

Banding Analysis Technical Guidance:

This paper describes the indicators contained in the PPSG banding analysis, the methods used to grade the indicators, both for delivery compared to the most similar group (MSG) and direction of travel compared to the previous year's data. It also describes how the grades for each domain are aggregated to give totals for the three domains: confidence, local crime and policing, and protection from serious harm.

Interpreting the report

'Ladders' are shown for each performance indicator included in the banding analysis. Grades for individual indicators are then combined to give overall domain grades for each domain.

Domain grades are shown on the blue ladders, and that domain's supporting indicators shown to the right. Where insufficient or no data (see low volume section) is available for a particular indicator, the ladder is not shown and the indicator does not contribute to either the direction or delivery domain grades for that particular force.

Data sets within the banding are as follows:

- British Crime Survey (BCS) Confidence
- User Satisfaction
- Crimes and Detections
- OBTJ
- Road Traffic data
- Gun crime
- Knife crime

Delivery assessment

A box with three or four levels, depending on the type of indicator, is used to show the assessment of delivery (typically relative to peers). The top compartment (dark green) shows the better performers, and so on down to the worst performers coloured as pink.

The relative position within the band of the force's performance on each non-survey indicator is shown by a solid horizontal line. Indicators using British Crime Survey (BCS) data and some of the user satisfaction indicators are shown on three bands; the two top bands of green are combined. Assessment of delivery is based, where possible, on 12 months' performance data using the most recent data available.

Assessments are shown with reference to a force's current position in their MSG, so that over time a downward movement in level can result from an improvement in peers as well as a decline in the performance of the force concerned.

Direction assessment

Direction of travel within the banding is gauged with reference to the position a year previously. An improvement in performance (e.g. crime reduction) is shown by an upwards arrow, deterioration by a downwards arrow. The diamond symbol indicates either no change or small change compared to the spread of changes observed in the 43 forces. Assessment of direction uses data for this period compared with the 12 months a year earlier.

The confidence domain

The confidence domain contains indicators that are based on the results of the BCS and User Satisfaction surveys carried out by forces.

In general, the BCS survey results have large confidence intervals; therefore these indicators are graded for delivery purely on whether the difference from the MSG are statistically significant. This means they can only be graded on a three point scale of 'poor', 'fair' and 'good/excellent'. The "User Satisfaction with whole experience" results have much smaller confidence intervals, and therefore absolute differences are considered – resulting in the usual four point scale.

Comparative Satisfaction

The comparative satisfaction for White versus BME direction score is calculated differently from other indicators within the confidence domain. The following criteria are used:

- Good/excellent - if there is no significant difference between the rates at the 90% level of significance
- Fair - significant difference at the 90% level
- Poor – if there is a statistical difference at the 95% level of significance.

For forces with very large sample sizes, these grade boundaries produce very small absolute differences in rates. Though this may represent a *statistically* significant difference, in reality the differences are negligible. We have set a minimum absolute difference in rates of $\pm 2.5\%$. Differences below this cut-off will be graded Good/Excellent.

For forces with small sample sizes, the confidence intervals are too large to determine reliably that a Good/Excellent grade should be awarded. We have set a maximum confidence interval of $\pm 10\%$. Forces that do not have a significant difference at the 90% level but have a confidence interval greater than this cut-off will not be graded for this indicator. These forces will also not be graded for direction of travel unless the change from last year is significant (at the 95 % level).

The local crime and policing and protection from serious harm domains

Indicators in this domain are based on recorded crime.

Delivery

Scores in these domains are calculated as the ratio of the force performance to its MSG average. These are transformed using the logarithmic function to make the distribution more regular.

The formula used to transform the data is as follows:

$$=\text{LOG}(\text{Force Performance}/\text{MSG Average})$$

This means that regardless of the initial level of performance an equal percentage change in performance will result in the same change in score relative to the MSG average

The distribution of force scores is then used to determine the grade boundaries. The distribution is assumed to be normal with a mean of 100% (i.e. average performance against the MSG comparator) and a standard deviation calculated from force scores. These are therefore obviously different for each iteration of the banding.

