The situation involved two data sources: Source X, which provided complete trade data but with a significant lag, and Source Y, which offered incomplete data for certain products but with no lag.

Our task was to compare the data from Source Y with Source X to identify which products had complete numbers from Source Y. For the products where Source Y provided incomplete numbers, we needed to determine if the difference between the data from Source X and Source Y was stable. If the difference was stable, we could develop a proxy based on this stable relationship.

To determine if a proxy is good, we need to calculate the maximum possible error based on historical data. For example, based on 10 years of data, in May, the average difference between Y and X is 10%. The maximum difference observed during May was 20%. So if we use the 10% average for May, based on the historical data, the maximum error that we would expect would be 10% (20% maximum – 10% average).