

**Faculdades Metropolitanas Unidas – FMU**

**Impacto das Variáveis Operacionais na Satisfação do Cliente: Uma Análise do NPS em  
Seguradoras**

**Lucas da Cruz Teixeira Lima**

**Econometria II**

## Sumário

|         |                                       |
|---------|---------------------------------------|
| 1.....  | Apresentação                          |
| 2.....  | Objetivo                              |
| 3.....  | Hipóteses                             |
| 4.....  | Base de Dados                         |
| 5.....  | Análise Descritiva dos Dados          |
| 6.....  | Distribuições com curvas de Densidade |
| 7.....  | Correlação Pearson das Variáveis      |
| 8.....  | Modelagem NPS                         |
| 9.....  | Análise do Modelo                     |
| 10..... | Conclusões Finais                     |

# Apresentação

## Net Promoter Score (NPS)

O Net Promoter Score (NPS) consiste em uma métrica amplamente utilizada para mensurar o grau de lealdade do cliente em relação a uma marca, bem como sua experiência durante o período de utilização de determinado serviço. Essa avaliação está diretamente relacionada à percepção que o cliente desenvolveu a partir de interações anteriores com a empresa.

## Aplicação do NPS

Para a aplicação do NPS, a organização deve disponibilizar uma pesquisa, geralmente por meio de um formulário, aos clientes que utilizaram o serviço, permitindo que estes realizem sua avaliação. O objetivo da pesquisa é extrair insights relevantes acerca dos sentimentos e percepções do consumidor. A pergunta central do método é:

*“Em uma escala de 0 a 10, como você avalia o nosso serviço?”*

## Classificação dos Perfis de Clientes

Com base nas respostas, o NPS categoriza os clientes em três grupos distintos:

- **Promotores (9 a 10):** Clientes altamente satisfeitos, com grande probabilidade de recomendar a marca.
- **Neutros (7 a 8):** Clientes satisfeitos, porém pouco engajados, podendo ser influenciados por concorrentes.
- **Detratores (0 a 6):** Clientes insatisfeitos, que tendem a prejudicar a reputação da marca.

## Objetivo do Projeto

O presente projeto tem como objetivo analisar de que forma determinadas variáveis operacionais, presentes no contexto de uma seguradora, podem influenciar o resultado final da métrica Net Promoter Score (NPS) ao longo de um período mensal. Para a realização das análises e construção do modelo proposto, foi empregada a linguagem de programação **Python**, devido à sua robustez e aplicabilidade em processos de modelagem estatística e análise de dados.

As variáveis consideradas no estudo incluem:

- **Tempo de resposta da seguradora** após a abertura do sinistro;
- **Número de interações necessárias** para a resolução do sinistro (0 = não houve reabertura; 1 = houve reabertura);
- **Quantidade de etapas do processo**;
- **SLA (Service Level Agreement)** estabelecido;
- **Existência de reabertura** do sinistro;
- **Valor total do sinistro**.

A análise dessas variáveis visa compreender sua correlação com a percepção do cliente e, consequentemente, com a nota atribuída ao NPS, permitindo identificar pontos críticos e oportunidades de melhoria nos processos internos da seguradora.

## Hipóteses

O presente estudo propõe a verificação de hipóteses relacionadas ao impacto de variáveis operacionais sobre a métrica Net Promoter Score (NPS) em uma seguradora. As hipóteses foram formuladas da seguinte maneira:

