



Trabalho decente e mercado de trabalho: análise temporal da taxa de desocupação no Brasil

Projeto Aplicado IV
Ciências de Dados - 2025.2 - 201825166.000.05A

Luis Fernando do Lago Attarian

RA - 10089158

Kelly Graziely dos Santos Pena

RA - 10416108

Universidade Presbiteriana Mackenzie
Faculdade de Computação e Informática
Tecnologia em Ciência de Dados

São Paulo, 24 de novembro de 2025

Sumário

0.1	Motivação e Justificativa	2
0.2	Objetivo geral e específicos	3
1	Descrição da base de dados	4
1.1	Brasil (Tabela 4099)	4
1.2	UFs (Tabela 4099)	4
1.3	Dicionário de Colunas	4
2	Referencial Teórico	5
3	Diagrama da Solução	6
3.1	Análise dos dados empregados: qualidade, estrutura, limitações e recortes adotados	11
4	Modelos	12
4.1	Seasonal-Trend decomposition using Loess	12
4.2	ACF e PACF	14
4.3	Comparação entre UFs	17
5	Resultados	19
5.1	Diagnóstico STL (Brasil)	19
5.2	Diagnóstico ACF/PACF (Brasil)	20
5.3	Comparação entre Bahia e Santa Catarina	20
6	Discussão	21
7	Conclusão e recomendações para o poder público	22
8	Melhorias futuras	24
9	Repositório do Projeto	25
10	Referências Bibliográficas	25

c A taxa de desocupação é um dos principais indicadores para avaliar a situação do mercado de trabalho e a estabilidade econômica de um país. No Brasil, esse índice tem oscilado significativamente desde 2012, refletindo tanto momentos de expansão quanto de retração econômica. Entre os fatores que marcaram a trajetória do desemprego ao longo desse período estão a recessão econômica de 2015–2016, os impactos da pandemia de COVID-19 a partir de 2020 e, mais recentemente, a retomada do crescimento acompanhada pela formalização de postos de trabalho [3].

Em 2024, a taxa média de desocupação no Brasil foi de 6,6%, o menor valor desde o início da série histórica da PNAD Contínua em 2012 [1]. No segundo trimestre de 2025, esse indicador caiu ainda mais, alcançando 5,8%, o menor patamar já registrado, representando queda de 1,2 ponto percentual em relação ao trimestre anterior e de 1,1 ponto percentual frente ao mesmo período de 2024 [2]. Esse resultado reflete não apenas o aumento da população ocupada, mas também a expansão do emprego formal, com o número de trabalhadores com carteira assinada no setor privado chegando a 39 milhões, o maior já registrado no país [2].

Esses avanços estão diretamente alinhados ao Objetivo de Desenvolvimento Sustentável 8 (ODS 8), que visa “promover o crescimento econômico sustentado, inclusivo e sustentável, o emprego pleno e produtivo e o trabalho decente para todas e todos” [15]. Nesse contexto, acompanhar a evolução da taxa de desocupação não é apenas relevante para mensurar o desempenho conjuntural da economia, mas também para monitorar o progresso do Brasil frente às metas globais estabelecidas pela ONU.

A análise de séries temporais da taxa de desocupação entre 2012 e 2025 permite identificar tendências estruturais, ciclos econômicos e efeitos de choques externos sobre o mercado de trabalho. Além disso, possibilita avaliar a eficácia de políticas públicas voltadas à geração de emprego e à formalização da ocupação. A disponibilidade de dados trimestrais consistentes torna essa série adequada para análises estatísticas, modelagens preditivas e visualizações que podem subsidiar a tomada de decisão tanto no setor público quanto no privado [5].

0.1 Motivação e Justificativa

A escolha da taxa de desocupação como objeto de estudo justifica-se pela centralidade desse indicador na compreensão da dinâmica econômica e social de um país. O desemprego afeta diretamente a renda das famílias, a capacidade de consumo, a arrecadação tributária e, em última instância, a qualidade de vida da população. No caso brasileiro, a taxa de desocupação apresentou oscilações marcantes desde 2012, refletindo períodos de recessão, instabilidade política, crise sanitária durante a pandemia de COVID-19 e, mais recentemente, sinais consistentes de recuperação econômica (IBGE, 2024; BRASIL, 2025).

Do ponto de vista dos Objetivos de Desenvolvimento Sustentável (ODS), este estudo dialoga especialmente com o ODS 8 – Trabalho decente e crescimento econômico, que tem como meta promover o emprego pleno, produtivo e de qualidade para todos. Analisar a evolução da desocupação por meio de séries temporais contribui para identificar gargalos estruturais do mercado de trabalho e apontar avanços que aproximam o país das metas globais estabelecidas pela ONU (NAÇÕES UNIDAS, 2023).

O tema também é relevante para gestores públicos, formuladores de políticas e empresas, uma vez que a compreensão das flutuações do desemprego auxilia na definição de estratégias para geração de empregos, qualificação profissional e estímulo ao crescimento

econômico sustentável. Para a sociedade em geral, a redução da taxa de desocupação está associada à melhoria nas condições de vida, maior inclusão social e diminuição da vulnerabilidade de grupos populacionais historicamente mais afetados pelo desemprego.

Adicionalmente, o uso de técnicas de análise de séries temporais oferece não apenas a possibilidade de avaliar tendências passadas, mas também de projetar cenários futuros. Isso amplia a aplicabilidade da solução em campos como economia do trabalho, planejamento governamental e estudos acadêmicos voltados ao desenvolvimento sustentável. Assim, este trabalho se mostra relevante tanto pela sua contribuição científica quanto pela sua utilidade prática, reforçando a importância de alinhar análises quantitativas a objetivos sociais mais amplos, como os propostos pelos ODS.

0.2 Objetivo geral e específicos

Analisar a evolução da taxa de desocupação no Brasil entre 2012 e 2025 por meio de técnicas de séries temporais, a fim de identificar tendências, padrões sazonais e impactos de eventos econômicos relevantes, relacionando os resultados ao Objetivo de Desenvolvimento Sustentável 8 (ODS 8) — trabalho decente e crescimento econômico.

Objetivos Específicos:

- Coletar e organizar os dados trimestrais da taxa de desocupação disponibilizados pelo IBGE (PNAD Contínua), estruturando-os em formato adequado para análise estatística.
- Construir representações gráficas que evidenciem a evolução do desemprego no período, destacando ciclos, sazonalidades e pontos de ruptura na série histórica.
- Aplicar técnicas de análise de séries temporais (como decomposição de tendência e sazonalidade) para compreender o comportamento do indicador ao longo do tempo.
- Comparar a taxa de desocupação entre diferentes períodos (pré-crise de 2015–2016, pandemia de COVID-19 e recuperação recente) para avaliar variações estruturais.
- Realizar análise desagregada por Unidade da Federação (UF), identificando disparidades regionais na taxa de desocupação e suas causas associadas.
- Identificar padrões sazonais do desemprego, de modo a orientar o planejamento de ações governamentais e políticas públicas direcionadas para os períodos de maior vulnerabilidade.
- Relacionar os resultados da análise com as metas do ODS 8, destacando avanços e desafios do Brasil na promoção do emprego pleno e produtivo.
- Gerar subsídios para a formulação de políticas públicas, indicando como os padrões detectados podem orientar medidas de combate ao desemprego e incentivo à formalização do trabalho.
- Propor cenários e projeções exploratórias sobre a evolução do desemprego, de forma a contribuir com decisões estratégicas de governos, empresas e sociedade civil.

1 Descrição da base de dados

O presente trabalho utiliza como base empírica a **Tabela 4099** do Sistema IBGE de Recuperação Automática (SIDRA), derivada da Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua), conduzida pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

Essa tabela apresenta a **taxa de desocupação** da população de 14 anos ou mais de idade, em periodicidade trimestral, cobrindo o período de 2012 até 2025. Os dados estão disponíveis para diferentes níveis geográficos, incluindo Brasil, Grandes Regiões e Unidades da Federação (UF).

Características do conjunto de dados

- **Variável principal:** Taxa de desocupação (%)
- **Periodicidade:** Trimestral
- **Unidade de análise:** Pessoas com 14 anos ou mais na força de trabalho
- **Abrangência geográfica:** Brasil, Grandes Regiões e Unidades da Federação (UF)
- **Período de coleta:** 2012 (início da PNAD Contínua) até 2025 (última atualização disponível)
- **Fonte oficial:** IBGE – SIDRA – Tabela 4099: Taxa de desocupação da população de 14 anos ou mais de idade. Disponível em: <https://sidra.ibge.gov.br/Tabela/4099>

Esse conjunto de dados é particularmente adequado para a análise de séries temporais, pois combina longa série histórica, regularidade trimestral e abrangência nacional e regional, permitindo avaliar tendências, sazonalidades e choques no mercado de trabalho brasileiro.

1.1 Brasil (Tabela 4099)

- **Número de linhas:** 55
- **Colunas:** [NC, NN, MC, MN, V, D1C, D1N, D2C, D2N, D3C, D3N]

1.2 UFs (Tabela 4099)

- **Número de linhas:** 1459
- **Colunas:** [NC, NN, MC, MN, V, D1C, D1N, D2C, D2N, D3C, D3N]

1.3 Dicionário de Colunas

NC: Nível de cobertura dos dados (geralmente não usado diretamente).

D1C: Código da primeira dimensão (ex.: código do Brasil, Região ou UF).

D1N: Nome da primeira dimensão (ex.: Brasil, Nordeste, São Paulo).

D2C: Código da segunda dimensão (período).

D2N: Nome do período (ex.: 2012 1^o trimestre, 2025 2^o trimestre).

V: Valor da variável selecionada (neste caso, taxa de desocupação em %).

2 Referencial Teórico

A taxa de desocupação mede a proporção de pessoas economicamente ativas que buscam trabalho, mas não encontram ocupação, sendo um dos principais indicadores de desempenho macroeconômico e social [3]. Esse indicador reage diretamente às oscilações do ciclo econômico, às crises e às políticas públicas de estímulo ao emprego [6]. Estudos sobre o mercado de trabalho brasileiro mostram que períodos de recessão e eventos inesperados, como a pandemia de COVID-19, provocam elevações expressivas nas taxas de desemprego, seguidas de recuperações graduais, o que reflete a rigidez estrutural e a inércia desse mercado [11]. Assim, compreender a dinâmica temporal da desocupação requer o uso de modelos estatísticos capazes de capturar tendências, sazonalidades e padrões cíclicos, permitindo tanto a análise descritiva quanto a previsão de cenários futuros [8].

