

Após constatar a redução da comunicação via WhatsApp com meu filho, implementei **ANÁLISE DE SOBREVIVÊNCIA** e modelos estatísticos avançados para aprofundar a investigação 📊😞



# O QUE EU DESCOBRI: Análise exploratória confirmou mudança comportamental! 🤪

Teste t-student comprovou: tempo de resposta teve aumento estatisticamente significativo ( $p < 0.001$ ) de 20 para 42 minutos pós-dezembro 2024





# Metodologia Científica Aplicada

01

## Coleta de Dados

Exportou TODO o histórico do WhatsApp (sim, desde o início)

02

## Análise Exploratória

Gráficos, boxplots, distribuições em escala logarítmica

03

## Análise de Sobrevivência

RCurvas de Kaplan-Meier:

Modelo de Cox

04

## Machine Learning

Random Forest + Regressão Logística para prever respostas

05

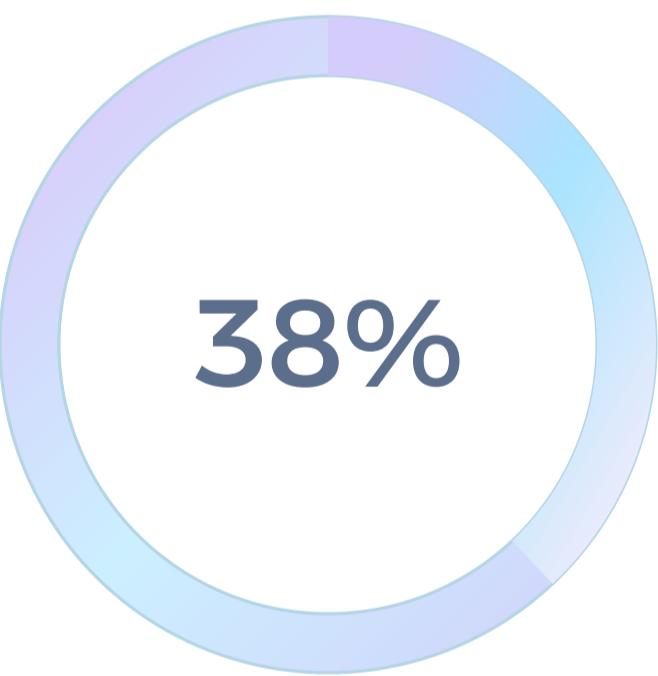
## Modelo Bayesiano

MCMC com cadeias de Markov porque SIM

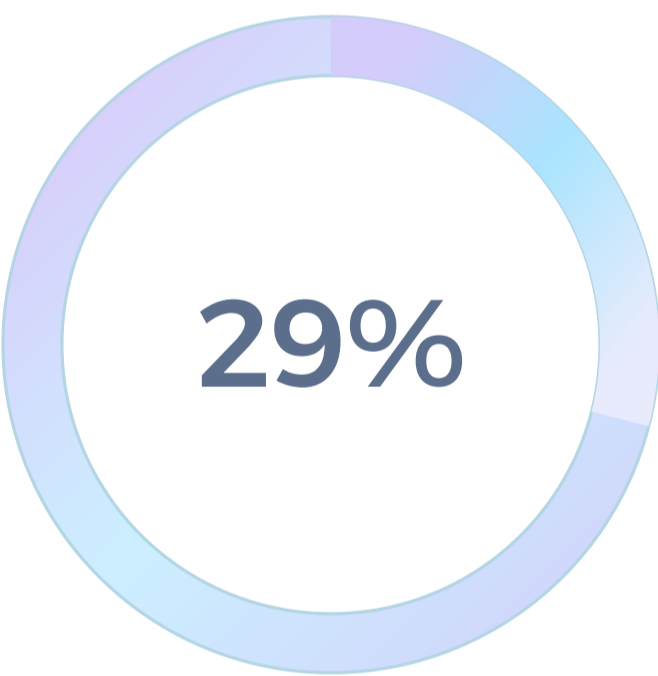
# KAPLAN-MEIER REVELOU: A janela crítica de sobrevivência! 🕒