- **Hipótese I – Tempo de Resposta da Seguradora após a abertura do sinistro** o **H<sub>0</sub>**: O tempo de resposta não afeta o NPS.  
o **H<sub>1</sub>**: Um maior tempo de resposta reduz o NPS.
- **Hipótese II – Número de Interações necessárias para resolução do sinistro** o **H<sub>0</sub>**: O número de interações não afeta o NPS.  
o **H<sub>1</sub>**: Quanto maior for o número de interações, menor será o NPS.
- **Hipótese III – Quantidade de Etapas do Processo** o **H<sub>0</sub>**: O número de etapas não afeta o NPS.  
o **H<sub>1</sub>**: Um maior número de etapas reduz o NPS.
- **Hipótese IV – Cumprimento do SLA (Service Level Agreement)** o **H<sub>0</sub>**: Cumprir o SLA não influencia o NPS.  
o **H<sub>1</sub>**: Cumprir o SLA influencia positivamente o NPS.
- **Hipótese V – Reabertura do Sinistro** o **H<sub>0</sub>**: A reabertura do sinistro não influencia o NPS.  
o **H<sub>1</sub>**: Chamados reabertos apresentam NPS inferior.

Essas hipóteses serão testadas por meio de análises estatísticas, visando identificar relações significativas entre as variáveis operacionais e a percepção do cliente, expressa pela nota NPS.

## Base de Dados

A base de dados utilizada para a análise e modelagem deste projeto é composta por **400 registros** e **6 variáveis**, cada uma representando aspectos operacionais relevantes ao processo de sinistro em uma seguradora. As variáveis foram selecionadas com base em critérios previamente definidos na Ata 2, por sua potencial influência na métrica Net Promoter Score (NPS). São elas:

- **Tempo de resposta da seguradora** após a abertura do sinistro (em horas);
- **Número de interações necessárias para resolução do sinistro**, categorizado como:
  - o *0*: não houve reabertura; o
  - o *1*: houve reabertura;
- **Quantidade de etapas do processo** até a conclusão do sinistro;
- **Cumprimento do SLA (Service Level Agreement)**, indicando se os prazos acordados foram respeitados;
- **Existência de reabertura do sinistro**, variável binária (sim/não);
- **Valor total do sinistro**, expresso em moeda corrente.

Essa base servirá como insumo para análises estatísticas e para a construção do modelo, permitindo avaliar a correlação entre as variáveis operacionais e a percepção do cliente, expressa pela nota NPS.

## Análise Descritiva dos Dados

A análise descritiva foi conduzida com o objetivo de compreender o comportamento das variáveis presentes na base de dados e identificar padrões relevantes para o estudo. Todas as operações estatísticas foram realizadas por meio da linguagem de programação **Python**, utilizando bibliotecas especializadas em análise de dados, como *pandas*, *numpy* e *matplotlib*. Essa etapa é fundamental para garantir uma visão preliminar da estrutura dos dados e subsidiar as fases posteriores de modelagem.

Entre as técnicas aplicadas, destacam-se:

- **Medidas de Contagem:** Frequência absoluta e relativa das variáveis categóricas, permitindo avaliar a distribuição dos registros.
- **Medidas de Tendência Central:** Cálculo da média, mediana e moda para variáveis numéricas, visando identificar valores representativos.
- **Medidas de Variação e Dispersão:** Análise do desvio padrão, variância e amplitude, essenciais para compreender a heterogeneidade dos dados.
- **Correlação entre Variáveis:** Aplicação do coeficiente de correlação de Pearson para mensurar a relação linear entre as variáveis operacionais e a nota NPS.
- **Distribuição dos Dados:** Representação gráfica por meio de histogramas e curvas de densidade, possibilitando a visualização da forma e dispersão dos dados.

Essa abordagem permitiu identificar possíveis outliers, padrões de concentração e relações preliminares entre as variáveis, fornecendo subsídios para a etapa de modelagem estatística.

