

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS**

Escola Politécnica  
Engenharia de Computação

**Igor Parreira de Paula**

# Análise de Regressão Linear: Fatores de Influência no Consumo de Cerveja

**Disciplina:** Inteligência Artificial (2025.2-CMP1110/A01)

Goiânia

2025

**Igor Parreira de Paula**

# Análise de Regressão Linear: Fatores de Influência no Consumo de Cerveja

Relatório apresentado como requisito de avaliação na disciplina de Inteligência Artificial, referente ao estudo de caso de análise exploratória e regressão linear em dados de consumo.

Goiânia

2025

## **Sumário**

<b>1</b>	<b>INTRODUÇÃO</b>	<b>2</b>
1.1	OBJETIVOS . . . . .	2
<b>2</b>	<b>ANÁLISE EXPLORATÓRIA DOS DADOS</b>	<b>2</b>
2.1	ESTATÍSTICAS DESCRIPTIVAS . . . . .	2
2.2	VISUALIZAÇÃO DE DADOS . . . . .	2
<b>3</b>	<b>MODELAGEM E DIAGNÓSTICO</b>	<b>3</b>
3.1	DIAGNÓSTICO DE MULTICOLINEARIDADE . . . . .	3
3.2	RESULTADOS DO MODELO DE REGRESSÃO (OLS) . . . . .	3
3.3	AVALIAÇÃO DE DESEMPENHO . . . . .	5
<b>4</b>	<b>ANÁLISE DE RESÍDUOS</b>	<b>5</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>5</b>

# 1 INTRODUÇÃO

A previsão de demanda é crucial para o planejamento estratégico em diversos setores industriais. No contexto de uma cervejaria, entender as variáveis que influenciam o consumo diário permite otimizar a produção e a logística.

Este trabalho apresenta uma Análise Exploratória de Dados (AED) seguida pela construção de um modelo de Regressão Linear Múltipla para prever o consumo de cerveja (em litros) baseando-se em variáveis climáticas e temporais.

## 1.1 OBJETIVOS

- Realizar a análise descritiva das variáveis: temperatura média, precipitação e ocorrência de final de semana.
- Diagnosticar possíveis problemas de multicolinearidade entre as variáveis preditoras.
- Treinar e avaliar um modelo OLS (*Ordinary Least Squares*) para quantificar a influência de cada variável no consumo.
- Analisar os resíduos para validar as premissas estatísticas do modelo.

## 2 ANÁLISE EXPLORATÓRIA DOS DADOS

### 2.1 ESTATÍSTICAS DESCRIPTIVAS

A base de dados contém 365 observações (um ano de registros). O consumo médio de cerveja é de aproximadamente 25.40 litros, com uma variação padrão de 4.39 litros. As variáveis independentes analisadas foram Temperatura Média, Precipitação e Indicador de Final de Semana.

### 2.2 VISUALIZAÇÃO DE DADOS

As visualizações iniciais indicam fortes correlações. O gráfico de dispersão (Figura 1) sugere uma relação linear positiva entre a temperatura média e o consumo de cerveja. Em contrapartida, a precipitação parece ter uma correlação negativa leve ou dispersa.