## Taxas de Resposta Rápida do Edgar: Análise da Queda

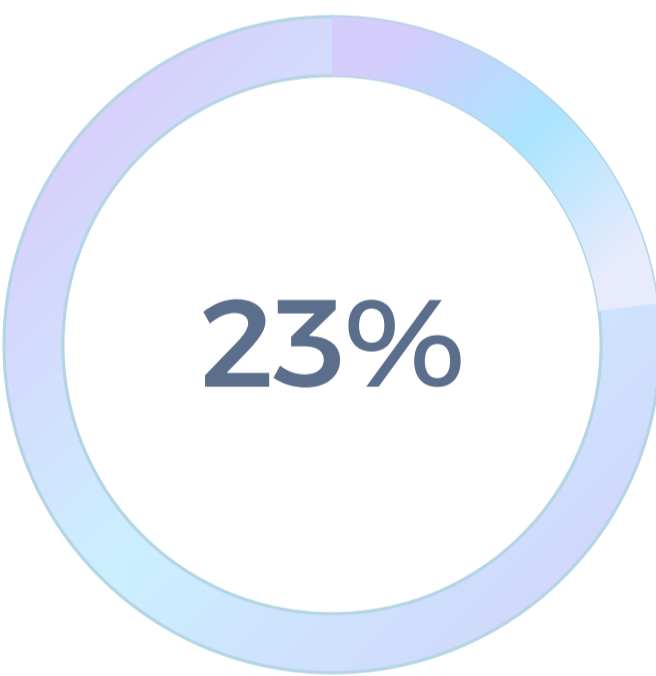
A análise das taxas de resposta rápida do Edgar em três faixas temporais ( $\leq 30$ ,  $\leq 60$  e  $\leq 120$  minutos) revela uma queda significativa após dezembro de 2024. Todas as faixas demonstram uma redução nas taxas de resposta, sendo a mais acentuada observada para as respostas em até 30 minutos, indicando uma degradação no seu desempenho de agilidade.