## Análise Descritiva no Python

```
===== Estatísticas Gerais =====
```

|                | count | mean        | std         | min        | 25%         |
|----------------|-------|-------------|-------------|------------|-------------|
| nps            | 400.0 | 4.200465    | 2.126631    | 0.000000   | 2.810748    |
| tempo_resposta | 400.0 | 11.611862   | 10.985371   | 0.060893   | 3.313475    |
| interacoes     | 400.0 | 2.227500    | 1.456381    | 0.000000   | 1.000000    |
| etapas         | 400.0 | 2.927500    | 1.645909    | 0.000000   | 2.000000    |
| sla            | 400.0 | 0.865000    | 0.342152    | 0.000000   | 1.000000    |
| reabertura     | 400.0 | 0.082500    | 0.275470    | 0.000000   | 0.000000    |
| valor_sinistro | 400.0 | 6910.130379 | 7723.195516 | 449.030684 | 3023.909780 |

|                | 50%         | 75%         | max           |
|----------------|-------------|-------------|---------------|
| nps            | 4.548445    | 5.793795    | 8.469812      |
| tempo_resposta | 8.602343    | 16.517487   | 55.884059     |
| interacoes     | 2.000000    | 3.000000    | 8.000000      |
| etapas         | 3.000000    | 4.000000    | 8.000000      |
| sla            | 1.000000    | 1.000000    | 1.000000      |
| reabertura     | 0.000000    | 0.000000    | 1.000000      |
| valor_sinistro | 4773.498024 | 8389.492090 | 113662.654026 |

```
===== Estatísticas Adicionais =====
```

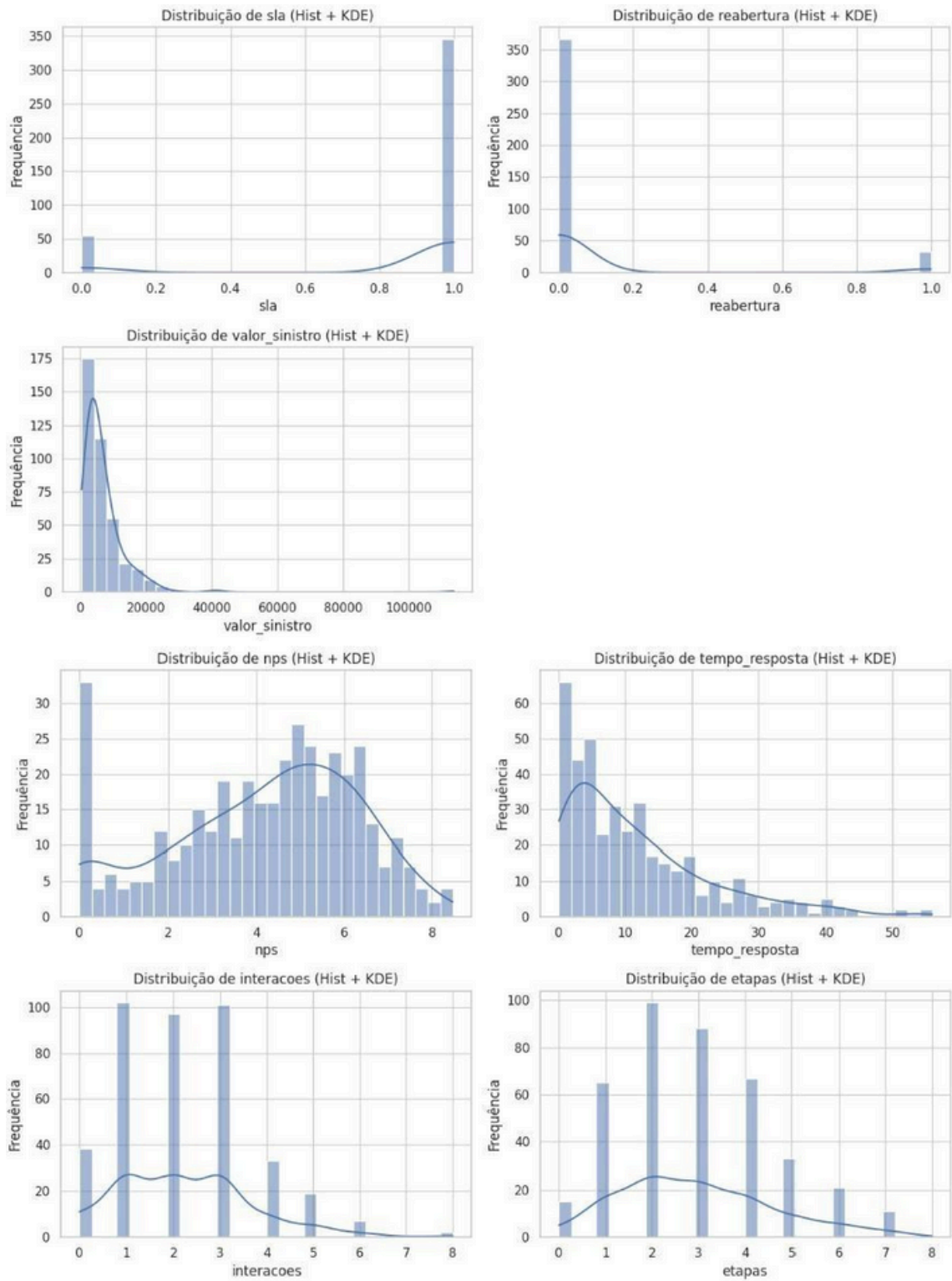
|                | missing (%) | unique_values | most_frequent | variance     |
|----------------|-------------|---------------|---------------|--------------|
| nps            | 0.0         | 373           | 0.000000      | 4.522560e+00 |
| tempo_resposta | 0.0         | 400           | 0.060893      | 1.206784e+02 |
| interacoes     | 0.0         | 9             | 1.000000      | 2.121046e+00 |
| etapas         | 0.0         | 9             | 2.000000      | 2.709016e+00 |
| sla            | 0.0         | 2             | 1.000000      | 1.170677e-01 |
| reabertura     | 0.0         | 2             | 0.000000      | 7.588346e-02 |
| valor_sinistro | 0.0         | 400           | 449.030684    | 5.964775e+07 |

