

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

Scenario: 1-1jp__ng

Scale factor: 10 GB

The 1000-query pack was generated on: 2023-11-08

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, 40] 10 / 9
SELECT_n_of_non_aggr_func__LOG	
0	899 (90%)
1	82 (8.2%)
2	16 (1.6%)
3	3 (0.3%)
SELECT_n_of_non_aggr_func__SQRT	
0	840 (84%)
1	127 (13%)
2	25 (2.5%)
3	6 (0.6%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__TRUNC	
0	943 (94%)
1	55 (5.5%)

Query parameter or value	Descriptive statistics
2	2 (0.2%)
SELECT_n_of_non_aggr_func__LOWER	
0	769 (77%)
1	183 (18%)
2	35 (3.5%)
3	10 (1.0%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func__RTRIM	
0	773 (77%)
1	172 (17%)
2	47 (4.7%)
3	8 (0.8%)
SELECT_n_of_non_aggr_func__FLOOR	
0	932 (93%)
1	61 (6.1%)
2	7 (0.7%)
SELECT_n_of_non_aggr_func__ROUND	
0	929 (93%)
1	67 (6.7%)
2	3 (0.3%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__UPPER	
0	802 (80%)
1	162 (16%)
2	27 (2.7%)
3	7 (0.7%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func__LTRIM	
0	797 (80%)
1	161 (16%)
2	34 (3.4%)
3	7 (0.7%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	
0	794 (79%)
1	161 (16%)
2	31 (3.1%)
3	11 (1.1%)
4	3 (0.3%)
SELECT_n_of_non_aggr_func__DAY	
0	923 (92%)
1	61 (6.1%)
2	12 (1.2%)
3	2 (0.2%)
4	2 (0.2%)

Query parameter or value	Descriptive statistics
SELECT_n_of_non_aggr_func__MONTH	
0	852 (85%)
1	104 (10%)
2	28 (2.8%)
3	16 (1.6%)
SELECT_n_of_non_aggr_func__YEAR	
0	862 (86%)
1	98 (9.8%)
2	31 (3.1%)
3	9 (0.9%)
SELECT_n_of_non_aggr_func__DOW	
0	877 (88%)
1	92 (9.2%)
2	27 (2.7%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__ABS	
0	955 (96%)
1	41 (4.1%)
2	4 (0.4%)
SELECT_n_of_all_non_aggr_func	[0.0, 1.0, 24.0] 2.6 / 3.7
SELECT_n_of_aggr_func__MIN	
0	931 (93%)
1	28 (2.8%)
2	22 (2.2%)
3	15 (1.5%)
4	1 (0.1%)
5	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_aggr_func__COUNT_DISTINCT	
0	932 (93%)
1	31 (3.1%)
2	19 (1.9%)
3	9 (0.9%)
4	6 (0.6%)
5	3 (0.3%)
SELECT_n_of_aggr_func__MAX	
0	941 (94%)
1	24 (2.4%)
2	17 (1.7%)
3	13 (1.3%)
4	3 (0.3%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_aggr_func__COUNT	
0	947 (95%)

Query parameter or value	Descriptive statistics
1	31 (3.1%)
2	10 (1.0%)
3	9 (0.9%)
4	2 (0.2%)
6	1 (0.1%)
SELECT_n_of_aggr_func__AVG	
0	993 (99%)
1	7 (0.7%)
SELECT_n_of_aggr_func__SUM	
0	989 (99%)
1	9 (0.9%)
2	1 (0.1%)
3	1 (0.1%)
SELECT_n_of_all_aggr_func	
0	775 (78%)
1	82 (8.2%)
2	60 (6.0%)
3	49 (4.9%)
4	19 (1.9%)
5	10 (1.0%)
6	5 (0.5%)
FROM_n_of_join_paths	
1	1,000 (100%)
FROM_n_of_joins__INNER	
0	508 (51%)
1	320 (32%)
2	136 (14%)
3	32 (3.2%)
4	4 (0.4%)
FROM_n_of_joins__RIGHT	
0	513 (51%)
1	314 (31%)
2	139 (14%)
3	32 (3.2%)
4	2 (0.2%)
FROM_n_of_processed_rows	[5, 10,001,103, 76,480,513] 28,555,495 / 31,229,835
WHERE_n_of_predicates	
0	258 (26%)
1	233 (23%)
2	205 (21%)
3	165 (17%)
4	94 (9.4%)
5	36 (3.6%)
6	9 (0.9%)
WHERE_n_of_attris_of_type__character	
0	835 (84%)

