

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

Scenario: 1-3jp_rjpp__g__h0ss_h0nss

Scale factor: 1 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, 109] 10 / 14
SELECT_n_of_non_aggr_func__LOWER	
0	715 (72%)
1	224 (22%)
2	46 (4.6%)
3	7 (0.7%)
4	3 (0.3%)
5	4 (0.4%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	702 (70%)
1	231 (23%)
2	51 (5.1%)
3	9 (0.9%)
4	3 (0.3%)

Query parameter or value	Descriptive statistics
5	4 (0.4%)
SELECT_n_of_non_aggr_func__YEAR	
0	874 (87%)
1	102 (10%)
2	12 (1.2%)
3	6 (0.6%)
4	4 (0.4%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	
0	721 (72%)
1	223 (22%)
2	39 (3.9%)
3	12 (1.2%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	902 (90%)
1	77 (7.7%)
2	15 (1.5%)
3	1 (0.1%)
4	4 (0.4%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	
0	853 (85%)
1	111 (11%)
2	19 (1.9%)
3	15 (1.5%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__FLOOR	
0	957 (96%)
1	36 (3.6%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	696 (70%)
1	240 (24%)
2	44 (4.4%)
3	14 (1.4%)
4	6 (0.6%)
SELECT_n_of_non_aggr_func__TRUNC	
0	941 (94%)
1	52 (5.2%)
2	6 (0.6%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	709 (71%)

Query parameter or value	Descriptive statistics
1	225 (23%)
2	52 (5.2%)
3	12 (1.2%)
4	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__ROUND	
0	940 (94%)
1	58 (5.8%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func__SQRT	
0	854 (85%)
1	116 (12%)
2	20 (2.0%)
3	8 (0.8%)
4	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	949 (95%)
1	46 (4.6%)
2	4 (0.4%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	917 (92%)
1	70 (7.0%)
2	10 (1.0%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__LOG	
0	912 (91%)
1	76 (7.6%)
2	9 (0.9%)
3	2 (0.2%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 37.00] 3.03 / 4.18
SELECT_n_of_aggr_func__COUNT_DISTINCT	
0	753 (75%)
1	80 (8.0%)
2	57 (5.7%)
3	42 (4.2%)
4	33 (3.3%)
5	22 (2.2%)
6	6 (0.6%)
7	6 (0.6%)
8	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	942 (94%)
1	38 (3.8%)

Query parameter or value	Descriptive statistics
2	13 (1.3%)
3	4 (0.4%)
4	3 (0.3%)
SELECT_n_of_aggr_func__MAX	
0	769 (77%)
1	88 (8.8%)
2	43 (4.3%)
3	46 (4.6%)
4	18 (1.8%)
5	13 (1.3%)
6	13 (1.3%)
7	7 (0.7%)
8	3 (0.3%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 10.00] 0.64 / 1.51
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 10.00] 0.62 / 1.44
SELECT_n_of_aggr_func__AVG	
0	951 (95%)
1	36 (3.6%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.66 / 2.30
FROM_n_of_join_paths	
1	319 (32%)
2	314 (31%)
3	367 (37%)
FROM_n_of_super_joins__RIGHT	
0	742 (74%)
1	231 (23%)
2	27 (2.7%)
FROM_n_of_super_joins__LEFT	
0	780 (78%)
1	202 (20%)
2	18 (1.8%)
FROM_n_of_super_joins__FULL	
0	750 (75%)
1	228 (23%)
2	22 (2.2%)
FROM_n_of_joins__INNER	
0	229 (23%)
1	251 (25%)
2	230 (23%)
3	137 (14%)
4	76 (7.6%)
5	49 (4.9%)
6	17 (1.7%)
7	10 (1.0%)

Query parameter or value	Descriptive statistics
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	241 (24%)
1	239 (24%)
2	238 (24%)
3	145 (15%)
4	76 (7.6%)
5	37 (3.7%)
6	19 (1.9%)
7	5 (0.5%)
FROM_n_of_processed_rows	[5, 6,820,641, 22,951,858] 6,772,098 / 6,057,657
WHERE_n_of_predicates	[0.00, 2.00, 9.00] 2.83 / 2.21
WHERE_n_of_attribs_of_type__character_varying	
0	434 (43%)
1	313 (31%)
2	162 (16%)
3	70 (7.0%)
4	19 (1.9%)
5	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	
0	426 (43%)
1	331 (33%)
2	172 (17%)
3	49 (4.9%)
4	17 (1.7%)
5	4 (0.4%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	654 (65%)
1	248 (25%)
2	80 (8.0%)
3	16 (1.6%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	829 (83%)
1	159 (16%)
2	12 (1.2%)
WHERE_n_of_attribs_of_type__character	
0	724 (72%)
1	222 (22%)
2	48 (4.8%)
3	6 (0.6%)
WHERE_n_of_pkey_attribs	
0	612 (61%)
1	300 (30%)
2	72 (7.2%)

Query parameter or value	Descriptive statistics
3	14 (1.4%)
4	1 (0.1%)
6	1 (0.1%)
WHERE_n_of_connect_OR	
0	467 (47%)
1	175 (18%)
2	131 (13%)
3	85 (8.5%)
4	71 (7.1%)
5	43 (4.3%)
6	21 (2.1%)
7	4 (0.4%)
8	3 (0.3%)
WHERE_n_of_operators__between	
0	603 (60%)
1	285 (29%)
2	75 (7.5%)
3	26 (2.6%)
4	10 (1.0%)
5	1 (0.1%)
WHERE_n_of_operators__greater_or_less	
0	393 (39%)
1	331 (33%)
2	157 (16%)
3	79 (7.9%)
4	32 (3.2%)
5	7 (0.7%)
6	1 (0.1%)
WHERE_n_of_operators__in	
0	624 (62%)
1	262 (26%)
2	93 (9.3%)
3	17 (1.7%)
4	4 (0.4%)
WHERE_n_of_operators__like	
0	830 (83%)
1	159 (16%)
2	10 (1.0%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	979 (98%)
1	21 (2.1%)
WHERE_n_of_non_aggr_func__ABS	
0	968 (97%)
1	32 (3.2%)
WHERE_n_of_non_aggr_func__DOW	

Query parameter or value	Descriptive statistics
0	972 (97%)
1	28 (2.8%)
WHERE_n_of_non_aggr_func__SQRT	
0	890 (89%)
1	104 (10%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func__LOG	
0	942 (94%)
1	57 (5.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__TRUNC	
0	975 (98%)
1	24 (2.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__FLOOR	
0	969 (97%)
1	30 (3.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	973 (97%)
1	27 (2.7%)
WHERE_n_of_non_aggr_func__YEAR	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_non_aggr_func__DAY	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_all_non_aggr_func	
0	700 (70%)
1	229 (23%)
2	59 (5.9%)
3	12 (1.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 10.00] 2.34 / 2.12
HAVING_n_of_main_predicates	
0	440 (44%)
1	194 (19%)
2	182 (18%)
3	184 (18%)
ORDER_BY_n_of_columns	
0	315 (32%)
1	373 (37%)
2	223 (22%)
3	62 (6.2%)
4	15 (1.5%)
5	9 (0.9%)
6	3 (0.3%)

Query parameter or value	Descriptive statistics
limit	[2, 514, 997] 511 / 291
offset	[0, 69, 997] 268 / 327