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

Scenario: 1-3jp_ujjp__g__h0ss_h0nss

Scale factor: 2 GB

The 1000-query pack was generated on: 2023-11-09

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 / 12
SELECT_n_of_non_aggr_func___LOWER	
0	682 (68%)
1	242 (24%)
2	64 (6.4%)
3	8 (0.8%)
4	2 (0.2%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func___SQRT	
0	846 (85%)
1	123 (12%)
2	21 (2.1%)
3	6 (0.6%)
4	4 (0.4%)
SELECT_n_of_non_aggr_func___DOW	

Query parameter or value	Descriptive statistics
0	882 (88%)
1	104 (10%)
2	10 (1.0%)
3	4 (0.4%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	686 (69%)
1	239 (24%)
2	58 (5.8%)
3	9 (0.9%)
4	6 (0.6%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	871 (87%)
1	100 (10%)
2	22 (2.2%)
3	7 (0.7%)
SELECT_n_of_non_aggr_func___LTRIM	
0	692 (69%)
1	241 (24%)
2	46 (4.6%)
3	12 (1.2%)
4	4 (0.4%)
5	2 (0.2%)
6	3 (0.3%)
SELECT_n_of_non_aggr_func___RTRIM	
0	688 (69%)
1	226 (23%)
2	69 (6.9%)
3	7 (0.7%)
4	8 (0.8%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	
0	710 (71%)
1	217 (22%)
2	50 (5.0%)
3	16 (1.6%)
4	6 (0.6%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	946 (95%)
1	47 (4.7%)
2	5 (0.5%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___MONTH	

Query parameter or value	Descriptive statistics
0	844 (84%)
1	119 (12%)
2	26 (2.6%)
3	5 (0.5%)
4	2 (0.2%)
5	3 (0.3%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	917 (92%)
1	70 (7.0%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	947 (95%)
1	49 (4.9%)
2	4 (0.4%)
SELECT_n_of_non_aggr_func___LOG	
0	917 (92%)
1	71 (7.1%)
2	7 (0.7%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	942 (94%)
1	53 (5.3%)
2	5 (0.5%)
SELECT_n_of_non_aggr_func___TRUNC	
0	942 (94%)
1	54 (5.4%)
2	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 36.0] 3.2 / 4.2
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 10.00] 0.67 / 1.52
SELECT_n_of_aggr_func___AVG	
0	928 (93%)
1	51 (5.1%)
2	16 (1.6%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_aggr_func___MAX	
0	776 (78%)
1	67 (6.7%)
2	53 (5.3%)
3	41 (4.1%)
4	27 (2.7%)
5	19 (1.9%)
6	7 (0.7%)

Query parameter or value	Descriptive statistics
7	5 (0.5%)
8	5 (0.5%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.62 / 1.42
SELECT_n_of_aggr_func___COUNT	
0	765 (77%)
1	87 (8.7%)
2	51 (5.1%)
3	30 (3.0%)
4	27 (2.7%)
5	19 (1.9%)
6	15 (1.5%)
7	5 (0.5%)
8	1 (0.1%)
SELECT_n_of_aggr_func___SUM	
0	948 (95%)
1	39 (3.9%)
2	9 (0.9%)
3	2 (0.2%)
4	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.69 / 2.28
FROM_n_of_join_paths	
1	330 (33%)
2	313 (31%)
3	357 (36%)
FROM_n_of_super_joins___LEFT	
0	767 (77%)
1	210 (21%)
2	23 (2.3%)
FROM_n_of_super_joins___RIGHT	
0	765 (77%)
1	212 (21%)
2	23 (2.3%)
FROM_n_of_super_joins___FULL	
0	772 (77%)
1	204 (20%)
2	24 (2.4%)
FROM_n_of_joins___INNER	
0	236 (24%)
1	272 (27%)
2	228 (23%)
3	135 (14%)
4	65 (6.5%)
5	43 (4.3%)
6	12 (1.2%)
7	8 (0.8%)
9	1 (0.1%)

