

Overview of changes estimating more accurate selectivity using frequency counts of String type columns. This is an extension of Histogram changes documented [here](#).

Frequency Count

For equality predicates on char and varchar column types in a table we will be using [ItemsSketch](#) from Apache Data Sketches maintaining frequency count of top 'N' number of values, much like Top-K queries.

If the predicate of the query has constants on either side of the equality, the selectivity will be computed using an estimate of the number of rows for that value.

The ItemsSketch gives a lower and upper bound of an estimated number of values for a given string constant with accuracy guarantees proportional to the *maxMapSize* and total number of rows in the partition or when merged, total number of rows in the table.

It is thus imperative to periodically adjust the *maxMapSize* per column as the table grows.

First Implementation

- Median of the upper and lower bounds of the estimation will be taken for selectivity.
- It is important to note that selectivity estimation should never cache the value outside of the sketch (at least that is maintained per partition) and should always do a lookup because depending on the fill ratio of the underlying hashmap the sketch decrements the counters by its median value.
- Therefore, for table level merged sketches (maintained in metastore?), error behavior needs to be backed with experimentation and most likely require more frequent merge compared to other statistics.
- *maxMapSize* will be a global configuration parameter across all string columns across all tables.
- *epsilon* (error bounds) can be a statistics of the ItemsSketch statistic that can be of interest for better fine tuning of the *maxMapSize* parameter.
- We will use error type NO_FALSE_POSITIVES for better accuracy and avoid type-I error. Meaning, if the item is in the list, it is indeed a highly frequent item and chances of being wrong is the least, however vice versa is more likely i.e., we may not have all the most frequent items in the list.
- Error type can be switched to NO_FALSE_NEGATIVES once the code stabilizes (deployed) and we see acceptable results of the selectivity.

Code Changes

The changes will be inlined to histogram changes of numeric column types [here](#). Mainly, following sections will be done (all except vectorization).

- Displaying Statistics
- Hive Config
- Metastore
- Direct SQL
- Statistics in-memory representation
- Statistics Aggregation
- ItemsSketch Estimator (*new*)

<TBD> How to capture per partition and table level sketch error (*epsilon*) statistic ?

Future Work

- Depending on the data (zipfian or normal), the sketch can show different error consistencies and thus tracking the epsilon can help in fine tuning the *maxMapSize* per column depending on the desired *alpha* on the stream of data.
- The above is particularly interesting for auto-tuning of *maxMapSize* for an unknown column distribution before and after presenting the data to the sketch. More relevant for transaction tables and transactions during bulk inserts/updates.
- Right now there is no way to accommodate bulk deletes or continuous delete overtime large enough violating the error bounds but invalidating the sketch and rebuilding it using full scan.
- Also, impact of changing the maxMapSize value (esp. reducing it) is not clear and thus might need a fallback to invalidate the sketch and rebuild it.
- Vectorization to build the sketch should be implemented before invalidating the sketch completely.