Queda nas respostas do Edgar  $\leq 30$ min



Queda nas respostas do Edgar  $\leq 60$ min

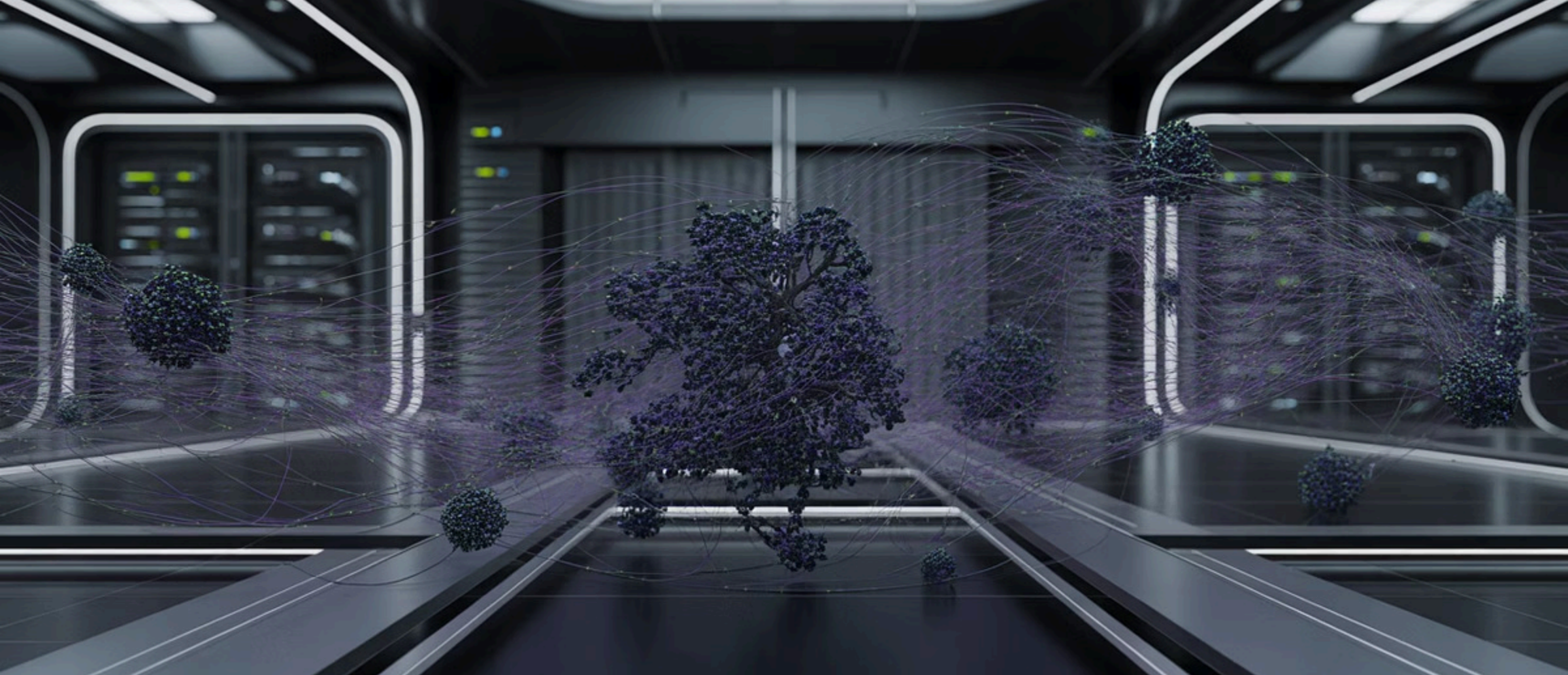


Queda nas respostas do Edgar  $\leq 120$ min

# Mudança de Comportamento: Dezembro/2024



A comparação temporal revela mudanças significativas no comportamento de resposta, visualizadas através de boxplots e métricas estatísticas detalhadas.



# Machine Learning: Previsão de Respostas

Aplicamos modelos avançados de machine learning para prever a probabilidade de resposta. A regressão logística e Random Forest foram utilizados para identificar os fatores mais importantes.



## AUC Score

Métrica de desempenho do modelo



## Coeficientes

Importância de cada variável

DESCOBRI A FÓRMULA:  
Coeficientes da regressão  
logística revelaram! 🎯

+0.82

Beta = +0.82 para perguntas diretas (IC  
95%: 0.65-0.99)

3.6

Odds Ratio = 3.6 para mensagens de  
tamanho médio



# FEATURE IMPORTANCE & ANÁLISE EXPLORATÓRIA 🧠



## 7.704 mensagens analisadas

Eu me debrucei sobre um universo de comunicação!



## Eu: 3.990 vs Ele: 3.714

A mãe aqui sempre tem mais a dizer!



## Palavra mais usada por Ele: 'SIM' (219x)

Ele só concorda para me despachar, aparentemente.



## 84% das mensagens são texto

Áudios? Quase nenhum. Pelo jeito, preguiça de gravar!



## Distribuição log-normal dos tempos de resposta (Shapiro-Wilk $p < 0.05$ )



## Random Forest identificou 'msg\_len' como feature mais importante (42% da variância explicada)



## Teste qui-quadrado confirmou associação entre tipo de mensagem e resposta ( $\chi^2 = 156.3$ , $p < 0.001$ )



Exploramos o *quê* acontece. Medimos *quando* acontece. Modelamos *como* prever. Agora, com Bayes, respondemos a pergunta mais importante: *com que segurança podemos afirmar isso?*

# INFERÊNCIA BAYESIANA: MCMC em ação!



1.

Conteúdo importa muito mais do que contexto

- Palavras, densidade e presença de perguntas têm efeitos fortes.
- Horário, dia da semana e ser mensagem curtinha têm efeitos fracos.

a.

