Weighted Average Bid Price with Outlier Detection

Weighted Average Bid Price: This calculates the average bid price, accounting for the quantities of each bid. Each bid price is multiplied by its corresponding quantity, summed across all bids, and divided by the total sum of quantities.

Formula:

$$Weighted\ Average = \frac{\sum(p_i \cdot q_i)}{\sum q_i}$$

where p_i is the bid price and q_i is the quantity for the $(i) - th$ bid.

Explanation: This gives a fair average that reflects the influence of each bid based on its quantity.

Weighted Standard Deviation: This measures the spread of bid prices from the weighted average, weighted by quantities. For each bid, the squared difference between the bid price and the weighted average is multiplied by the quantity, summed across all bids, divided by the total quantity, and the square root is taken.

Formula:

$$Weighted\ Standard\ Deviation = \sqrt{\frac{\sum q_i \cdot (p_i - \mu)^2}{\sum q_i}}$$

where μ is the weighted average, p_i is the bid price and q_i is the quantity for the $(i) - th$ bid.

Explanation: This shows how much bid prices typically vary from the weighted average, considering quantities.

Outlier Detection: A bid is flagged as an outlier if its price deviates from the weighted average by more than the weighted standard deviation.

Formula:

$$Outlier\ if\ |p_i - \mu| > \sigma$$

where p_i is the bid price, μ is the weighted average, and σ is the weighted standard deviation. If true, the bid is labeled an "Outlier"; otherwise, it's "Not Outlier."

Explanation: This identifies bids with prices unusually far from the weighted average, based on the typical variation.

This method ensures bids are evaluated with quantities in mind, highlighting prices that are significantly higher or lower than the typical pattern.

Locally Estimated Scatterplot Smoothing – LOESS: This is the method used to create a smooth curve through a scatterplot of data points. LOESS is useful when the relationship between variables is non-linear or complex. Since bid prices don't always have a linear relationship with quantity, i.e., don't always follow a straight line as quantity increases, LOESS helps understand the general trends between bid prices and quantity without assuming a specific mathematical model/formula.

How it works:

- LOESS fits a local/simple model (a line or a curve) that best matches the nearby points (i.e., bid price and quantity).
- Data points closer to the target point have more weight, thus influencing the model more than distant points.
- Local/simple models are then combined to create a smooth curve across the entire dataset.

Confidence interval:

- A 95% confidence interval bands show the range of expected variation in bid price for a given quantity based on historical data and observed trends.
- It indicates that 95% of similar bids are likely expected to be within this interval band, making the prices outside the band statistically unusual.
- For any quantity, the LOESS model estimates a central trend and determines a range around it where 95% of similar bids are expected to fall, implying a 95% confidence that the actual bid price for that quantity lies within this interval, assuming the data follows the same pattern.
- When a bid price falls outside the confidence interval, it's considered unusual, either too high (overpriced) or too low (underpriced) compared with the typical price for that quantity.