The boundaries were selected such that the expected distribution of Excellent, Good, Fair and Poor grades would be as follows: 20%:25%:45%:10%. For example the boundary between Poor and Fair is the boundary between the bottom 10% of the distribution and the upper 90% if the distribution were perfectly normal.

Direction

The method for direction grading is the same as delivery grading.

Scores for most recorded data are calculated as the ratio of the force performance to last year's performance. These are transformed using the logarithmic function to make the distribution more regular.

The formula used to transform the data is as follows:

$$=\text{LOG}(\text{current RY Performance}/\text{Previous RY Performance})$$

This means that regardless of the initial level of performance, an equal percentage change in performance will result in the same change in score relative to last year,

The distribution of force scores is then used to determine the grade boundaries. The distribution is assumed to be normal with a mean of 100%. A standard deviation is then calculated from force scores.

The boundaries have been selected such that you would expect the distribution of improved / stable / deteriorated to be 33% each. The boundaries for each indicator will be different, being dependent on the standard deviation, or the spread of force scores.

Low volume indicators

It is important to note that indicators within local crime and policing and protection from serious hard domains are subject to low volume checks. An indicator may not be scored for either direction or delivery if the volumes for the indicator are particularly low. An example of this is frequently the knife crime indicator.

On every indicator in these two domains a result is calculated for each force by taking the logarithm of the force crime rate divided by the force's MSG average crime rate (or rate according to indicator type i.e. sanction detection, OBTJ etc). The standard deviation of these results indicates the spread of the data.

Separately to this the Poisson error is calculated for each indicator for every force. It is calculated on both direction and delivery scores. The error for each indicator is calculated as the square root of the number of crimes (or sanction detection, OBTJ etc).

Finally if the Poisson error is more than two times the spread of the scores, then the grade / direction is found to be a 'low volume' indicator and the force remains ungraded for that particular indicator.

Grade aggregation to domain scores

A numeric value is given to each component grade in the domain.

The average value for all the indicators in a domain is then calculated and used to determine the domain grade.

Delivery

Grade	Indicator Value	Criteria for domain grade
Excellent	3	>2
Good/Excellent	2.5	n/a (confidence domain only)
Good	2	>1.5 < 2
Fair	1	> 2/3 <1.5
Poor	-1	<2/3

Direction

Grade	Indicator Value	Criteria for domain grade
Improved	1	> 0.25
Stable	0	> -0.25 < 0.25
Deteriorated	-1	< -0.25

Weighting

Not every indicator has an equal weighting on the domain total. The breakdown is as follows:

Indicator	Delivery	Direction
Confidence		
Confidence in dealing with concerns – agencies	5 (Half domain weighting)	5 (Half domain weighting)
User Satisfaction (whole experience)	1	1
Comparative Satisfaction (white/BME)	1 If graded	1 If graded
User Satisfaction (racist incidents)	1	1
Confidence in dealing with concerns (police)	1	1
ASB perceptions	1	1
Local crime and policing		
BCS comparator crime	1	1
Serious acquisitive crime	1	1
Domestic burglary	No	No
Robbery	No	No
Vehicle crime	No	No
Other wounding	1	1
SAC OBTJ	0.5	0.5
SAC Sanction detections	0.5	0.5
Racially religiously aggravated sanction detections	1 if passes low volume test	1 if passes low volume test
Road traffic casualties	1	1
Protection from serious harm		
Violence against the person with injury	1	1
Gun crime	1, if passes low volume test	1, if passes low volume test
Knife crime	1	1
MSV OBTJ	No	No
Violence against the person sanction detections	1	1
Serious sexual OBTJ	0.5	0.5
Serious sexual sanction detections	0.5, if passes low volume test	0.5, if passes low volume test