The Pie Chart Evaluation Metric

Authors:
Christine Task, Knexus Research Corporation
Issac Slavitt, DrivenData Inc.

Executive Summary

Please provide a 1-2 page, easily readable review of the main ideas. This is likely to be especially useful for people reading multiple submissions during the public voting phase. The executive summary should be readily understood by a technical layperson and include:

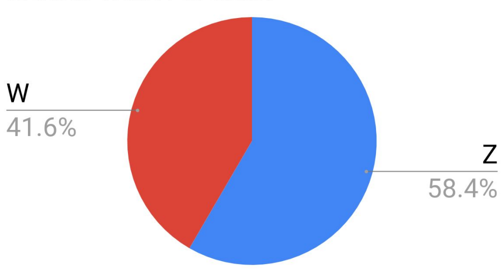
- *The high-level explanation of the proposed metric, reasoning and rationale for why it works*
- *An example real world use case*

Metric Overview:

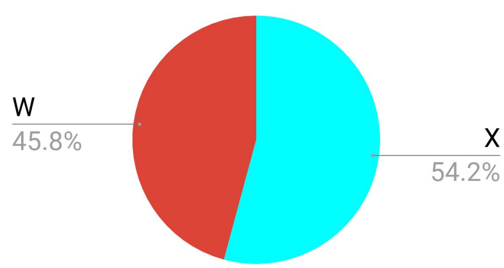
The objective of the pie chart is to measure how faithfully the privatization algorithm preserves the most significant patterns in the data, within each map/time segment. It does this by only considering the record types that make up at least k% of the total records (the 'sufficiently thick pie slices'). The less frequent events types are dropped from the data, and then a 'pie chart' is drawn up on the remaining frequent events. A toy example is given below, with event types X, Y, Z, W, and frequency threshold 5%.

Data-set	Raw Event Counts	Event Percentages	Frequent Event Counts
Ground Truth	[X:0, Y:2, Z:28, W:20]	[X:0%, Y:4%, Z:56%, W:40%]	[Z:28, W:20]
Privatized	[X:26, Y:0, Z:2, W:22]	[X:52%, Y:0%, Z:4%, W:44%]	[X:26, W:22]

Ground Truth Pie Chart



Privatized Pie Chart



Pie charts are drawn up for both the ground truth and privatized data in every map/time segment. These pie charts are next compared using the **Jensen-Shannon distance metric (JSD)**, which is essentially a symmetric version of K-L divergence (see technical background section for more details). A JS distance of 0 means the distributions match precisely; a JS distance of 1 means they differ completely. In this example we see that the distributions are similar for label W (41.6% vs

45.8%), but differ completely over labels Z and X.   The total JSD score (with base=2) for these two distributions is 0.75

However, the JS distance by itself doesn't capture the practical usability of the privatized pie chart. We include two additional score penalties for deviations that could impact decision making over the privatized results: Misleading Presence Penalty, and Bias Penalty.

In our toy example, label X appears in the privatized pie chart but not in the ground truth pie chart. Added privatization noise can cause an event that was infrequent in the ground truth data to appear as frequent in the privatized data, and these spurious label counts are misleading for decision makers. We add a **Misleading Presence Penalty (MPP)** to the JS distance for each label that exists in the privatized pie chart but not in the ground truth pie chart.

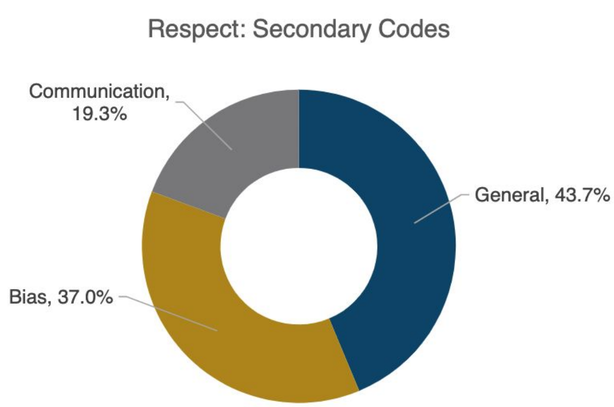
Pie charts are a good tool to confirm that the proportion of record types is maintained, but they hide the actual count of records.   It is possible for added privatization noise to increase the apparent number of events that occur, even while maintaining the relative proportions of frequent events.   To penalize positive bias in record counts, we include a **Bias Penalty (BP)** for map/time segments whose total record count varies significantly from the true record count.

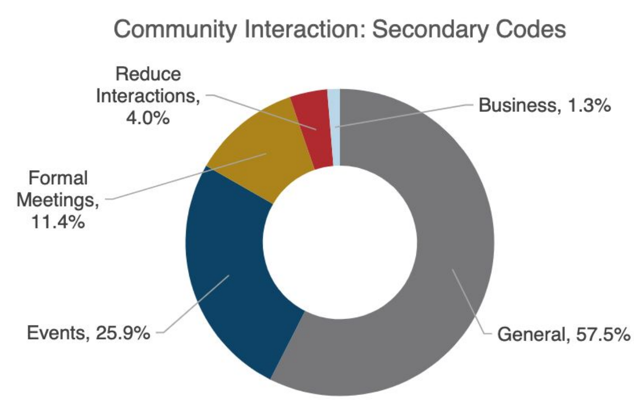
Because higher scores are nice to have for motivating competitions, the complete score for one pie chart is computed as **(1 - min(JSD + MPP + BP, 1)),** where 1.0 is a perfect score and 0.0 indicates a pie chart with so little utility as to be unusable**.**   Our total score for one full temporal map data set is the simple sum of pie-chart scores across all map and time segments.

