

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

## Scenario: 1-3jp\_ujpp\_\_g\_\_h0ss\_h0nss

## Scale factor: 10 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, 7, 98] 11 / 14
SELECT_n_of_non_aggr_func__DOW	
0	887 (89%)
1	88 (8.8%)
2	20 (2.0%)
3	5 (0.5%)
SELECT_n_of_non_aggr_func__FLOOR	
0	929 (93%)
1	62 (6.2%)
2	7 (0.7%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__LOG	
0	893 (89%)
1	88 (8.8%)
2	14 (1.4%)

Query parameter or value	Descriptive statistics
3	5 (0.5%)
SELECT_n_of_non_aggr_func__LOWER	
0	708 (71%)
1	206 (21%)
2	61 (6.1%)
3	17 (1.7%)
4	4 (0.4%)
5	3 (0.3%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	718 (72%)
1	210 (21%)
2	43 (4.3%)
3	20 (2.0%)
4	7 (0.7%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__MONTH	
0	844 (84%)
1	105 (11%)
2	33 (3.3%)
3	11 (1.1%)
4	4 (0.4%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__ROUND	
0	950 (95%)
1	44 (4.4%)
2	4 (0.4%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func__RTRIM	
0	689 (69%)
1	248 (25%)
2	46 (4.6%)
3	9 (0.9%)
4	5 (0.5%)
5	2 (0.2%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func__SQRT	
0	863 (86%)
1	99 (9.9%)
2	23 (2.3%)
3	8 (0.8%)
4	4 (0.4%)
5	1 (0.1%)
6	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	862 (86%)

Query parameter or value	Descriptive statistics
1	106 (11%)
2	22 (2.2%)
3	5 (0.5%)
4	3 (0.3%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__UPPER	
0	691 (69%)
1	229 (23%)
2	61 (6.1%)
3	13 (1.3%)
4	4 (0.4%)
5	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__SUBSTR	
0	677 (68%)
1	246 (25%)
2	59 (5.9%)
3	12 (1.2%)
4	4 (0.4%)
6	2 (0.2%)
SELECT__n_of__non__aggr__func__DAY	
0	902 (90%)
1	83 (8.3%)
2	14 (1.4%)
3	1 (0.1%)
SELECT__n_of__non__aggr__func__TRUNC	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)
SELECT__n_of__non__aggr__func__ABS	
0	937 (94%)
1	56 (5.6%)
2	6 (0.6%)
3	1 (0.1%)
SELECT__n_of__all__non__aggr__func	[0.0, 2.0, 41.0] 3.3 / 4.8
SELECT__n_of__aggr__func__COUNT_DISTINCT	[0.00, 0.00, 10.00] 0.55 / 1.33
SELECT__n_of__aggr__func__MAX	[0.00, 0.00, 10.00] 0.63 / 1.45
SELECT__n_of__aggr__func__COUNT	[0.00, 0.00, 12.00] 0.67 / 1.58
SELECT__n_of__aggr__func__MIN	[0.00, 0.00, 10.00] 0.71 / 1.54
SELECT__n_of__aggr__func__SUM	
0	941 (94%)
1	42 (4.2%)
2	14 (1.4%)
3	3 (0.3%)
SELECT__n_of__aggr__func__AVG	

Query parameter or value	Descriptive statistics
0	955 (96%)
1	37 (3.7%)
2	6 (0.6%)
3	2 (0.2%)
SELECT_n_of_all_aggr_func	[0.00, 2.00, 12.00] 2.69 / 2.34
FROM_n_of_join_paths	
1	328 (33%)
2	340 (34%)
3	332 (33%)
FROM_n_of_super_joins__RIGHT	
0	775 (78%)
1	204 (20%)
2	21 (2.1%)
FROM_n_of_super_joins__LEFT	
0	780 (78%)
1	195 (20%)
2	25 (2.5%)
FROM_n_of_super_joins__FULL	
0	779 (78%)
1	207 (21%)
2	14 (1.4%)
FROM_n_of_joins__INNER	
0	241 (24%)
1	283 (28%)
2	196 (20%)
3	142 (14%)
4	86 (8.6%)
5	22 (2.2%)
6	21 (2.1%)
7	8 (0.8%)
8	1 (0.1%)
FROM_n_of_joins__RIGHT	
0	255 (26%)
1	253 (25%)
2	204 (20%)
3	141 (14%)
4	87 (8.7%)
5	38 (3.8%)
6	15 (1.5%)
7	5 (0.5%)
8	2 (0.2%)
FROM_n_of_processed_rows	[5, 68,179,156, 229,441,534] 68,142,511 / 61,094,347
WHERE_n_of_predicates	[0.00, 2.00, 10.00] 2.78 / 2.19
WHERE_n_of_attribs_of_type__integer	
0	483 (48%)
1	296 (30%)