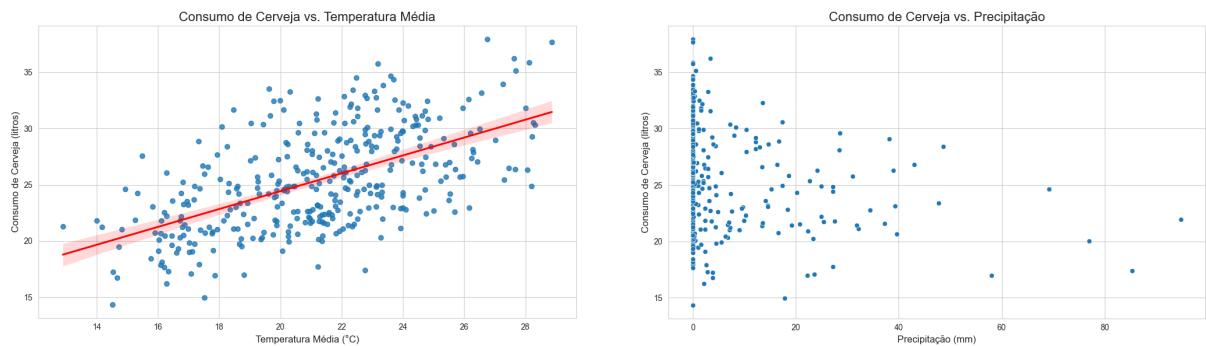


Figura 1: Esquerda: Consumo vs. Temperatura Média. Direita: Consumo vs. Precipitação.

Além disso, a análise por tipo de dia (Figura 2) através de Boxplots demonstra claramente que o consumo é significativamente maior aos finais de semana.

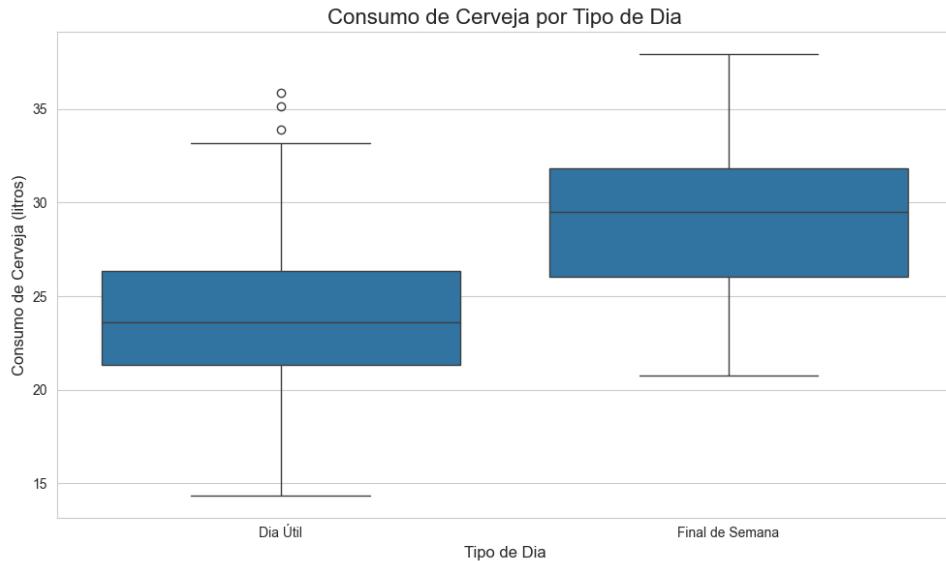


Figura 2: Boxplot: Comparaçāo do Consumo em Dias Úteis vs. Finais de Semana.

### 3 MODELAGEM E DIAGNÓSTICO

#### 3.1 DIAGNÓSTICO DE MULTICOLINEARIDADE

Para garantir a estabilidade dos coeficientes da regressão, calculou-se o Fator de Inflação da Variância (VIF). Os valores encontrados (todos abaixo de 5) indicam ausência de multicolinearidade severa, permitindo o uso seguro das variáveis no modelo.

#### 3.2 RESULTADOS DO MODELO DE REGRESSĀO (OLS)

O modelo de regressão linear foi ajustado utilizando o método dos Mínimos Quadrados. O sumário estatístico gerado pelo software pode ser visualizado na Figura 3.

```

OLS Regression Results
=====
Dep. Variable: consumo_cerveja R-squared: 0.661
Model: OLS Adj. R-squared: 0.658
Method: Least Squares F-statistic: 234.8
Date: Thu, 18 Sep 2025 Prob (F-statistic): 1.79e-84
Time: 17:25:49 Log-Likelihood: -860.61
No. Observations: 365 AIC: 1729.
Df Residuals: 361 BIC: 1745.
Df Model: 3
Covariance Type: nonrobust
=====
            coef    std err      t   P>|t| [0.025
0.975]
-----
const      6.4735    0.920    7.040    0.000    4.665
8.282
temperatura_media 0.8397    0.042   19.782    0.000    0.756
0.923
precipitacao -0.0742    0.011   -6.835    0.000   -0.096
-0.053
final_de_semana 5.2279    0.299   17.511    0.000    4.641
5.815
-----
Omnibus: 19.647 Durbin-Watson: 1.853
Prob(Omnibus): 0.000 Jarque-Bera (JB): 10.385
Skew: 0.227 Prob(JB): 0.00556
Kurtosis: 2.309 Cond. No. 153.
-----
Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.

