

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

Scenario: 1-3jp_rjpp__g__h0ss_h0nss

Scale factor: 200 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, 120] 10 / 13
SELECT_n_of_non_aggr_func__DOW	
0	882 (88%)
1	97 (9.7%)
2	14 (1.4%)
3	5 (0.5%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__LOWER	
0	699 (70%)
1	231 (23%)
2	58 (5.8%)
3	8 (0.8%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	

Query parameter or value	Descriptive statistics
0	855 (86%)
1	110 (11%)
2	20 (2.0%)
3	8 (0.8%)
4	5 (0.5%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	707 (71%)
1	227 (23%)
2	48 (4.8%)
3	16 (1.6%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___UPPER	
0	696 (70%)
1	230 (23%)
2	61 (6.1%)
3	10 (1.0%)
4	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	708 (71%)
1	227 (23%)
2	46 (4.6%)
3	13 (1.3%)
4	5 (0.5%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	726 (73%)
1	213 (21%)
2	41 (4.1%)
3	9 (0.9%)
4	7 (0.7%)
5	3 (0.3%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	953 (95%)
1	43 (4.3%)
2	3 (0.3%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	899 (90%)
1	94 (9.4%)
2	3 (0.3%)
3	2 (0.2%)
5	1 (0.1%)
8	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func__ROUND	
0	946 (95%)
1	49 (4.9%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func__DAY	
0	905 (91%)
1	82 (8.2%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__SQRT	
0	852 (85%)
1	120 (12%)
2	24 (2.4%)
3	2 (0.2%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__YEAR	
0	870 (87%)
1	104 (10%)
2	18 (1.8%)
3	2 (0.2%)
4	4 (0.4%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	955 (96%)
1	42 (4.2%)
2	2 (0.2%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	954 (95%)
1	42 (4.2%)
2	3 (0.3%)
3	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.00, 2.00, 39.00] 3.07 / 4.04
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 9.00] 0.64 / 1.46
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.74 / 1.56
SELECT_n_of_aggr_func__AVG	
0	958 (96%)
1	32 (3.2%)
2	8 (0.8%)
3	2 (0.2%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 9.00] 0.62 / 1.39
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.70 / 1.54
SELECT_n_of_aggr_func__SUM	
0	936 (94%)
1	46 (4.6%)

Query parameter or value	Descriptive statistics
2	12 (1.2%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.84 / 2.33
FROM_n_of_join_paths	
1	329 (33%)
2	320 (32%)
3	351 (35%)
FROM_n_of_super_joins__LEFT	
0	770 (77%)
1	212 (21%)
2	18 (1.8%)
FROM_n_of_super_joins__FULL	
0	758 (76%)
1	213 (21%)
2	29 (2.9%)
FROM_n_of_super_joins__RIGHT	
0	780 (78%)
1	195 (20%)
2	25 (2.5%)
FROM_n_of_joins__INNER	
0	247 (25%)
1	270 (27%)
2	205 (21%)
3	145 (15%)
4	76 (7.6%)
5	40 (4.0%)
6	10 (1.0%)
7	6 (0.6%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	239 (24%)
1	258 (26%)
2	220 (22%)
3	133 (13%)
4	94 (9.4%)
5	35 (3.5%)
6	14 (1.4%)
7	6 (0.6%)
8	1 (0.1%)
FROM_n_of_processed_rows	[5, 1,362,018,489, 4,590,055,392] 1,342,151,423 / 1,229,706,369
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.79 / 2.21
WHERE_n_of_attribs_of_type__character	
0	734 (73%)
1	207 (21%)

Query parameter or value	Descriptive statistics
2	46 (4.6%)
3	11 (1.1%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	
0	459 (46%)
1	309 (31%)
2	149 (15%)
3	63 (6.3%)
4	16 (1.6%)
5	3 (0.3%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	643 (64%)
1	264 (26%)
2	76 (7.6%)
3	15 (1.5%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__character_varying	
0	442 (44%)
1	320 (32%)
2	146 (15%)
3	72 (7.2%)
4	19 (1.9%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	833 (83%)
1	141 (14%)
2	25 (2.5%)
3	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	615 (62%)
1	296 (30%)
2	70 (7.0%)
3	16 (1.6%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_connect_OR	
0	469 (47%)
1	172 (17%)
2	138 (14%)
3	105 (11%)
4	62 (6.2%)
5	27 (2.7%)
6	18 (1.8%)
7	8 (0.8%)
8	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_operators__between	
0	605 (61%)
1	296 (30%)
2	80 (8.0%)
3	14 (1.4%)
4	5 (0.5%)
WHERE_n_of_operators__greater_or_less	
0	404 (40%)
1	304 (30%)
2	159 (16%)
3	89 (8.9%)
4	28 (2.8%)
5	10 (1.0%)
6	5 (0.5%)
7	1 (0.1%)
WHERE_n_of_operators__in	
0	609 (61%)
1	294 (29%)
2	74 (7.4%)
3	18 (1.8%)
4	5 (0.5%)
WHERE_n_of_operators__like	
0	846 (85%)
1	138 (14%)
2	15 (1.5%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__ABS	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func__TRUNC	
0	967 (97%)
1	33 (3.3%)
WHERE_n_of_non_aggr_func__MONTH	
0	963 (96%)
1	36 (3.6%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__LOG	
0	940 (94%)
1	58 (5.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__DOW	
0	978 (98%)
1	21 (2.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__YEAR	
0	975 (98%)

Query parameter or value	Descriptive statistics
1	25 (2.5%)
WHERE_n_of_non_aggr_func__FLOOR	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	914 (91%)
1	79 (7.9%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func__ROUND	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_non_aggr_func__DAY	
0	964 (96%)
1	36 (3.6%)
WHERE_n_of_all_non_aggr_func	
0	689 (69%)
1	231 (23%)
2	61 (6.1%)
3	19 (1.9%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.42 / 2.16
HAVING_n_of_main_predicates	
0	429 (43%)
1	182 (18%)
2	194 (19%)
3	195 (20%)
ORDER_BY_n_of_columns	
0	281 (28%)
1	383 (38%)
2	237 (24%)
3	73 (7.3%)
4	18 (1.8%)
5	4 (0.4%)
6	2 (0.2%)
7	2 (0.2%)
limit	[1, 504, 999] 509 / 285
offset	[0, 43, 997] 244 / 310