Using default parameter values of MPP = 0.2, BP threshold = 500, and BP = 0.25, our toy example score is (1 - (0.75 + 0.2 + 0)) = .05

Real World Use Case:

Pie charts are often used by analysts and decision makers, when drawing up reports on public policy issues.   Relevant to the NIST Challenge's Sprint 1 use-case of Baltimore 911 Call and Police Incident Data, below we include examples from the Police Foundation's document: [Baltimore Police and Community Input Reports on Community Policing and Engagement](). This report, on potentially sensitive survey data about community engagement, primarily uses the pie chart format for reporting its results. The pie chart reporting format is common across studies performed by the Police Foundation for their member organizations.





Metric Definition

Please provide the following:

- *Any technical background information needed to understand the metric. Note that these metric write-ups should be accessible to technical experts from a diverse variety of disciplines. Please provide clear definitions of any terms/tools that are specific to your field, and provide a clear explanation for any properties that will be relevant to your metric definition or defense.*
- *A written definition of the metric, including English explanation and pseudocode or clear step-by-step instructions that have been clearly written and annotated with comments. Code can also be included (optionally) with the submission.*
- *Explanation of parameters and configurations. Note that this includes feature-specific configurations. For instance, a metric could reference "demographic features" or "financial features" for specific treatment, and given a new data set with a new schema, the appropriate features could be specified in a configuration file without loss of generalizability.*
- *Walk-through examples of metric use in snapshot mode (quickly computable summary score) and/or deep dive mode (generates reports locating significant points of disparity between the real and synthetic data distributions) as applicable to the metric.*

Technical Background

Jensen Shannon Distance:

The Jensen-Shannon distance (metric) is the square root of the Jensen-Shannon divergence. Given two probability vectors ***p*** and ***q***, the Jensen-Shannon distance is defined as,

$$\sqrt{\frac{D(p \parallel m)+D(q \parallel m)}{2}}$$

where ***m*** is the pointwise mean of ***p*** and ***q*** and ***D*** is the Kullback-Leibler divergence.

The Jensen Shannon Distance is based on the more commonly used KL divergence; however, unlike KL divergence it measures the symmetric distance between two distributions. A symmetric metric is important in our use case because due to positive privatization noise, the differentially private pie chart may include labels that do not appear in the ground truth data. KL divergence measures difference from a specified baseline distribution (ie, distance from the ground truth in our case) and is undefined at points where the baseline distribution is 0. This effectively means that the KL divergence is infinite whenever the privatized pie chart includes an extra label. Our metric penalizes spurious labels with the Misleading Presence Penalty… but an infinite penalty might be too harsh.

Additional resources:

- Scipy Library Documentation:
  https://scipy.github.io/devdocs/generated/scipy.spatial.distance.jensenshannon.html

- Detailed accessible explanation of KL divergence, Jensen Shannon Distance, and the relationship between the two: https://medium.com/datalab-log/measuring-the-statistical-similarity-between-two-samples-using-jensen-shannon-and-kullback-leibler-8d05af514b15

Formal Metric Definition

In this section we provide a detailed definition of the pie chart metric and a step-by-step walkthrough of how to compute it.

One row of the *differentially private* submission file $\mathbf{dp}_i$ represents a vector of non-negative incident counts for a particular neighborhood-month in the standard order (ascending by incident type code), while the corresponding *ground truth* vector $\mathbf{gt}_i$ represents the true counts. Here's a simplified example for one time period with only four incident counts:

gt	0	2	28	20
	0%	4%	56%	40%

dp	26	0	2	22
	52%	0%	4%	44%

We use a custom "pie chart" loss metric to compare the sets of counts. This metric is inspired by a typical public policy use case for aggregated data where insignificant entries below some percentage threshold are dropped and then the resultant entries are shown as a breakdown summing to 100%.

After removing insignificant counts, this metric is essentially based on the information-theoretic Jensen Shannon distance plus some penalties that try to capture undesirable properties from Differential Privacy subject matter expertise: (1) a misleading presence penalty (MPP) which quantifies the harm done to analyses where seemingly important counts appear but are purely an artifact of privatization, and (2) a bias penalty (BP) where the proportions of counts may be correct but the overall number are unreasonably far from the truth.