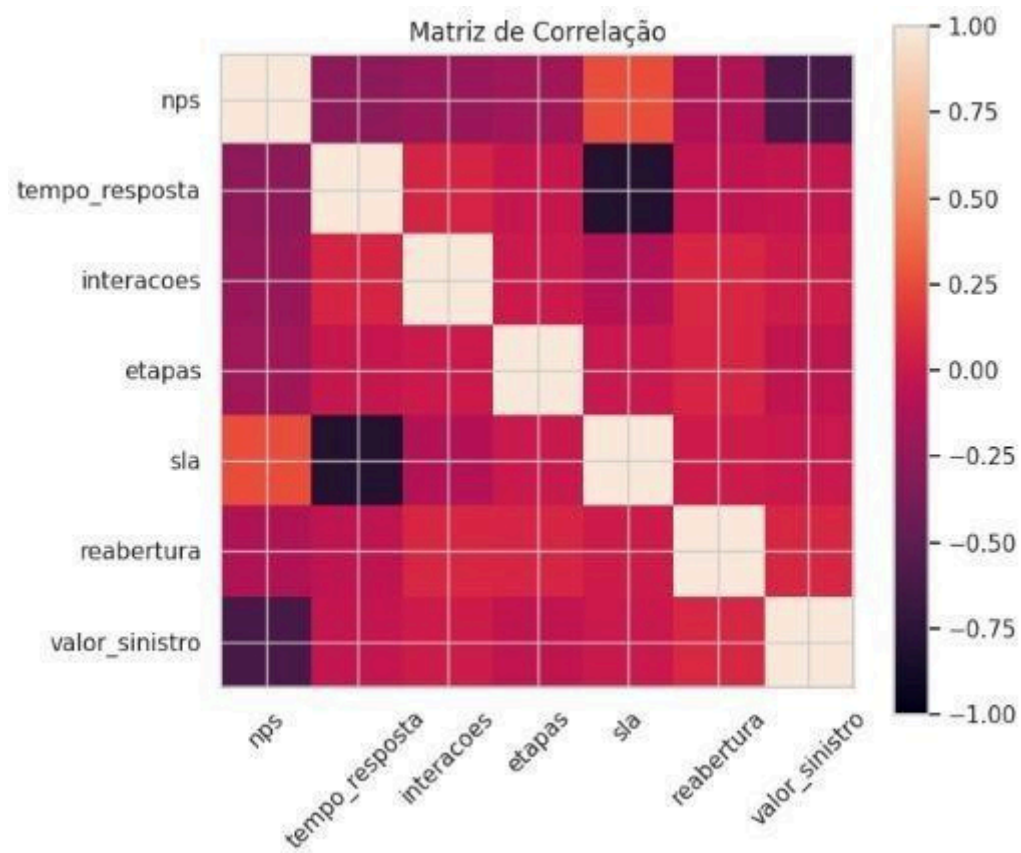
|                | coef_var (%) |
|----------------|--------------|
| nps            | 50.628470    |
| tempo_resposta | 94.604732    |
| interacoes     | 65.381875    |
| etapas         | 56.222338    |
| sla            | 39.555090    |
| reabertura     | 333.902453   |
| valor_sinistro | 111.766278   |