Query parameter or value	Descriptive statistics
FROM_n_of_joins___RIGHT	
0	218 (22%)
1	283 (28%)
2	205 (21%)
3	148 (15%)
4	89 (8.9%)
5	37 (3.7%)
6	11 (1.1%)
7	7 (0.7%)
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 13,615,796, 45,887,236] 12,896,100 / 11,946,391
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.87 / 2.22
WHERE_n_of_attribs_of_type___character_varying	
0	428 (43%)
1	309 (31%)
2	176 (18%)
3	67 (6.7%)
4	17 (1.7%)
5	2 (0.2%)
6	1 (0.1%)
WHERE_n_of_attribs_of_type___character	
0	736 (74%)
1	209 (21%)
2	46 (4.6%)
3	8 (0.8%)
4	1 (0.1%)
WHERE_n_of_attribs_of_type___date	
0	834 (83%)
1	142 (14%)
2	21 (2.1%)
3	3 (0.3%)
WHERE_n_of_attribs_of_type___integer	
0	437 (44%)
1	308 (31%)
2	179 (18%)
3	55 (5.5%)
4	14 (1.4%)
5	7 (0.7%)
WHERE_n_of_attribs_of_type___numeric	
0	637 (64%)
1	267 (27%)
2	78 (7.8%)
3	14 (1.4%)
4	4 (0.4%)
WHERE_n_of_pkey_attribs	
0	609 (61%)

Query parameter or value	Descriptive statistics
1	285 (29%)
2	84 (8.4%)
3	19 (1.9%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_connect_OR	
0	427 (43%)
1	212 (21%)
2	120 (12%)
3	91 (9.1%)
4	80 (8.0%)
5	44 (4.4%)
6	18 (1.8%)
7	8 (0.8%)
WHERE_n_of_operators___greater_or_less	
0	401 (40%)
1	309 (31%)
2	161 (16%)
3	85 (8.5%)
4	33 (3.3%)
5	10 (1.0%)
6	1 (0.1%)
WHERE_n_of_operators___in	
0	598 (60%)
1	286 (29%)
2	96 (9.6%)
3	17 (1.7%)
4	3 (0.3%)
WHERE_n_of_operators___like	
0	822 (82%)
1	156 (16%)
2	22 (2.2%)
WHERE_n_of_operators___between	
0	619 (62%)
1	267 (27%)
2	85 (8.5%)
3	21 (2.1%)
4	8 (0.8%)
WHERE_n_of_non_aggr_func___YEAR	
0	967 (97%)
1	32 (3.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DOW	
0	978 (98%)
1	22 (2.2%)
WHERE_n_of_non_aggr_func___LOG	

Query parameter or value	Descriptive statistics
0	935 (94%)
1	61 (6.1%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func___ABS	
0	977 (98%)
1	21 (2.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	896 (90%)
1	98 (9.8%)
2	6 (0.6%)
WHERE_n_of_non_aggr_func___MONTH	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_non_aggr_func___FLOOR	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func___TRUNC	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func___DAY	
0	971 (97%)
1	29 (2.9%)
WHERE_n_of_non_aggr_func___ROUND	
0	974 (97%)
1	26 (2.6%)
WHERE_n_of_all_non_aggr_func	
0	710 (71%)
1	220 (22%)
2	54 (5.4%)
3	12 (1.2%)
4	4 (0.4%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.42 / 2.19
HAVING_n_of_main_predicates	
0	451 (45%)
1	185 (19%)
2	182 (18%)
3	182 (18%)
ORDER_BY_n_of_columns	
0	307 (31%)
1	372 (37%)
2	238 (24%)
3	57 (5.7%)
4	16 (1.6%)
5	8 (0.8%)
6	2 (0.2%)

Query parameter or value	Descriptive statistics
limit	[1, 488, 999] 499 / 291
offset	[0, 28, 999] 255 / 320