Query parameter or value	Descriptive statistics
<i>1</i>	142 (14%)
<i>2</i>	21 (2.1%)
<i>3</i>	2 (0.2%)
WHERE_n_of_attribs_of_type__integer	
<i>0</i>	574 (57%)
<i>1</i>	325 (33%)
<i>2</i>	81 (8.1%)
<i>3</i>	18 (1.8%)
<i>4</i>	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
<i>0</i>	746 (75%)
<i>1</i>	211 (21%)
<i>2</i>	41 (4.1%)
<i>3</i>	2 (0.2%)
WHERE_n_of_attribs_of_type__character_varying	
<i>0</i>	563 (56%)
<i>1</i>	299 (30%)
<i>2</i>	115 (12%)
<i>3</i>	20 (2.0%)
<i>4</i>	3 (0.3%)
WHERE_n_of_attribs_of_type__date	
<i>0</i>	900 (90%)
<i>1</i>	91 (9.1%)
<i>2</i>	9 (0.9%)
WHERE_n_of_pkey_attribs	
<i>0</i>	723 (72%)
<i>1</i>	241 (24%)
<i>2</i>	34 (3.4%)
<i>3</i>	2 (0.2%)
WHERE_n_of_connect_OR	
<i>0</i>	622 (62%)
<i>1</i>	199 (20%)
<i>2</i>	113 (11%)
<i>3</i>	53 (5.3%)
<i>4</i>	10 (1.0%)
<i>5</i>	3 (0.3%)
WHERE_n_of_operators__between	
<i>0</i>	730 (73%)
<i>1</i>	220 (22%)
<i>2</i>	47 (4.7%)
<i>3</i>	3 (0.3%)
WHERE_n_of_operators__greater_or_less	
<i>0</i>	559 (56%)
<i>1</i>	274 (27%)
<i>2</i>	125 (13%)
<i>3</i>	33 (3.3%)

Query parameter or value	Descriptive statistics
4	9 (0.9%)
WHERE_n_of_operators__in	
0	739 (74%)
1	217 (22%)
2	42 (4.2%)
3	2 (0.2%)
WHERE_n_of_operators__like	
0	886 (89%)
1	105 (11%)
2	8 (0.8%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func__SQRT	
0	935 (94%)
1	64 (6.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func__FLOOR	
0	985 (99%)
1	15 (1.5%)
WHERE_n_of_non_aggr_func__LOG	
0	958 (96%)
1	42 (4.2%)
WHERE_n_of_non_aggr_func__ABS	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func__ROUND	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func__MONTH	
0	977 (98%)
1	23 (2.3%)
WHERE_n_of_non_aggr_func__TRUNC	
0	982 (98%)
1	18 (1.8%)
WHERE_n_of_non_aggr_func__DOW	
0	987 (99%)
1	13 (1.3%)
WHERE_n_of_non_aggr_func__YEAR	
0	991 (99%)
1	9 (0.9%)
WHERE_n_of_all_non_aggr_func	
0	776 (78%)
1	192 (19%)
2	31 (3.1%)

Query parameter or value	Descriptive statistics
<i>3</i>	1 (0.1%)
GROUP_BY_n_of_columns	
<i>0</i>	1,000 (100%)
HAVING_n_of_main_predicates	
<i>0</i>	1,000 (100%)
ORDER_BY_n_of_columns	
<i>0</i>	402 (40%)
<i>1</i>	306 (31%)
<i>2</i>	167 (17%)
<i>3</i>	85 (8.5%)
<i>4</i>	27 (2.7%)
<i>5</i>	13 (1.3%)
limit	[2, 522, 1,000] 506 / 283
offset	[0, 0, 1,000] 254 / 327