```

Figura 3: Sumário Estatístico do Modelo OLS (Statsmodels).

O modelo apresenta um  $R^2$  de **0.661**, explicando aproximadamente 66.1% da variabilidade do consumo. A equação estimada é:

$$Consumo = 6.47 + 0.84 \times (Temp) - 0.07 \times (Precip) + 5.23 \times (FDS)$$

#### Interpretação dos Coeficientes:

- **Temperatura:** Para cada aumento de 1°C, o consumo aumenta em média 0.84 litros.
- **Chuva:** Para cada mm de chuva, o consumo cai 0.07 litros.
- **Final de Semana:** Nos fins de semana, o consumo aumenta em média 5.23 litros.

### 3.3 AVALIAÇÃO DE DESEMPENHO

- RMSE (Raiz do Erro Quadrático Médio): 2.708
- $R^2$ : 0.674 (na avaliação)

## 4 ANÁLISE DE RESÍDUOS

A validação das premissas da regressão linear foi feita através da análise gráfica dos resíduos (Figura 4).

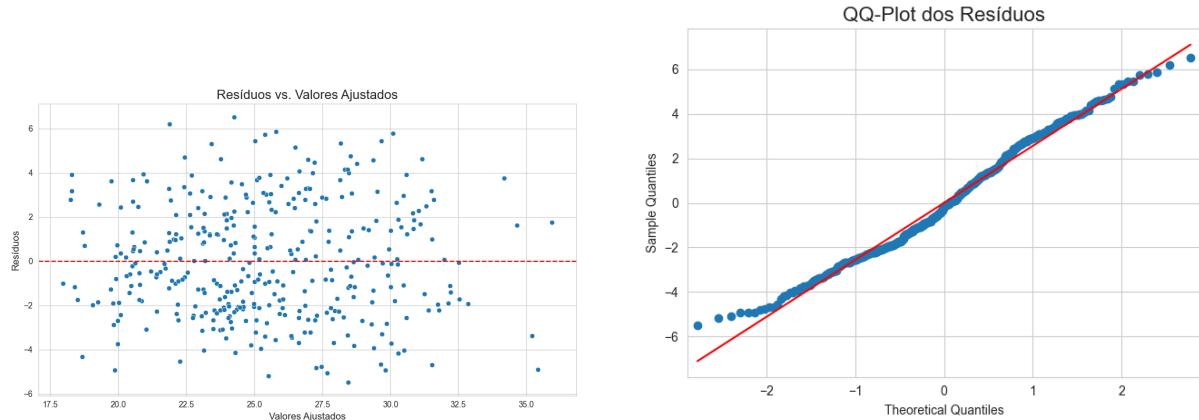


Figura 4: Esquerda: Resíduos vs. Valores Ajustados. Direita: QQ-Plot dos Resíduos.

- **Resíduos vs. Valores Ajustados:** Os pontos estão dispersos aleatoriamente em torno da linha zero, sugerindo **homoscedasticidade**.
- **QQ-Plot:** Os resíduos seguem a linha vermelha teórica, indicando que os erros seguem uma **distribuição normal**.

## 5 CONCLUSÃO

O estudo confirmou estatisticamente as hipóteses intuitivas sobre o consumo de cerveja. O modelo desenvolvido demonstra que a temperatura média e o fato de ser final de semana são impulsionadores significativos do consumo, enquanto a precipitação atua como um fator inibidor. Com um  $R^2$  superior a 0.66, o modelo é considerado robusto para o problema proposto.