

The following document provides details on the methodology used to calculate the median and mean series for the H2 table.

#### **Outlier exclusion:**

The data series in H2 table are currently not treated for outliers. We will review our analysis methodology once we have collected more data on these series and decide whether an outlier removal process is appropriate.

The distribution of responses is widely spread for some of the data series. Therefore there is a substantial difference between the published mean and median values as the mean data is skewed by a relatively small number of highly positive responses.

#### **Median calculations:**

The median estimates are the weighted median, obtained using the positive weights  $w_i$  associated with the responses  $x_i$ .

The weighted median is calculated by ordering the responses  $x_i$ , for  $1 < i < n$ , in ascending order so that

$$x_1 \leq x_2 \leq \dots \leq x_n$$

where each response is associated with a positive weights  $w_i$  such that  $\sum_{i=1}^n w_i = 1$ . Weights  $w_i$  are obtained using the RIM (Random Iterated Method) weightings of each respondent.

The weighted median is then the response  $x_k$  where  $k$  satisfies the following two conditions:

- (i)  $\sum_{i=1}^{k-1} w_i \leq \frac{1}{2}$
- (ii)  $\sum_{i=k+1}^n w_i \leq \frac{1}{2}$

If two responses satisfy conditions (i) and (ii) then the average of the two values is taken.

#### **Mean calculations:**

The mean estimates are the weighted average, obtained using the positive weights  $w_i$  associated with the responses  $x_i$ .

The weighted mean  $\bar{X}_w$  is calculated as:

$$\bar{X}_w = \frac{\sum_i w_i x_i}{\sum_i w_i}$$

where  $x_i$  and  $w_i$  are the response and weight, respectively, of respondent  $i$ .