A análise de séries temporais é uma abordagem consolidada na estatística e na econometria, cujo objetivo é identificar dependências temporais e padrões de comportamento nos dados observados ao longo do tempo [7]. Entre os modelos tradicionais, destacam-se os modelos autorregressivos integrados de médias móveis (ARIMA), que assumem que o valor atual de uma variável depende linearmente de seus valores passados e de termos de erro [9]. Esses modelos são eficientes em séries estacionárias e bem comportadas, mas apresentam limitações quando há mudanças estruturais abruptas, como as observadas em períodos de crise econômica [8]. Nesses casos, a decomposição aditiva da série — separando-a em tendência, sazonalidade e resíduo — oferece uma leitura mais interpretável dos componentes subjacentes ao fenômeno [12, 13].

A técnica STL (Seasonal-Trend decomposition using Loess), proposta e popularizada por Cleveland et al. [10] e recomendada por Hyndman e Athanasopoulos [13], utiliza regressão local para decompor séries temporais de forma robusta e flexível, permitindo que os componentes sazonais mudem de forma ao longo do tempo. No contexto da taxa de desocupação brasileira, a aplicação do STL permite identificar uma tendência de longo prazo associada às fases de crescimento e recessão da economia, bem como uma sazonalidade leve e estável, relacionada a flutuações típicas do mercado de trabalho entre trimestres — como a queda no desemprego no final do ano e o aumento no início do ciclo anual. A análise da componente residual, por sua vez, evidencia choques transitórios como o aumento abrupto da desocupação em 2020, durante a pandemia, e a posterior normalização gradual [17].

Complementarmente, o uso dos gráficos de autocorrelação (ACF) e autocorrelação parcial (PACF) permite investigar a presença de dependência temporal e identificar a estrutura autorregressiva da série. A ACF mede a correlação entre o valor atual e seus atrasos (lags), enquanto a PACF identifica as correlações diretas com cada defasagem, removendo a influência intermediária [7]. Em séries econômicas, é comum observar altos valores de autocorrelação nos primeiros lags, o que indica persistência temporal e dependência de curto prazo. Esse padrão foi identificado também na série da taxa de desocupação do Brasil, sugerindo um comportamento compatível com um modelo autorregressivo de primeira ordem (AR(1)), em que o valor atual é fortemente influenciado pelo

trimestre anterior [8, 13]. Quando há tendência, a diferenciação da série torna o processo estacionário, conduzindo ao modelo ARIMA(1,1,0), adequado para representar séries que exibem inércia e ajuste lento, como o desemprego [12, 9].

Essa persistência temporal revela uma memória longa do mercado de trabalho: aumentos ou reduções na taxa de desocupação tendem a se propagar ao longo de vários trimestres, refletindo tanto o ritmo lento de criação de empregos quanto o tempo necessário para que políticas de estímulo produzam efeito [6]. A leve sazonalidade observada, por outro lado, reforça a ideia de que o componente cíclico tem impacto marginal frente à tendência estrutural, predominando fatores macroeconômicos de longo prazo [11]. O diagnóstico ACF/PACF corrobora essa interpretação, mostrando que a taxa de desocupação apresenta alta autocorrelação positiva de curto prazo e ausência de picos sazonais marcantes, o que reforça o caráter gradual e persistente das variações no emprego.

Pesquisas recentes destacam que a combinação de métodos clássicos de séries temporais e técnicas modernas de aprendizado de máquina melhora a capacidade preditiva e interpretativa das análises econômicas [16]. No caso da decomposição STL, sua capacidade de isolar componentes e revelar comportamentos não lineares torna-a uma ferramenta poderosa para comunicação de resultados a gestores públicos e formulação de políticas de emprego. Em âmbito global, organismos internacionais como a Organização Internacional do Trabalho (ILO) [14] e a Organização das Nações Unidas [15] reforçam a importância de acompanhar continuamente o desemprego, dentro da agenda do ODS 8 – Trabalho decente e crescimento econômico, como forma de orientar políticas sustentáveis de geração de emprego e renda.

Em síntese, o uso combinado das abordagens STL e ACF/PACF oferece uma compreensão mais profunda da dinâmica da taxa de desocupação, permitindo identificar simultaneamente a estrutura de longo prazo, as variações sazonais e a força de dependência temporal do fenômeno. Essa abordagem se mostra adequada para contextos macroeconômicos complexos como o brasileiro, pois alia robustez estatística, clareza interpretativa e relevância empírica, contribuindo para análises que podem apoiar decisões estratégicas em políticas públicas e planejamento econômico.

3 Diagrama da Solução

1. Coleta (SIDRA/IBGE – PNAD Contínua, t.4099)

Recuperamos a taxa de desocupação para Brasil e UFs via API (2012–2025, frequência trimestral). Mantemos metadados de período (código e rótulo) para rastreabilidade.

2. Padronização temporal e estrutural

Convertemos o código do período (ex.: 201201) em ano, trimestre, período (Q-DEC) e data (1º mês do trimestre). Uniformizamos os atributos (*local*, *cod_local*, *ano*, *trimestre*, *periodo*, *data*, *valor*) e ordenamos as observações por tempo e local. (Opcional: grade trimestral completa, sem imputação.)

3. EDA (analítica e gráfica)

Estatísticas descritivas e visualizações para entender o fenômeno:

- Série temporal do Brasil;
- Boxplot por trimestre (sazonalidade);

- Comparações entre UFs (médias e volatilidade);
- Comparação entre períodos macroeconômicos (pré-crise, crise, pandemia, recuperação).

Essa etapa gera os “achados” iniciais e hipóteses para a modelagem.

4. Decomposição STL (tendência, sazonalidade, resíduo)

Usamos STL (`period = 4`) para separar a série em componentes e obter a série ajustada sazonalmente (`valor_aj_sazonal`). Entregas: colunas `stl_trend`, `stl_season` e `valor_aj_sazonal`, que permitem comparar níveis sem viés sazonal e quantificar a amplitude da sazonalidade.

5. Diagnóstico ACF/PACF

Aplicamos ACF/PACF na série ajustada (e, se necessário, nas versões diferenciadas $\Delta 1/\Delta 4$) para avaliar memória temporal e sinais sazonais residuais.

Resultado esperado desta etapa: um *notebook* com dados padronizados, painéis de EDA, componentes STL e gráficos/tabelas de ACF/PACF, prontos para alimentar a análise final e futuras projeções exploratórias, alinhadas ao ODS 8.

Visão geral

[Coleta] → [Padronização] → [EDA] → [Decomposição STL] → [Diagnóstico ACF/PACF] → [Comparação entre UFs] → [Síntese dos Achados]

Pré-processamento e Análise Exploratória de Dados (EDA)

1. Coleta e organização dos dados

Recuperação da taxa de desocupação para Brasil e UFs a partir da PNAD Contínua (SIDRA/IBGE, tabela 4099), via API, para o período de 2012 a 2025 (frequência trimestral). Armazenamento em formato tabular, com manutenção dos metadados de período (código e rótulo).

... Brasil (raw) – amostra:

	NC	NN	MC	MN	V	D1C	D1N	D2C		D2N	D3C	D3N
0	Nível Territorial (Código)	Nível Territorial	Unidade de Medida (Código)	Unidade de Medida	Valor	Brasil (Código)	Brasil	Variável (Código)		Variável	Trimestre (Código)	Trimestre
1	1	Brasil	2	%	8.0	1	Brasil	4099		Taxa de desocupação, na semana de referência, ...	201201	1º trimestre 2012
2	1	Brasil	2	%	7.6	1	Brasil	4099		Taxa de desocupação, na semana de referência, ...	201202	2º trimestre 2012
3	1	Brasil	2	%	7.1	1	Brasil	4099		Taxa de desocupação, na semana de referência, ...	201203	3º trimestre 2012
4	1	Brasil	2	%	6.9	1	Brasil	4099		Taxa de desocupação, na semana de referência, ...	201204	4º trimestre 2012

Figura 1: Amostra dos dados.

2. Padronização temporal e estrutural

Conversão do código do período (por exemplo, 201201) em ano, trimestre, período (Q-DEC) e data (1º mês do trimestre). Uniformização dos atributos (*local*, *cod_local*, *ano*, *trimestre*, *periodo*, *data*, *valor*) e ordenação das observações por tempo e local, sem imputação de lacunas trimestrais.

3. Tratamento de tipos e consistência

Conversão explícita dos campos de data para o formato `datetime` e dos valores da taxa de desocupação para tipo numérico, com remoção de registros com datas ou valores inválidos. Verificação da continuidade temporal e da ausência de valores faltantes relevantes.

Brasil (tidy) – amostra:

	cod_local	local	cod_perodo	perodo_nome	valor	ano	trimestre	perodo	data
0	1	Brasil	201201	1º trimestre 2012	8.0	2012	1	2012Q1	2012-01-01
1	1	Brasil	201202	2º trimestre 2012	7.6	2012	2	2012Q2	2012-04-01
2	1	Brasil	201203	3º trimestre 2012	7.1	2012	3	2012Q3	2012-07-01
3	1	Brasil	201204	4º trimestre 2012	6.9	2012	4	2012Q4	2012-10-01
4	1	Brasil	201301	1º trimestre 2013	8.0	2013	1	2013Q1	2013-01-01

Figura 2: Amostra dos dados após tratamento.

4. Cálculo de médias e medidas descritivas

Cálculo de médias, medianas, desvios-padrão, valores máximos e mínimos da taxa de desocupação por Brasil e por UF, bem como por períodos agregados (pré-crise, crise, pandemia, recuperação), visando caracterizar o nível e a variabilidade do desemprego.

=== EDA – Brasil (estatísticas descritivas) ===

valor

count	55.000000
mean	9.785455
std	2.776917
min	5.600000
25%	7.300000
50%	9.000000
75%	12.100000
max	14.900000

Figura 3: Estatísticas Descritivas

5. Série temporal do Brasil

Construção da série trimestral da taxa de desocupação (2012–2025) para o Brasil, incluindo a média móvel de 4 trimestres, com o objetivo de suavizar oscilações de curto prazo e destacar a tendência de longo prazo.