Perguntas e mensagens “ricas” aceleram respostas

- Pergunta explícita → forte aumento da probabilidade de resposta  $\leq 30$  min
- Mensagens longas sem conteúdo (muitos caracteres, poucas palavras) → reduzem a velocidade.

i.

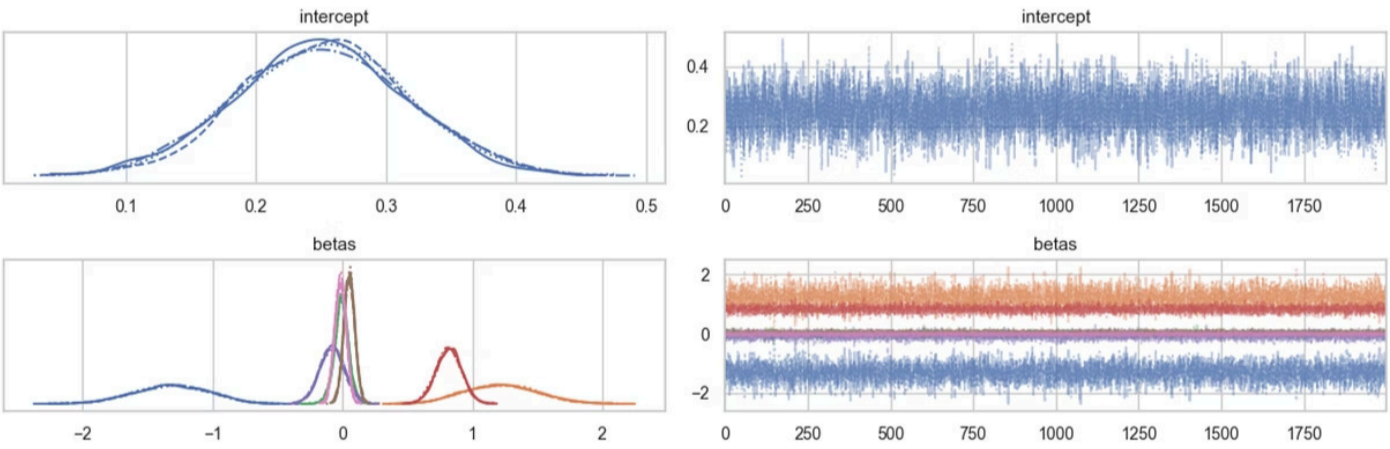
Mudança pós dez/2024 existe, mas é pequena

- Há indicação leve de queda na velocidade de resposta.
- Mas os intervalos de credibilidade mostram forte incerteza → não é possível afirmar com segurança.