# Distribuições com curvas de Densidade



### Gráfico Heatmap de Correlação Pearson das Variáveis



## Modelagem NPS

O presente estudo tem como objetivo analisar a influência de diferentes variáveis sobre a pontuação do Net Promoter Score (NPS), por meio da construção de um modelo econométrico. A formulação proposta busca compreender como fatores operacionais e de experiência do cliente impactam diretamente a avaliação do NPS.

O modelo pode ser representado pela seguinte equação:

$$NPS_i = \beta_0 + \beta_1 \text{tempo\_resposta}_i + \beta_2 \text{interacoes}_i + \beta_3 \text{etapas}_i + \beta_4 \text{SLA}_i \\ + \beta_5 \text{reabertura}_i + \beta_6 \text{valor\_sinistro}_i + \varepsilon_i$$

Em que:

- $NPS_i$ : Nota atribuída pelo cliente  $i$ ;
- $\beta_0$ : Intercepto do modelo;
- $\beta_k$ : Coeficientes associados às variáveis explicativas;
- $\varepsilon_i$ : Termo de erro aleatório.

As variáveis independentes representam aspectos relevantes do processo de atendimento, tais como:

- **tempo\_resposta**: Tempo médio para resposta ao cliente;
- **interacoes**: Número de interações realizadas durante o atendimento;
- **etapas**: Quantidade de etapas necessárias para resolução;
- **SLA**: Cumprimento do Acordo de Nível de Serviço;
- **reabertura**: Ocorrência de reabertura do chamado;
- **valor\_sinistro**: Valor monetário associado ao sinistro.

Esse modelo permite avaliar a magnitude e a direção do impacto de cada variável sobre a percepção do cliente, fornecendo subsídios para estratégias de melhoria contínua no atendimento.

## Análise do Modelo

A implementação do modelo foi realizada utilizando a linguagem de programação **Python**, devido à sua ampla aplicação em análise estatística e aprendizado de máquina, bem como à disponibilidade de bibliotecas robustas para manipulação de dados e modelagem preditiva, como **pandas**, **numpy**, **statsmodels** e **scikitlearn**.