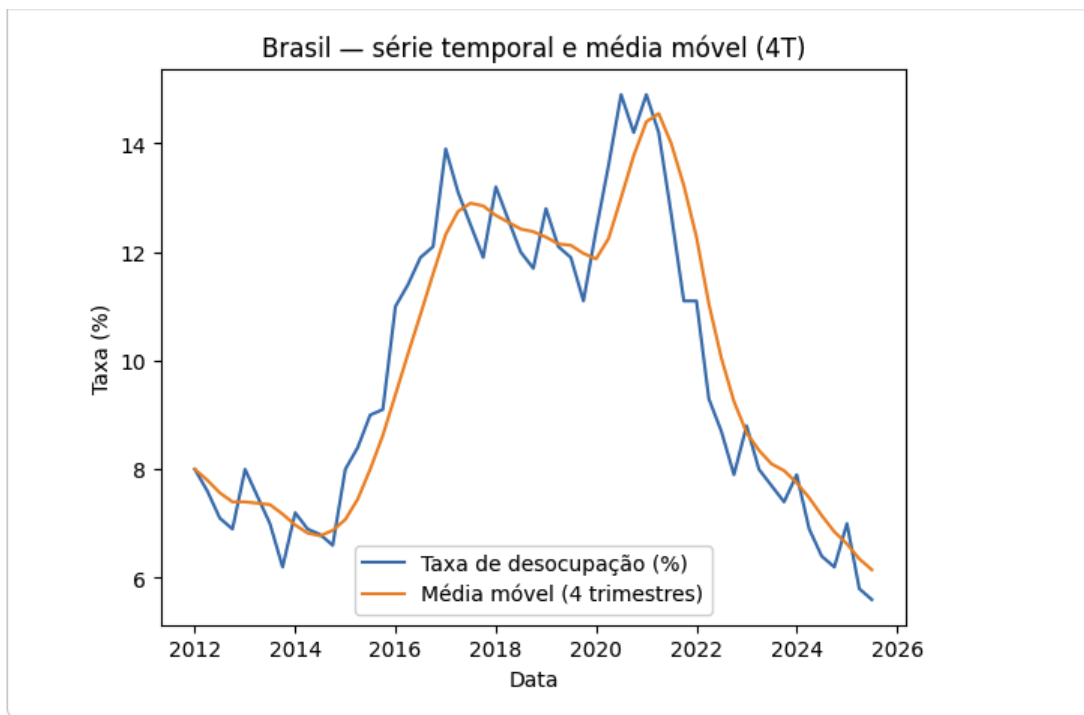


Figura 4: Série temporal - Brasil

6. Análise de sazonalidade via boxplots

Elaboração de boxplots da taxa de desocupação por trimestre (T1, T2, T3, T4), para identificar padrões sazonais, diferenças de nível entre trimestres e eventuais valores extremos ao longo do ano.

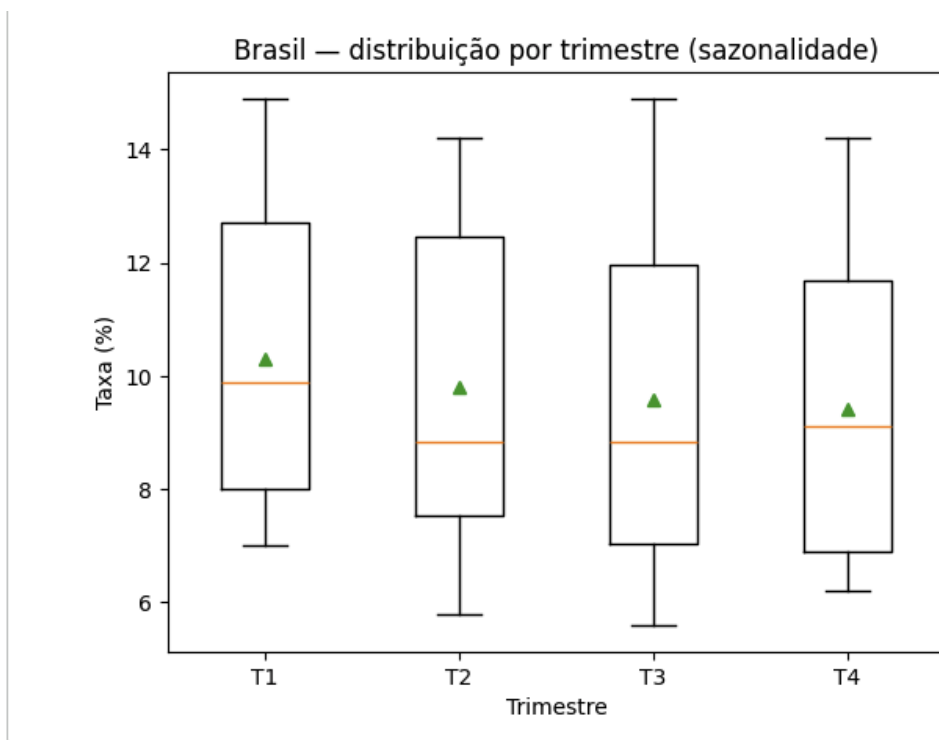


Figura 5: Distribuição por trimestre

7. Comparação entre UFs (médias e volatilidade)

Cálculo das médias trimestrais da taxa de desocupação por UF e identificação dos estados com maiores e menores taxas médias. Visualização gráfica das trajetórias temporais para comparar níveis e volatilidade entre as unidades da federação.

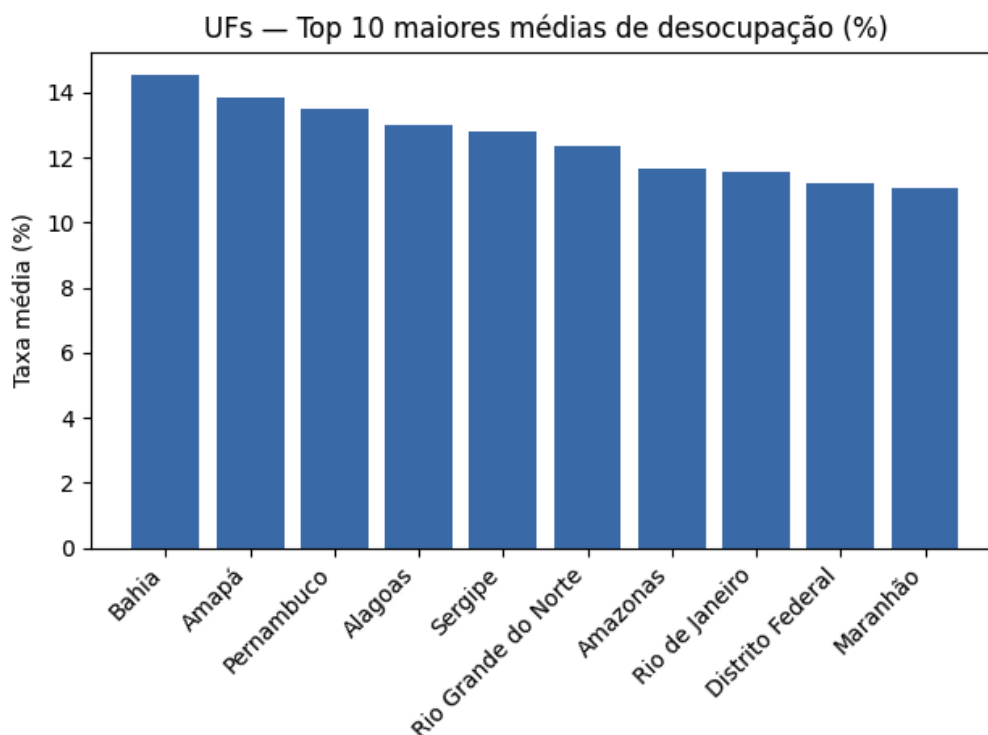


Figura 6: Maiores médias de desocupação

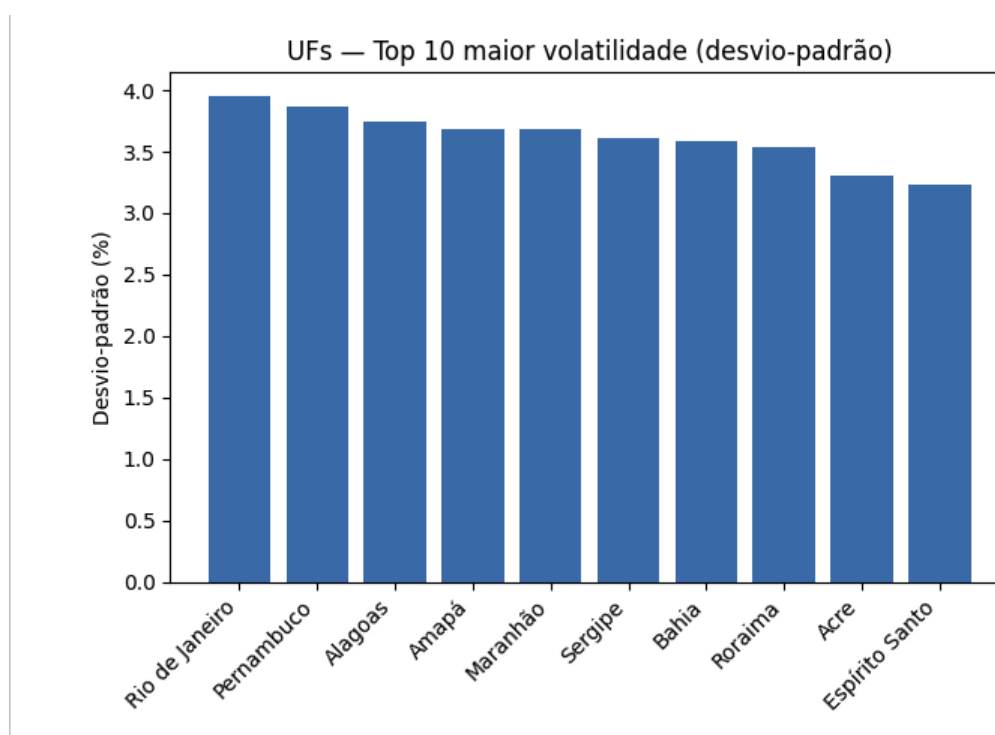


Figura 7: Maiores volatilidades

8. Comparação entre períodos macroeconômicos

Agrupamento dos trimestres em quatro regimes (pré-crise, crise 2015–2016, pandemia de COVID-19 e recuperação) e cálculo das taxas médias de desocupação em cada período, com gráficos comparativos para evidenciar mudanças de patamar ao longo do tempo.

