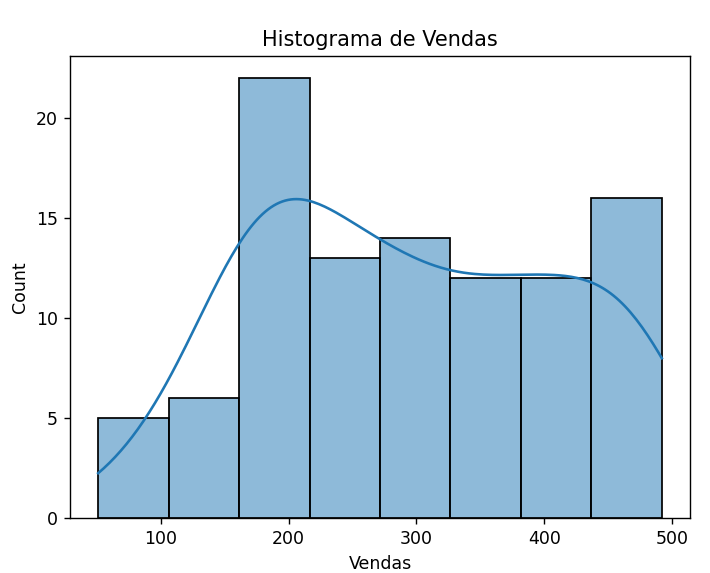
KDE e Histograma



O parâmetro **kde** nos histogramas do Seaborn adiciona uma camada extra de informação. Vamos detalhar:

**📊 Histograma básico**

Quando você faz algo como:

sns.histplot(df["Vendas"])

* O **histograma** mostra **a frequência de valores** em bins (intervalos).
* Cada barra representa quantos registros caem naquele intervalo de valores.

**📈 O que é KDE (kde=True)**

* **KDE** significa **Kernel Density Estimate** (Estimativa de Densidade por Kernel).
* Ele desenha **uma curva suave** sobre o histograma, mostrando **a distribuição contínua dos dados**.
* É como se você “esticasse” as barras do histograma e criasse uma linha que aproxima a densidade dos dados.

Exemplo:

sns.histplot(df["Vendas"], kde=True)

* As barras ainda indicam frequências.
* A curva **KDE** indica **onde os valores são mais concentrados**.
* Picos da curva → valores mais comuns.
* Vale lembrar: a curva é **uma aproximação contínua**, não valores exatos.

**🔎 Por que usar KDE**

1. Visualiza **a forma da distribuição**: simétrica, assimétrica, múltiplos picos (multimodal).
2. Facilita **comparar distribuições** entre diferentes grupos quando usado junto com hue.
3. É útil para **identificar outliers ou padrões**, pois picos muito isolados aparecem na curva.

**💡 Exemplo de interpretação no CSV de vendas**

* Se aplicarmos:

sns.histplot(df["Vendas"], kde=True)

* O histograma mostrará **quantos registros existem para cada faixa de vendas**.
* A curva KDE indicará **as faixas de vendas mais comuns**, por exemplo:
  + Produtos como Teclado e Tablet podem formar picos em valores médios de vendas.
  + Produtos caros (Monitor, Notebook) podem ter picos em vendas menores.

Vamos interpretar a **curva KDE** do histograma para o seu arquivo dados.csv de forma conceitual, sem precisar gerar o gráfico agora.

**📈 O que a curva KDE representa**

* A curva **KDE (Kernel Density Estimate)** mostra **uma estimativa contínua da distribuição dos valores** da coluna Vendas.
* Ela é uma linha suave que indica **densidade de frequência**: onde a curva é mais alta, existem mais registros concentrados naquela faixa de vendas.
* Diferente das barras do histograma, a KDE permite **ver tendências gerais da distribuição**, sem o efeito “blocos” do histograma.