| OLS Regression Results |                  |                     |          |       |        |          |
|------------------------|------------------|---------------------|----------|-------|--------|----------|
| Dep. Variable:         | nps              | R-squared:          | 0.529    |       |        |          |
| Model:                 | OLS              | Adj. R-squared:     | 0.521    |       |        |          |
| Method:                | Least Squares    | F-statistic:        | 28.02    |       |        |          |
| Date:                  | Sun, 07 Dec 2025 | Prob (F-statistic): | 7.35e-28 |       |        |          |
| Time:                  | 19:10:39         | Log-Likelihood:     | -718.47  |       |        |          |
| No. Observations:      | 400              | AIC:                | 1451.    |       |        |          |
| Df Residuals:          | 393              | BIC:                | 1479.    |       |        |          |
| Df Model:              | 6                |                     |          |       |        |          |
| Covariance Type:       | HC3              |                     |          |       |        |          |
|                        | coef             | std err             | z        | P> z  | [0.025 | 0.975]   |
| const                  | 6.5010           | 1.091               | 5.956    | 0.000 | 4.362  | 8.640    |
| tempo_resposta         | -0.0359          | 0.010               | -3.423   | 0.001 | -0.056 | -0.015   |
| interacoes             | -0.2423          | 0.076               | -3.189   | 0.001 | -0.391 | -0.093   |
| etapas                 | -0.2631          | 0.060               | -4.399   | 0.000 | -0.380 | -0.146   |
| sla                    | 0.7012           | 0.295               | 2.377    | 0.017 | 0.123  | 1.279    |
| reabertura             | -0.2215          | 0.513               | -0.432   | 0.666 | -1.227 | 0.784    |
| valor_sinistro         | -0.0002          | 0.000               | -1.311   | 0.190 | -0.000 | 8.32e-05 |
| Omnibus:               | 265.537          | Durbin-Watson:      | 2.073    |       |        |          |
| Prob(Omnibus):         | 0.000            | Jarque-Bera (JB):   | 9373.769 |       |        |          |
| Skew:                  | 2.255            | Prob(JB):           | 0.00     |       |        |          |
| Kurtosis:              | 26.283           | Cond. No.           | 8.10e+04 |       |        |          |

## Coeficientes Estimados e Interpretação

A análise dos coeficientes revela a magnitude e direção do impacto das variáveis sobre o NPS:

- **Constante ( $\beta_0 = 6.5510$ )**

Representa o NPS médio esperado quando todas as variáveis independentes assumem valor zero. Serve apenas como referência. **Variáveis Significativas**

- **Tempo de Resposta ( $\beta = -0.0396, p < 0.01$ )**

Impacto negativo e estatisticamente significativo. Cada hora adicional no tempo de resposta reduz o NPS em aproximadamente **0,04 pontos**. Um atraso de 24h implica queda próxima de **1 ponto**, o que é coerente com a experiência prática.

- **Interações ( $\beta = -0.2423, p < 0.01$ )**

Também negativo e significativo. Cada interação extra com o cliente reduz o NPS em cerca de **0,24 pontos**, indicando que processos mais complexos prejudicam a percepção do cliente.

- **Etapas ( $\beta = -0.2631, p < 0.001$ )**

Forte impacto negativo. Cada etapa adicional no processo reduz o NPS em **0,26 pontos**, sendo uma das variáveis mais influentes do modelo.

- **Cumprimento do SLA ( $\beta = 0.7012, p < 0.05$ )**

Impacto positivo e significativo. Cumprir o SLA aumenta o NPS em aproximadamente **0,70 pontos**, reforçando a importância do cumprimento de prazos.