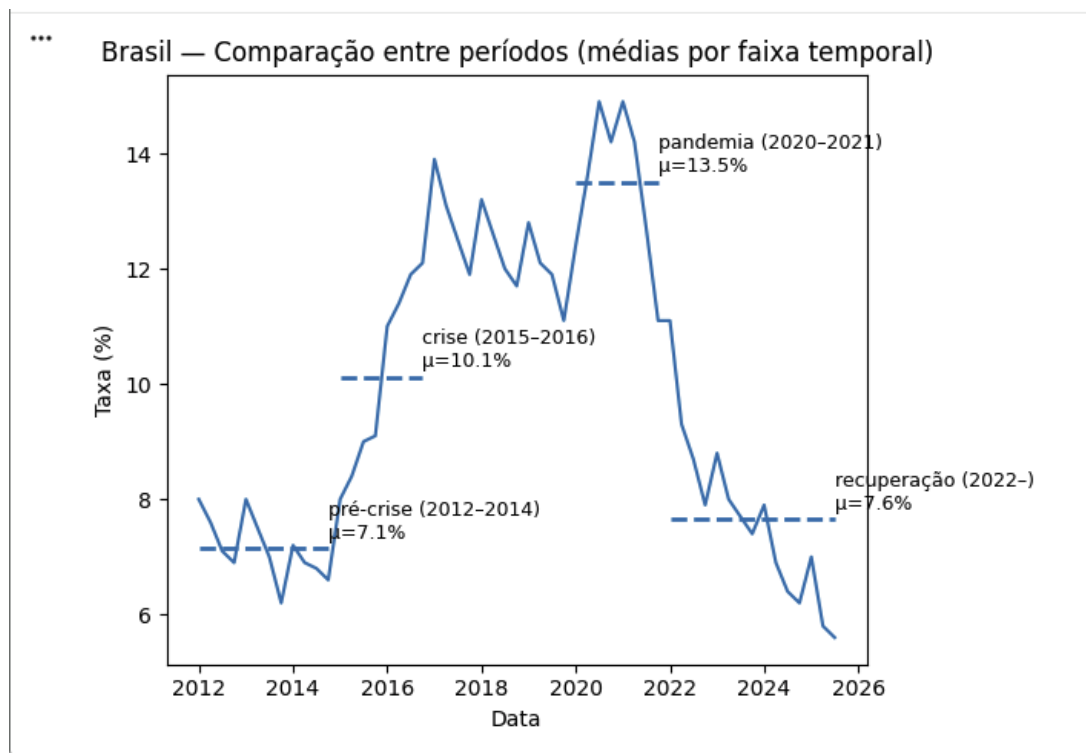


Figura 8: Comparação entre períodos

9. Identificação de achados iniciais

A partir das estatísticas e visualizações, elaboração de hipóteses sobre tendência, sazonalidade, choques e disparidades regionais, que orientam as etapas seguintes de decomposição STL e diagnóstico ACF/PACF.

3.1 Análise dos dados empregados: qualidade, estrutura, limitações e recortes adotados

Os dados foram padronizados e estruturados para garantir consistência temporal e comparabilidade entre períodos e regiões. O código do período foi convertido em colunas explícitas de ano, trimestre, período e data, permitindo ordenação cronológica correta e o uso de funções de janela e modelagem. Essa transformação é ilustrada na Figura 2, em que a base é organizada em formato longo, com as colunas principais *local*, *ano*, *trimestre* e *valor*, sem imputação de lacunas, preservando a integridade dos dados originais.

Os dados da PNAD Contínua (IBGE) apresentam alta qualidade e continuidade temporal, sem valores ausentes ou *outliers* relevantes. A Figura 4 evidencia uma trajetória coerente: aumento da desocupação na crise de 2015–2016, pico durante a pandemia (2020–2021) e queda gradual no período de recuperação (2022–2025), confirmando a plausibilidade econômica da série.

A análise da sazonalidade (ver Figura 5) mostrou variações suaves entre trimestres, com médias próximas ($T1 \approx 10,3\%$ e $T4 \approx 9,4\%$). O boxplot indica sazonalidade leve, sugerindo que a tendência de longo prazo é o principal componente da série. As disparidades regionais são evidenciadas nas figuras seguintes: a Figura 6 destaca Bahia, Amapá e Pernambuco como estados com maiores taxas médias (cerca de 13–15%), enquanto a Figura 7 mostra que Rio de Janeiro e Pernambuco apresentam maior volatilidade, refletindo instabilidade conjuntural e possível dependência de setores mais sensíveis ao ciclo econômico.

Foram adotadas simplificações metodológicas para manter o foco exploratório: análise apenas da taxa de desocupação agregada, sem desagregações por sexo, idade ou escolaridade, e uso de ferramentas descritivas e de decomposição (média móvel e STL) em vez de modelos paramétricos completos (como SARIMA). Essas escolhas, ilustradas ao longo das figuras de EDA, priorizam clareza interpretativa em detrimento de maior complexidade modelística.

Entre as limitações, destacam-se o erro amostral da PNAD (mais relevante em estados de menor população), a periodicidade trimestral (que suaviza choques muito rápidos) e a ausência de variáveis explicativas externas (como PIB, inflação ou políticas públicas específicas). Ainda assim, as evidências apresentadas nas figuras de EDA indicam que se trata de uma base robusta, pública e representativa, adequada para análises exploratórias do mercado de trabalho e para apoiar etapas posteriores de modelagem temporal e projeções alinhadas ao ODS 8.

A dinâmica entre pré-crise, crise, pandemia e recuperação também pode ser vista na Figura 8, que resume as taxas médias por regime macroeconômico.

4 Modelos

4.1 Seasonal-Trend decomposition using Loess

A decomposição da série temporal foi realizada utilizando o método STL (*Seasonal-Trend Decomposition using Loess*), aplicado à taxa trimestral de desocupação do Brasil (2012–2025). O STL permite decompor uma série temporal em três componentes fundamentais: tendência, sazonalidade e resíduo, fornecendo uma visão clara das variações estruturais e cíclicas do fenômeno ao longo do tempo. Foi adotado um período sazonal de 4 trimestres, correspondente à frequência anual.

O procedimento envolveu os seguintes passos:

1. **Interpolação e regularização temporal:** assegurou que os intervalos fossem trimestrais e contínuos, condição necessária para o ajuste do modelo.
2. **Aplicação do STL:** separação da série original em componentes independentes — *trend* (tendência), *season* (sazonalidade) e *resid* (resíduo).

...

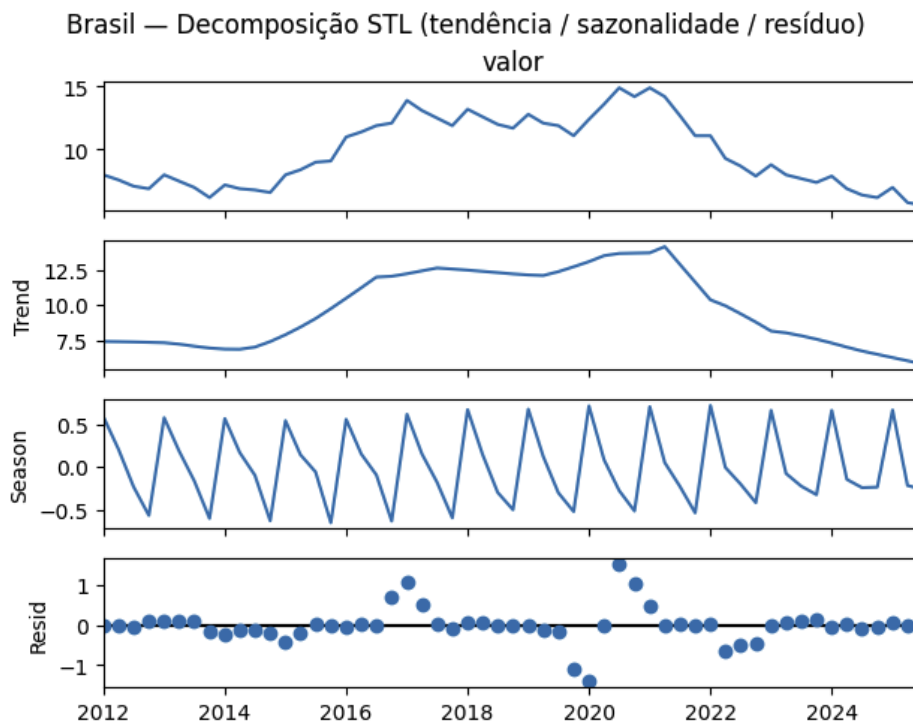


Figura 9: Brasil decomposição STL

3. **Construção da série ajustada sazonalmente:** obtenção da série à vista da subtração da componente sazonal do valor original ($\text{valor_aj_sazonal} = \text{valor} - \text{stl_season}$).
4. **Visualização e validação:** elaboração de gráficos comparando a série original e a série ajustada, permitindo avaliar a influência e a estabilidade das oscilações sazonais.