The difference between two vectors is evaluated using the following steps:

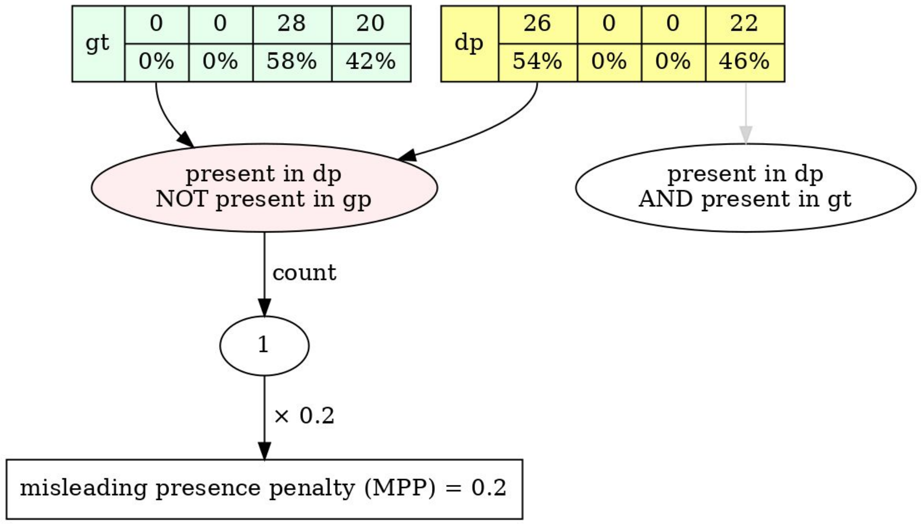
- **Zero out non-significant counts in each vector and re-normalize.** Any incident count in either $\mathbf{dp}_i$ or $\mathbf{gt}_i$ that accounts for less than FT% (Frequency Threshold) of the overall incidents in this vector is set to zero. In this example FT is set to 5%. After zeroing insignificant counts, each vector is divided by its new sum to get the proportions instead of raw counts.

gt	0	0	28	20
	0%	0%	58%	42%

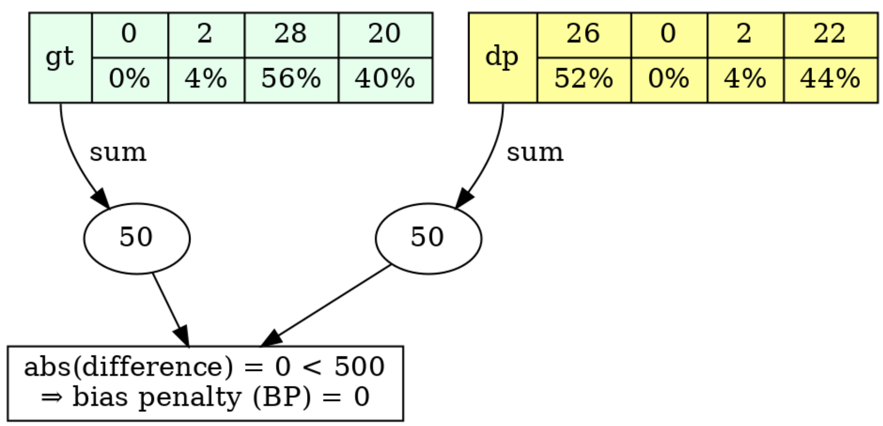
dp	26	0	0	22
	54%	0%	0%	46%

- **Calculate the individual score components:**
    - **Calculate the Jensen–Shannon distance (JSD)** between the normalized vectors. Divide each vector by its sum to get probability vectors and then calculate JSD. Using base-2, the JSD for these vectors is: 0.7505
    - **Add a misleading presence penalty (MPP)** for each time a category of incident shows up as significant in the privatized data but is actually zero in the ground truth

vector. In this example, we've set the additive penalty to 0.2 per misleading presence.



○ **Add a bias penalty (BP)** (in this example, BP = 0.25), if the sum of the ***original, pre-thresholding*** counts in **dp_i** is more than BPT (Bias Penalty Threshold) off from **gt_i**. For the example below, we've set BPT = 500.



- **Sum up the scoring components and clip to [0, 1].** Putting the pieces together, **$PieChart(dp_i, gt_i) = 1 - min(JSD(dp_i, gt_i) + MPP(dp_i, gt_i) + BP(dp_i, gt_i), 1.0)$** which means that the minimum score for each map segment and time period is 0.0, and given that each component is non-negative the overall score will therefore be between 0.0 and 1.0. In our toy example, the total score is 1 - min((0.75 + 0.2 + 0), 1.0) = 0.05.

The overall score for a submission will be the sum of all of these individual pie chart losses across all map and time segments. For the police incident data set we are using in Sprint 1 of the NIST Differential Privacy Temporal Map Challenge, this totals a maximum possible score of 1,008. However, the maximum score possible while preserving privacy will be lower.

Explanation of Metric Parameters:

In this section, we provide a list and quick explanation for each of the parameters that affect the performance of the metric. We first list settings that are specific to configuring the metric to work on a given input data schema. We then cover the tuning parameters that can impact coverage and discriminative power of the metric (independent of the schema).

Data Configuration Parameters:

Record Type:

The pie chart metric is based on a histogram of record types-- it uses a set of the counts of each 'type' of record, in each map and time segment. The precise definition of record type should be configured based on the input data schema.

In the example data-set for Sprint 1, the Baltimore police incident and 911 data, it's natural to define the record types as the call/incident descriptions (ex: Burglary, Traffic Accident, Elopement, etc).