Query parameter or value	Descriptive statistics
2	150 (15%)
3	58 (5.8%)
4	12 (1.2%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	711 (71%)
1	223 (22%)
2	59 (5.9%)
3	7 (0.7%)
WHERE_n_of_attribs_of_type__numeric	
0	625 (63%)
1	279 (28%)
2	75 (7.5%)
3	18 (1.8%)
4	3 (0.3%)
WHERE_n_of_attribs_of_type__character_varying	
0	456 (46%)
1	322 (32%)
2	150 (15%)
3	54 (5.4%)
4	14 (1.4%)
5	4 (0.4%)
WHERE_n_of_attribs_of_type__date	
0	802 (80%)
1	160 (16%)
2	34 (3.4%)
3	4 (0.4%)
WHERE_n_of_pkey_attribs	
0	648 (65%)
1	253 (25%)
2	84 (8.4%)
3	14 (1.4%)
4	1 (0.1%)
WHERE_n_of_connect_OR	
0	464 (46%)
1	170 (17%)
2	140 (14%)
3	85 (8.5%)
4	79 (7.9%)
5	40 (4.0%)
6	19 (1.9%)
7	1 (0.1%)
8	2 (0.2%)
WHERE_n_of_operators__in	
0	614 (61%)
1	261 (26%)

Query parameter or value	Descriptive statistics
2	97 (9.7%)
3	26 (2.6%)
4	2 (0.2%)
WHERE_n_of_operators___greater_or_less	
0	429 (43%)
1	284 (28%)
2	167 (17%)
3	87 (8.7%)
4	25 (2.5%)
5	7 (0.7%)
6	1 (0.1%)
WHERE_n_of_operators___between	
0	635 (64%)
1	257 (26%)
2	82 (8.2%)
3	18 (1.8%)
4	8 (0.8%)
WHERE_n_of_operators___like	
0	840 (84%)
1	144 (14%)
2	16 (1.6%)
WHERE_n_of_non_aggr_func___SQRT	
0	912 (91%)
1	83 (8.3%)
2	5 (0.5%)
WHERE_n_of_non_aggr_func___FLOOR	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DOW	
0	966 (97%)
1	34 (3.4%)
WHERE_n_of_non_aggr_func___DAY	
0	959 (96%)
1	40 (4.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	959 (96%)
1	41 (4.1%)
WHERE_n_of_non_aggr_func___YEAR	
0	968 (97%)
1	31 (3.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___ROUND	
0	966 (97%)
1	34 (3.4%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func__ABS	
0	960 (96%)
1	39 (3.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__MONTH	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func__LOG	
0	945 (95%)
1	54 (5.4%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	667 (67%)
1	242 (24%)
2	74 (7.4%)
3	15 (1.5%)
4	2 (0.2%)
GROUP_BY_n_of_columns	[0.00, 2.00, 10.00] 2.30 / 2.12
HAVING_n_of_main_predicates	
0	466 (47%)
1	197 (20%)
2	159 (16%)
3	178 (18%)
ORDER_BY_n_of_columns	
0	320 (32%)
1	370 (37%)
2	225 (23%)
3	58 (5.8%)
4	16 (1.6%)
5	8 (0.8%)
6	2 (0.2%)
7	1 (0.1%)
limit	[1, 517, 1,000] 503 / 292
offset	[0, 0, 993] 237 / 314