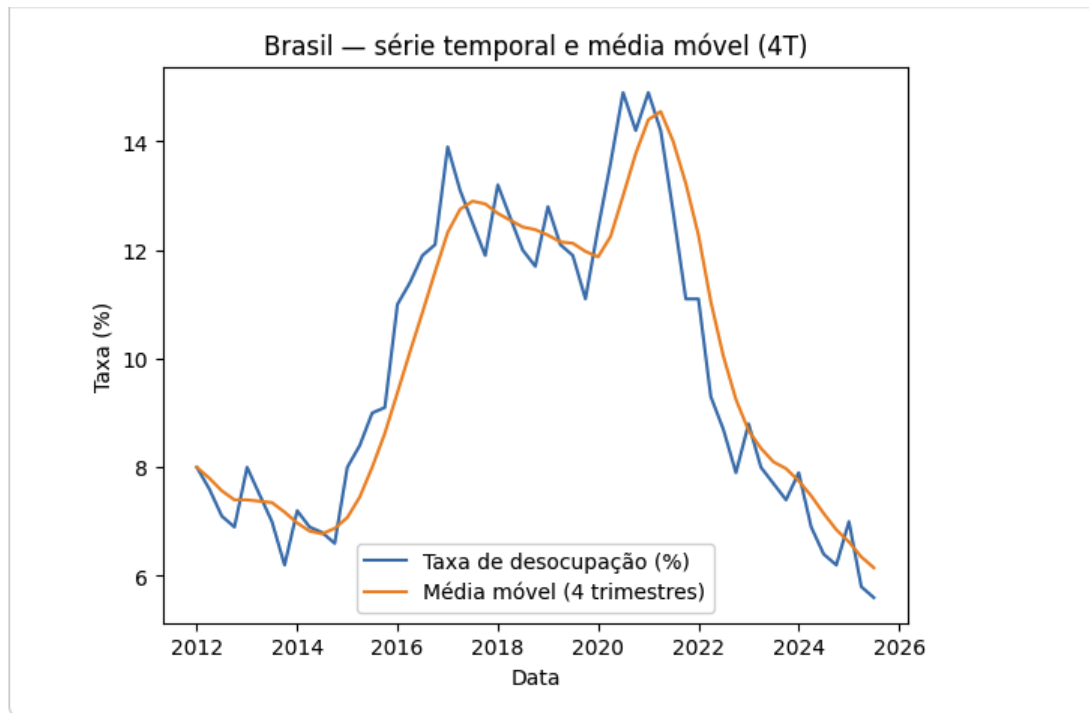


Figura 10: Brasil série original vs ajustada sazonalmente

4.2 ACF e PACF

O procedimento consistiu em:

1. **Conversão e ordenação temporal:** a coluna *data* foi transformada para o formato de data e a série ordenada cronologicamente.
2. **Cálculo da ACF e PACF:**
 - A ACF (*Autocorrelation Function*) mede a correlação da série com suas defasagens (*lags*), indicando o quanto o valor atual depende de períodos anteriores.
 - A PACF (*Partial Autocorrelation Function*) mostra o impacto direto de cada defasagem após remover a influência das intermediárias.
3. **Comparação entre a série original e a ajustada sazonalmente:** avaliou-se o comportamento antes e depois da remoção da sazonalidade, para verificar se os padrões de correlação mudam.

Os gráficos apresentam, no eixo x , as defasagens trimestrais (*lags*) e, no eixo y , os coeficientes de correlação entre -1 e 1 . A faixa azul representa o **intervalo de confiança de 95%**.

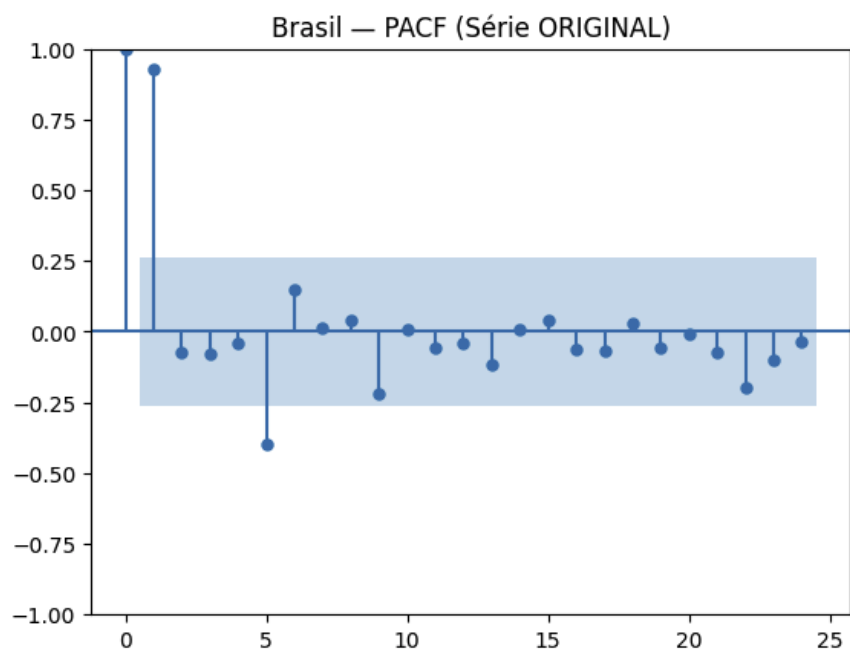
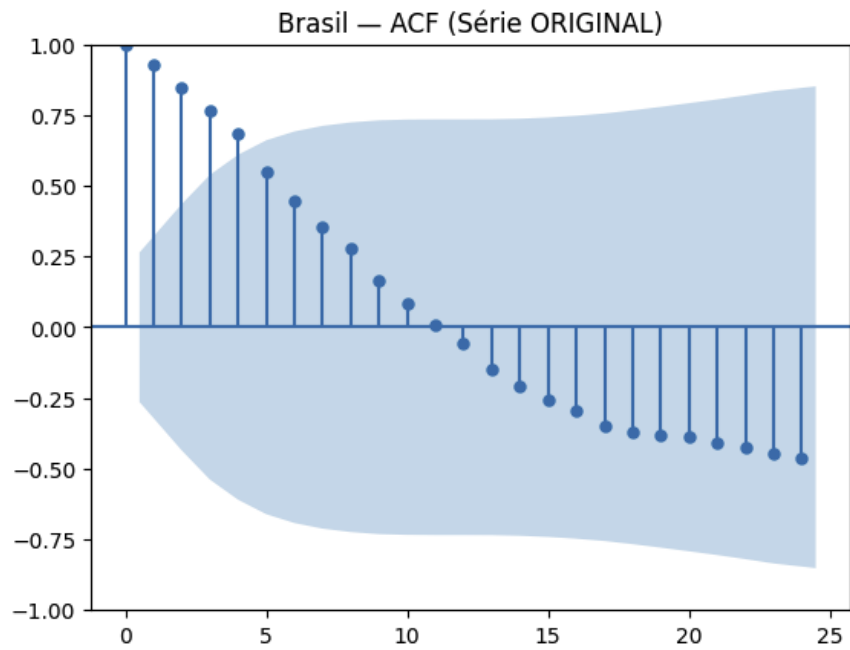
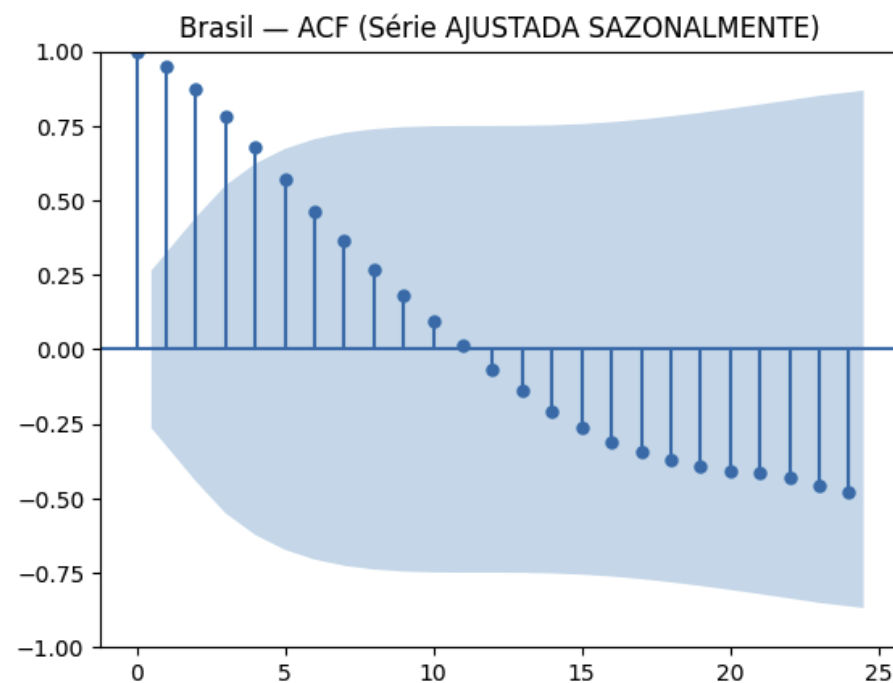


Figura 11: Brasil ACF/PACF - série original



<Figure size 640x480 with 0 Axes>

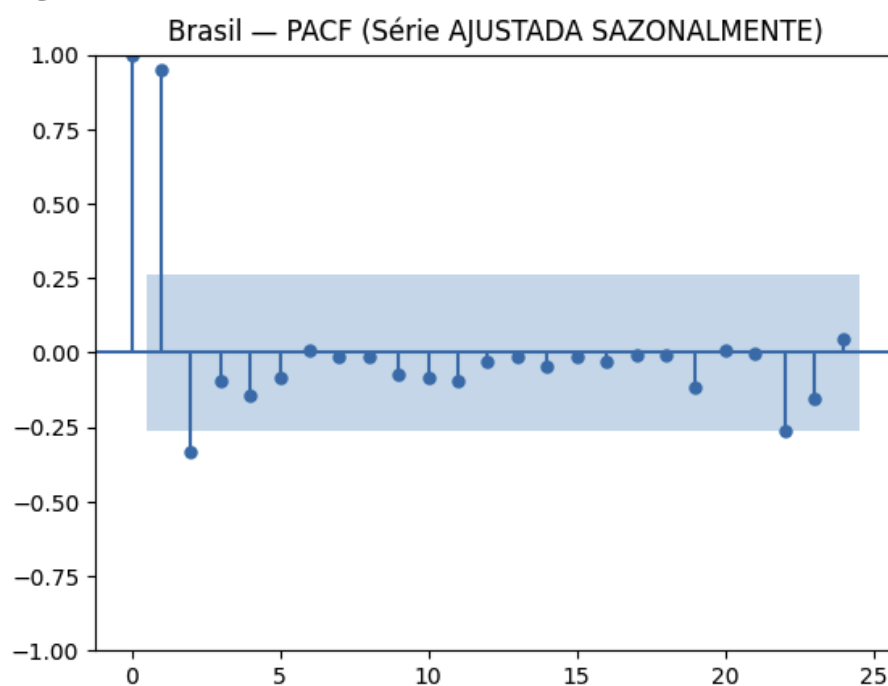


Figura 12: Brasil ACF/PACF - série ajustada

4.3 Comparação entre UFs

UFs com maior e menor média da taxa de desocupação (2012-2025)
Bahia x Santa Catarina