However in a more complex data domain, there may be more options for configuring the record type. For example, given demographic data which includes information on Sex, Age and Race, the record type could be configured as a combination of all three features (ie, a record type label could be specified as: [F, 36, White]). This may not be the best approach, however. The pie chart metric focuses on patterns of high frequency record types, and if the data is spread out too thinly across many possible record types (i.e: 2 Sexes X 100 Ages x 6 Races = 1,200 possible record types), the pie chart metric may not function effectively. Instead the metric could be configured to use record types that reference fewer features (ie, marginal counts). Selecting Sex and Race (ie, a record type would be [F, White]) will result in a total of 12 possible record types and will likely capture meaningful frequent patterns in the data. Numerical variables can be binned (collected into ranges), to reduce the number of possible values. Selecting Age, but binning into 5-year bits (ie, a record type would be [35-40]), results in a total of 20 possible record types. Depending on the choice of record type definition, the pie chart metric will evaluate different aspects of the privatized data quality.

Tuning Parameters:

Frequency Threshold: FT

The pie chart metric begins by removing all record types that occur 'infrequently'. The frequency threshold is a parameter that can be set; smaller thresholds will tend to have higher discriminative power (scoring more harshly the differences between the privatized and ground truth data), higher thresholds will tend to focus more narrowly on maintaining the most significant patterns in the data. In the example above (and in the Challenge) the Frequency Threshold has been set to 5%. In the model defense we explore the impact of other possible threshold values.

MPP Penalty Amount: MPP

The metric penalizes privatized pie charts which include labels that don't appear in the ground truth pie chart (Misleading Presence Penalty). The amount of this penalty is a parameter that can be set, with higher penalties scoring more harshly for deviations from the ground truth data. For the example above, and the Challenge, we've set MPP = 0.2, intuitively docking a "letter grade" for every spurious label in the privatized pie chart.

Bias Penalty Threshold: BPT

In addition to evaluating the record type ratios that appear in the pie chart, the metric penalizes data where the total record count in a given time/map segment varies 'significantly' (beyond a specified threshold) from the ground truth data. This threshold is a parameter that can be set; lower values will be more sensitive to added privacy noise.   For the first Sprint of the Challenge, we've generously selected BPT = 500, based on the distribution of positive bias on the baseline solution at epsilon = 2.   *This may be changed for later sprints.*

Bias Penalty Amount: BP

Finally, the amount of the penalty incurred by violating the Bias Penalty Threshold is a parameter that can be set.   Higher penalties will score more harshly for the impact of privacy noise on record totals; this may be especially noticeable in privacy algorithms that independently add noise to a large number of record types, or ground truth data sets with many counts at or near zero (if the algorithm simply rounds up negative privatized values to zero). For this challenge we've set BP = 0.25.

Snapshot and Deep Dive Modes:

In this section we describe how the Pie Chart metric can be used to either give a single total data quality score for a privatized temporal map (Snapshot Mode), or to investigate and pinpoint sources of disparity between the privatized and ground truth data (Deep Dive Mode).

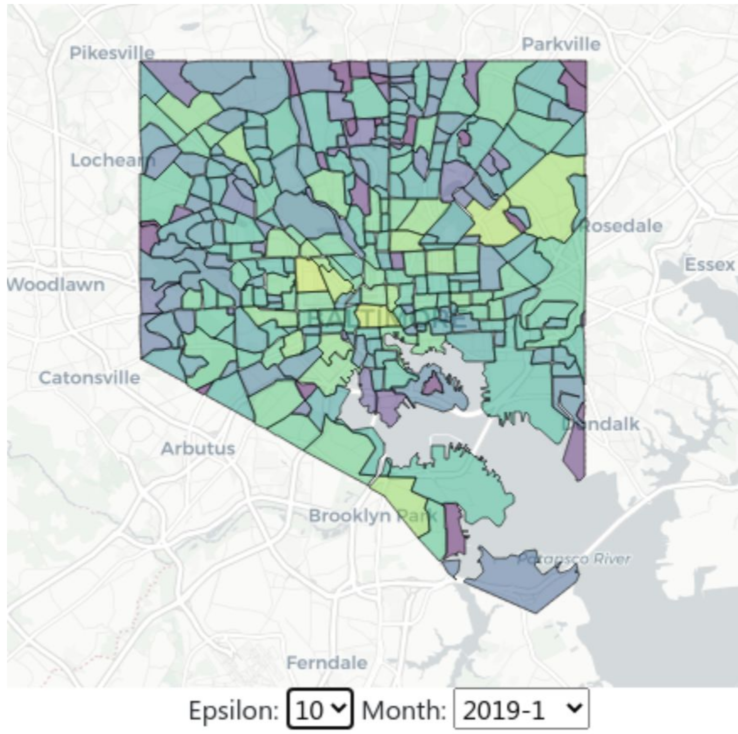
Snapshot Mode:

The basic pie chart metric gives a score between 0.0 and 1.0 for a given map/time segment. To get a single score for aloof the map/time segments in a full temporal map, we propose simply summing the scores for the individual pie charts.   This will result in a minimum score of 0.0 and a maximum score equal to [total number of map segments] x [total number of time segments]. The score, intuitively, is proportional (*but not equal*) to the number of map/time segments that performed 'well'.