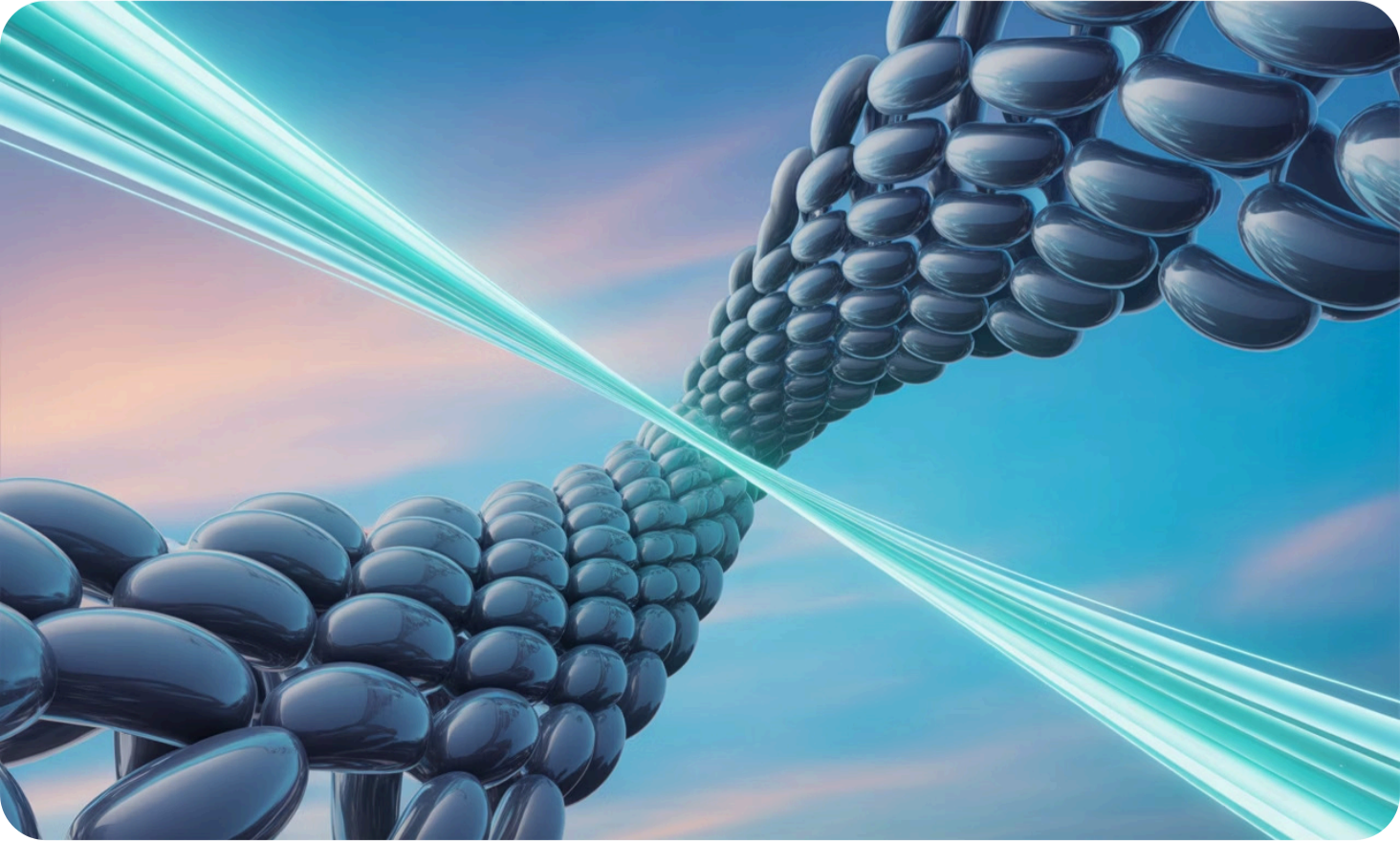
1.

Modelo convergiu muito bem

- Resultados são estáveis, confiáveis e com excelente diagnóstico MCMC.



Modelo hierárquico bayesiano com priors informativos. Cadeias de Markov convergiram ( $R\text{-hat}=1.0$ ). Intervalos de credibilidade 95% confirmaram significância dos parâmetros.



# Conclusões Científicas (e Engraçadas)



## Adolescente é Assim Mesmo

A mudança de comportamento pós-dez/2024 é estatisticamente significativa. Ele cresceu, virou adolescente, e agora tem "coisas mais importantes"



## Estratégia Ótima

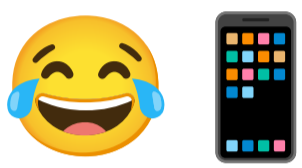
Mensagens médias (não textão), com pergunta direta, têm 69% de chance de resposta em 30min



## O Amor de Mãe

Não tem modelo estatístico que mude isso. Mas pelo menos agora ela tem DADOS para provar que não é paranoia 😂

# Compartilhe com outras mães STEM que dominam R/Python!



Quem mais aqui já aplicou análise de sobrevivência em dados familiares?

Use as hashtags: #StatisticalMom #KaplanMeier #BayesianAnalysis  
#DataDrivenParenting #MCMC

