

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

Scenario: 1-3jp_rjpp__g__h0ss_h0nss

Scale factor: 50 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, 107] 10 / 12
SELECT_n_of_non_aggr_func__SQRT	
0	857 (86%)
1	108 (11%)
2	21 (2.1%)
3	7 (0.7%)
4	5 (0.5%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__LOWER	
0	686 (69%)
1	230 (23%)
2	64 (6.4%)
3	13 (1.3%)
4	6 (0.6%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
SELECT__n_of__non_aggr_func__LTRIM	
0	732 (73%)
1	210 (21%)
2	35 (3.5%)
3	17 (1.7%)
4	5 (0.5%)
5	1 (0.1%)
SELECT__n_of__non_aggr_func__DAY	
0	918 (92%)
1	68 (6.8%)
2	10 (1.0%)
3	2 (0.2%)
4	2 (0.2%)
SELECT__n_of__non_aggr_func__FLOOR	
0	953 (95%)
1	44 (4.4%)
2	2 (0.2%)
3	1 (0.1%)
SELECT__n_of__non_aggr_func__MONTH	
0	854 (85%)
1	116 (12%)
2	19 (1.9%)
3	5 (0.5%)
4	3 (0.3%)
5	1 (0.1%)
6	2 (0.2%)
SELECT__n_of__non_aggr_func__RTRIM	
0	678 (68%)
1	250 (25%)
2	50 (5.0%)
3	14 (1.4%)
4	5 (0.5%)
5	2 (0.2%)
8	1 (0.1%)
SELECT__n_of__non_aggr_func__SUBSTR	
0	701 (70%)
1	228 (23%)
2	51 (5.1%)
3	12 (1.2%)
4	5 (0.5%)
5	3 (0.3%)
SELECT__n_of__non_aggr_func__UPPER	
0	655 (66%)
1	270 (27%)
2	59 (5.9%)
3	10 (1.0%)

Query parameter or value	Descriptive statistics
4	4 (0.4%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	874 (87%)
1	99 (9.9%)
2	18 (1.8%)
3	7 (0.7%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__ABS	
0	947 (95%)
1	49 (4.9%)
2	3 (0.3%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	905 (91%)
1	73 (7.3%)
2	16 (1.6%)
3	5 (0.5%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	892 (89%)
1	90 (9.0%)
2	16 (1.6%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__ROUND	
0	951 (95%)
1	47 (4.7%)
2	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	944 (94%)
1	52 (5.2%)
2	2 (0.2%)
3	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 2.0, 38.0] 3.2 / 4.3
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 9.00] 0.59 / 1.39
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 11.00] 0.70 / 1.51
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.62 / 1.41
SELECT_n_of_aggr_func__AVG	
0	959 (96%)
1	27 (2.7%)
2	12 (1.2%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_aggr_func__SUM	
0	926 (93%)
1	60 (6.0%)

Query parameter or value	Descriptive statistics
2	10 (1.0%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 9.00] 0.64 / 1.42
SELECT_n_of_all_aggr_func	[0.00, 2.00, 11.00] 2.70 / 2.26
FROM_n_of_join_paths	
1	325 (33%)
2	341 (34%)
3	334 (33%)
FROM_n_of_super_joins__RIGHT	
0	750 (75%)
1	227 (23%)
2	23 (2.3%)
FROM_n_of_super_joins__LEFT	
0	783 (78%)
1	206 (21%)
2	11 (1.1%)
FROM_n_of_super_joins__FULL	
0	766 (77%)
1	216 (22%)
2	18 (1.8%)
FROM_n_of_joins__INNER	
0	241 (24%)
1	260 (26%)
2	205 (21%)
3	140 (14%)
4	84 (8.4%)
5	48 (4.8%)
6	19 (1.9%)
7	2 (0.2%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	247 (25%)
1	261 (26%)
2	218 (22%)
3	151 (15%)
4	76 (7.6%)
5	35 (3.5%)
6	11 (1.1%)
8	1 (0.1%)
FROM_n_of_processed_rows	[5, 340,491,940, 1,147,549,266] 314,739,166 / 295,240,791
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.87 / 2.28
WHERE_n_of_attribs_of_type__integer	
0	444 (44%)
1	304 (30%)

Query parameter or value	Descriptive statistics
2	174 (17%)
3	56 (5.6%)
4	17 (1.7%)
5	5 (0.5%)
WHERE_n_of_attribs_of_type__character	
0	721 (72%)
1	216 (22%)
2	50 (5.0%)
3	11 (1.1%)
4	1 (0.1%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__numeric	
0	657 (66%)
1	262 (26%)
2	64 (6.4%)
3	15 (1.5%)
4	2 (0.2%)
WHERE_n_of_attribs_of_type__character_varying	
0	441 (44%)
1	295 (30%)
2	176 (18%)
3	62 (6.2%)
4	20 (2.0%)
5	4 (0.4%)
6	1 (0.1%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	828 (83%)
1	142 (14%)
2	24 (2.4%)
3	6 (0.6%)
WHERE_n_of_pkey_attribs	
0	605 (61%)
1	281 (28%)
2	90 (9.0%)
3	21 (2.1%)
4	3 (0.3%)
WHERE_n_of_connect_OR	[0.00, 1.00, 9.00] 1.44 / 1.74
WHERE_n_of_operators__in	
0	610 (61%)
1	280 (28%)
2	80 (8.0%)
3	25 (2.5%)
4	4 (0.4%)
5	1 (0.1%)
WHERE_n_of_operators__between	

Query parameter or value	Descriptive statistics
0	629 (63%)
1	255 (26%)
2	92 (9.2%)
3	18 (1.8%)
4	6 (0.6%)
WHERE_n_of_operators___greater_or_less	
0	401 (40%)
1	299 (30%)
2	152 (15%)
3	102 (10%)
4	34 (3.4%)
5	11 (1.1%)
6	1 (0.1%)
WHERE_n_of_operators___like	
0	838 (84%)
1	147 (15%)
2	12 (1.2%)
3	2 (0.2%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	966 (97%)
1	33 (3.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	976 (98%)
1	24 (2.4%)
WHERE_n_of_non_aggr_func___LOG	
0	935 (94%)
1	63 (6.3%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	906 (91%)
1	86 (8.6%)
2	8 (0.8%)
WHERE_n_of_non_aggr_func___DOW	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func___DAY	
0	962 (96%)
1	38 (3.8%)
WHERE_n_of_non_aggr_func___ROUND	
0	975 (98%)
1	25 (2.5%)
WHERE_n_of_non_aggr_func___YEAR	
0	977 (98%)
1	22 (2.2%)

Query parameter or value	Descriptive statistics
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	980 (98%)
1	18 (1.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__ABS	
0	984 (98%)
1	16 (1.6%)
WHERE_n_of_all_non_aggr_func	
0	707 (71%)
1	214 (21%)
2	69 (6.9%)
3	9 (0.9%)
4	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 9.00] 2.33 / 2.08
HAVING_n_of_main_predicates	
0	435 (44%)
1	182 (18%)
2	197 (20%)
3	186 (19%)
ORDER_BY_n_of_columns	
0	303 (30%)
1	372 (37%)
2	238 (24%)
3	52 (5.2%)
4	19 (1.9%)
5	12 (1.2%)
6	2 (0.2%)
7	2 (0.2%)
limit	[1, 453, 1,000] 469 / 289
offset	[0, 0, 998] 236 / 322