Deep Dive Mode:

Because the pie chart metric scores each map/time segment separately, it provides clear utility for deep dive investigation. The visualizer we provide for the challenge leverages this.

The **Interactive Map** allows you to see your scores geographically (across all map segments). Here we see that dense urban neighborhoods closer to the city center, which generally contain more records, have better scores than rural and suburban neighborhoods where records may be more sparse. These are challenges that will need to be creatively overcome to achieve good performance on the Sprint 1 task.



The **Temporal Scores Chart** allows you to select a given neighborhood and see the change in your pie chart scores in that neighborhood over each of the time segments. Here we see the scores are relatively uniform across months for our baseline privacy algorithm. However, a privacy algorithm that leverages the temporal aspect of the problem, for example by aggregating counts across multiple time segments, might see more interesting variation here.

Remington (213)

year	month	score
2019	1	0.6073
2019	2	0.6631
2019	3	0.6635
2019	4	0.6849
2019	5	0.6235
2019	6	0.6508
2019	7	0.6536
2019	8	0.5685
2019	9	0.5944
2019	10	0.6263
2019	11	0.6558
2019	12	0.6480

Metric Defense

Please provide the following:

- *The metric's tuning properties that control the focus, breadth, and rigor of evaluation.*
- *The discriminative power of the proposed metric: how well it identifies points of disparity between the ground truth and privatized data*
- *The coverage properties of the proposed metric: how well it abstracts/covers a breadth of uses for the data*
- *The feasibility of implementing the proposed metric. For instance, what is the computation time and resource requirements for the metric when running on data? How does the metric scale with an increase in variables, map segments, time segments, and records? This information may include empirical results (e.g. runtime) or theoretical results (e.g. mathematical properties). Feel free to provide assumptions about hardware (e.g. CPU model, memory, operating system) and feature constraints.*
- *Examples of 2-3 very different data applications where metric can be used.*

Exploration of Parameter Tuning:

Data:

The data set used for these analyses was the 2019 Baltimore Police Incident and 911 data-set provided as publicly available data in Sprint 1 of the NIST Differential Privacy Temporal Map Challenge. We look at the distribution of pie chart scores on a subsample of individual neighborhoods/months, selected uniformly randomly from the ground truth data. We generate privatized data using the baseline privacy algorithm at 3 scales of noise addition. The 'scale' refers to the mean absolute privacy noise value sampled from the laplacian distribution (ie, the scale parameter of the laplace distribution). Larger scales imply noisier data, and better privacy (smaller epsilons): We consider scale = 2 (epsilon = 10), Scale = 10 (epsilon = 2), Scale = 20 (epsilon = 1).

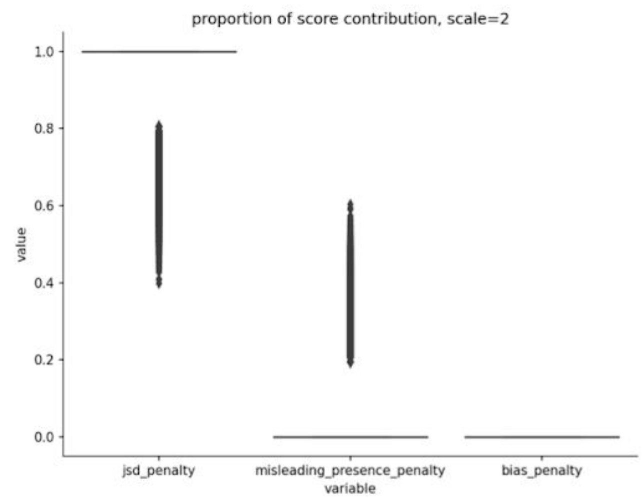
Metric Parameters:

Except where otherwise specified, the pie chart metric parameters have been set to the defaults given in the Metric Definition (FT = 0.05, MPP = 0.2, BPT = 500, BP = 0.25).

Score Composition:

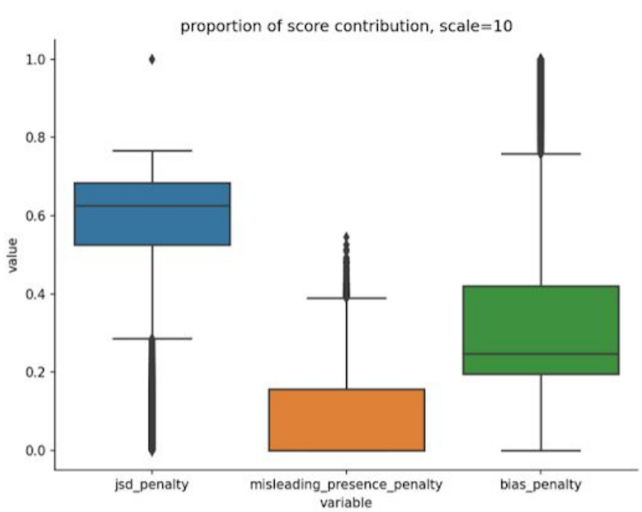
The Pie Chart metric has three components (as described in the Metric Definition section above): JSD, MPP, BP. Here we look at the impact each component has on the total score, dependent on the quality of the data.

