr_func__LTRIM	
0	790 (79%)
1	186 (19%)
2	18 (1.8%)
3	6 (0.6%)
SELECT_n_of_non_aggr_func__SQRT	
0	896 (90%)
1	99 (9.9%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__TRUNC	
0	967 (97%)
1	33 (3.3%)
SELECT_n_of_non_aggr_func__UPPER	
0	785 (79%)

Query parameter or value	Descriptive statistics
1	189 (19%)
2	24 (2.4%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	797 (80%)
1	174 (17%)
2	26 (2.6%)
3	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__LOWER	
0	803 (80%)
1	166 (17%)
2	30 (3.0%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	925 (93%)
1	66 (6.6%)
2	8 (0.8%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	939 (94%)
1	56 (5.6%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__MONTH	
0	888 (89%)
1	100 (10%)
2	7 (0.7%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func__ROUND	
0	967 (97%)
1	32 (3.2%)
2	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	927 (93%)
1	68 (6.8%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	785 (79%)
1	182 (18%)
2	31 (3.1%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	916 (92%)
1	74 (7.4%)
2	7 (0.7%)
3	3 (0.3%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non_aggr_func___ABS	
0	967 (97%)
1	28 (2.8%)
2	5 (0.5%)
SELECT__n_of__non_aggr_func___FLOOR	
0	970 (97%)
1	28 (2.8%)
2	2 (0.2%)
SELECT__n_of__all__non_aggr_func	[0.00, 1.00, 16.00] 1.90 / 2.16
SELECT__n_of__aggr_func___COUNT	
0	762 (76%)
1	103 (10%)
2	65 (6.5%)
3	42 (4.2%)
4	17 (1.7%)
5	7 (0.7%)
6	3 (0.3%)
7	1 (0.1%)
SELECT__n_of__aggr_func___SUM	
0	955 (96%)
1	37 (3.7%)
2	8 (0.8%)
SELECT__n_of__aggr_func___COUNT_DISTINCT	
0	766 (77%)
1	108 (11%)
2	66 (6.6%)
3	33 (3.3%)
4	16 (1.6%)
5	7 (0.7%)
6	3 (0.3%)
7	1 (0.1%)
SELECT__n_of__aggr_func___MAX	
0	760 (76%)
1	129 (13%)
2	65 (6.5%)
3	27 (2.7%)
4	9 (0.9%)
5	9 (0.9%)
6	1 (0.1%)
SELECT__n_of__aggr_func___MIN	
0	804 (80%)
1	92 (9.2%)
2	56 (5.6%)
3	26 (2.6%)
4	16 (1.6%)
5	4 (0.4%)

Query parameter or value	Descriptive statistics
6	2 (0.2%)
SELECT_n_of_aggr_func__AVG	
0	963 (96%)
1	31 (3.1%)
2	5 (0.5%)
3	1 (0.1%)
SELECT_n_of_all_aggr_func	
0	188 (19%)
1	290 (29%)
2	239 (24%)
3	145 (15%)
4	70 (7.0%)
5	45 (4.5%)
6	17 (1.7%)
7	6 (0.6%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins__INNER	
0	493 (49%)
1	334 (33%)
2	132 (13%)
3	34 (3.4%)
4	7 (0.7%)
FROM_n_of_joins__RIGHT	
0	495 (50%)
1	314 (31%)
2	149 (15%)
3	39 (3.9%)
4	3 (0.3%)
FROM_n_of_processed_rows	[5, 4,050,353, 38,249,944] 13,783,686 / 15,591,727
WHERE_n_of_predicates	
0	250 (25%)
1	237 (24%)
2	193 (19%)
3	181 (18%)
4	91 (9.1%)
5	36 (3.6%)
6	12 (1.2%)
WHERE_n_of_attrbs_of_type__numeric	
0	761 (76%)
1	207 (21%)
2	28 (2.8%)
3	4 (0.4%)
WHERE_n_of_attrbs_of_type__integer	
0	563 (56%)
1	316 (32%)

Query parameter or value	Descriptive statistics
2	100 (10%)
3	20 (2.0%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	818 (82%)
1	157 (16%)
2	23 (2.3%)
3	1 (0.1%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type__character_varying	
0	548 (55%)
1	326 (33%)
2	105 (11%)
3	19 (1.9%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	898 (90%)
1	88 (8.8%)
2	14 (1.4%)
WHERE_n_of_pkey_attribs	
0	722 (72%)
1	243 (24%)
2	32 (3.2%)
3	3 (0.3%)
WHERE_n_of_connect_OR	
0	612 (61%)
1	209 (21%)
2	115 (12%)
3	44 (4.4%)
4	17 (1.7%)
5	3 (0.3%)
WHERE_n_of_operators__greater_or_less	
0	532 (53%)
1	312 (31%)
2	110 (11%)
3	39 (3.9%)
4	6 (0.6%)
5	1 (0.1%)
WHERE_n_of_operators__in	
0	747 (75%)
1	204 (20%)
2	44 (4.4%)
3	4 (0.4%)
4	1 (0.1%)
WHERE_n_of_operators__between	
0	736 (74%)

Query parameter or value	Descriptive statistics
1	212 (21%)
2	47 (4.7%)
3	4 (0.4%)
4	1 (0.1%)
WHERE_n_of_operators_like	
0	878 (88%)
1	116 (12%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func_TRUNC	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func_DOW	
0	986 (99%)
1	14 (1.4%)
WHERE_n_of_non_aggr_func_ABS	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func_MONTH	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func_ROUND	
0	983 (98%)
1	17 (1.7%)
WHERE_n_of_non_aggr_func_YEAR	
0	986 (99%)
1	14 (1.4%)
WHERE_n_of_non_aggr_func_LOG	
0	959 (96%)
1	41 (4.1%)
WHERE_n_of_non_aggr_func_DAY	
0	978 (98%)
1	21 (2.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_SQRT	
0	940 (94%)
1	55 (5.5%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func_FLOOR	
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_all_non_aggr_func	
0	781 (78%)
1	190 (19%)
2	26 (2.6%)
3	3 (0.3%)
GROUP_BY_n_of_columns	

Query parameter or value	Descriptive statistics
0	259 (26%)
1	271 (27%)
2	232 (23%)
3	134 (13%)
4	59 (5.9%)
5	35 (3.5%)
6	8 (0.8%)
7	2 (0.2%)
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.36 / 2.61
HAVING_n_of_main_predicates__non_scalar_subquery	
0	435 (44%)
1	185 (19%)
2	185 (19%)
3	195 (20%)
HAVING_n_of_main_predicates__scalar_subquery	
0	465 (47%)
1	196 (20%)
2	162 (16%)
3	177 (18%)
HAVING_n_of_subqueries__non_scalar_subquery	
0	435 (44%)
1	185 (19%)
2	185 (19%)
3	195 (20%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 25.0] 2.2 / 3.6
HAVING_n_of_processed_rows_by_subqueries	[0, 2,250,075, 879,748,597] 49,075,126 / 102,893,073
ORDER_BY_n_of_columns	
0	311 (31%)
1	464 (46%)
2	194 (19%)
3	25 (2.5%)
4	3 (0.3%)
5	1 (0.1%)
6	2 (0.2%)
limit	[1, 498, 1,000] 509 / 292
offset	[0, 23, 1,000] 257 / 323