### Variáveis Não Significativas

- **Reabertura ( $\beta = -0.2215, p = 0.666$ )**

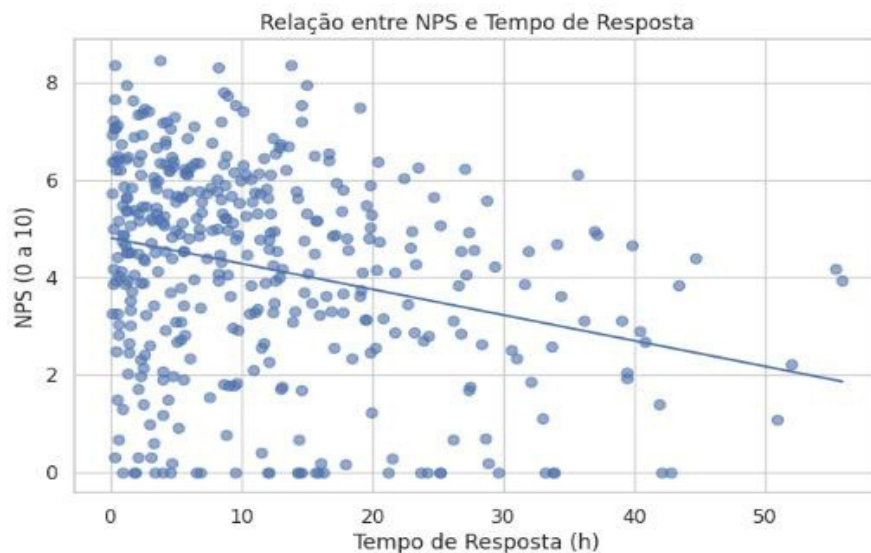
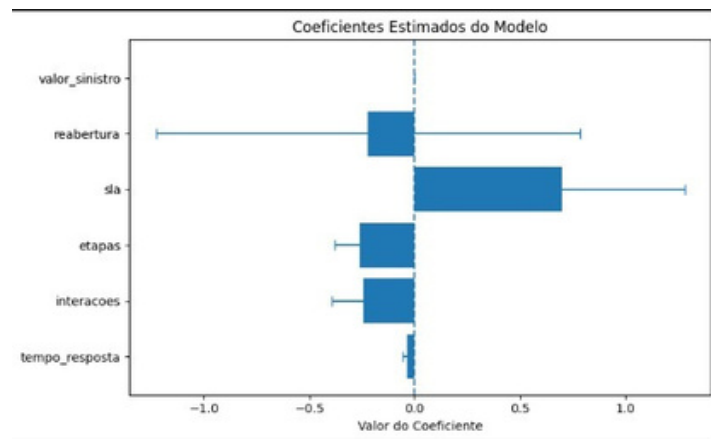
Apesar do sinal negativo, não apresentou significância estatística. O efeito é absorvido por variáveis correlacionadas (interações e etapas).

- **Valor do Sinistro ( $\beta = -0.0002, p = 0.190$ )**

Também não significativo. A alta variabilidade do valor do sinistro e o baixo impacto na satisfação explicam esse resultado.

## Qualidade do Ajuste

- $R^2=0.529$ : O modelo explica cerca de 52,9% da variação do NPS, considerado bom para estudos de satisfação.
- **F-statistic**: Altamente significativo, confirmando relevância do conjunto de variáveis.
- **Durbin-Watson = 2.07**: Ausência de autocorrelação dos resíduos.
- **Jarque-Bera elevado**: Indica não normalidade dos resíduos, comum em modelos de NPS. Correções robustas (HC3) garantem validade da inferência.
- **Cond. No. =  $8.1e+04$** : Sugere multicolinearidade fraca/moderada, mas  $VIF < 5$  indica ausência de problema grave.



## Conclusão

Os resultados obtidos indicam que o modelo proposto apresenta bom poder explicativo para a variação do Net Promoter Score (NPS), com um  $R^2$  de 0,529, valor considerado satisfatório em estudos de satisfação. A análise dos coeficientes revela que variáveis relacionadas à experiência operacional do cliente como **tempo de resposta**, **número de interações** e **quantidade de etapas** exercem impacto negativo e estatisticamente significativo sobre o NPS, confirmando que processos mais complexos e demorados reduzem a percepção positiva do serviço.

Por outro lado, o **cumprimento do SLA** mostrou-se um fator relevante e positivo, aumentando significativamente a avaliação do cliente, o que reforça a importância do cumprimento de prazos como estratégia para melhoria da satisfação. Variáveis como **reabertura de chamados** e **valor do sinistro** não apresentaram significância estatística, sugerindo que aspectos financeiros ou ocorrências pontuais têm menor influência na percepção global do cliente quando comparados à eficiência operacional.

Apesar da presença de multicolinearidade moderada e da não normalidade dos resíduos características comuns em modelos de satisfação, os testes robustos aplicados garantem a validade das inferências. Assim, conclui-se que o modelo é adequado para identificar os principais determinantes do NPS e pode servir como base para ações estratégicas voltadas à redução de etapas e interações, bem como à melhoria do tempo de resposta e cumprimento de SLA.