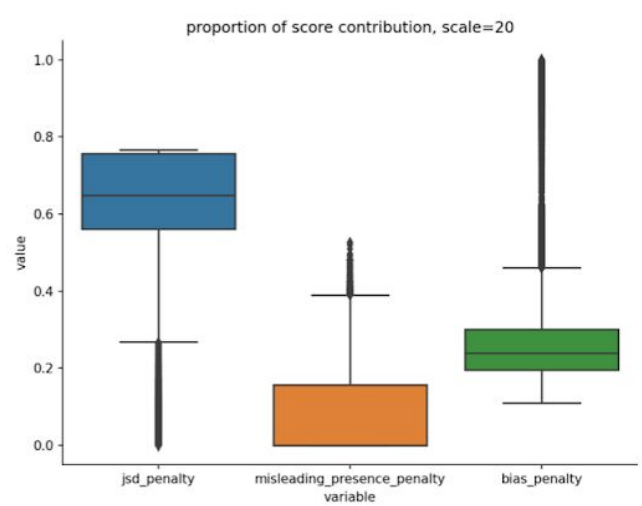
At scale 2 (where most record counts will have noise addition in the approximate range +/- 2), we see that the vast majority of the score is determined by the Jensen-Shannon distance metric that compares the privatized and ground truth pie charts. There are relatively few instances of a pie chart label appearing in the privatized data when it was not in the ground truth data (MPP), and, unsurprisingly no instances of the total count of records in the privatized data varying from the ground truth data by more than 500 (BP).



At scale = 10 (epsilon = 2), we have a stronger level of privacy protection, but the baseline privacy algorithm is showing its failure points. We are seeing more frequent occurrence of labels appearing in the privatized pie charts that are not in the ground truth pie charts (MPP) as larger added noise values have a better chance of pushing counts above the frequency threshold. We also see a significant increase in the bias penalty for privatized data that differs from the ground truth by more than 500 records. (Note to privacy algorithm competitors-- because we are adding many noise values in the +/- 10 range to counts across nearly 200 different record types, this is unsurprising. Although the Laplace noise is symmetric around 0, so we are adding both positive and negative noise, the baseline privacy solution simply rounds up negative counts to 0. This leaves a noise distribution with a significant positive bias).

At scale = 20 (epsilon = 1), we see that there are still issues with both spurious labels (MPP) and positive bias (BP), but the Jenson-Shannon distance itself JSD) is now making up a larger portion of the total score. This is likely occurring as large noise values begin to impact the pie slice widths (ie, the normalized distribution of high frequency record types) even among the more frequent record types and in map/time segments with more total records.





Effect of Frequency Threshold:

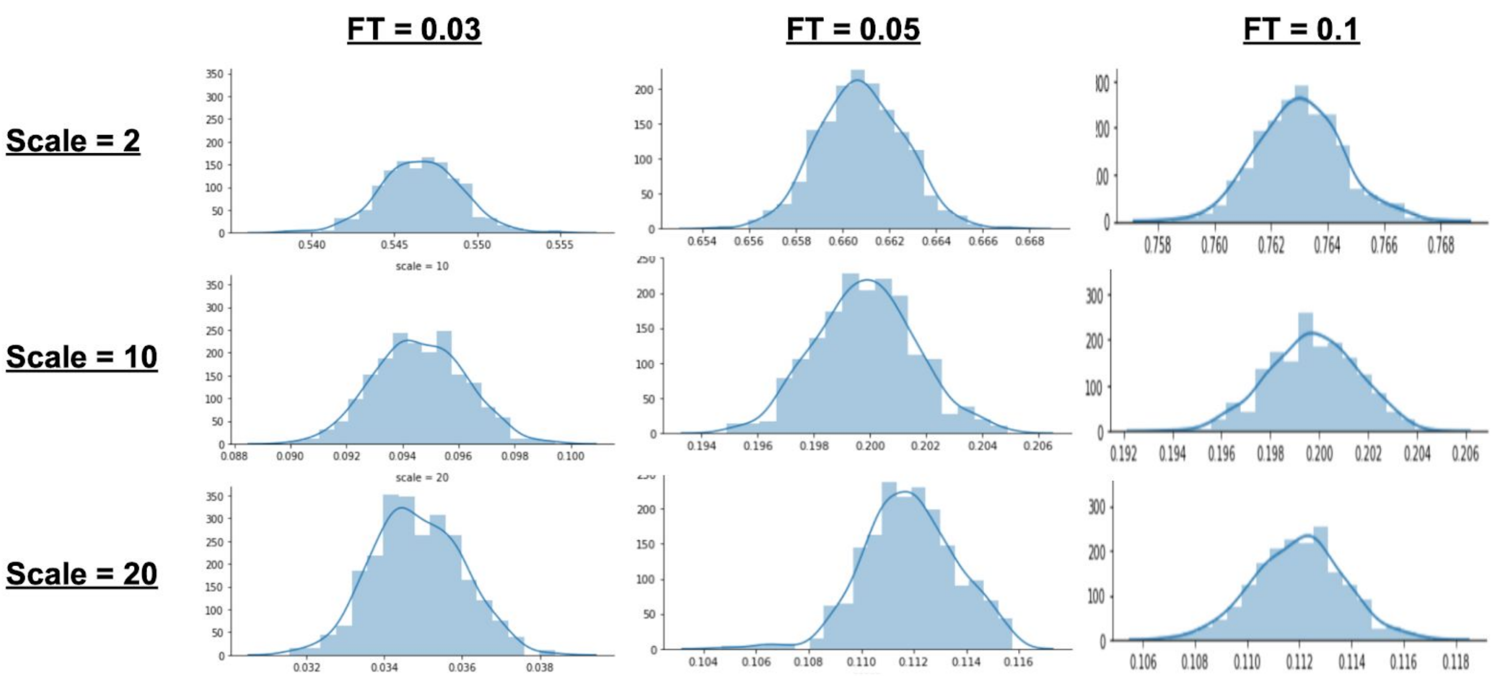
We now briefly explore the effect of increasing and decreasing the frequency threshold from the default value 5% (FT = 0.05).