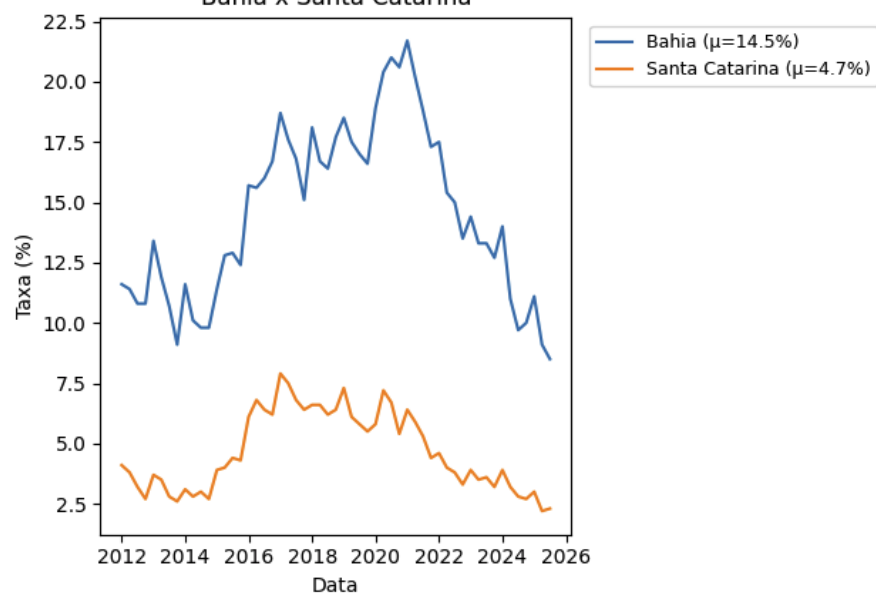


Figura 13: Maior e menor média de desemprego

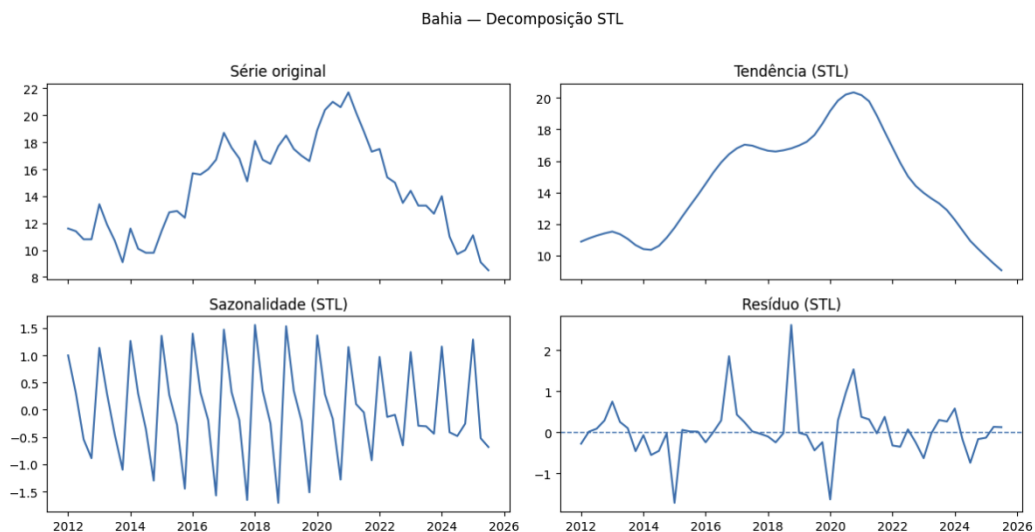


Figura 14: Bahia STL

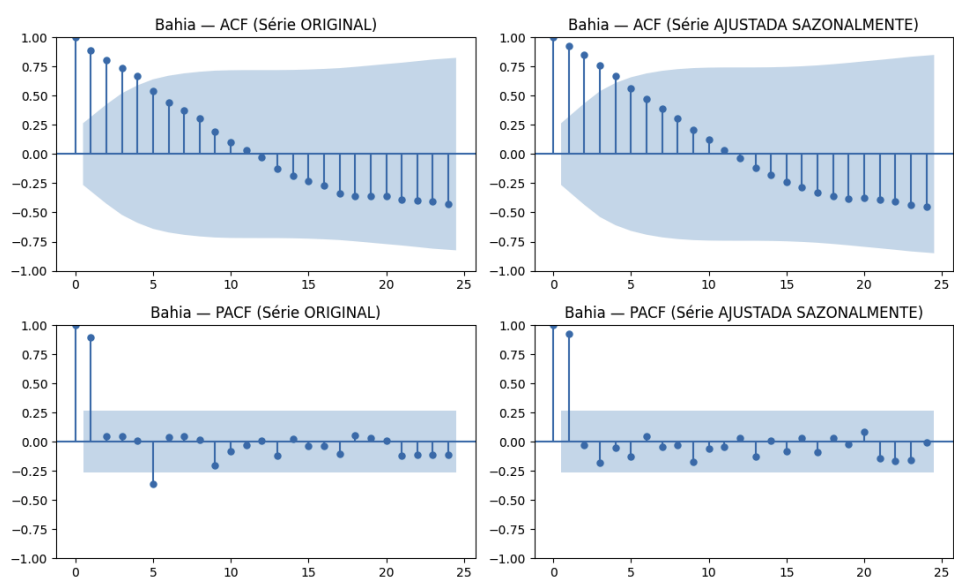


Figura 15: Bahia ACF/PACF

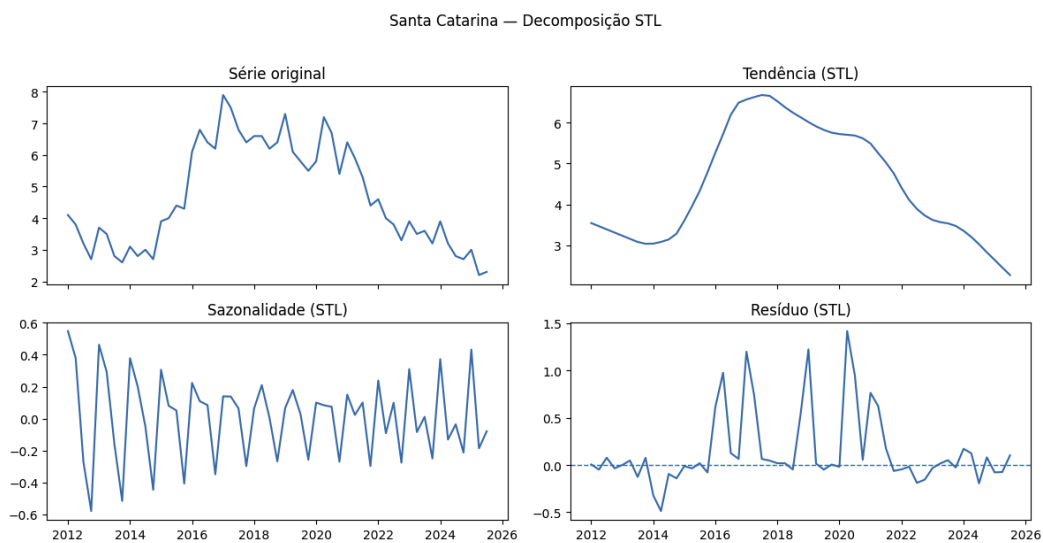


Figura 16: Santa Catarina STL

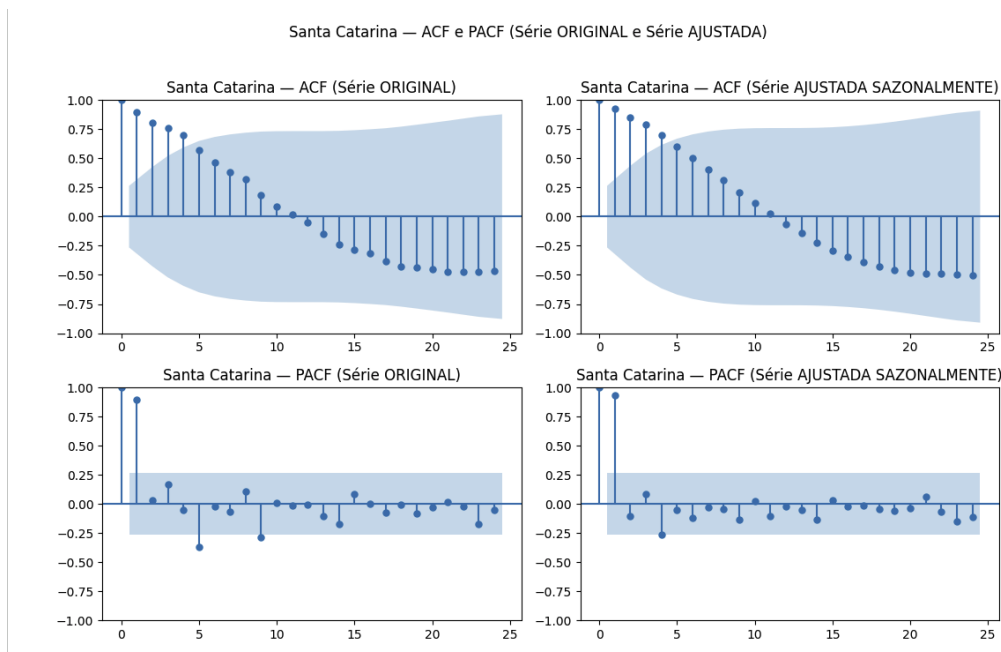


Figura 17: Santa Catarina ACF/PACF

5 Resultados

5.1 Diagnóstico STL (Brasil)

A decomposição revelou que a tendência (*trend*) da taxa de desocupação no Brasil apresenta três fases bem definidas (Figura 10). No primeiro momento, de **crescimento (2014–2017)**, há um aumento contínuo da taxa, refletindo o impacto da crise econômica e política iniciada em 2015. Em seguida, observa-se uma fase de **estabilidade com leve queda (2018–2019)**, compatível com um período de recuperação moderada do mercado de trabalho. Por fim, ocorre um **novo pico e posterior queda (2020–2025)**, com forte elevação durante a pandemia de COVID-19, quando a taxa se aproxima de 15%, seguida de redução expressiva a partir de 2022, chegando a valores próximos de 6% em 2025.

Esse padrão é sintetizado no gráfico de **comparação entre períodos** (Brasil — Comparação entre períodos), que resume a trajetória em quatro regimes macroeconômicos, conforme ilustrado na Figura 8: pré-crise (2012–2014), com média em torno de 7,1%; crise (2015–2016), com média aproximada de 10,1%; pandemia (2020–2021), com média em torno de 13,5%; e recuperação (2022–), com média próxima de 7,6%. Esses patamares médios destacam o salto no desemprego durante a crise e, principalmente, na pandemia, bem como a queda consistente na fase recente de recuperação.