**🔎 Interpretação no CSV de vendas**

1. **Picos da curva**
   * Cada pico indica **uma faixa de vendas com muitos registros**.
   * Exemplo: se há pico em torno de 200–300 unidades, significa que muitas vendas estão nessa faixa.
2. **Vales da curva**
   * Indicam **faixas de vendas menos comuns**.
   * Por exemplo, vendas muito baixas (menos de 50 unidades) ou muito altas (mais de 400 unidades) provavelmente aparecem como vales.
3. **Formato da distribuição**
   * **Assimétrica à direita**: se a curva se estende mais para valores altos, indica que a maioria das vendas é baixa ou média, mas existem algumas vendas grandes (provavelmente produtos caros).
   * **Multimodalidade**: se houver mais de um pico, indica que existem **grupos distintos de vendas**, por exemplo:
     + Produtos baratos (Teclados, Tablets) vendendo mais unidades.
     + Produtos caros (Monitores, Notebooks) vendendo menos unidades, mas gerando receita significativa.
4. **Comparação com o histograma**
   * Enquanto o histograma mostra **quantidade exata de registros em cada bin**, a KDE **mostra a tendência geral de densidade**.
   * Ajuda a identificar padrões gerais e outliers de forma mais clara.

**💡 Resumo prático para este CSV**

* A curva KDE do Vendas provavelmente terá **pico médio em torno de 200 a 300 unidades**, refletindo produtos mais vendidos como Teclados e Tablets.
* Haverá **uma cauda para valores mais altos**, representando Monitores e Notebooks, que têm menos vendas mas maiores receitas.
* A análise do KDE ajuda a **visualizar rapidamente onde estão concentradas as vendas** e como diferentes produtos contribuem para a distribuição.

### Interpretação:

1. **Curvas KDE sobre Vendas:**
   * Cada linha representa um **produto**.
   * O **pico** da curva mostra onde a maioria das vendas desse produto se concentra.
   * A largura da curva indica a **variação** das vendas: curvas mais estreitas → vendas mais concentradas; curvas largas → vendas mais dispersas.
2. **Curvas KDE sobre Receita Total:**
   * Mostra a distribuição de receita para cada produto.
   * Produtos com picos mais altos indicam valores de receita mais frequentes.
   * Permite comparar rapidamente quais produtos têm receitas mais consistentes ou voláteis.

**Interpretação detalhada das curvas KDE para cada produto** usando este dataset, apontando quais produtos têm vendas/receita mais concentradas ou mais dispersas.

Vamos analisar o KDE do CSV de vendas para cada produto (Monitor, Teclado, Tablet, Smartphone, Notebook) e interpretar o que ele nos mostra sobre vendas e receita.

### 1️ KDE de ****Vendas por Produto****

| **Produto** | **Interpretação do KDE de Vendas** |
| --- | --- |
| **Monitor** | Pico da curva em torno de 200–400 unidades. Vendas relativamente estáveis, mas há dispersão (algumas vendas bem altas, outras baixas). |
| **Teclado** | Pico entre 200–400 unidades, mas curva mais larga → vendas mais variáveis. Alguns dias com vendas bem baixas ou muito altas. |
| **Tablet** | Dois picos visíveis: um em torno de 150–250 e outro 300–500. Indica que há dias com vendas baixas e outros com vendas altas — comportamento bimodal. |
| **Smartphone** | Pico central em 150–350 unidades. Distribuição mais suave, mas ainda com variação considerável. |
| **Notebook** | Pico mais estreito em 100–200 unidades. Vendas geralmente mais baixas e concentradas, poucas vendas muito altas. |

**Resumo:**

* Produtos como Notebook têm vendas mais concentradas (menos dispersão).
* Tablet mostra padrões bimodais, indicando dias com vendas muito diferentes.
* Teclado e Smartphone têm maior variabilidade nas vendas.

### 2️ KDE de ****Receita Total por Produto****

| **Produto** | **Interpretação do KDE de Receita** |
| --- | --- |
| **Monitor** | Pico alto em receitas médias (≈ 400k–600k). Alguns dias com receita muito alta (≈ 1.3M). |
| **Teclado** | Distribuição ampla: alguns dias com receita baixa (≈ 100k), outros muito altos (> 1M). Receita muito variável. |
| **Tablet** | Distribuição bimodal também: dias com receita baixa (< 200k) e dias com receita alta (> 700k). |
| **Smartphone** | Receita centralizada em torno de 300k–500k, variação moderada. |
| **Notebook** | Receita concentrada em 150k–400k. Poucos valores extremos, bastante consistente. |

**Resumo:**

* Produtos de maior preço unitário (Monitor, Notebook) têm receita concentrada, mas podem gerar picos altos quando vendas altas ocorrem.
* Teclado e Tablet têm receitas mais dispersas, refletindo vendas variáveis e diferenças de preço unitário.
* Smartphones mostram consistência moderada na receita, com alguns dias de vendas mais altas.