At the 5% (FT = 0.05) threshold, most scores for single neighborhood/month pie charts are around 0.66 for scale 2; they lower to around 0.2 for scale 10 (once the MPP and BP penalties begin to be incurred more often), and they fall to around 0.11 for scale 20 as the quality degrades somewhat further.

However, if we lower the frequency threshold to 3% (FT = 0.03), the ground truth pie chart will include 'thinner' pie slices, and thus the metric becomes more sensitive to smaller amounts of added privacy noise. At scale 2, scores fall around 0.55, at scale 10 they drop further to 0.095, and at scale 20 the privatized pie charts begin to very little utility, with scores closer to 0.03.

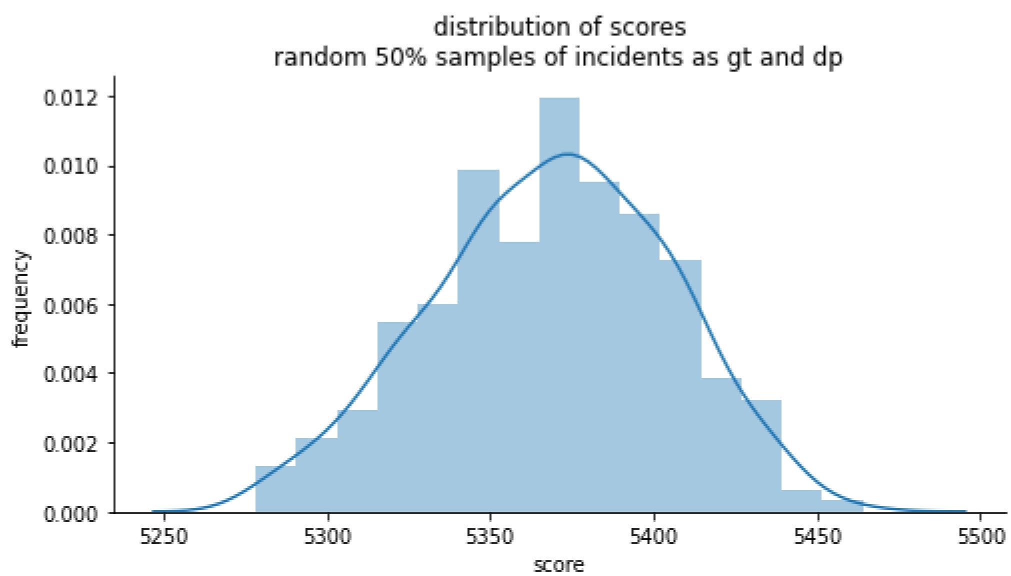
Raising the frequency threshold to 10% (FT = 0.10) will narrow the focus to only the more frequent event types (the thicker pie slices). This decreases the sensitivity of the metric to added privacy noise and has the impact of improving the Jensen Shannon distance for the smaller noise scales. At scale 2, these scores range around 0.762, significantly better than the scores at the 5% threshold. For the larger noise scales, though, where MPP and BP penalties impact the score, the scores are essentially the same as the 5% threshold (ranging around 0.2 for scale 10, and 0.112 for scale 20).

As 5% is more a common threshold than 10% in real world pie charts, we chose the 5% default.



50% Sampling Error Benchmark:

Finally we apply the pie chart metric to a 50% sampling error benchmark. When a metric has high discriminative power and is very sensitive to discrepancies between the privatized and real data, it may be difficult to get a sense of what a "good" metric score is. We can get a benchmark score for good performance by comparing the noise in our privatized results to sampling error. In this case, we take two uniform random 50% subsamples of the full incident record data, and then arbitrarily choose one to treat as ground truth and another as the privatized data, and we compute the pie chart metric score (across all neighborhood/months). This effectively gives us the difference between two views of the same ground truth data. If the added privacy error is less than the sampling-error benchmark, the added privacy noise is comparable to variation typically encountered due to sampling. The distribution below was created with repetitions of the 50% subsampling procedure. We see that it actually performs worse on average than the 5% and 10% threshold pie chart scores at scale 2. This is likely due to the sensitivity of the pie chart metric to small differences in distributions in neighborhood/months with few event records.