A componente **sazonal** (*season*) manteve um padrão regular de flutuação trimestral, com amplitude aproximada de $\pm 0,6$ ponto percentual, o que indica **sazonalidade leve e estável**. Esse comportamento sugere que, embora existam variações sazonais típicas — como aumento no início do ano e redução no último trimestre — o efeito é pequeno diante das mudanças estruturais na tendência.

O **resíduo** (*resid*) mostrou dispersão concentrada em torno de zero, sem picos acen-tuados, indicando boa adequação do modelo e ausência de grandes choques fora dos já capturados pela tendência. Pequenas oscilações residuais em 2020–2021 estão associadas a impactos extraordinários da pandemia, mas não comprometem a estabilidade geral da decomposição.

Ao comparar as séries **original e ajustada sazonalmente**, como ilustrado na Figura 10, observa-se que a linha ajustada é mais suave, eliminando variações periódicas e permitindo avaliar melhor a evolução estrutural do desemprego. Isso facilita a análise de políticas públicas e eventos macroeconômicos, uma vez que o comportamento ajustado reflete apenas movimentos de longo prazo, sem interferência das oscilações cíclicas.

5.2 Diagnóstico ACF/PACF (Brasil)

O gráfico da ACF da série original (Figura 11) revela correlações muito elevadas nos primeiros *lags* (aproximadamente 0,9 no *lag* 1) e um decaimento lento e gradual ao longo das defasagens. Esse padrão indica forte **persistência temporal**: o nível de desemprego em um trimestre está fortemente condicionado aos valores observados nos trimestres anteriores. Além disso, o fato de a autocorrelação permanecer positiva por vários *lags* antes de cruzar a linha zero sugere que a série não é estacionária em nível, sendo compatível com a presença de tendência ou de raiz unitária.

A PACF da mesma série, também apresentada na Figura 11, mostra um pico expressivo e estatisticamente significativo apenas no *lag* 1, enquanto os demais *lags* se mantêm próximos de zero e dentro da faixa de confiança. Esse comportamento é típico de um processo autorregressivo de primeira ordem (AR(1)), em que o valor corrente é fortemente explicado apenas pelo trimestre imediatamente anterior, e contribui para a escolha de modelos do tipo ARIMA com componente autorregressiva simples (por exemplo, ARIMA(1,1,0) após a diferenciação).

Após a remoção da componente sazonal identificada pela decomposição STL, a ACF da série ajustada sazonalmente (Figura 12) preserva o formato geral observado na série original, com correlações altas nos primeiros *lags* e decaimento progressivo. No entanto, nota-se uma leve redução das autocorrelações em *lags* intermediários, indicando que parte da estrutura de dependência estava associada ao padrão sazonal trimestral, agora isolado. Ainda assim, o comportamento confirma que a série ajustada continua exibindo memória longa, dominada por fatores estruturais (tendência e ciclos econômicos) e não apenas por sazonalidade.

Na PACF da série ajustada, também mostrada na Figura 12, o padrão se mantém: o *lag* 1 permanece como o único claramente significativo, enquanto os demais *lags* se distribuem em torno de zero. Isso reforça o diagnóstico de uma estrutura autorregressiva simples, em que um termo AR(1) é suficiente para capturar a dependência de curto prazo, tanto na série original quanto na série sem sazonalidade. Em conjunto, os gráficos de ACF e PACF sugerem que, uma vez tratadas a tendência e a sazonalidade, um modelo parcimonioso com componente AR(1) sobre a série diferenciada é uma especificação adequada para representar a dinâmica do desemprego brasileiro.

5.3 Comparação entre Bahia e Santa Catarina

A comparação direta das séries trimestrais da Bahia e de Santa Catarina (Brasil — UFs com maior e menor média da taxa de desocupação, 2012–2025), ilustrada na Figura 13, evidencia um contraste acentuado de níveis e trajetórias. A Bahia apresenta uma taxa média em torno de 14,5%, enquanto Santa Catarina gira em torno de 4,7%, o que indica que, ao longo de todo o período, o desemprego baiano se manteve estruturalmente elevado, mais que o triplo do patamar catarinense na maior parte do tempo. Ambas as séries acompanham o ciclo macroeconômico nacional — aumento na crise e na pandemia, queda

na recuperação recente — mas partindo de “platôs” muito diferentes.

A decomposição STL da Bahia (Figura 14) mostra uma tendência com forte elevação entre 2014 e 2019, atingindo picos superiores a 20% no período da pandemia, seguida de queda mais lenta a partir de 2021–2022. Isso sugere que choques negativos no mercado de trabalho baiano geram efeitos persistentes e demoram a ser revertidos. A componente sazonal apresenta amplitude relativamente alta (cerca de $\pm 1,5$ p.p.), com oscilações regulares entre trimestres, o que indica que a estrutura produtiva do estado é bastante sensível a variações sazonais, possivelmente associadas a setores com forte rotatividade ou forte concentração de atividades em determinados períodos do ano. Os resíduos exibem picos relevantes em alguns trimestres, sinalizando choques idiossincráticos adicionais além do ciclo agregado.

Em Santa Catarina, a decomposição STL (Figura 16) revela uma tendência muito mais moderada: mesmo nos momentos de maior estresse, a taxa de desocupação permanece em níveis relativamente baixos (próximos de 7–8%), com queda consistente após o auge da pandemia. A sazonalidade tem amplitude bem menor (cerca de $\pm 0,5$ p.p.), o que sugere que o mercado de trabalho catarinense sofre menos com flutuações de curto prazo e apresenta um padrão de contratação mais estável ao longo do ano. Os resíduos se mantêm, em geral, próximos de zero, com poucos desvios extremos, reforçando a leitura de maior estabilidade e resiliência.

Os gráficos de ACF e PACF para a Bahia e para Santa Catarina (Figuras 15 e 17, respectivamente) confirmam a presença de alta persistência temporal: a autocorrelação é elevada nos primeiros *lags* e decai lentamente, indicando que choques na taxa de desocupação se propagam por vários trimestres. A PACF apresenta um pico dominante no primeiro *lag*, com os demais coeficientes próximos de zero, padrão característico de um processo autorregressivo de primeira ordem (AR(1)). Em conjunto, esses resultados sugerem que, do ponto de vista da dinâmica temporal, ambos os estados podem ser descritos por estruturas AR(1), mas com diferenças marcantes em nível, amplitude sazonal e volatilidade residual.

Em síntese, a análise conjunta dos gráficos de série temporal, decomposição STL e ACF/PACF mostra que Bahia e Santa Catarina compartilham o mesmo “formato” de ciclo (crise–pandemia–recuperação), porém inseridos em realidades muito distintas: a Bahia combina **desemprego estruturalmente alto, sazonalidade forte e maior exposição a choques**, enquanto Santa Catarina apresenta **níveis baixos, sazonalidade moderada e rápida absorção de distúrbios**. Essa assimetria reforça a importância de políticas de emprego com foco regional, especialmente em estados com maior vulnerabilidade estrutural.

6 Discussão

Os resultados obtidos mostram que a trajetória da taxa de desocupação no Brasil entre 2012 e 2025 está fortemente alinhada aos grandes ciclos macroeconômicos do período: pré-crise, crise de 2015–2016, choque da pandemia de COVID-19 e fase de recuperação recente. A decomposição STL evidenciou uma tendência de alta até o auge da pandemia, seguida de queda consistente a partir de 2022, o que é coerente com a redução histórica da taxa de desocupação observada em 2024–2025. A sazonalidade se mostrou leve e estável, indicando que o comportamento do desemprego é dominado por fatores estruturais, e não apenas por flutuações de curto prazo. Em termos de ODS 8, esse resultado sugere que o país vem avançando em direção ao objetivo de “trabalho decente e crescimento econômico”,

mas ainda de forma desigual e vulnerável a choques externos.

A comparação entre Bahia (maior média de desocupação) e Santa Catarina (menor média) reforça o caráter estrutural das assimetrias regionais do mercado de trabalho brasileiro. Enquanto a Bahia apresenta um nível médio em torno de 14,5%, com picos superiores a 20% e sazonalidade mais intensa, Santa Catarina mantém taxas em torno de 4,7%, com flutuações mais moderadas e rápida recuperação após choques. As decomposições STL de ambos os estados mostram um ciclo qualitativo semelhante (alta na crise e na pandemia, queda na recuperação), porém em patamares muito distintos. Os gráficos de ACF e PACF indicam forte persistência temporal e estrutura autorregressiva de primeira ordem (AR(1)) em ambos os casos, o que significa que choques negativos tendem a se propagar ao longo de vários trimestres. Na prática, isso implica que estados com níveis iniciais mais altos, como a Bahia, tendem a levar mais tempo para retornar a patamares mais baixos de desemprego, ampliando desigualdades regionais e tensionando o cumprimento do ODS 8 em todo o território nacional.

Do ponto de vista de política pública, as evidências sugerem que **não basta uma estratégia macro uniforme**. Em estados com baixa desocupação e maior resiliência, como Santa Catarina, o foco pode estar na qualificação profissional contínua, na transição para empregos de maior produtividade e na formalização de vínculos, reforçando a dimensão “trabalho decente” do ODS 8. Já em estados com desemprego estrutural elevado, como a Bahia, é necessário combinar políticas de estímulo à atividade econômica (infraestrutura, incentivos setoriais, atração de investimentos) com políticas ativas de emprego (programas de intermediação de mão de obra, subsídios temporários à contratação, qualificação focada em setores dinâmicos locais). A sazonalidade mais intensa na Bahia sugere espaço para políticas específicas em períodos críticos do ano, como programas temporários, frentes de trabalho ou reforço de benefícios de proteção social nos trimestres em que a taxa sobe.

A persistência temporal identificada pelos gráficos de ACF/PACF também tem implicações importantes. Como a taxa de desocupação apresenta forte memória, políticas de curto prazo tendem a ter efeitos limitados se não forem mantidas ou complementadas por ações estruturais de médio e longo prazo. Isso vale tanto para as respostas à crise (como durante a pandemia) quanto para a consolidação da recuperação. A presença de um comportamento compatível com modelos AR(1) indica que, uma vez iniciado um ciclo de redução do desemprego, é fundamental evitar interrupções abruptas em políticas de geração de emprego, sob risco de “reverter” avanços e prolongar períodos de desocupação elevada, especialmente em regiões mais vulneráveis.

