As séries temporais são conjuntos de dados observados sequencialmente ao longo do tempo. Elas podem surgir em inúmeros contextos e podem ser analisadas para realização de previsão de tendências futuras. Entretanto, essas séries podem conter outliers, que são observações anômalas que não seguem o padrão geral dos dados. Esses outliers podem distorcer a análise e prejudicar a precisão das previsões, portanto, é essencial tratá-los adequadamente.

Durante a aula, utilizamos o intervalo de confiança da previsão da série temporal, a partir de uma linha de mínimo e máximo, e removemos os valores discrepantes que estavam além destas linhas, tornando os dados de aprendizado mais consistentes. Este tipo de tratamento de outliers é conhecido como **remoção de outliers**.

Podemos utilizar diversos critérios como ponto de partida para a remoção de outliers, como por exemplo, se estiverem afastados a 3 desvios-padrões da média ou a partir de um boxplot, removendo os dados que estão representados por pontos a 1,5 vezes a distância interquartílica, que seria a largura da caixa do boxplot. Cada um desses critérios devem ser selecionados de acordo com a análise. De todo modo, esses tipos de tratamento se encaixam no mesmo tipo de remoção de outliers. Será que essa é a única estratégia possível para lidar com valores discrepantes?

## **Substituição por estimativas**

Uma opção possível para o tratamento de outliers é a substituição deles por estimativas. Isso faz com que a informação não seja completamente perdida. Os outliers podem ser substituídos por valores considerados mais adequados, como a média ou a mediana dos dados não influenciados pelos outliers.

Outra abordagem é substituir os outliers por estimativas mais razoáveis com base em modelos estatísticos, como regressão linear ou modelos ARIMA (modelo auto-regressivo integrado de médias móveis).

## **Truncagem**

A truncagem consiste na definição de um valor limite inferior e superior em que o valor discrepante será substituído. Esse tipo de estratégia também consiste na substituição do valor ao invés da remoção, mas considera esses valores de mínimo e máximo que podem ser atingidos pelos dados para que fiquem mais consistentes. Os mesmos critérios da remoção podem ser adotados, como limites sendo 3 desvios-padrões da média ou análise por boxplot.

## **Transformações**

As transformações de dados permitem que os dados fiquem menos esparsos e mais agrupados, o que pode fazer com que o impacto dos outliers seja reduzido. Transformações matemáticas nos dados, como logarítmica, raiz quadrada ou inversa, podem reduzir a influência dos outliers e tornar os dados mais adequados para análise.

## **Modelos robustos**

Além de todas as estratégias que podem ser adotadas, o uso de modelos robustos, que são menos sensíveis a outliers, pode ser uma estratégia eficaz para lidar com esses pontos discrepantes sem a necessidade de tratamento explícito.

Portanto, o tratamento de outliers em análises de séries temporais é muito importante para garantir a precisão das previsões, mas é importante destacar que não existe uma abordagem única que sirva para todos os casos. Essa escolha vai depender da natureza dos dados e contexto do problema. Além disso, é importante ter cuidado ao lidar com outliers, uma vez que removê-los pode levar à perda de informações importantes. Com isso, é essencial realizar uma análise cuidadosa antes de decidir sobre a melhor estratégia de tratamento de outliers em séries temporais.