Metric Testing

Expected Output Tests

Discernability Metric (DM)

Test Case 1

Input data:

10	a	x	red
10	a	y	red
10	b	x	blue
20	b	y	red
20	b	z	green
30	a	z	green

Test:

Calculate the DM penalty for the input data set, with attrib1, attrib2 & attrib3 as QIDs. Note this is a non-anonymised data set, thus it contains no equivalence classes

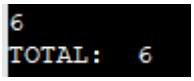
Pass Conditions:

- Actual DM penalty matches expected DM penalty

Expected output:

6

Actual output:



Passed: Yes

Test Case 2

Input data:

10	['a', 'b']	['x', 'y']	blue
10	['a', 'b']	['x', 'y']	orange
10	['a', 'b']	['x', 'y']	red
[20, 30]	['a', 'b']	['y', 'z']	red
[20, 30]	['a', 'b']	['y', 'z']	green
[20, 30]	['a', 'b']	['y', 'z']	yellow

Test:

Calculate the DM penalty for the anonymised input data set, with attrib1, attrib2 & attrib3 as QIDs.

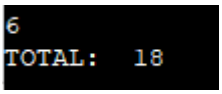
Pass Conditions:

- Actual DM penalty matches expected DM penalty

Expected output:

18

Actual output:



Passed: Yes

ILoss Metric

Test Case 3

Input data:

10	a	x	red
10	a	y	red
10	b	x	blue
20	b	y	red
20	b	z	green
30	a	z	green

Test:

Calculate the ILoss penalty for the input data set, with attrib1, attrib2 & attrib3 as QIDs. Sanity check – the input data set is non-anonymised and therefore has no generalisations to penalise in the calculation

Pass Conditions:

- Actual ILoss penalty matches expected ILoss penalty

Expected output:

0

Actual output:

```
TOTAL:  0
```

Passed: Yes

Test Case 4

Input data:

10	['a', 'b']	['x', 'y']	blue
10	['a', 'b']	['x', 'y']	orange
10	['a', 'b']	['x', 'y']	red
[20, 30]	['a', 'b']	['y', 'z']	red
[20, 30]	['a', 'b']	['y', 'z']	green
[20, 30]	['a', 'b']	['y', 'z']	yellow

Test:

Calculate the ILoss penalty for the anonymised input data set, with attrib1, attrib2 & attrib3 as QIDs

Pass Conditions:

- Actual ILoss penalty matches expected ILoss penalty

Expected output:

6

Actual output:

```
TOTAL:  6.0
```

Passed: Yes

Classification Metric (CM)

Test Case 5

Input data:

10	a	x	red
10	a	y	red
10	b	x	blue
20	b	y	red
20	b	z	green
30	a	z	green

Test:

Calculate the CM penalty for the input data set, with attrib1, attrib2 & attrib3 as QIDs. Sanity check – the input data set is non-anonymised and therefore has no equivalence classes to check for homogeneity.

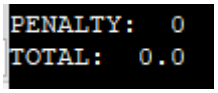
Pass Conditions:

- Actual CM penalty matches expected CM penalty

Expected output:

0

Actual output:



Passed: Yes

Test Case 6

Input data:

10	['a', 'b']	['x', 'y']	blue
10	['a', 'b']	['x', 'y']	blue
10	['a', 'b']	['x', 'y']	red
[20, 30]	['a', 'b']	['y', 'z']	red
[20, 30]	['a', 'b']	['y', 'z']	red
[20, 30]	['a', 'b']	['y', 'z']	blue

Test:

Calculate the CM penalty for the anonymised input data set, with attrib1, attrib2 & attrib3 as QIDs

Pass Conditions:

- Actual CM penalty matches expected CM penalty

Expected output:

0.3333

Output:

```
PENALTY:  2
TOTAL:  0.3333333333333333
```

Passed: Yes