7 Conclusão e recomendações para o poder público

Os resultados deste estudo, obtidos por meio de EDA, decomposição STL e análise ACF/PACF, mostram que a taxa de desocupação no Brasil avançou em direção às metas do ODS 8, principalmente a partir de 2022, mas ainda apresenta **marcantes desigualdades regionais e forte inércia temporal**. As séries da Bahia e de Santa Catarina ilustram dois extremos de um mesmo sistema: enquanto uma parte do país vive o desafio do desemprego estrutural, outra opera próxima de pleno emprego, ainda que sujeita a choques conjunturais. Isso exige políticas públicas territorializadas, baseadas em evidências e monitoradas continuamente.

Com base nos diferentes períodos analisados, algumas recomendações são:

- **Período pré-crise (2012–2014)**

Utilizar ferramentas de decomposição e acompanhamento de séries (como STL e médias móveis) para construir **painéis de monitoramento antecipado**. Mesmo em fases de relativa estabilidade, é possível detectar mudanças sutis na tendência que indiquem deterioração futura do mercado de trabalho, permitindo que medidas anticíclicas sejam desenhadas antes da materialização de uma crise.

- **Crise de 2015–2016**

Empregar modelos de séries temporais (ARIMA/AR(1) sobre série diferenciada) para simular cenários alternativos de desemprego sob diferentes pacotes de políticas (investimento público, redução de tributos setoriais, programas de emprego). Essas simulações podem apoiar decisões sobre **intensidade e duração** das políticas de estímulo, ajudando a encurtar o período de alta desocupação.

- **Pandemia de COVID-19 (2020–2021)**

Em contextos de choque abrupto, o uso de séries ajustadas sazonalmente e de indicadores em tempo quase real permite avaliar com mais clareza o impacto de medidas emergenciais (auxílios, manutenção de renda, programas de preservação de empregos). A análise residual e de quebras estruturais pode ajudar o poder público a distinguir o que é efeito transitório da crise e o que se torna uma nova tendência, orientando a transição de políticas emergenciais para políticas estruturais.

- **Período de recuperação (2022–2025)**

Ferramentas como STL, ACF/PACF e *dashboards* com séries regionais permitem acompanhar se a recuperação está sendo **inclusiva** ou concentrada em poucos estados. Nos casos em que a tendência de queda do desemprego é lenta (como na Bahia), os indicadores servem de alerta para a necessidade de reforçar políticas de emprego e desenvolvimento produtivo local, evitando que a recuperação nacional oculte bolsões persistentes de vulnerabilidade.

Quanto ao uso da ferramenta analítica, o *notebook* desenvolvido neste projeto pode ser facilmente adaptado para um ambiente institucional (por exemplo, uma secretaria de trabalho ou um observatório de políticas públicas). Ele já organiza a coleta automatizada de dados (SIDRA/IBGE), padroniza séries, gera gráficos de tendência, sazonalidade, comparação regional e diagnósticos ACF/PACF. Em formato de painel interativo ou rotinas periódicas de análise, esse tipo de solução permite ao poder público:

- **Monitorar continuamente** a taxa de desocupação por UF, região e faixas de tempo;
- **Detectar rapidamente mudanças** na tendência ou na volatilidade do desemprego;
- **Comparar estados extremos** (como Bahia e Santa Catarina) para calibrar políticas específicas;
- **Avaliar o impacto de políticas** ao observar como a tendência e a componente residual evoluem após intervenções.

Em suma, a combinação de dados oficiais (PNAD Contínua) com técnicas de séries temporais fornece uma base sólida para decisões alinhadas ao ODS 8. Ela permite que

gestores públicos saiam de uma visão puramente descritiva e passem a trabalhar com **evidências dinâmicas**, capazes de orientar a formulação, o monitoramento e a correção de políticas de emprego em tempo hábil, com foco na redução das desigualdades regionais e na promoção de trabalho decente para toda a população brasileira.

8 Melhorias futuras

Embora os resultados obtidos com EDA, decomposição STL e análise ACF/PACF tenham se mostrado consistentes e informativos, há diversas possibilidades de aprimoramento metodológico para trabalhos futuros:

- **Incorporação de modelos não lineares e redes neurais**

Explorar arquiteturas de aprendizado profundo para séries temporais, como redes recorrentes (RNN), LSTM/GRU ou modelos baseados em *Transformers*, capazes de capturar relações não lineares e padrões de longo alcance na dinâmica do desemprego. Esses modelos poderiam ser comparados a especificações ARIMA tradicionais em termos de capacidade preditiva e robustez a choques.

- **Maior cruzamento com variáveis de contexto macroeconômico e social**

Integrar ao modelo variáveis explicativas como PIB, inflação, massa salarial, nível de formalização, indicadores de políticas públicas (por exemplo, gasto social ou programas de emprego) e choques setoriais. Essa abordagem, via modelos ARIMAX, VAR ou redes neurais multivariadas, permitiria avaliar explicitamente como fatores de contexto influenciam a taxa de desocupação e fortaleceria a capacidade de simulação de cenários de política econômica.

- **Análise desagregada por grupos populacionais**

Estender a análise para recortes por sexo, faixa etária, escolaridade ou cor/raça, utilizando os microdados da PNAD Contínua. Isso permitiria investigar se a dinâmica temporal e a resposta a choques (crise, pandemia, recuperação) são homogêneas entre grupos ou se determinados segmentos sofrem desemprego mais persistente, o que é relevante para o ODS 8 na perspectiva de “não deixar ninguém para trás”.

- **Modelos hierárquicos e espaciais para as UFs**

Empregar modelos hierárquicos bayesianos ou modelos de espaço de estado multirregionais que permitam compartilhar informação entre UFs, levando em conta correlações espaciais e similaridades estruturais. Isso pode melhorar estimativas em estados com amostras menores e possibilitar inferências mais precisas sobre padrões regionais de desemprego.

- **Integração com painéis de monitoramento em tempo quase real**

Automatizar o fluxo de coleta, atualização e previsão em um *dashboard* interativo (por exemplo, em Python/R + ferramentas de visualização) que permita ao poder público acompanhar, em tempo quase real, a evolução do desemprego e de cenários projetados. Isso fortaleceria o uso da ferramenta como suporte operacional contínuo para políticas alinhadas ao ODS 8.

9 Repositório do Projeto

O código e materiais complementares deste trabalho estão disponíveis no repositório público do GitHub, acessível pelo link abaixo:

https://github.com/lattarian/projeto_aplicado_iv

10 Referências Bibliográficas

Referências

- [1] AGÊNCIA GOV. Taxa de desemprego fica em 6,2% em dezembro e média do ano fica em 6,6%, menor patamar da série histórica. Agência Gov, 31 jan. 2025. Disponível em: <https://agenciagov.ebc.com.br/noticias/202501/taxa-de-desocupacao-fica-em-6-2-em-dezembro-e-media-do-ano-fica-em-6-6-menor-p>. Acesso em: 7 set. 2025.
- [2] BRASIL. Desemprego cai a 5,8% no segundo trimestre, menor nível da série histórica. Secretaria de Comunicação Social, 31 jul. 2025. Disponível em: <https://www.gov.br/secom/pt-br/assuntos/noticias/2025/07/desemprego-cai-a-5-8-no-segundo-trimestre-menor-nivel-da-serie-historica>. Acesso em: 7 set. 2025.
- [3] IBGE. PNAD Contínua: Pesquisa Nacional por Amostra de Domicílios Contínua. Rio de Janeiro: IBGE, 2024. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/trabalho/9171-pesquisa-nacional-por-amostra-de-domicilios-continua-trimestral.html>. Acesso em: 7 set. 2025.
- [4] NAÇÕES UNIDAS. Objetivo 8: Trabalho decente e crescimento econômico. ONU Brasil, 2023. Disponível em: <https://brasil.un.org/pt-br/sdgs/8>. Acesso em: 7 set. 2025.
- [5] IBGE. Tabela 4099: Taxas de desocupação e de subutilização da força de trabalho, na semana de referência, das pessoas de 14 anos ou mais de idade. SIDRA – Sistema IBGE de Recuperação Automática, [s. l.], 2025. Disponível em: <https://sidra.ibge.gov.br/Tabela/4099>. Acesso em: 7 set. 2025.
- [6] BARROS, R.; CORSEUIL, C. Desemprego e dinâmica do mercado de trabalho no Brasil. *Revista Brasileira de Economia*, v. 75, n. 3, p. 321–345, 2021.
- [7] BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. *Time Series Analysis: Forecasting and Control*. 5. ed. Hoboken: John Wiley & Sons, 2016.
- [8] CARVALHO, M.; MOREIRA, P. Previsão da taxa de desocupação brasileira com modelos ARIMA. *Economia Aplicada*, v. 24, n. 2, p. 245–266, 2020.
- [9] CHATFIELD, C. *The Analysis of Time Series: An Introduction*. 6. ed. Boca Raton: Chapman & Hall/CRC, 2003.

- [10] CLEVELAND, R. B.; CLEVELAND, W. S.; MCRAE, J. E.; TERPENNING, I. STL: A seasonal-trend decomposition procedure based on Loess. *Journal of Official Statistics*, v. 6, n. 1, p. 3–73, 1990.
- [11] FIRPO, S.; GONZAGA, G.; MONTEIRO, J. Crises econômicas e mercado de trabalho no Brasil. *Estudos Econômicos*, v. 52, n. 1, p. 5–30, 2022.
- [12] HARVEY, A. C. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press, 1990.
- [13] HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: Principles and Practice*. 2. ed. Melbourne: OTexts, 2018.
- [14] ILO. *World Employment and Social Outlook: Trends 2023*. Genebra: International Labour Office, 2023.
- [15] NAÇÕES UNIDAS. ODS 8: Trabalho decente e crescimento econômico. 2023. Disponível em: <https://brasil.un.org/pt-br/sdgs/8>. Acesso em: 7 set. 2025.
- [16] SANTOS, L.; PORSSE, A. Previsão do desemprego regional com aprendizado de máquina. *Revista de Economia Aplicada*, v. 26, n. 1, p. 89–112, 2022.
- [17] SILVA, J.; PORSSE, A. Modelos de espaço de estado para séries de desemprego. *Estudos Econômicos*, v. 51, n. 4, p. 753–780, 2021.