Further Questions:

At our challenge launch deadline rapidly approaches, our analyses on the tuning properties of the Pie Chart metric are limited to the above. However, there are other interesting questions that could be explored with respect to this metric:

- How does the pie chart metric score change as the number of possible record types increases, given a similar power-law distribution of data?
- If we increase the MPP to 1.0 (effectively having zero tolerance for spurious labels), how does that affect the score distribution? How does it affect the 50% sampling error benchmark?
- What is the Jensen Shannon Distance score for distributions that closely match on all but one record label? What is the maximum error a single record label can have while maintaining a total Jensen Shannon score above the 50% sampling error benchmark?

- How sensitive is the Pie Chart metric to relatively small differences in map/time segments that have very few records? To what extent does setting the FT higher mitigate this effect?
- Can we define a minimum score threshold for “trustworthy” privatized pie charts, and then classify map/time segments as “feasible to privatize” by whether the sampling error benchmark achieves a trustworthy score on those segments?

Submissions to the metric challenge (with ample time between the challenge launch and the January submission deadline), should feel welcome to more fully explore interesting properties of their own metrics, either theoretically or empirically as appropriate. Questions relating to edge cases, impact on practical use cases, and useful or unexpected properties are all of interest.

Description of Discriminative Power:

Every metric will have both capabilities and limitations; no single metric will capture all possible definitions of utility. The objective of this section is to clearly understand and summarize the properties of the metric.

In this section we briefly outline some of the capabilities and limitations of the pie chart metric with respect to its discriminative power-- how well it can distinguish between the ground truth and privatized data.

Capabilities

- The pie chart metric identifies disparities between the distribution of high frequency record types
- The pie chart metric specifically penalizes positive (or negative) bias in total record counts
- The pie chart metric specifically penalizes when rare large noise values are sampled from the laplace distribution, causing spurious record types to appear to be high frequency in the privatized data.
- We’ve demonstrated that the pie chart metric responds to small changes in the value of epsilon (or sampling error), and allows us to meaningfully understand the impact of those changes on the data.

Limitations

- The pie chart metric does not measure the impact of privatization noise on patterns in less frequent record types; those record types are discarded during the frequency thresholding.
- The pie chart metric does not capture the relative ranking of high frequency record types. In general, Jensen-Shannon distance summarizes the total difference in distributions; two similarly-sized pie slices may change order without significantly impacting the JS score.
- The pie chart metric does not measure trends across time. Scores are summed across all time/map segments without attention to any broader patterns.
- The pie chart metric focuses on record *types*, effectively categorical information; if the data includes numerical features these must be partitioned into bins, and the width of those bins will impact the discriminative power of the metric.

Description of Coverage:

Every metric will have both capabilities and limitations; no single metric will capture all possible definitions of utility. The objective of this section is to clearly understand and summarize the properties of the metric.

In this section we briefly outline some of the capabilities and limitations of the pie chart metric with respect to its coverage-- how well it represents a breadth of possible use cases.

Capabilities

- Pie charts are a commonly used tool for communicating survey results in public policy and other decision making contexts. They provide the reader with a quick sense of the 'most significant' simple features of the data. The pie chart metric evaluates whether the privatized data will be suitable for these common, basic applications.
- Because the pie chart metric evaluates the relative proportion of frequent record types, a high pie chart score is an indicator that the privatized that will maintain utility for policy decisions on questions like funding, fairness, and staff/resource placement (although we have not explicitly studied the relationship between the pie chart score and funding disbursement).

Limitations

- The pie chart metric doesn't cover more complex analytics or machine learning tasks (ex: regression, classification) which may have different sensitivities to added noise.
- As mentioned above, the pie chart metric doesn't cover applications that study larger patterns/trends across time or geography
- The pie chart metric focuses on categorical information (record types), and does not evaluate for many criteria specific to numerical data, such as whether the privatized data maintains long tailed distributions or maintains accuracy across repeated numerical operations over privatized data ("differences of differences").
- The pie chart metric doesn't cover applications that reference infrequent record types.

Scalability/Feasibility:

The pie chart metric is a constant time operation in terms of the number of record types, map segments and time segments. The pie chart metric on the 196 record types and 10,008 map/time segments in the Sprint 1 Baltimore Police Data can be computed within seconds on a typical laptop.

Generalizability, Alternate Use Cases:

In this document, we've demonstrated how the pie chart metric can apply to event records. Here are a few additional examples how the pie chart metric could be applied:

- The pie chart metric can apply to demographic data, financial data, etc, by using marginals (and binning numerical features) when defining the "record type", as outlined in the Parameters-Configuration, in the Metric Definition.

- Although by default the pie chart metric does not capture trends across time or geography, it could be applied to capture more temporal information. By setting the record types to be "Events that happened between timestamp x_1 and timestamp x_2", a pie chart can be drawn up to measure clusters of records in time the same way we've used it above to measure distributions in feature space. Similarly, the pie chart metric could be used to capture frequent patterns geographically, by setting the record types to be "Events that happened in region X". Combinations of time and geography are also possible.

.

...And that's it!
We hope this challenge gets you thinking creatively and deeply about the interplay of data, space and time, and the possibility of capturing the meaningful patterns in the data without being overly sensitive to any individual data point.