ANOMALY DETECTION

Anomalies of a time series data can be defined as outliers of the **remainders** once the linear trend and seasonal periodicity are taken into account.

We use library "anomalize" in R to identify anomalies. We call the function `time_decompose()` with the option `method="stl"` (factoring the linear and seasonal components), and the function `anomalize()` with the option `method="iqr"`. By default, this method defines outliers as observations lying below $Q1 - 3 \cdot IQR$ or above $Q3 + 3 \cdot IQR$, where $Q1$ is the first quartile (25th percentile), $Q3$ is the third quartile (75th percentile), and the interquartile range is $IQR = Q3 - Q1$.

The default setting can be changed by specifying a value for alpha other than 0.05 (option `"alpha="`). Outliers are defined as values that lie $0.15/\alpha \cdot IQR$ distance away from the quartiles. The default value of $\alpha = 0.05$, thus resulting in the multiplicative constant of 3. If alpha is increased, more observations become outliers. If alpha is decreased, fewer observations are labeled as outliers.